A Framework on Data-Driven Business Process Complexity Analysis for Decision-Making Support

vorgelegt von
M. Sc.
Aleksandra Revina
ORCID: 0000-0002-8405-0018

an der Fakultät VII – Wirtschaft und Management
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktorin der Wirtschaftswissenschaften
- Dr. rer. oec. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Søren Salomo
Gutachter: Prof. Dr. Rüdiger Zarnekow
Gutachterin: Prof. Dr. Vera G. Meister

Tag der wissenschaftlichen Aussprache: 27. Januar 2023

Berlin 2023

Dedicated to my family

Acknowledgments

This dissertation was started during my employment as a research assistant at the Telekom Innovation Laboratories (T-Labs), a private-public partnership between the Technical University of Berlin and Deutsche Telekom AG, at the end of 2017. First and foremost, I would like to thank those who inspired and made this project possible. My deepest gratitude goes to my academic supervisors, Prof. Dr. Rüdiger Zarnekow and Prof. Dr. Vera G. Meister, who constructively guided me through all the years of my doctoral studies providing space for freedom and creativity. Thank you for your trust in my personality and skills and your support in all the questions and problems I encountered during my studies.

My warmest thank you goes to Dr. Nina Rizun, who shared a wealth of knowledge and motivation and could find encouraging words at every step of this challenging way. Further, I would like to thank my T-Labs supervisors, Dr.-Ing. Michal Dunaj and Dr. Heiko Lehmann, who encouraged this work and gave me all the needed practical guidance and network at the starting point. My thanks also go to the Change Management team of Slovak Telekom, which provided a case study context and necessary ground to begin my research project.

In 2020, I continued my doctoral research as a part of the academic staff at the Brandenburg University of Applied Sciences. I would like to express my sincere thanks to Prof. Dr. Vera G. Meister for enabling and allowing me to make this move and for providing the necessary resources to continue my research.

A special thank you goes to all my co-authors, who collaborated with me and contributed their time and expertise to the project. Nina, thank you for a successful project kick-off, research topics elaboration, and your genuine enthusiasm and energy you always gratefully share. Krisztian, I thank you for your reliability and the best co-working experience. Ünal, thank you for giving me an opportunity to learn from you, for your time and patience, as well as for guiding me in Seville and my whole research.

I am grateful to my colleagues at T-Labs, Jörn, Elmar, Elena, Alexandra, and Brandenburg University of Applied Sciences, Dennis, Wenxin, and Irena, who supported my project, shared their knowledge, and made my experience more delightful. I am also thankful to my friends, Natalia, Anna, Elena, Regina, Victoria, Max, and Sergey, for their willingness to listen. Your advice has always given me the strength to finalize the research.

Finally, I express my gratefulness to my family and parents. Я благодарю мою семью и родителей, Ирину и Сергея Ревиных, за воспитание и образование, которое Вы мне дали. Я безгранично благодарна Вам за все свои достижения.

Aleksandra Revina,
Berlin, September 2022

Contents

Acknowledgments **iii**

List of Figures **ix**

List of Tables **xi**

List of Acronyms **xiv**

Summary **xv**

Zusammenfassung **xvii**

1 Introduction **1**
1.1 Research Motivation and Context 3
1.2 Research Design 7
1.2.1 Research Methodology 7
1.2.2 Research Questions 10
1.2.3 Research Methods and Techniques 11
1.3 Research Contributions 13
1.4 Publications 15
1.5 Dissertation Outline 18
1.6 Preliminaries 19

2 Comprehensive Literature Review on Complexity in Organizations **25**
2.1 Introduction 28
2.2 Related Work 30
2.3 Research Approach 30
2.3.1 Systematic Literature Review 31
2.3.2 Literature Classification 34
2.3.3 Goal Question Metric 36
2.4 Systematic Literature Review 36
2.4.1 Complexity Concepts 37
2.4.2 Complexity Metrics and Analysis 38
2.4.3 Complexity Research Motivations and Novelty 39
2.4.4 Complexity Research Focus Areas and Application Cases . . . 41
2.4.5 Complexity Research Future Research Directions 41

2.5 Literature Classification 42
2.5.1 Morphological Box 42
2.5.2 Integrated Multi-Dimensional Complexity Framework 44
2.6 Method to Address Complexity 46
2.7 Discussion 49
2.7.1 Demand for an Extensive Literature Analysis 49
2.7.2 Demand for a Structured Overview 49
2.7.3 Demand for an Integration of Findings 50
2.7.4 Demand for a Practical Guidance 50
2.8 Conclusions and Future Work 51
2.9 Appendix A. List of Papers Analyzed in This Study 52

3 Unveiling the Potential of Textual Data for Business Process Complexity Analysis 69
3.1 Introduction 72
3.2 Related Work 74
3.2.1 Business Process Complexity 74
3.2.2 Research Closely Related to Business Process Complexity . . . 75
3.2.3 Business Process Management and Natural Language Processing 79
3.3 Theoretical and Practical Background 80
3.3.1 Conceptual Model Development Based on the Theory of Situation Awareness 80
3.3.2 ITIL Change Management IT Ticket Processing Motivating Example 82
3.4 Concept of Business Process Complexity 82
3.4.1 Definition 1: Objective Knowledge 82
3.4.2 Definition 2: Subjective Knowledge 83
3.4.3 Definition 3: Meta-knowledge 85
3.4.4 Definition 4: Business Process Complexity 86
3.5 Illustrative Application 87
3.5.1 Step 1: Data Collection and Preprocessing 87
3.5.2 Step 2: Objective Knowledge Extraction 87
3.5.3 Step 3: Subjective Knowledge Extraction 88
3.5.4 Step 4: Meta-knowledge Extraction 88
3.5.5 Step 5: BP Complexity Identification 89
3.6 Evaluation 92
3.7 Discussion, Contribution, and Limitations 95
3.8 Conclusion and Future Work 96

4 Predicting Business Process Complexity Using Textual Data 105
4.1 Introduction 108
4.2 Related Work 109
4.2.1 Text Representation 109
4.2.2 Text and Ticket Classification 110
4.3 Methods 111

4.3.1 Feature Extraction 111
4.3.2 SUCCESS 113
4.3.3 Other Classifiers 114
4.4 Experimental Evaluation 114
4.4.1 Datasets 114
4.4.2 Experimental Settings 115
4.4.3 Comparison of SUCCESS and QuickSUCCESS 115
4.4.4 Results 115
4.5 Discussion 117
4.6 Conclusion 118
4.7 Appendix I. Taxonomy of Decision-Making Logic Levels 119
4.8 Appendix II. Business Sentiment Lexicon with Assigned Valences . . 120

5 Enriching Textual Data-Driven Business Process Complexity Analysis with Event Log-Driven Analysis 125
5.1 Introduction 128
5.2 Related work 130
5.3 Background 131
5.3.1 Linguistic features 131
5.3.2 Event log complexity metrics 132
5.4 Approach development 133
5.4.1 Inputs 133
5.4.2 Running example 134
5.4.3 Calculating textual data-based complexity 135
5.4.4 Calculating event log-based complexity 135
5.4.5 Correlation analysis 137
5.4.6 Significant difference analysis of event logs 138
5.5 Case Study 138
5.5.1 Data collection 139
5.5.2 Calculating textual data-based complexity 139
5.5.3 Calculating event log-based complexity 140
5.5.4 Analyzing correlations 140
5.5.5 Analyzing significant differences in business process executions 143
5.6 Discussion 145
5.6.1 Analyzing results and deriving observations 145
5.6.2 Outlining benefits and limitations 147
5.7 Conclusion and future work 148
5.8 Appendix A. DML Taxonomy 148

6 Research Agenda and Future Work 151
6.1 Introduction 154
6.2 Traditional BPM Logics and Their Problematization 155
6.2.1 Business Processes 155
6.2.2 IT Infrastructure 156
6.2.3 Process Actors 156
6.3 Framework Exploring Integrated Data Perspective in BPM 157

6.3.1 Foundations for RQ1 159
6.3.2 Foundations for RQ2 159
6.3.3 Foundations for RQ3 160
6.3.4 Foundations for RQ4 160
6.4 Conclusions 161

7 Conclusion **167**
7.1 Contributions and Validation 169
7.2 Implications 172
7.2.1 Implications for Practice 172
7.2.2 Implications for Research 179
7.3 Limitations and Open Issues 181
7.4 Future Work 182

Appendix A - Evaluation Questionnaire **185**

Bibliography **189**

Curriculum Vitae **199**

List of Figures

1.1.1 Research area and framework positioning 7
1.3.1 An approach for analyzing business process execution complexity based on textual data and event log 14

2.4.1 Publication trend per complexity type 37
2.4.2 Complexity concepts mind map 37
2.4.3 Metrics origin distribution per complexity type 38
2.4.4 Types of inputs per complexity type 39
2.4.5 Types of outputs per complexity type 39
2.4.6 Motivations over time per complexity type 40
2.4.7 Motivations in relation to metrics origin per complexity type 40
2.4.8 Novelty per complexity type 40
2.4.9 Focus areas per complexity type 41
2.4.10 Application cases per complexity type 41
2.5.1 Future research directions over time per complexity type 42

3.3.1 Extended situation awareness model of BP complexity 81
3.4.1 Objective knowledge extraction and DML level identification 84
3.4.2 Subjective knowledge extraction and BS identification 84
3.4.3 Meta-knowledge extraction and readability identification 86
3.5.1 BP complexity for decision support 87
3.6.1 Evaluation process 92

4.4.1 When sampling the data, the size r of the sample should be chosen carefully so that the sampled data is representative 114
4.4.2 Overview of the experimental pipeline 115
4.6.1 kNN performance with various k values (in the left, using Euclidean distance) and distance functions (in the right, k=1) on Data2 . . . 118

5.1.1 Our approach for analyzing business process execution complexity based on textual data and event log 129
5.4.1 Textual data-based complexity calculation 135
5.4.2 Event log-based complexity calculation 136
5.4.3 Determining event log-based complexity of an event log subset . . 137

5.5.1 Statistically significant differences between SR001 (with B) and SR002 (with R) 138
5.5.2 Textual data-based complexity calculation by means of prediction . 140
5.5.3 Correlations between textual data- and event log-based complexities 143
5.5.4 Significant differences between more interactive channels (with B) and less interactive channels (with R) 144
5.5.5 Significant differences betweencategories 3, 4, 5, and 7 (with B) and others (with R) 144
5.6.1 Significant differences between duration longer than 10 days (with B) and up to 10 days (with R) 145
5.6.2 Significant differences between organizational unit2 (with B) and others (with R) 146

6.3.1 Framework and research agenda of integrated process data perspective in BPM 157

List of Tables

1.2.1 Overview of the research questions, methods, techniques, and target qualities 11
1.6.1 Example of an event log of an incident management process 21

2.3.1 Inclusion and exclusion criteria 34
2.3.2 Identified aspects and features for morphological box 35
2.3.3 Complexity dimensions and notions 35
2.5.1 Morphological box describing three complexity aspects 43
2.6.1 Relative distribution of the complexity notions per complexity type 46
2.6.2 An example illustrates the use of the proposed method 48
2.9.1 Analyzed papers in this study 53

3.2.1 Related work review of BP complexity 76
3.5.1 Threshold rules of objective knowledge extraction (DML) 88
3.5.2 Threshold rules of subjective knowledge extraction (BS) 89
3.5.3 Threshold rules of meta-knowledge extraction (readability) 89
3.5.4 Expert-defined decision rules for BP complexity identification . . . 90
3.5.5 BP complexity concept application on the motivating example . . . 90
3.5.6 Generalized summary of the five-step application scenario of BP complexity prediction 91
3.6.1 Evaluation results of IT ticket complexity prediction 93
3.6.2 Distribution statistics of DML, BS, readability and IT ticket complexity 94

4.2.1 Text representation techniques 110
4.3.1 Text classification techniques 111
4.3.2 Overview of linguistic features illustrated by a ticket example . . . 112
4.4.1 Datasets 114
4.4.2 Execution time (in seconds) and accuracy of SUCCESS and Quick-SUCCESS 115
4.4.3 Evaluation results using linguistic features and TF-IDF 116
4.5.1 Evaluation results using the five best performing linguistic features 117
4.5.2 Evaluation results using linguistic features without "cognitive words" feature 117
4.6.1 Evaluation results using expert decision rule 118
4.7.1 Taxonomy of Decision-Making Logic levels 119

4.8.1 Business Sentiment lexicon with assigned valences 120

5.3.1 Linguistic features 132
5.4.1 An example of an event log of a Service Request Management process 133
5.4.2 Metrics for event log-based complexity calculation 134
5.4.3 Running example IT tickets 134
5.4.4 Textual data-based complexity calculation of the running example IT tickets 136
5.4.5 Event log-based complexity calculation of the running example IT tickets 137
5.5.1 Textual data- and event log-based complexities of the running example IT tickets 138
5.5.2 Identified ticket attributes 139
5.5.3 LIWC features 141
5.5.4 Evaluation of prediction models for textual data-based complexity 142
5.5.5 Correlations between textual data- and event log-based complexities 142
5.8.1 Decision-Making Logic taxonomy for Service Request Management process 148

7.2.1 BPM case-1 175
7.2.2 BPM case-2 176
7.2.3 BPM case-3 177
7.2.4 BPM case-4 178

List of Acronyms

AI	Artificial Intelligence
BI	Business Intelligence
BoW	Bag-of-Words
BP	Business Process
BPM	Business Process Management
BPMN	Business Process Model and Notation
BS	Business Sentiment
BSC	Balanced Score Card
CAB	Change Advisory Board
CART	Classification and Regression Tree
CHM	Change Management
CI	Configuration Item
CMMN	Case Management Model and Notation
COG	Cognitive Words feature
CRM	Customer Relationship Management
DCR	Dynamic Condition Response
DML	Decision-Making Logic
DMN	Decision Model and Notation
DMS	Document Management System
DSS	Decision Support System
EL	Event Log-Based
ERP	Enterprise Resource Planning
GAM	Goal Argument Metric
GQM	Goal Question Metric
GUI	Graphical User Interface
HPC	High-Performance Computing
HR	Human Resources
IoT	Internet of Things
IS	Information System
IT	Information Technology
ITIL	Information Technology Infrastructure Library
ITSM	Information Technology Service Management
KIP	Knowledge-Intensive Process

KM	Knowledge Management
kNN	K-Nearest Neighbor algorithm
KPI	Key Performance Indicator
LDA	Latent Dirichlet Allocation
LIWC	Linguistic Inquiry and Word Count
LSI	Latent Semantic Indexing
ML	Machine Learning
NLP	Natural Language Processing
PAIS	Process-Aware Information System
PM	Process Mining
PoS	Parts-of-Speech
PPT	People Process Technology framework
PSO	Projected Service Outage
Q&A	Questions and Answers
RAMON	Reference And Management Of Nomenclatures
RfC	Request for Change
RPA	Robotic Process Automation
RTCC	Resources Techniques Capacities Choices
SA	Situation Awareness
SLA	Service Level Agreement
SLR	Systematic Literature Review
SRM	Service Request Management
SUCCESS	Semi-sUpervised ClassifiCation of TimE SerieS algorithm
SVM	Support Vector Machines
TD	Textual Data-Based
TTF	Task Technology Fit theory
TF-IDF	Term Frequency-Inverse Document Frequency
TDQM	Total Data Quality Management framework
VADER	Valence Aware Dictionary and sEntiment Reasoner
XAI	Explainable Artificial Intelligence

Summary

Organizations become heavily dependent on diverse information systems to enable their business processes (BPs) due to continuous digitalization and increasing advancements in information technology. Data maintained in the information systems, such as event logs and textual data, grow exponentially, complicating decision-making and creating new challenges for Business Process Management (BPM). To exploit decision-making support opportunities inherent in these data volumes, several techniques are developed and adopted in BPM, for example, process mining and text mining. However, in these and other BP analysis techniques, oftentimes, the consideration of textual data massively generated in BPs is limited to process models and process descriptions. Moreover, these techniques do not offer extensive solutions using both event log and textual data to assist BP actors in making decisions. At the same time, as organizations get more complex, the interest in BP complexity analysis increases with an exclusive focus on a BP model and event log perspective, neglecting textual data generated in BPs. Despite the fact that texts often lack structure, they are believed to have a high business value remaining the most readily available source of knowledge. Accordingly, we identified three main concerns in the current BP complexity analysis approaches: *incomprehensiveness*, *inefficiency*, and *bias*.

The mentioned concerns motivated the research in this dissertation. Specifically, this dissertation focuses on investigating the complexity-related insights that can be derived from textual data created by BP actors in BP execution and the ways to combine such data with an event log. As a result, a framework on data-driven BP complexity analysis for decision-making support is suggested. The main contributions of this dissertation can be summarized in line with a gradual development of the framework, each of the stages addressing particular issues of current BP analysis approaches. Thus, a comprehensive literature review provides an overview of various types of complexities in organizations as well as a method to manage them. Further, linguistic features forming a textual data-based BP complexity are elaborated to show the value of textual data in BP execution and decision-making. Afterward, they are combined with an event log to identify patterns that yield complexity in BP execution and insights one can get from such a synergistic perspective. Finally, a research agenda reveals an extensive potential of an integrated process data perspective, including event log and textual data.

In general, the research conducted in this dissertation aims to draw the attention of the BPM community to the potential and high relevance of textual data generated in BP execution, making important contributions to comprehensive, efficient, and objective BP complexity analysis approaches.

Zusammenfassung

Aufgrund der zunehmenden Digitalisierung und Fortschritte in der Informationstechnologie sind Unternehmen in hohem Maße von verschiedenen Informationssystemen abhängig, um ihre Geschäftsprozesse (GP) zu unterstützen. Die in den Informationssystemen gespeicherten Daten, wie zum Beispiel Ereignisprotokolle und Textdaten, wachsen exponentiell, was die Entscheidungsfindung erschwert und neue Herausforderungen für das Geschäftsprozessmanagement (GPM) schafft. Um die mit diesen Datenmengen verbundenen Möglichkeiten der Entscheidungsunterstützung zu nutzen, wurden verschiedene Techniken entwickelt und im GPM eingesetzt, zum Beispiel Process Mining und Text Mining. Bei diesen und anderen GP-Analysetechniken beschränkt sich die Betrachtung der in GP massiv anfallenden Textdaten jedoch häufig auf Prozessmodelle und Prozessbeschreibungen. Darüber hinaus bieten diese Techniken keine umfassenden Lösungen, die sowohl Ereignisprotokoll- als auch Textdaten nutzen, um GP-Akteure bei der Entscheidungsfindung zu unterstützen. Gleichzeitig steigt mit der zunehmenden Komplexität von Unternehmen das Interesse an der Komplexitätsanalyse von GP, wobei der Fokus ausschließlich auf GP-Modellen und Ereignisprotokollen liegt und die in GP generierten Textdaten vernachlässigt werden. Trotz der Tatsache, dass Texte oft unstrukturiert sind, wird ihnen ein hoher Geschäftswert zugeschrieben, da sie die am leichtesten verfügbare Wissensquelle darstellen. Dementsprechend haben wir drei Hauptprobleme in den aktuellen Ansätzen zur Komplexitätsanalyse von GP identifiziert: *Unvollständigkeit*, *Ineffizienz* und *Bias*.

Die genannten Probleme motivierten die Forschung in dieser Dissertation. Konkret geht es in dieser Dissertation um die Untersuchung der komplexitätsbezogenen Erkenntnisse, die aus den von GP-Akteuren bei der GP-Ausführung erzeugten Textdaten abgeleitet werden können, sowie um die Möglichkeiten, solche Daten mit dem Ereignisprotokoll zu kombinieren. Als Ergebnis wird ein Framework zur datengesteuerten Komplexitätsanalyse von GP zur Entscheidungsunterstützung vorgeschlagen. Die wichtigsten Beiträge dieser Dissertation lassen sich entsprechend einer schrittweisen Entwicklung des Frameworks zusammenfassen, wobei jede der Phasen bestimmte Probleme aktueller GP-Analyseansätze adressiert. So bietet eine umfassende Literaturübersicht einen Überblick über verschiedene Arten von Komplexität in Organisationen sowie eine Methode zu deren Bewältigung. Des Weiteren werden linguistische Features, die eine auf Textdaten basierende GP-Komplexität abbilden, ausgearbeitet, um den Wert von Textdaten bei der GP-Ausführung und Entscheidungsfindung aufzuzeigen. Anschließend werden sie mit Ereignisprotokollen kombiniert, um Muster zu identifizieren, die Komplexität in der GP-Ausführung anzeigen, und um darzustellen, welche Erkenntnisse man aus einer solchen synergetischen Perspektive gewinnen kann. Abschließend wird eine Forschungsagenda aufgestellt, die das umfangreiche Potenzial einer integrierten Prozessdatenperspektive, einschließlich Ereignisprotokoll und Textdaten, darstellt.

Im Allgemeinen zielt die in dieser Dissertation durchgeführte Forschung darauf ab, die Aufmerksamkeit der GPM-Community auf das Potenzial und die hohe Relevanz von Textdaten zu lenken und einen wichtigen Beitrag zu umfassenden, effizienten und objektiven GP-Komplexitätsanalyseansätzen zu leisten.

Titel der Dissertation übersetzt auf Deutsch: Ein Rahmen für die datengestützte Analyse der Komplexität von Geschäftsprozessen zur Unterstützung der Entscheidungsfindung

CHAPTER 1

Introduction

Business processes (BPs) are a critical component of any organization. They are fairly considered arterial systems of organizations. Errors in BPs can cause the collapse of the entire ecosystem of an organization [1]. For example, after an erroneous software deployment process performed by a software engineer and a wrong risk assessment, Microsoft Azure faced a collective 4.46 billion hours of cloud service outage (10.5 hours of downtime affected 425 million customers) [2]. A faulty rail franchise bidding risk assessment process cost UK taxpayers 40 million pounds [3]. Similarly, human error, the top root cause of information technology (IT)-related problems [4], accounted for the IT downtime affecting 75,000 passengers of British Airways [5]. These are just a few examples of process flaws leading to far-reaching consequences, not to mention the daily losses that occur in nearly every organization due to poorly designed and/or supported processes [1, 6].

Business Process Management (BPM) is an established discipline containing the best practices to ensure successful and sustainable businesses and avoid errors. These best practices target fundamental activities, such as documenting, monitoring, improving, and managing processes [1]. In this regard, improvements are guided by context-specific organizational objectives and include reducing complexity, costs, human errors, execution time and gaining competitive advantage through innovation, to name a few [1].

To support efficient BP execution and business operations, today's organizations, both private and governmental, increasingly rely on information systems (IS) in their IT landscape. For example, one of the most commonly used IS for enabling production, distribution, and finance processes is an Enterprise Resource Planning (ERP) system. Customer Relationship Management (CRM) systems providing a key value to customers are another example of the widely used IS. Moreover, contemporary IS combine a wide range of functionalities to provide comprehensive support for BPs in organizations. Likewise, modern Document Management Systems (DMS) are used not only to digitize, organize, and store different kinds of documents but also to manage roles, access rights, and privacy, automate document creation and workflows, as well as facilitate employee management [7]. Latest IT Ticket Management systems, in addition to diverse IT service-specific features like Incident, Problem, Release or Change Management [8], also provide workflow management support, reporting and analysis for informed decision-making, service catalogs, and multiple channel communication support. Such a variety of features and wide coverage allows IS to record vast amounts of data, including event log, i.e., BP execution history, and textual data massively generated by BP actors and making up more than 80% of data in organizations [9].

1.1 Research Motivation and Context

BP complexity in terms of its execution and complexity metrics based on event log data have recently received considerable attention in BPM [10]. Despite the large amounts of textual data generated and stored in IS, there is limited research on textual data in the studies focusing on BP executions and their complexity [11]. These studies mainly consider BP descriptions, documentation, and texts included in BP models, such as labels [11]. The use of natural language processing (NLP) in BPM also presents a number of open problems, including the need for semantic improvements and domain- or organization-specific NLP modifications. Therefore, a closer relationship between these two subject areas reveals an unrealized opportunity to significantly enhance the BPM toolkit.

Textual data serving as an input to BPs have a significant impact on how they are carried out. For example, due to the pandemics, numerous organizations, like the governmental or health sector, are progressively offering various digital interfaces to remotely collect their customers' requests, such as requests for issuing a new passport or detailing a health concern. The speed and quality of request processing are heavily influenced by the clarity and completeness of textual descriptions of these requests [12]. Likewise, in the IT Service Management (ITSM) area, processing customer requests is highly dependent on textual descriptions provided by customers, especially setting the priority, urgency, and severity of requests, as well as defining the roles and teams to be involved in these requests [12].

As humans are limited in their ability to deal with large amounts of data, data deluge in BPs naturally brings complexity to any decision-making. Decision-making is defined as the process of choosing from a variety of options considering the potential outcomes and consequences of each option [13]. Accordingly, with the growing complexity of products, services, processes, internal and external structures, and, hence, the number of factors influencing the success and sustainability of businesses, decision-making becomes more sophisticated. Therefore, ever-increasing data volumes, specifically those generated by IS in the form of event log and textual data created by BP actors, open new opportunities for decision-making support. Further, texts generated by human actors in BPs are believed to have a high business value [11, 14, 15], despite the fact that such texts often lack structure. Although business-relevant knowledge can be obtained from various sources, unstructured texts remain the most readily available source of knowledge [16]. In this context, for the data-driven analysis of BPs, several techniques are developed and adopted in the BPM field, for example, process mining and text mining. Below, we describe prominent BP analysis approaches. While doing so, we put a special emphasis on textual data usage and decision-making support in the approaches. Moreover, we group the issues that we identified in them. In a broader scope, we also aim to set the research context.

BPM is operationalized in a model called the BPM lifecycle that is based on the following six phases: process identification, process discovery, process analysis, process redesign, process implementation, and process monitoring and controlling [17]. However, the central theme and, at the same time, limitation of the model as well as the BPM discipline in general is a strong focus on BP modeling support. Following the classical BPM logic, BPs should be thoroughly modeled, and it is unacceptable if an activity cannot be modeled [18]. Accordingly, process analysis, one of the important BPM lifecycle phases, primarily aims at BP modeling support. It encompasses various techniques to determine BP weaknesses and the ways to improve them.

BPM commonly differentiates between qualitative and quantitative BP analysis approaches. Typical examples of qualitative analyses are value-added, waste, and root cause analyses. Whereas the value-added approach focuses on the value-adding activities in BPs, waste analysis addresses the identification of activities negatively influencing BPs. Root cause analysis finds and interprets the underlying causes of problems and unfavorable outcomes [1]. Quantitative analyses comprise a set of techniques to gain systematic insights into a process, such as flow and queue analyses and simulation. Flow analysis bases on the quantitative BP characteristics like execution time of activities or number of decision gateways. One of the important approaches of flow analysis continuously gaining the attention of the research community is complexity analysis mainly focusing on BP models and event log [10]. Queue analysis is built on the queuing theory representing a set of mathematical tools to analyze systems with resource contention leading to queues. Finally, process simulation is the most extensively used approach in terms of a quantitative analysis of process models. In this regard, a process simulator is used to generate a large number of hypothetical instances of a process [1]. The described analyses rely on diverse data sources and techniques to collect the data, for example, data from questionnaires and interviews, expert knowledge, BP models and descriptions, as well as event log captured by IS.

Despite this variety of BP analysis techniques, they do not provide comprehensive solutions employing event log and textual data captured in BP execution to address decision-making support for BP actors. Further, with the proliferating complexity in businesses, we have seen an increasing interest in BP complexity analysis from the BP model and event log perspective [10] that neglects textual data generated in BPs. Based on the mentioned observations, we recognize the following concerns serving as a motivation for the present dissertation.

Concerns

1. **incomprehensiveness**: current data-driven approaches prevailingly focus on analyzing one data type, i.e., either event log or textual data, to suggest various decision-making support solutions and do not offer synergistic ways to combine both types in one approach. For instance, [19] investigate how to leverage textual information for improving decision-making in the BPM lifecycle. Similarly, recent work [20] studies the usage of process mining and event log analysis for better business strategy decision-making. However, these efforts are separate work streams. Further, the consideration of textual data in the scope of BP analysis is mostly limited to BP textual descriptions and BP models [11], disregarding abundant communication data produced by BP actors and containing information highly relevant for decision-making. This leads to incomprehensive analysis results that might miss important insights.
2. **inefficiency**: emerging technology, changing organizational structures, and dynamic work environments provide new opportunities to organizations while bringing challenges. Such developments, when combined with the increasing volume and variety of data generated and exchanged, mostly result in complexity. In organizations, complexity can be caused by a myriad of factors. In line with the dominant BP modeling logic, BPM complexity analysis approaches are mainly focused on the complexity of BP models and event log [10]. Considering the fact that 80% of data in organizations are textual and may contain important information regarding complexity [9], their neglection is at least inefficient.
3. **bias**: current BP analysis approaches are strongly focused on BP modeling support. Modeling makes up the dominant way of thinking and the main logic regarding the processes in BPM [21–23]. By relying on this BP logic, BPM enables BP owners,

actors, and stakeholders to redesign BPs. Accordingly, as much work performed in organizations as possible should be represented as explicit knowledge providing the basis for BP model creation, analysis, and redesign [24]. In those cases when the work bases on tacit knowledge, that part should be systematically modeled to be further analyzed. Even though such knowledge is not formalized, it still makes up an important part of daily work [24]. Naturally, not being able to consider all that knowledge generated in BPs and thus catch up with the BP dynamics, such a biased approach is known to cause multiple BPM project failures [18]. Moreover, the support for BP actors in making decisions while executing a particular BP is not addressed precisely.

To address the mentioned concerns, below we detail the target qualities acting as a basis for this dissertation:

Target qualities

1. **comprehensiveness**: combining event log and textual data generated in BPs in one approach provides more opportunities for performing comprehensive analysis and getting valuable BP insights. For example, insights obtained from textual data can complement the event log-based insights or contradict them, the latter demanding closer consideration. Such a comprehensive view enables a more accurate analysis and better BP improvement recommendations.
2. **efficiency**: BP complexity analysis focusing on BP models and event log lacks important complexity-relevant knowledge that can be extracted from textual data. For example, certain semantic, syntactic, and stylistic characteristics of textual data can increase the efficiency of BP complexity analysis. Moreover, such aspects of textual data have a great potential to improve the quality of BP complexity predictions. Based on the latter, other relevant estimations can be performed, such as effort estimation, prioritization, and work assignment. Only an inclusive consideration of data sources can enable accurate identification of complexity factors as well as the ways to deal with complexity in the form of various efficient decision-making support solutions for BP actors.
3. **objectivity**: in reality, BP models become outdated fast due to the business dynamics and workarounds of BP actors attempting to simplify their work. Limited focus on BP modeling in BP analyses does not yield significant decision-making support to BP actors. Hence, extending the current BP analysis approaches with the solutions aimed at BP actors' decision-making support can complement the BPM toolset synergistically and alleviate those problems triggered by the BP modeling focus.

In Figure 1.1.1, we present a BPM lifecycle and, in particular, its process analysis phase extended with our data-driven complexity analysis framework. This way, we position the framework in a broader research context.

The remainder of Section 1 is structured as follows. Section 1.2 describes the methodological approach selected to conduct the work in this dissertation. Section 1.3 lists the main contributions. Section 1.4 presents a complete list of publications resulted from this research and beyond. Section 1.5 provides an overview of the remaining chapters, i.e., publications presented in the scope of this dissertation. Fi-

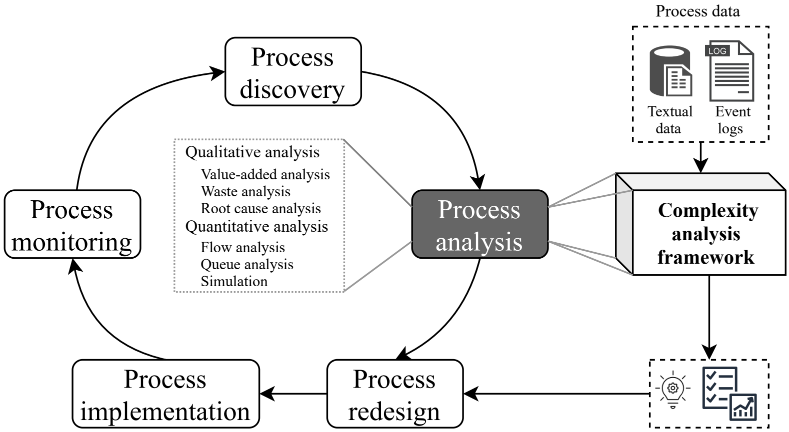


Figure 1.1.1: Research area and framework positioning

nally, in Section 1.6, we give the background information necessary to understand the research artifact, related subject fields and terms.

1.2 Research Design

In this section, the overall research methodology, guiding research questions, and specific research methods and techniques are presented.

1.2.1 Research Methodology

The research presented in this dissertation belongs to the IS research field. IS research seeks to advance the knowledge that assists in the effective use of IT in organizations [25]. Similar to BPM, IS research is a highly multidisciplinary field incorporating theories and knowledge from Social Sciences, Economics, and Computer Science. Accordingly, there are two closely interrelated but distinct perspectives in IS research: behavioral science and design science [25, 26].

Taking its roots in natural sciences, *behavioral science* sets the goal to create and test theories interpreting or predicting behavior in organizations. It starts with a hypothesis. Afterward, researchers gather data to either prove or reject the hypothesis, which eventually leads to theory building [26]. Having its origin in engineering and the sciences of artificial [25], *design science* is, on the contrary, a problem-solving paradigm with the objective to develop and evaluate a research artifact. An innovative concept, process, IT capability, or product enabling efficient IS design, implementation, administration, and application can constitute a research artifact [27]. Design science research is especially relevant for so-called *wicked* problems implying unstable requirements and constraints based upon poorly defined contexts [25]. That's why it is not possible to assess solutions to such wicked problems based on theories but based on their utility. However, the natural rules and behavioral theories obviously apply to these research artifacts. Their development is based on existing basic theories that are used, verified, improved, and enhanced with the help of the researcher's expertise, creativity, intelligence, and problem-solving skills [28]. As these perspectives are complementary, we use both of them to conduct research in this dissertation. Hereby, design science takes up a major role in our methodology. Behavioral science complements it specifically in the development and justification of the hypotheses behind the way people, i.e., BP actors, express their thoughts and communicate in BPs and the role of textual data in building the BP actors' awareness.

In this regard, [25] propose a set of seven guidelines for effective IS research, which are especially relevant for design science-focused studies. These guidelines serve as a methodological foundation for the work in this dissertation. Below, we explain how the seven guidelines are addressed:

1. Design as an artifact: The goal of design science research is the creation of artifacts improving organizational performance. Such artifacts should be presented comprehensively to be implemented and used in relevant domains. A construct, model, method, framework, or their adaptations can all be considered artifacts [25].

In this dissertation, we address the problem that BP analysis approaches neglect the potential of textual data massively generated by BP actors. These approaches show incomprehensiveness and bias as they focus on textual data contained in BP models, BP descriptions, and event log. As a result, the suggested decision-making support solutions mainly limited to BP modeling activities lack efficiency as they miss that large amount of knowledge inherent in the textual data created in the BP execution. As a response to these concerns, we gradually develop a framework on BP analysis from a BP complexity perspective to provide decision-making support for BP actors, for example, in the form of effort estimation or prioritization. Hereby, we base our approach on both types of data, i.e., textual data generated by BP actors and event log generated by IS, addressing *incomprehensiveness*, *inefficiency*, and, consequently, *bias* of current approaches. We also use a real-life setting to demonstrate the applicability of the framework.

2. Problem relevance: Addressing challenges that are important to a certain community determines the relevance of design science research. Professionals who plan, develop, implement, and manage IS technologies constitute this community for IS research [25].

The relevance of the work reported in this dissertation arises from the popularity and extensive practice of BPM in organizations on the one hand [17] and high rates of BPM project failures on the other [18]. Hence, BPM research aiming to address the existing issues and facilitate successful BPM practice in organizations is relevant for both BPM practitioners and researchers. In particular, the relevance of concerns addressed in this dissertation has been derived from the literature and limitations of the state-of-the-art solutions. Subsequently, in publications [12] in Chapter 3 and [29] in Chapter 5, we elaborate on the importance and relevance of textual data consideration in BPM. Further, we highlight related shortcomings of current approaches in general [11] and specifically in terms of BPM complexity analysis [10].

3. Design evaluation: The usability, quality, and efficiency of a design artifact must be thoroughly proven by means of consistent evaluation methods. The requirements for artifact evaluation are determined by its business context. It demands an establishment of relevant metrics and collection as well as analysis of relevant data [25].

The research contributions of this dissertation are developed and evaluated using two case studies: (1) an industrial case study [12] in Chapter 3 and [30] in Chapter 4 and (2) an academic case study [29] in Chapter 5. In separate conference papers [31–33] merged into one journal publication, i.e., already mentioned [12] in Chapter 3, we exhaustively investigate the possibilities of textual data in the BP complexity analysis context. We explore and develop a set of linguistic features that may impact BP complexity. In an industrial case study of a Change Management (CHM) IT ticket processing [30] in Chapter 4, we select those linguistic features advantageous for BP complexity prediction. In [29] in Chapter 5, we combine textual data in the form of selected linguistic features and event log. Hereby, another case study of Service Request IT ticket processing from academia is used. The methods and techniques including qualitative and quantitative approaches we used in this research are described in Section 1.2.3 following below.

4. Research contributions: Design science research should contribute to the knowledge base in an innovative, meaningful, and broad way. This excludes routine design defined by a mere application of existing knowledge and practices. Scientific design, on the contrary, implies the use of novel artifacts to tackle unsolved issues or improve current solutions [25].

The research contributions of this dissertation are addressed in the clearly stated research questions, methodology, as well as continuous comparison with the knowledge base. We deal with the unsolved issues and provide improvements to the existing solutions. Thus, the introduced framework can be considered the first approach combining two data types for BP complexity analysis. Herewith, we also extend the analysis of textual data in BPM with the data generated by BP actors in the BP execution enabling enriched decision-making support. We outline our contributions in Section 1.3 in detail.

5. Research rigor: In the development and evaluation of artifacts, design science demands the use of rigorous techniques. Hereby, rigor is defined as an appropriate application of knowledge base, including theoretical foundations and research methodologies. The researcher's ability to select suitable foundations and methodologies as well as evaluation strategy is critical to research success [25].

In the development of research artifacts, we use the knowledge base from diverse research fields, including theories and practical instruments. For instance, while developing the set of linguistic features, we take as a basis linguistic theoretical assumptions, such as the theory of least effort [32, 34]. When justifying the application of linguistic features for decision-making support of a BP worker, we refer to the Theory of Situational Awareness [12]. Additionally, we apply standard practical instruments from NLP, process mining, event log analysis, and machine learning. For example, we analyze textual data using common NLP-related techniques, such as topic modeling [35], sentiment analysis [33], and taxonomies, adjusting them to the research objectives. When performing the complexity analysis of event log, we build on the approaches described in [10]. Furthermore, we formalize the designed artifacts so that they can be reproduced by other researchers [29].

6. Design as search process: In design science, problem-solving is referred to appropriate techniques to achieve desired outcomes while considering the rules of the environment, which is a standard iterative develop/test cycle [36]. However, due to the wicked nature of IS problems, it is not possible to specify all conditions, methods, purposes, and rules related to the problem [37]. As a result, design science research frequently seeks a satisfying and generalizing solution without having to explore all options [36].

To a large extent, the research presented in this dissertation deals with textual data generated by BP actors in the BP execution. Hence, it becomes impossible to consider all the aspects of the problem due to the versatility of natural language. Theoretically, one can define an infinite number of approaches to investigate the issues that our work focuses on. Under such conditions, we aim to create a solid solution and draw the attention of the BPM community to the potential and high relevance of textual data. As natural language is characterized by high diversity, variability, and multidimensionality, it is hard to create a solution precisely capturing all the problem-relevant aspects inherent in natural language. Accordingly, our framework aims to capture a significant portion of knowledge focusing on complexity and decision-making and shows the importance of textual data consideration in this context.

7. Communication of research: Design science research must address both technical and managerial audiences [25]. The technical audience requires the necessary details to be able to deploy the designed artifacts. The managerial audience needs an adequate level of information to assess organizational resources that should be devoted to the management and usage of the artifact. Further, the problem relevance as well as the solution novelty and efficacy should be emphasized for the managerial audience [38].

The works resulted from the research conducted in the scope of this dissertation have been published in peer-reviewed international academic journals with high impact factor and at renowned international conferences (see Section 1.4). Furthermore, we complemented the textual data analysis with open source codes and supplementary documentation on Github[1]. Interested scholars and practitioners may easily access and comprehend the implemented approaches.

Overall, based on the above-discussed design science research criteria, we can conclude that our research meets internationally acknowledged research standards and makes an important contribution to the IS body of knowledge.

1.2.2 Research Questions

Using the research methodology presented in Section 1.2.1 and based on the discussion of concerns (incomprehensiveness, inefficiency, and bias) and target qualities (comprehensiveness, efficiency, and objectivity) in Section 1.1, the main research question (MRQ) of this dissertation is stated as follows:

MRQ: *How can process-related data be combined in a comprehensive, efficient, and objective way to predict BP complexity and provide decision-making support to BP actors?*

In line with the main research question, we derive five sub-research questions (RQ):

RQ1: *What concepts and measures of complexity are available in the literature?*

[1] https://github.com/IT-Tickets-Text-Analytics

RQ2: *How can a novel BP complexity concept be defined for efficiently measuring BP complexity based on textual data serving as an input to BP?*

RQ3: *How can the BP complexity concept introduced in the answer to RQ2 be employed objectively and efficiently to predict BP complexity?*

RQ4: *How can the BP complexity concept introduced in the answer to RQ2 be enhanced with BP execution data, i.e., event log, for a comprehensive and objective BP complexity analysis?*

RQ5: *How can such an integrated process data perspective addressed in RQ4 impact BPM research and practice?*

1.2.3 Research Methods and Techniques

To deal with the research questions presented above, in this dissertation, a framework on data-driven business process complexity analysis has been gradually developed based on the fundamental design science research guidelines and a number of specific research methods and techniques. These are literature review, conceptual modeling, and case study including workshops, questionnaires, and expert interviews, which we explain below. In Table 1.2.1, we provide an overview of this dissertation research in the form of the mapping between research questions, chapters, research methods and techniques, as well as addressed target qualities. The column with addressed target qualities displays the target quality in the primary focus in black color and target quality in the secondary focus in gray color.

Table 1.2.1: Overview of the research questions, methods, techniques, and target qualities

Research questions	**Chapters**	**Research methods and techniques**						**Addressed target qualities**
		Literature review	*Conceptual modeling*	*Case study*	*Workshop*	*Questionnaire*	*Expert interview*	
MRQ	all	+	+	+	+		+	all
RQ1	2	+						comprehensiveness
RQ2	3	+	+	+	+	+	+	efficiency
RQ3	4		+	+		+		objectivity efficiency
RQ4	5		+	+			+	comprehensiveness objectivity
RQ5	6	+						comprehensiveness

Literature Review

A literature review is an important research method to use when conducting research [39–41]. It identifies relevant theories, methods, and gaps in any discipline, laying a solid ground for knowledge advancement [42]. Accordingly, every literature review must be performed in a thorough and rigorous way to gather and evaluate the relevant information. Hereby, following a systematic approach ensures trustworthy results. The steps in such a process include defining research questions serving as guidance for the literature review, creating search criteria, searching, and evaluating the collected works to extract knowledge [39–41].

In this dissertation, the guidelines outlined in [40] are followed to acquire the necessary knowledge and provide the foundation for the research. The collected relevant information is discussed in separate literature review articles as well as in the related work and background sections of the corresponding publications presented in this dissertation.

Conceptual Modeling

Conceptual modeling is a method that is frequently implemented to represent certain characteristics of a domain in an abstract way [43] using graphical or mathematical representations. Hereby, phenomena, their entities, and relationships between them are identified and represented [44, 45]. This method is used to depict the concept of BP complexity and its components in Chapter 3 and the comprehensive framework on data-driven BP complexity analysis in Chapter 5.

Case Study

Case study research is a way of investigating the particularity and complexity of phenomena in a real-life context by studying a single case or a limited number of cases [46, 47]. Case study research focuses on a detailed investigation of research artifacts to develop an understanding, build and test theories [47]. In this dissertation, case studies are performed to gradually develop and evaluate the research artifact, i.e., framework on data-driven business process complexity analysis. Whereas the first case study is used to conceptualize the textual data-based complexity (see Chapter 3), the second serves to evaluate and enrich the latter with event log data (see Chapter 5). In this research, we use the case study protocols established by [48].

Workshop

Workshops are research techniques aiming to provide trustworthy and credible information in a certain area. Workshops can be defined as scheduled short-term activities aimed at experts sharing one or more of the following characteristics: common domain, such as IT Service Management, working in the same field, such as IT Service Desk, or having common goals such as BP automation or redesign. The participation group is, as a rule, of a small size to ensure everyone receives individual attention and has the opportunity to contribute. Participants are encouraged to actively engage in and influence the workshop direction. Workshop also implies a certain result, such as new ideas, insights, (re)designs of a product, process, or innovation [49]. [50, 51] differentiate between four workshop participation types: (1) contractual, in which participants are paid by researchers to participate in studies and experiments, (2) consultative, whereby participants are asked for their opinions, (3) collaborative, in which researchers and participants collaborate, but the researchers guide the process, and (4) collegiate, in which researchers and participants collaborate but the participants guide the process. In the beginning, a collegiate workshop in the first case study has contributed to the formulation of the topic of this dissertation. Further, collaborative workshops have been used to develop an understanding of BP complexity in the setting of the first case study (see Chapter 3).

Questionnaire

Questionnaire is one of the most commonly known techniques to get information from people and answer research questions [52]. Questionnaires are often, but not always, used in situations where the purpose is to collect data from a relatively large number of people. This technique provides input to different kinds of research, such as profiling and descriptive research, where the goal is to create a characteristic sample profile, predictive and analytical research with the goal to investigate relationships between variables, and conceptualizing and evaluating measurement scales or a set of statements to quantify a complex variable like service quality, trust, or innovation. Hereby, various types of questions can be used, i.e., open or closed (list, category, ranking) [52]. Questionnaires can be distributed online or physically, both ways having their advantages and disadvantages. In this dissertation, questionnaire is used to support the process of conceptualization and evaluation of the BP complexity concept presented in Chapter 3. Based on the problems and motivations identified in terms of the above-mentioned collegiate workshop in the first case study, the questionnaire has been designed to specify and evaluate the research artifact in the case study context. Corresponding evaluation documentation, including the questionnaire, is available on Github[2] as well as in Appendix A - Evaluation Questionnaire of this dissertation. The results obtained from the questionnaire have been further detailed in the expert interviews.

Expert Interview

An expert interview is typically defined as a qualitative interview based on a topical guide that focuses on the expertise of a field expert [46, 53]. Expert interviews, as a form of research technique, have been widely explored and applied in a variety of fields. The goal of this research technique is to investigate and gather data from subject matter experts based on the interactions with them, which are often led by questions of different types, for example, structured, semi-structured, or open-ended. Semi-structured interviews are most commonly used in this dissertation to motivate and enhance the dialogue with the experts, particularly in Chapter 3 and Chapter 5 in terms of BP complexity concept operationalization in the case study contexts.

1.3 Research Contributions

This dissertation aims at raising the attention of the BPM community to the importance and potential related to the consideration of textual data produced by BP participants. This data can enlarge the BP analysis toolset providing a more comprehensive way to identify BP redesign and improvement opportunities. As a result, we develop

[2] https://github.com/IT-Tickets-Text-Analytics/Evaluation

a framework on data-driven BP complexity analysis, which is the main contribution of this work. First, a set of linguistic features specifically aimed at complexity-related decision-making support is designed (see Chapter 3) and selected (see Chapter 4) based on the first case study and its IT ticket descriptions, providing solutions to the aforementioned concerns of inefficiency and bias. Afterward, the second case study (see Chapter 5) is used to validate the set of linguistic features and enhance the textual data-based BP complexity prediction with event log data addressing the incomprehensiveness and bias concerns. In Figure 1.3.1, we summarize the outcome, i.e., the concept of the framework, which is explained in detail in Chapter 5.

The framework is gradually developed starting from a comprehensive literature analysis, detailed elaboration of BP complexity concept in the form of linguistic features and its machine learning-based operationalization, followed by a publication unifying the linguistic and process mining approaches, as well as a final study offering an overview of research agenda in a wider context. Below, we provide a summary of contributions structured in accordance with the gradual framework development and the dissertation outline.

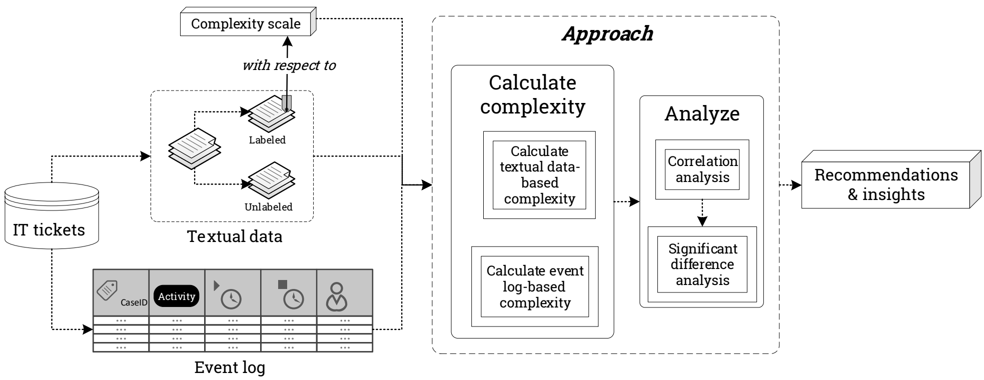


Figure 1.3.1: An approach for analyzing business process execution complexity based on textual data and event log

1. *Comprehensive literature analysis on complexity in organizations*: The main goal of this analysis is to develop an understanding and extend the body of knowledge on complexity research and practice in organizations addressing its high fragmentation. As a result of this analysis, a morphological box containing three aspects and ten features is suggested and framed with a multi-dimensional complexity basis to synthesize and standardize the results. Afterward, a method for complexity management to provide key insights and decision support in the form of extensive guidelines is proposed. This way, we highlight the shortcomings of current literature on the topic and the importance of a comprehensive approach to complexity in organizations. The analysis and its results are explained in Chapter 2.
2. *Textual data-driven BP complexity analysis*: The primary objective of this analysis is to investigate those complexity- and decision-making-related knowledge aspects that can be extracted from BP textual data. Accordingly, three commonly known

knowledge types [54], i.e., objective [31], subjective [33], and meta-knowledge [32], are studied in three separate publications. Thus, in [31], a method to measure a BP cognition is introduced and evaluated in the first case study setting. In [33], a concept of business sentiment as a measurement of a perceived anticipated effort reflecting an emotional component of a BP is developed. In [32], certain stylistic patterns influencing the readability and understandability of the BP texts are suggested. Afterward, the elaborated techniques are summarized in a concept of textual data-driven BP complexity outlined in one journal publication [12] presented in Chapter 3. Application scenarios of this concept, such as the multi-criteria knowledge-based recommender system [55] and context awareness support [56], are also studied in separate publications. However, they remain beyond the scope of this dissertation. Further, in Chapter 4, the textual data-driven BP complexity concept is implemented using machine learning algorithms whereby the features most beneficial for BP complexity prediction are selected [30]. These are cognition level [31] and stylistic patterns [32]. These publications are discussed as a part of the journal publication [12] in Chapter 3. Business sentiment [33] has shown not to have a sufficient impact on the prediction. Moreover, the advantages of the selected linguistic features are proven by their comparison with a well-established text analysis technique LIWC[3] in Chapter 5 as a part of [29] publication. Hence, using the elaborated textual data-driven BP complexity concept, it becomes possible to establish objective and efficient BP analysis approaches.

3. *Textual and event log data-driven BP complexity analysis*: The major emphasis is on enrichment and alignment of a textual data-driven BP complexity analysis with the BP model and event log data common in BPM. This is achieved by identifying a set of metrics for both complexities taking existing works as a basis (for textual data-based complexity, particularly those presented in Chapter 3 and selected in Chapter 4) and studying the relation between textual and event log data-based complexities in Chapter 5. More specifically, the developed approach investigates how textual data can contribute to event log complexity prediction. Then, it is adapted and illustrated in a real-world setting of the second case study. Such a comprehensive approach presented in Figure 1.3.1 and explained in detail in Chapter 5 comprises the core of the framework in this dissertation.
4. *Research agenda and future work*: Important concluding part of the dissertation in Chapter 6 presents an overall research agenda for innovative future studies and theory building in the comprehensive consideration of machine- and human-generated process data in BPM. The agenda provides researchers and managers with a structured overview of the possible generic research questions and success factors concerning (i) IS design, (ii) BPM as a discipline, and (iii) business environment.

1.4 Publications

The research in the scope of this dissertation suggests the framework on the data-driven business process complexity analysis for decision-making support, this way

[3] https://www.liwc.app/

making the BPM toolset more comprehensive, efficient, and objective.
The chapters of this dissertation include peer-reviewed scientific publications, specifically four journal articles and one conference paper, which are listed below:

- [57] Revina, A., Aksu, Ü. & Meister, V. G., "Method to address complexity in organizations based on a comprehensive overview", *Information*, vol. 12, no. 10, 2021, DOI: 10.3390/info12100423, Creative Commons Attribution (CC BY) License.
- [12] Rizun, N., Revina, A. & Meister, V. G., "Assessing business process complexity based on textual data: evidence from ITIL IT ticket processing", *Business Process Management Journal*, vol. 27, no. 7, 2021, pp. 1966–1988, DOI: 10.1108/BPMJ-04-2021-0217, Emerald Publishing Limited, all rights reserved.
- [30] Revina, A., Buza, K. & Meister, V. G., "IT ticket classification: the simpler, the better", *IEEE Access*, vol. 8, 2020, pp. 193380–193395, DOI: 10.1109/ACCESS.2020.3032840, Creative Commons Attribution (CC BY) License.
- [29] Revina, A. & Aksu, Ü., "An approach for analyzing business process execution complexity based on textual data and event log", *Information Systems*, vol. 114, 2023, p. 102184, DOI: 10.1016/j.is.2023.102184, Creative Commons Attribution (CC BY) License.
- [58] Revina, A., "Business process management: integrated data perspective. A framework and research agenda", in: *Proceedings of the 29th International Conference on Information Systems Development: Crossing Boundaries between Development and Operations (DevOps) in Information Systems (ISD2021), Valencia, Spain, September 8-10, 2021*, 2021, authors retain copyright for the paper with an agreement to allow AIS a non-exclusive license to publish it.

The works performed in the scope of this dissertation but not included in its outline comprise further ten conference publications, three out of which as the only author, one journal article, and one book chapter:

- [59] Revina, A., "Assessing process suitability for AI-based automation. Research idea and design", in: *Proceedings of the 21st International Conference on Business Information Systems (BIS2018), Berlin, Germany, July 18-20, 2018*, Springer, 2018, pp. 697–706
- [31] Rizun, N., Revina, A. & Meister, V. G., "Method of decision-making logic discovery in the business process textual data", in: *Proceedings of the 22nd International Conference on Business Information Systems (BIS2019), Seville, Spain, June 26-28, 2019*, Springer, 2019, pp. 70–84
- [32] Rizun, N., Meister, V. G. & Revina, A., "Discovery of stylistic patterns in business process textual descriptions: IT ticket case", in: *Proceedings of the 33rd International Business Information Management Association Conference (IBIMA2019), Education Excellence and Innovation Management through Vision 2020, Granada, Spain, April 10-11, 2019*, International Business Information Management Association (IBIMA), 2019, pp. 2103–2113
- [33] Rizun, N. & Revina, A., "Business sentiment analysis. Concept and method for perceived anticipated effort identification", in: *Proceedings of the 28th International Conference on Information Systems Development: Information Systems Beyond 2020 (ISD2019), Toulon, France, August 28-30, 2019*, 2019
- [55] Revina, A. & Rizun, N., "Multi-criteria knowledge-based recommender system for decision support in complex business processes", in: *Proceedings of the Workshop on Recommendation in Complex Scenarios co-located with 13th ACM Conference on Recommender Systems (RecSys2019), Copenhagen, Denmark, September 20, 2019*, vol. 2449, CEUR, 2019, pp. 16–22

- [60] Buza, K. & Revina, A., "Speeding up the SUCCESS approach for massive industrial datasets", in: *Proceedings of the 2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA2020), Novi Sad, Serbia, August 24-26, 2020*, IEEE, 2020, pp. 1–6
- [61] Revina, A., "Considering business process complexity through the lens of textual data", in: *Proceedings of the the 16th International Multi-Conference on Computing in the Global Information Technology (ICCGI2021), Nice, France, July 18-22, 2021*, ThinkMind, 2021, pp. 12–14
- [62] Revina, A., "Exploring dashboards as socio-technical artifacts: literature review-based insights", in: *Proceedings of the 18th International Conference on E-Business (ICE-B2021), Online, July 7-9, 2021*, ScitePress, 2021
- [63] Revina, A. & Aksu, Ü., "Towards a business process complexity analysis framework based on textual data and event logs", in: *Proceedings of the 17th International Conference on Wirtschaftsinformatik (WI2022), Nürnberg, Germany, February 21-23, 2022*, AIS eLibrary, 2022
- [56] Revina, A., Rizun, N. & Aksu, Ü., "Towards a framework for context awareness based on textual process data", in: *Proceedings of the 26th International Enterprise Design, Operations and Computing Workshop (EDOCW2022), Bolzano, Italy, October 3-7, 2022*, Springer, in press
- [64] Rizun, N., Revina, A. & Meister, V. G., "Analyzing content of tasks in business process management. Blending task execution and organization perspectives", *Computers in Industry*, vol. 130, 2021, p. 103463
- [65] Revina, A., Buza, K. & Meister, V. G., "Designing explainable text classification pipelines: insights from IT ticket complexity prediction case study", in: *Interpretable Artificial Intelligence: A Perspective of Granular Computing*, vol. 937, Springer, 2021, pp. 293–332

Additionally, the author has worked on several other research projects beyond the scope of this dissertation:

- [66] Meister, V. G., Hu, W., Revina, A., Cikus, M. & Müller, J., "Fachanwendung für digitale Modulkataloge. Eine Untersuchung zu Graph-basierter Daten-Modellierung und Navigation", in: *Proceedings of the 50. Jahrestagung Der Gesellschaft Für Informatik, Hochschule 2030, Karlsruhe, Germany, September 29 - October 1, 2020*, Gesellschaft für Informatik, 2020, pp. 467–480
- [67] Kalabina, E., Belyak, O., Meister, V. G. & Revina, A., "What kind of employees' team is necessary for industrial digital transformation? Theoretical and practical analysis", in: *Proceedings of the Digital Transformation in Industry: Trends, Management, Strategies (DTI2020), Ekaterinburg, Russia, November 27, 2020*, Springer, 2021, pp. 183–193
- [68] Belyak, O., Kalabina, E., Revina, A. & Meister, V. G., "Comparative analysis of the knowledge sharing process in cross-functional teams: management practices in Russia and Germany", in: *Proceedings of the 7th International GSOM Emerging Markets Conference, Saint Petersburg, Russia, November 11-21, 2020*, 2020, pp. 167–171
- [69] Salminen, J., Mustak, M., Rizun, N., Revina, A., Nikiforova, A., Almerekhi, H., et al., "Integrating AI into customer service: improving the actionability of customer feedback analysis using machine learning", in: *Proceedings of the AIRSI2022: Technologies 4.0 in Tourism, Services and Marketing, University of Zaragoza, Spain, July 11-13, 2022*, 2022, pp. 175–179

- [70] Salminen, J., Mustak, M., Rizun, N., Revina, A., Nikiforova, A., Almerekhi, H., et al., “Integrating AI into customer analysis deep learning experiments on customer feedback prioritization in social media”, *Technological Forecasting & Social Change*, 2022, submitted to the journal
- [71] Revina, A., Kalabina, E. & Belyak, O., “Employee engagement and soft skills in the digitalization of the economy. Preliminary study results”, in: *Proceedings of the Digital Transformation in Industry: Trends, Management, Strategies (DTI2022), Ekaterinburg, Russia, October 28, 2022*, 2022
- [72] Revina, A., Kalabina, E. & Belyak, O., “Do employee soft skills influence their engagement in organizations? Insights from industry and academia”, in: *Proceedings of the 9th International GSOM Emerging Markets Conference, Saint Petersburg, Russia, October 5-8, 2022*, in press, 2022
- [73] Rizun, N., Revina, A. & Edelmann, N., “Application of text analytics in public service co-creation: literature review and research framework”, in: *Proceedings of the 24th Annual International Conference on Digital Government Research - Together in the unstable world: Digital government and solidarity, Gdańsk, Poland, July 11-14, 2023*, submitted to the conference, 2023

1.5 Dissertation Outline

The remainder of the dissertation is structured into the six following chapters:

Chapter 2 — *Comprehensive Literature Review on Complexity in Organizations*. This chapter is based on the journal article [57] and serves to introduce the topic of complexity in an organizational context, existing approaches, related challenges, and research gaps. Hereby, the focus lies on the diverse data types existing in organizations. As an outcome of the literature analysis, diverse structuring approaches are presented, leading to a comprehensive method to address complexity developed based on a prominent Goal Question Metric. The method uses various data types as input to derive complexity types and approaches to deal with them. Finally, a simple, relevant example illustrates the method application in a typical organizational setting.

Chapter 3 — *Unveiling the Potential of Textual Data for Business Process Complexity Analysis*. This chapter is devoted to the conceptual development of the BP complexity analysis based on BP textual data. It deals with the extraction of complexity-relevant knowledge from textual data and bases on the journal article [12]. This article presents the textual data-based BP complexity concept while consolidating the work on the BP complexity-relevant knowledge extraction from textual data performed in the three conference publications [31–33].

Chapter 4 — *Predicting Business Process Complexity Using Textual Data*. This chapter presents a machine learning-based prediction of BP complexity using the linguistic features conceptualized in Chapter 3. Hereby, various machine learning algorithms are tested, and, in a feature selection process, the features most beneficial for BP complexity prediction are identified. As a result, a classification pipeline for BP complexity prediction based on textual data is designed. All the data and experiments in Chapter 3 and Chapter 4 originate from the industrial case study of a German telecommunication provider, specifically, ITIL Change Management [8] IT ticket processing.

Chapter 5 — *Enriching Textual Data-Driven Business Process Complexity Analysis with Event Log-Driven Analysis*. Based on the journal article [29], this chapter aims to fulfill three purposes: (1) to evaluate the approach for measuring textual data-based BP complexity presented in Chapter 4 in another case study, (2) to compare the performance of suggested linguistic features with the linguistic features from a well-accepted text analysis technique LIWC [74], (3) to enrich the textual data-based BP complexity with the event log-based complexity, and (4) to study their relation. Accordingly, an approach for analyzing BP execution complexity based on two data types is introduced. By means of the second case study of Service Request IT ticket processing from an academic institution in the Netherlands, the mentioned approach and obtained results are demonstrated and discussed in a real-life setting.

Chapter 6 — *Research Agenda and Future Work*. This chapter bases on a single conference publication [58] that aims to put the research in a wider context of BPM, IS design, and business environment. Based on a solid theoretical framing, problematization methodology for generating novel research questions, and dialectical interrogation, the traditional BPM logics regarding BPs, IT infrastructure, and BP actors are used to derive the research agenda. The interweaving of process mining and NLP is considered an essential determinant of BPM future research under the growing importance of natural language and NLP maturity.

Chapter 7 — *Conclusion*. In the last chapter, research questions are discussed by synthesizing the contributions of this dissertation, followed by the threats to validity. Further, implications for research and practice are examined. Finally, we offer a perspective on the research limitations and possible directions for future work.

1.6 Preliminaries

In this section, we summarize and provide background information and definitions necessary for understanding the research work conducted in this dissertation.

Business process is specified as a series of interconnected events, activities, and decision points, including a number of actors and objects that contribute to a valuable outcome for at least one customer [1]. For instance, incident management is a typical business process performed in most organizations to provide support to their customers about the malfunctions in products and services offered to them.

Business Process Management is broadly defined as a subject field bringing together the knowledge on the (re-)design of individual BPs, building a core BPM capacity in organizations, and serving a wide range of objectives and circumstances [21]. BPM discipline has emerged as a response to growing needs for globalization, standardization, innovation, agility, and operational efficiency, as well as the opportunities opened by the advancements in digital technologies. It combines the techniques and tools from various fields, such as Industrial Engineering, Information Systems, Computer Science, Human Resources, and Corporate Governance. BPM addresses a complete lifespan of a BP, offering a systematic representation of how a certain BP can be managed. Hereby, BPM serves to develop and suggest a comprehensive set of methods, techniques, and tools to discover, analyze, redesign, execute, and monitor BPs and optimize their performance [1].

Business Process Management lifecycle is defined as a continuous cycle including the following six core phases [1]: (1) *process identification* related to identifying a business problem and relevant processes; (2) *process discovery*, whereby the present state of relevant processes is recorded, usually in the form of one or more as-is process models; (3) in the *process analysis* phase, issues with the current process are detected, recorded, and quantified if possible; (4) the goal of *process redesign* phase is to improve the process based on the issues identified in the previous phase. For this purpose, a variety of change options is analyzed and compared to develop a to-be process model. These two phases (3) and (4) are often combined, as every time a new change is suggested, it should be analyzed using the process analysis techniques; (5) *process implementation* phase serves to prepare and implement a set of steps to move from as-is to to-be process model which commonly involves organizational changes and automation; (6) *process monitoring* aims to control how well the process is running based on the defined performance measures and objectives.

Business process textual data are, as a rule, understood as BP descriptions officially recognized in organizations, BP definitions recorded based on the interviews and questionnaires, as well as textual data contained in BP models or event log. In fact, textual data are used in all the phases of the lifecycle. For example, in the process discovery phase, the BP analysts might need to infer processes from such textual data sources as ethnographic studies, interviews, and existing process documentation [18, 75, 76]. Further, in the process analysis, redesign, and monitoring phases, all the changes and adjustments must follow legal requirements, corporate policies, and compliance documents existing in a textual form [77]. We are aware of the fact that the linguistic complexities of these textual data sources are different. Whereas interviews and process descriptions are likely to follow a sequential writing style, the compliance documents are written in a declarative manner. In this dissertation, we explicitly set the focus on the textual data generated by BP actors in the BP execution, which is usually written in an interactive manner. The linguistic complexities of aforementioned sources can be studied as a part of future work.

Event log is BP execution data typically containing information such as what activity was performed, when, and who performed an activity. In other words, it is a collection of events, each of which refers to an activity performed in a process supplemented with additional information related to the activities [78]. For example, receiving, registering, prioritizing, assigning, resolving, and closing an IT ticket are typical series of activities executed in an incident management process. In Table 1.6.1, an exemplary event log of an incident management process is shown. As can be seen, each row contains what activity is performed, when, for which incident, and other additional information like who performed (resource) and its priority.

Table 1.6.1: Example of an event log of an incident management process

Incident	Activity	Time stamp	Resource	Priority	Attr.
m001	Register	01-12-2021 10:11:10	BPActor1	Low	...
m001	Assign	01-12-2021 13:10:08	BPActor2	Low	...
m002	Register	01-12-2021 16:01:03	BPActor3	High	...
m002	Assign	01-12-2021 17:10:40	BPActor4	High	...
m002	Resolve	01-12-2021 17:25:49	BPActor5	High	...
m002	Close	02-12-2021 08:40:24	BPActor6	High	...
m001	Hand-over	02-12-2021 09:30:30	BPActor7	Low	...
m003	Register	02-12-2021 10:01:02	BPActor3	Medium	...
m003	Assign	02-12-2021 10:22:33	BPActor2	Medium	...
m003	Reject	02-12-2021 11:25:36	BPActor6	Medium	...
m001	Hand-over	02-12-2021 13:24:16	BPActor8	Low	...
m001	Resolve	02-12-2021 14:47:05	BPActor4	Low	...
m001	Close	02-12-2021 14:52:02	BPActor1	Low	...

Process mining is a set of approaches aimed at the data-driven analysis of processes to understand how processes are carried out in reality. For this, event logs are used as the main input. Process mining enables businesses to assess the efficiency and compliance of real processes compared to the assumed ones based on event logs. As a result, important insights into understanding BPs and recommendations on how to improve them can be provided [78].

Linguistics is a scientific study of human language in all of its aspects, including organization, usage, and history [79]. It is generally distinguished between theoretical and applied linguistics [80]. While the first is considered the foundation of linguistics focusing on language nature and structure, the latter aims at solving real-world problems based on natural language [81], using, for example, natural language processing and text mining offering a collection of approaches to analyze natural language. Hereby, text mining is known as the knowledge extraction from text addressing such tasks as information retrieval, text classification and clustering, entity, relation, and event extraction. NLP is an attempt to extract a more complete meaning representation from text, such as who did what, when, how, why, and to whom. It usually employs linguistic concepts like parts-of-speech and grammatical structure such as phrases, dependency relations, and word order. Further, it addresses word ambiguity and semantics [82]. In this dissertation, we use the term **linguistic features**. We define it as quantified characteristics of text extracted based on various linguistic assumptions regarding the ways humans express their thoughts in different contexts, for example, the theory of least effort and Zipf's Law declaring that humans tend to express their thoughts as concise as possible [83].

Data-driven process analysis implies the primer usage and reliance on the data obtained in the BP execution for the analysis. These data are also referred to as **process data** in the MRQ definition, implying the two most typical data types, i.e., event log and textual data, generated in the BPs. The meaning of data-driven analysis becomes clear in its comparison with the process-centric approach. Although these approaches are seen as complementary and interrelated, they base on a completely different logic. Whereas a modeled process constitutes the core of process-centric analysis, in the data-driven approach, the data generated and collected within the BP execution plays this role [78]. Considering the current organizational dynamics, large streams of data and their critical value for decision-making support, we reasonably give a preference to a data-driven approach.

Decision-making in an organizational context can be specified as the process of choosing a plan of action from a set of two or more options in order to find a solution to a given problem or task [84]. Especially in complex organizational settings, decision-making support gains importance [85]. Hence, complexity and decision-making often appear together in one context. In the scope of this dissertation, the decision-making support is based on complexity analysis. Being aware of the complexity of a process, BP actors can better prioritize the tasks, assess the urgency, analyze, minimize, and avoid errors, reduce rework and redundancy, increase efficiency, accelerating the processing time. Further, an improved BP actors' assignment, including automation of simple routine tasks, and time management can be envisioned. Such decision-making support will likely result in better processes and, thus, increased BP actors' and customers' satisfaction.

Complexity has a wide range of meanings and interpretations. In terms of this dissertation, we consider complexity from the process perspective, which also implies activities and tasks a BP actor needs to accomplish within a process. While gradually developing the framework, we exemplarily differentiate among the processes of low, medium, and high complexity having the following meaning: (1) low complexity implies clear rules, fast and easy execution by BP actors, possibility to be programmed and performed by IS at economically feasible costs; (2) medium complexity implies no exact rule set, a clear need of information acquisition and evaluation. IS supporting the process cannot substitute but assists BP actors and increases their productivity; (3) high complexity means complex problem-solving in addition to information acquisition and evaluation. IS can offer very limited support.

Information Technology Service Management is a branch of Service Science that focuses on IT operations like service delivery and support. Opposed to conventional technology-oriented approaches, ITSM is a process-oriented discipline for managing IT operations as a service, making up 60% to 90% of total IT ownership costs [86]. ITSM is frequently related to ITIL, the British government's Information Technology Infrastructure Library framework of best practices aimed at high-quality IT service delivery at a reasonable cost [8].The core ITIL components on the service support operational level include the following: (1) Service Desk, user single point of contact with the service provider responsible for the management of problems and service requests as well as user communication; (2) Incident Management responsible for incidents' lifecycle to get the IT service back to users as soon as feasible; (3) Problem Management in charge of overseeing the lifecycle of all problems to prevent problems or reduce their effect; (4) Change Management responsible for changes' lifecycle and aimed to perform changes with the lowest possible IT service interruption; (5) Release Management including hardware, software, documentation, processes, and other components needed to implement the changes; (6) Configuration Management keeping track of configuration items and their connections needed to deliver an IT service [8]. In this dissertation, two case studies are used, in particular, an ITIL Change Management IT ticket processing from a prominent telecommunication provider in Germany (see Chapter 3 and Chapter 4) and a Service Request IT ticket processing from an academic institution in the Netherlands (see Chapter 5).

CHAPTER 2

Comprehensive Literature Review on Complexity in Organizations

Summary

This chapter introduces the concept of complexity in the context of organizations, existing approaches, related challenges, and research gaps. The focus is on the diverse data types existing in organizations. As a result of the literature review, diverse structuring approaches are provided, leading to the development of a comprehensive method for dealing with complexity based on a well-known Goal Question Metric. The method uses various data types as input to derive relevant complexity types and approaches to deal with them. Finally, a generic example illustrates the method application in a typical organizational setting.

This chapter is based on the following publication:

Revina, A., Aksu, Ü. & Meister, V. G., "Method to address complexity in organizations based on a comprehensive overview", *Information*, vol. 12, no. 10, 2021, DOI: 10.3390/info12100423.

Title Method to Address Complexity in Organizations Based on a Comprehensive Overview

Authors Aleksandra Revina[1,2], Ünal Aksu[3], and Vera G. Meister[2]
[1] Technical University of Berlin
[2] Brandenburg University of Applied Sciences
[3] Utrecht University

Published in Information, 2021, MDPI, Open Access, journal rank 2021: CiteScore Q2 (Information Systems).

Abstract Digitalization increasingly enforces organizations to accommodate changes and gain resilience. Emerging technologies, changing organizational structures and dynamic work environments bring opportunities and pose new challenges to organizations. Such developments, together with the growing volume and variety of the exchanged data, mainly yield complexity. This complexity often represents a solid barrier to efficiency and impedes understanding, controlling, and improving processes in organizations. Hence, organizations are prevailingly seeking to identify and avoid unnecessary complexity, which is an odd mixture of different factors. Similarly, in research, much effort has been put into measuring, reviewing, and studying complexity. However, these efforts are highly fragmented and lack a joint perspective. Further, this negatively affects the complexity research acceptance by practitioners. In this study, we extend the body of knowledge on complexity research and practice addressing its high fragmentation. In particular, a comprehensive literature analysis of complexity research is conducted to capture different types of complexity in organizations. The results are comparatively analyzed, and a morphological box containing three aspects and ten features is developed. In addition, an established multi-dimensional complexity framework is employed to synthesize the results. Using the findings from these analyses and adopting the Goal Question Metric, we propose a method for complexity management. This method serves to provide key insights and decision support in the form of extensive guidelines for addressing complexity. Thus, our findings can assist organizations in their complexity management initiatives.

Keywords organizational complexity; technological complexity; textual complexity; morphological box; goal question metric information

MDPI

Systematic Review

Method to Address Complexity in Organizations Based on a Comprehensive Overview

Aleksandra Revina [1,2,*,†], Ünal Aksu [3] and Vera G. Meister [2]

[1] Chair of Information and Communication Management, Faculty of Economics and Management, Technical University of Berlin, 10623 Berlin, Germany
[2] Faculty of Economics, Brandenburg University of Applied Sciences, 14770 Brandenburg an der Havel, Germany; vera.meister@th-brandenburg.de
[3] Department of Information and Computing Sciences, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands; u.aksu@uu.nl
* Correspondence: revina@th-brandenburg.de
† Current address: Faculty of Economics, Brandenburg University of Applied Sciences, 14770 Brandenburg an der Havel, Germany.

Abstract: Digitalization increasingly enforces organizations to accommodate changes and gain resilience. Emerging technologies, changing organizational structures, and dynamic work environments bring opportunities and pose new challenges to organizations. Such developments, together with the growing volume and variety of the exchanged data, mainly yield complexity. This complexity often represents a solid barrier to efficiency and impedes understanding, controlling, and improving processes in organizations. Hence, organizations are prevailingly seeking to identify and avoid unnecessary complexity, which is an odd mixture of different factors. Similarly, in research, much effort has been put into measuring, reviewing, and studying complexity. However, these efforts are highly fragmented and lack a joint perspective. Further, this negatively affects the complexity research acceptance by practitioners. In this study, we extend the body of knowledge on complexity research and practice addressing its high fragmentation. In particular, a comprehensive literature analysis of complexity research is conducted to capture different types of complexity in organizations. The results are comparatively analyzed, and a *morphological box* containing three aspects and ten features is developed. In addition, an established *multi-dimensional complexity framework* is employed to synthesize the results. Using the findings from these analyses and adopting the *Goal Question Metric*, we propose a *method for complexity management*. This method serves to provide key insights and decision support in the form of extensive guidelines for addressing complexity. Thus, our findings can assist organizations in their complexity management initiatives.

Keywords: organizational complexity; technological complexity; textual complexity; morphological box; goal question metric

Citation: Revina, A.; Aksu, Ü.; Meister, V. G. Method to Deal with Complexity in Organizations Based on a Comprehensive Overview. *Information* **2021**, *12*, 423. https://doi.org/10.3390/info12010423

Academic Editor: Kostas Vergidis

Received: 20 September 2021
Accepted: 12 October 2021
Published: 16 October 2021

1. Introduction

As organizations develop and digitize their businesses, the number of interactions and dependencies between their processes, information systems, and organizational units increases dramatically [1]. To deal with this growth, organizations often enhance the technology supporting their businesses. That can also affect the structure of these organizations [2]. Such dynamics are likely to bring significant challenges [3]. One prominent challenge organizations have to tackle is a complexity that blocks decision-making and leads to unreasonably high and mainly hidden costs. For example, according to the study conducted by The Hackett Group, a world-class strategic consultancy, in 2019 [4], the operating costs of low-complexity companies as a percentage of overall revenue were almost 60% lower than those of the companies operating in highly complex settings. The low-complexity companies employed 66% fewer staff and spent 30% less on technology.

There has always been much interest from both academia and industry in complexity research. Many disciplines have established their own complexity subfields adjusting the complexity concepts to the specific goals [5], for example, social complexity [6], software complexity [7], or managerial complexity [8], in addition to generic studies. While recognizing the contributions of these subfields, it is hard to find an agreed-upon definition. Missing conceptual alignment and differing complexity interpretations complicate the common acceptance of the complexity field [5]. Further, there remains much unclarity and fragmentation on what contributes to the complexity and how to not only measure and reduce it but, more importantly, benefit from it. Especially in the organizational domain, the discussion on complexity is either rather generic [9] or devoted to a narrow topic, like task complexity [10]. Hereby, the methods for measuring complexity are prevailingly designed for a particular type of complexity [11] and, therefore, are hardly transferable to other types [12]. However, the organizational domain can be characterized as a complex environment embracing different types of complexity, often even implicit, for example, related to people, structures, technology, or processes. Hence, in such settings, managers lack comprehensive guidelines on how to approach complexity management initiatives, including steps, measurements, and necessary data and information.

Additionally, ongoing technological developments and the use of diverse information systems increase the variety of data sources and types. Such advancements create a demand for a continual review of the existing studies in complexity research that can offer new perspectives, highlight essential gaps, and suggest approaches for complexity management. Moreover, in the literature, we observed a lack of structured outcome providing an integrated overview of the complexity research.

To address the mentioned shortcomings, in our study, we conduct a Systematic Literature Review (SLR) and create a comprehensive overview of complexity research relevant in organizations. To create such a summary, we take as the basis the People Process Technology (PPT) framework [13,14] as well as typical information and data sources existing in organizations. With this, we aim to cover the complexity that is related to the main components of organizations, namely *people*, *technology*, and *processes*. Hence, we focus on the typically observed complexities related to an organization as a whole (*organizational* complexity), technology (*technological* complexity), and people, in particular, communication. The latter we analyze through the lens of generated textual data serving different purposes (*textual* complexity).

In fact, complexity can be influenced by a myriad of factors and may arise in various forms. For example, in the case of organizational complexity ranging from complex projects to organizational structures, managers and executives face inefficiency, delays, performance decrease, or even the inability to handle their businesses. Similarly, unstructured textual data, like complex and even confusing instructions, project communication and documentation, textual task descriptions will cause unclarities, errors, and rework. Furthermore, complex technology will likely be not only the source of high costs but also an obstacle to successful business functioning, although it has the goal of supporting the processes in organizations. Hence, dynamic business processes, computing platforms, applications, and services can cause a massive complexity encompassing many concepts, technologies, data and information sources, which are interlinked together in diverse organizational interactions.

Afterwards, to address the confusion impeding a common acceptance of the field, we structure and classify the literature findings in the form of a morphological box and integrate them into a multi-dimensional complexity framework. The morphological box [15] aims at summarizing the fundamental aspects of complexity research, whereas the multi-dimensional complexity framework [5] incorporates the complexity types identified in the literature into one integrated structure contributing to the standardization efforts of the complexity research. Based on these two analyses, we derive our main contribution, that is, a method to apply our study findings by extending the Goal Question Metric (GQM) [16].

Hence, our study contributes to the common acceptance and understanding of complexity research in organizations by addressing its high fragmentation. It provides researchers with a comprehensive overview based on the PPT framework and structured outcome of research results in the form of a morphological box and multi-dimensional framework.

As a practical contribution, we address the lack of instructions for selecting of a specific approach to address complexity in organizations. Hence, the method can serve as practical guidance for organizations in their complexity management projects.

The remainder of the paper is structured as follows: Section 2 provides an overview of the related work in the literature and highlights the gaps. In Section 3, we present the research methods used for conducting the literature review and analyzing and classifying its results. Section 4 gives an overview of the results. Section 5 analyzes and classifies the results providing structured outcomes in the form of the morphological box and integrated complexity framework. In Section 6, we present the key contribution of our study, that is, the method to address complexity in organizations. We discuss our findings as well as their implications and limitations in Section 7. Finally, conclusions and future work are presented in Section 8.

2. Related Work

The term complexity has always received the attention of scholars in different fields, such as Computer Sciences, Organizational Sciences and Linguistics. For example, in Computer Sciences, the term complexity, as a rule, determines the complexity of an algorithm, that is, the number of resources required to execute the algorithm [17]. Organizational Sciences mostly adapt concepts from Complexity Theory and define an organization as a complex dynamic system consisting of elements interacting with each other and their environment [18]. In Linguistics, complexity is studied from the language perspective and comprises phonological, morphological, syntactic, and semantic complexities [19]. Hence, complexity reveals various definitions and implications. It can be expressed in exact terms, such as McCabe's software complexity defined as a number of possible paths in code [7]. Alternatively, it can cover broad ideas serving as an umbrella term for many concepts. For example, Edmonds defines complexity as a property of a language that makes it troublesome to formulate its behavior, even given nearly complete information regarding its atomic parts and their interrelations [20]. Such heterogeneity of complexity concepts hampers its wide recognition [5]. Besides, this high fragmentation becomes apparent when searching for literature review studies conducted in the field of complexity, like in the following reviews: [21] on task complexity, [22] on innovation complexity, and [23] on interorganizational complexity from employer and safety perspectives.

Similarly for the term complexity, there is much ambiguity and fragmentation on what contributes to the complexity and the ways to reduce and benefit from it. For example, [24], while conducting research on complexity drivers in manufacturing companies, observe that the studies mostly focus on one definition (perspective) of complexity drivers providing no comprehensive analyses. [25] states that the task complexity drivers or contributors found in the literature are the consequence of combining diverse settings and subjective preferences. As a result, they are ambiguous and difficult to grasp and analyze. [26] developed guidelines on how to reduce complexity focusing on a rather specific field of product and process complexity in a case study of a consumer products manufacturer. The authors justify the necessity of such recurring complexity studies by an increasingly changing complexity nature. Besides, another group of studies aiming to deal with process complexity in organizations consider it from the specific perspective of process models, workflows, and event logs [27–29]. Another rather unconventional approach to process complexity is suggested in [30]. The authors study the complexity of IT Service Management processes through the lens of textual data massively generated in the organizations and the related readability and understandability of textual work instructions. This represents a promising but rather narrow research direction. Further, addressing the process complexity of software development, [31] also emphasize a limited scope of the previous studies focusing on the code-related measures.

Moreover, the shortcoming of a narrow focus causes the drop back of certain industries in the complexity management field [12]. Likewise, despite their abundance, complexity measures, such as software complexity measures that are essential for managing complexity, have not been sufficiently described and evaluated in the literature [32]. This is also confirmed in the recent studies evidencing the poor usage of code complexity measures in the industry [33].

Further, a number of recent studies investigate the notion of complexity in companies empirically based on a bottom-up approach, the fact indicating that there is still no alignment and much dissonance on how to approach complexity in organizations. Accordingly, [34] study how project complexity has been perceived by practitioners in different industry sectors. In total, five sectors and more than 140 projects have been researched. It has been concluded that a comprehensive project complexity framework and guidelines could aid in the management of complex projects by raising awareness of the (anticipated) complexities [34]. [35] also declare the need for developing approaches to measure and manage complex projects. While studying the complexity of supply chains in four case studies, [36] highlight the necessity of frameworks that can assist managers in building overarching complexity management strategies and practices supporting them [37].

In fact, as fairly noted in the recent work [38], organizations comprise many different interconnected elements making the complexity study challenging and also having a "negative" context. Interestingly, the authors mention that this often inherent and inevitable complexity implies essential benefits, particularly in dynamic and unpredictable conditions. In addition, they provide some strategic leadership guidance on complexity management in organizations. Similarly, in other works [2,39–41], the attempts to suggest guidance to address organizational complexity are limited to high-level strategic leadership recommendations.

Hence, in our study aiming to develop a method to comprehensively address complexity in organizations, the following gaps (IG: Identified Gap) serve as a motivation:

- IG1: Scarce comprehensive analyses and high fragmentation of the complexity field impeding its common acceptance;
- IG2: Much unclarity and confusion on what contributes to complexity, how to measure, reduce, and benefit from it;
- IG3: Lack of extensive guidelines for managers on how to approach complexity;
- IG4: Organizational domain embracing different types of complexity, which are often implicit.

3. Research Approach

In this section, we present the research approach we follow to propose a method to deal with complexity. It includes the following: (i) an SLR on complexity in organizations and (ii) morphological box and multi-dimensional framework to classify, analyze, and synthesize the SLR findings. Based on (i) and (ii), we extend the GQM approach and propose practical guidelines for managers to address complexity in organizations.

3.1. Systematic Literature Review

As the starting point in the method development, we perform an SLR on complexity management in organizations, which we describe in the subsections below.

3.1.1. Scope

As complexity is a very broad subject area, conducting an SLR solely on complexity would have yielded an incomparably large set of papers, which is difficult to analyze and provide valuable insights. Hence, we set the focus on complexity management in organizations.

Organizations are commonly described by the three main components: *people*, *technology*, and *processes* connecting them [13]. This viewpoint is also known as the *People*

Process Technology (PPT) framework [14]. Despite being a popular concept, especially in industry, its origin is not straightforward. The oldest and most prominent source using the logic of these three elements is Leavitt's diamond model [13]. This model focuses on problem-solving in organizations and highlights three types of solutions — structure (by means of organizational chart or responsibilities), technology (by means of technologies), and people (by means of Human Resources). Later on, the PPT components make up the fundamentals of the prominent Information Technology Infrastructure Library (ITIL) framework launched in the 1980s [42]. Based on the underlying assumptions in ITIL, any technology solution is only as right as the processes it supports. Similarly, processes are only as good as the people who follow them. Hence, we focus on these three PPT components (people, technology, and process) and build up our complexity types accordingly. In particular, we take *organizational*, *technological*, and *textual* complexity types as the starting point. For each type, we explain our motivation below.

- Organizational complexity: In the first type of complexity, organizational, we consider a broad organizational perspective discussed by Mintzberg [43] and related studies [44]. Among others, the authors consider organizations from the viewpoint of allocation of tasks and resources. Such a configuration is enabled by major organizational components and common assets, such as tasks, projects, and processes in relation to projects and organizations [45]. In this regard, task complexity research going back to the 1980s [21,46] can be considered the most well-studied one in the organizational context [25]. Consequently, task, project, and process complexities are considered important constituents and subtypes of organizational complexity.
- Technological complexity: As the second type of complexity, we study technological complexity. The oldest and most popular example of technological complexity can be acknowledged software complexity, such as McCabe cyclomatic complexity [7], which is based on the control flows represented in the form of graphs. The McCabe cyclomatic complexity served as a basis for several other technological complexities related to process models and event logs [47];
- Textual complexity: To address the people component, we focus on communication, that is, on how people exchange information in organizations and receive their tasks. Subsequently, we pose a question about how we can obtain this information. Textual data generated inside and outside organizations remain one of the most valuable types of unstructured data [30,48–50]. Hence, we consider textual complexity as the third type of complexity, which, among others, reflects the people component in organizations. Indeed, beyond the structured program codes and event logs, analysts estimate that upward of 80% of enterprise data today is unstructured, whereby the lion's share is occupied by textual data [48]. There is a great variety of textual data types relevant for organizations. Emails, files, instant messages, posts and comments on social media are some examples.

To provide an integrated overview of the research on the complexity from organizational, technological, and textual perspectives, it is necessary to analyze how complexity is measured in each perspective. Accordingly, we cover the following topics: what complexity metrics are available in the related literature and whether any tool support is provided for them. To capture what the existing studies in the literature focus on, apart from complexity measurement, we investigate the motivations and declared novelty in the studies. Likewise, we set to unveil the practical contributions of the studies in the complexity research on organizations. Another objective is to check the future work avenues mentioned in the related literature so that potential contribution areas can be identified.

Based on the discussion of the scope above and the gaps identified in the related work section, we design five research questions (RQs):

- RQ1: What concepts of complexity are available in the literature from the organizational, technological, and textual perspectives?
- RQ2: What metrics are proposed to measure the complexity concepts in organizations?

- RQ3: What are common motivations and declared novelty areas of the complexity research on organizations?
- RQ4: Which focus areas and application cases are typically used in the related literature on complexity in organizations?
- RQ5: What future research directions are communicated in the complexity research on organizations?

To answer the RQs and identify and analyze the existing studies in the related literature, we conduct an SLR. For this purpose, we follow the well-established guidelines by Kitchenham [51]. The main reason is that these guidelines are rigorously applied in the literature, cover major steps of a typical SLR, and could be used to design a review protocol in various fields. For guiding and evaluating literature reviews, we point to the framework proposed in [52]. In addition, we emend our analysis with the tool-support guidelines as presented in [53]. Accordingly, the retrieval and selection mechanism of the papers used in the SLR is given in the subsection below.

3.1.2. Paper Retrieval and Selection

Search strategy: To retrieve papers in the literature, we create search strings that are generic enough to include the studies discussing at least one of the three complexity types. Specifically, based on the complexity types and subtypes that we observed while identifying the gaps in the related work, we outline the following set of search strings:

- for organizational complexity: "task complexity" **OR** "project complexity" **OR** "process complexity";
- for technological complexity: "software complexity" **OR** "process model complexity" **OR** "event log complexity" **OR** "workflow complexity" **OR** "control flow complexity";
- for textual complexity: "textual complexity" **OR** "readability" **OR** "understandability".

Having such separated search strings for each complexity type has advantages. One of them is balancing the granularity in the terms of complexity types. This way, we eliminate the risk of having an unequal number of papers for each complexity type. Another advantage is determining the papers that are mostly devoted to the study of a particular complexity type.

We applied the search strings in the search engine Google Scholar. The reason for this choice is twofold. Firstly, Google Scholar is the world's largest academic search engine and provides an integrated search environment [54] by encompassing other academic databases, like ACM Digital Library, IEEE Xplore, Scopus, and Web of Science. Secondly and more importantly, Google Scholar ranks search results by their relevance considering all major features of papers, such as full text, authors, published source, and how often each paper has been cited in academic databases. The papers that have at least one of the above search strings in their title, keywords, or in their main body are retrieved. For the three complexity types, our search resulted in a large number of papers, that is, more than 5000.

Inclusion and exclusion criteria: To exclude irrelevant papers, we defined and applied the inclusion and exclusion criteria listed in Table 1. Specifically, we aimed at those papers closely related to organizational, technological, and textual complexities in organizations. We added the citation minimum to select meaningful papers recognized by other researchers. Hereby, the citation minimum was not applied to the recently published papers. The papers dealing with how people in organizations are grouped were excluded, as they mostly focus on organizational structure. Likewise, papers without a scientific basis and containing assumptions or expectations, that is, theoretical speculations, were filtered out. Lastly, as the papers on the same topic of the same authors have a high overlap, the ones with extended content were considered more relevant.

As the result of filtering out based on the inclusion and exclusion criteria, in total, 130 papers were selected and combined in the final set. Due to the practical orientation of complexity and relevance, we included a limited number of technical reports and Ph.D. theses, which made up less than 5% of the final set of papers. Using the reference management tool Zotero (Check out Zotero, https://www.zotero.org/, accessed on 15 August 2021), paper metadata were obtained for each paper in the final set. Paper metadata included the reference data about a paper, for example, authors, publication date, authors' keywords, and abstract.

Table 1. Inclusion and exclusion criteria.

Type	Criterion
inclusion	• related to organizational, technological, and textual complexity in organizations; • cited at least five times (except for the publications later than 2018); • written in English.
exclusion	• with an exclusive focus on product complexity, system and architecture complexity, or network complexity; • complexity related to organizational structure; • theoretical speculations; • studies of the same author on the same topic (only the most relevant ones are selected).

3.2. Literature Classification

In this subsection, we explain the two approaches, that is, morphological box and multi-dimensional complexity framework, that we used to classify and synthesize the SLR findings obtained following the process described in the previous subsection.

Morphological box: To summarize the primary aspects of complexity research on organizations, we analyzed and classified the literature review results. For this, we used a morphological box, which is a commonly used (for example, in [55,56]) method to present a set of relationships inherent in the multi-dimensional non-quantifiable problem complexes [15]. Developing a morphological box to describe complexity aspects requires the identification of relevant features of papers and grouping these features. To do so, we focused on the three main themes: (1) common characteristics that can be observed in any research paper, (2) complexity quantification, and (3) tangible research artifacts for quantifying complexity. With the last two themes, we aimed to identify reusable and distributed complexity research outputs that we can exploit for developing our method on addressing complexity.

Regarding the first theme, we took the identified gaps IG1 and IG4 listed in the related work section. With this, we aimed to show the fragmented studies on complexity. Hence, the following features were identified: motivation, novelty, focus area, application cases, and future research. We grouped them and named this group *generic aspects*, as they can be observed in any research paper. Based on the common sense on quantification and the second gap IG2 in the related work, four features were identified in the second theme. These are metrics origin, input, output, and validation. To combine these features, we propose *complexity metrics and analysis aspects* as the second group. Aligned with the third theme and IG3, it might be critical to know the specific conditions of *implementation*, such as being openly accessible, distributed in a free or commercial manner. Moreover, the implementation of a complexity analysis and metrics should be of deep concern. It is the only procedure that allows the application of particular approaches in the sense of evaluation or a real-world scenario. Thus, we included tool support as another feature in the morphological box. In this feature, we considered the usage of existing tools, own development tools, or no tool support. Table 2 shows the defined three aspects with their features.

Based on the identified features, each paper in the final set was coded to extract topics necessary to answer the RQs. For this, deductive coding was applied [51]. A predefined set of codes observed while identifying the gaps in the literature was taken as the starting point for coding. While assigning the codes in each paper, newly observed codes were added to the code set. In other words, observed concepts in papers were used as codes. In deductive coding, it is important to avoid the researcher's bias, that is, the researcher's preferences are likely to influence the selection of papers [51]. Therefore, the topics in papers were identified and coded by two researchers independently. Afterwards, a discussion session was carried out to align on differences in coding. The coding was conducted using the qualitative data analysis tool NVivo (Check out Nvivo, https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software, accessed on 15 August 2021) to ensure consistency. The results obtained from coding served as the basis for the morphological box development.

Table 2. Identified aspects and features for morphological box.

Aspect	Feature	Definition
Generic	motivation	factors that encourage the researchers to conduct their research
	novelty	new and interesting contributions within the research, originality, distinctive research contribution
	focus area	a broad area where the research is conducted
	application case	a narrow and specific area where research artifacts are tested
	future research	potential directions for further studies
Complexity Metrics and Analysis	metrics origin	theories and disciplines laid the foundation of complexity metrics
	input	all required data and information to perform complexity analysis and calculation
	output	results of the complexity analysis and calculation
	validation	how proposed metrics are validated
Implementation	tool support	whether any tool is used or developed

Multi-dimensional complexity framework: In a variety of disciplines, the term complexity has been applied to the specific context that prompted the establishment of standalone complexity research subfields [5]. This variety impairs the general acceptance of the complexity research (see IG1 in the identified gaps). To address this problem, [5] suggest a four dimension framework unifying the most prevalent views on complexity. Hereby, each dimension comprises two opposing complexity notions derived based on the established Complexity Science literature: objective and subjective (D1, observer perspective), structural and dynamic (D2, time perspective), qualitative and quantitative (D3, measures perspective), and organized and disorganized (D4, perspective of dynamics and predictability). In Table 3, the dimensions and their complexity notions are listed. These dimensions allowed us to synthesize the SLR findings based on an agreed-upon vocabulary. As a result, we obtained an integrated multi-dimensional complexity framework.

Table 3. Complexity dimensions and notions.

Dimension	Perspective	Notions
D1	observer	objective, subjective
D2	time	structural, dynamic
D3	measures	qualitative, quantitative
D4	dynamics, predictability	organized, disorganized

3.3. Goal Question Metric

The morphological box and multi-dimensional complexity framework described in the previous subsection served to structure and standardize our efforts in the complexity research review. However, these structured outcomes of the SLR findings lack a practical application value for the managers having to deal with complexity in their daily businesses. Hence, two questions remain: (1) how to select a suitable approach to complexity analysis and measurement and (2) which type of complexity is relevant for a given problem or in a specific situation.

To address this kind of challenge, several approaches exist in the literature. For example, one popular approach originating from strategic management is Balanced Score Card (BSC) [57] embracing four perspectives, that is, financial (shareholders' view), customer (value-adding view), internal (process-based view), and learning and growth (future view). Initially developed in the business domain, BSC has been adapted to the software domain, specifically in relation to GQM [58,59]. GQM is one of the well-established and widely used approaches to determine metrics for goals, which is also known to be the most goal-oriented approach [16]. GQM was originally developed for the evaluation of defects in the NASA Goddard Space Flight Center environment in a series of projects. Despite this original specificity, its application has been extended to a broader context, software development among others [16].

GQM has a hierarchical structure beginning with a goal definition at the corporate, unit, or project level. It should contain the information on the measurement purpose, the object to be measured, the problem to be measured, and the point of view from which the measure is taken. The goal is refined into several questions that typically break down the problem into its main components. Next, various metrics and measurements are proposed to address each question. Finally, one needs to develop the data collection methods, including validation and analysis [16].

GQM opponents criticize it for being too flexible and generating a large number of metrics [60]. To address the GQM shortcomings, several approaches have been proposed, such as Goal Argument Metric (GAM) [61], or generic approaches as in [62]. Accordingly, GQM should be applied based on the organization's process maturity level to select the most acceptable metrics [62]. [63] suggests including a prioritization step into GQM to minimize the number of generated metrics. However, such an approach has a drawback that certain perspectives may remain neglected. To sum up, in the GQM related literature, we have observed a common practice of extending the GQM to adjust it to particular research needs, like in [64] for agile software development and [65] for data warehouses quality assessment.

Hence, in our study, we extend GQM for the purpose of the method development to comprehensively address the complexity in organizations. When adapting GQM to our study needs, we also set to overcome the above-mentioned limitations. Introducing a step-by-step guidance leading from the problem statement to a solution specification, we aim to assist organizations in identifying what is actually needed to deal with complexity.

In the following sections, we describe in detail the application of the discussed research approach. Accordingly, we start with reporting the SLR findings followed by their classification. The latter includes the morphological box development and integration of the results into a multi-dimensional complexity framework. Lastly, a GQM extension is proposed using an illustrative example from a real-world setting.

4. Systematic Literature Review

This section describes the results of our literature review. We address each RQ in a separate subsection and elaborate on our findings. In the visualization of our findings, we use a standard coloring for the three complexity types for the reasons of consistency. The assigned color for each complexity type is as follows: *organizational* ● (RGB: 199,199,199), *technological* ● (RGB: 127,127,127), and *textual* ● (RGB: 65,65,65) .

4.1. Complexity Concepts

We present our findings on the complexity concepts of the three complexity types that we focus on, that is, *organizational*, *technological*, and *textual*. In particular, we elaborate on the trend of papers over the years, the components of complexity types, and their distribution.

The majority of the analyzed 130 papers, 54 papers, discuss technological complexity, whereas 40 of the remaining papers are devoted to textual complexity and the rest 36 are about organizational complexity.

In Figure 1, the trend of publications over time is presented. As can be seen, the distribution of papers in all three types of complexity represents a similar trend with a peak at the period between 2011 and 2015. This indicates that, despite the different nature of the three complexities, they are all triggered by and based on today's information age developments. The most common examples of these developments are fast-expanding technology solutions, complex processes, and new organizational skills needed to deliver products and services faster, with higher quality, and at lower costs than before [66].

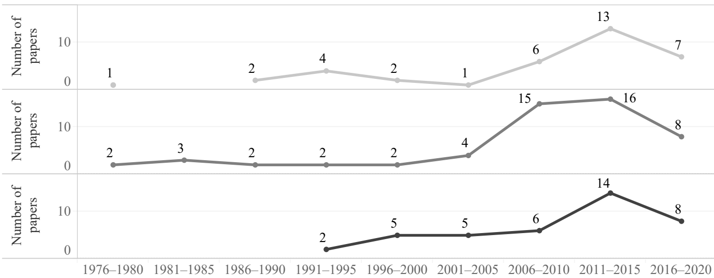


Figure 1. Publication trend per complexity type.

Based on the topics that appeared in the reviewed papers, we developed a mind map of complexity concepts, which is depicted below in Figure 2.

Figure 2. Complexity concepts mind map.

Key topics mentioned in the definitions of the organizational type of complexity are task, project, process, and product complexities. In the task complexity, which takes up the major part (36%) of all papers in the organizational complexity group, key topics are task complexity elements and models [11,46,67–69], cognitive resources necessary to perform the tasks [70], identification and utilization of complexity factors [71] and effects [72–75], performing reviews and analyses, as well as building frameworks [10,21,25]. Very often, one and the same study addresses a bundle of topics. For example, in the project complexity (25%), apart from reviews [76], the core topics are building frameworks and models [77–79]. Process complexity (14%) can be characterized by either the study of specific processes such as (IT) services [80–82] or generic topic of complexity factors and effects [83]. Product (and production) complexity (5.56%) is studied as part of project complexity in the context of new product development projects [84] or generic approaches to production complexity [85]. The product or production complexities are largely dependent on the type of the product itself. Hence, we include very few works in our analysis.

In the technological type of complexity, we found several topics as the key topics. For example, those related to business processes, business process models with 41%, such as in [86] and event logs and workflows taking up 9%, like [87]. Software programs in general (22%), for example, [7], and with a specific focus on user interfaces (13%), as [88], are the topics related to software. Furthermore, there are topics that relate to systems. To name the most common ones, rule-based systems (4%), for example, [89], enterprise systems (4%), like [90], and IT architectures (6%), as discussed in [5].

In the textual complexity type, the most studied topic appeared to be corporate and accounting narratives and legislative documentation in general, such as contracts or institutional mission statements, making up 68% [91,92]. This can be accounted for a strong need for clarity in communication between institutional management and various stakeholders. Other, rather rare topics relate to two groups: (1) textual complexity of webpages and online reviews (13%) [93,94] and (2) management textbooks [95], texts used for reading comprehension [96], and news articles (10%) [97].

4.2. Complexity Metrics and Analysis

In this subsection, to answer the RQ2 "What metrics are proposed to measure the complexity concepts in organizations?" and get a deep understanding of existing approaches on measuring complexity, we analyze the following guiding questions:

1. Which theories and disciplines laid the foundation of complexity metrics, that is, metrics origin?
2. What kind of information and data serve as input for complexity metrics?
3. What kind of output is expected?
4. Whether tool support is provided?
5. How are the proposed metrics validated?

In this analysis, we filtered out the literature and critical review papers (24 in total), since, as a rule, they do not offer any metric. In Figure 3, metrics origin is shown for each type of complexity. Organizational Sciences is the prominent origin of metrics in organizational complexity. In the technological complexity, Software Engineering is the dominating origin, whereas textual complexity metrics, as naturally, are mainly driven by Linguistics. In general, Cognitive Sciences can be considered as a significant common driver for the origin of the metric in the three complexity types.

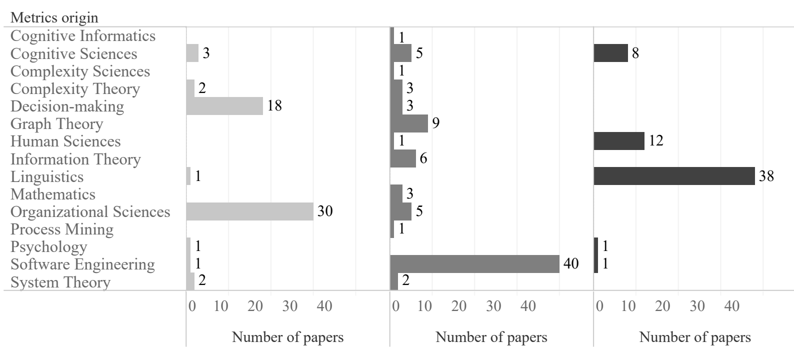


Figure 3. Metrics origin distribution per complexity type.

To understand how respective complexities are measured, we provide the analysis of inputs and outputs related to complexity metrics. In Figures 4 and 5, commonly used inputs and outputs in each complexity type are shown. Since business processes are one of the main assets of organizations, various data types related to business processes (for example, business process model, event log, and business process descriptions) are used to measure organizational and technological complexities.

Answering the question regarding validation, we identified that most of the complexity research (66%) is validated empirically. This is an indication of the practical value and applicability of complexity metrics. However, tool support can be fairly considered as a catalyst for the reproducibility and applicability of research findings.

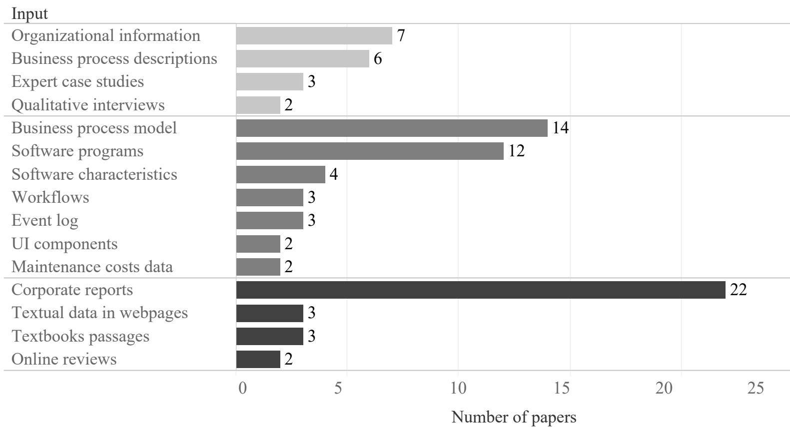


Figure 4. Types of inputs per complexity type.

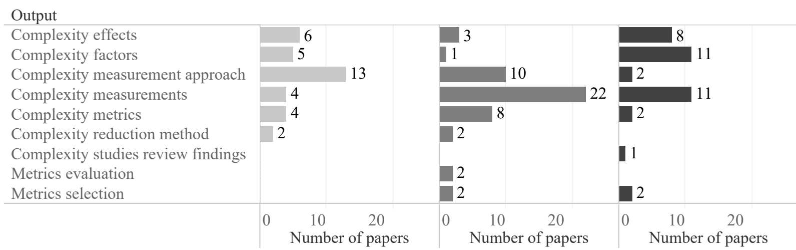


Figure 5. Types of outputs per complexity type.

4.3. Complexity Research Motivations and Novelty

For any novice but also advanced researchers, it may be useful to get an overview of those motivating considerations and successful research "selling points", or declared novelty, prominent in the field. In Figure 6, we present the motivations according to complexity type over time. We exclude literature and critical review papers. One can observe whether motivation gains importance over time for a complexity type. For example, in organizational complexity, there is a continuous interest in complexity factors and effects.

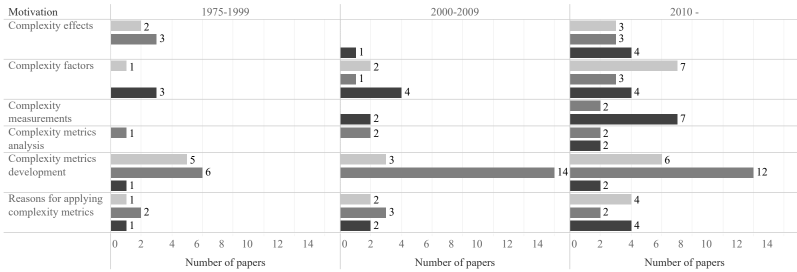


Figure 6. Motivations over time per complexity type.

In Figure 7, we present motivations in relation to the metrics origin for each complexity type. We excluded literature and critical review papers. One can observe those metrics origin areas that are essential for particular motivations related to complexity metrics. For example, in the development of new complexity metrics, Decision-making and Organizational Sciences are the two prominent areas to consider.

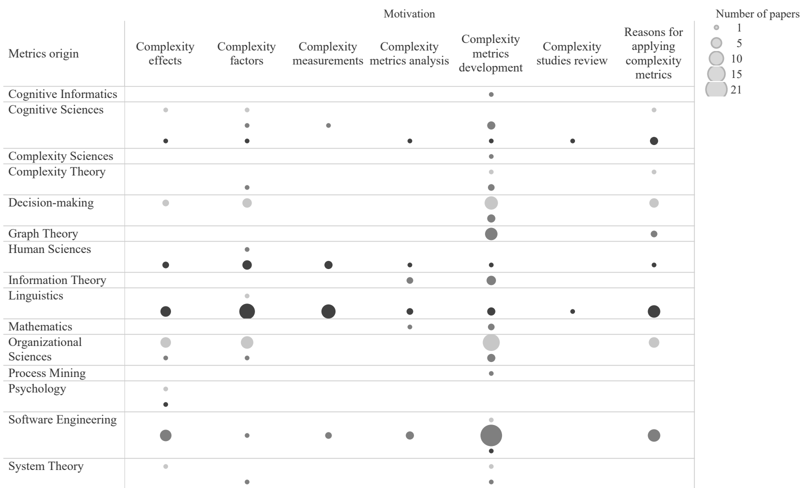


Figure 7. Motivations in relation to metrics origin per complexity type.

In Figure 8, we present the absolute numbers regarding the research novelty aspect. The most attractive topic is selecting a specific application area. Metrics evaluation and tool support are the two topics that attracted less interest in this context.

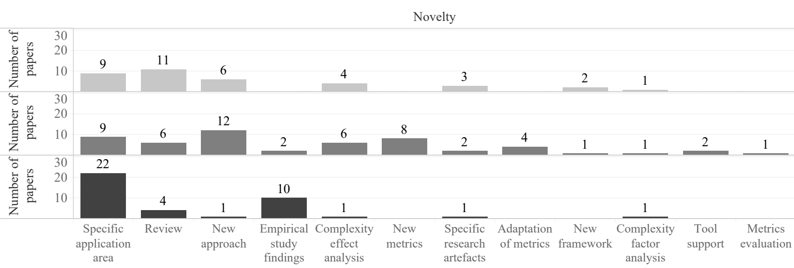


Figure 8. Novelty per complexity type.

4.4. Complexity Research Focus Areas and Application Cases

In Figures 9 and 10, the absolute distributions of specific focus areas and application cases according to complexity type are shown. The papers in which no application is specified are excluded from the application case distribution. To keep the most frequent values of application cases, the threshold for the minimum number of papers is set to two.

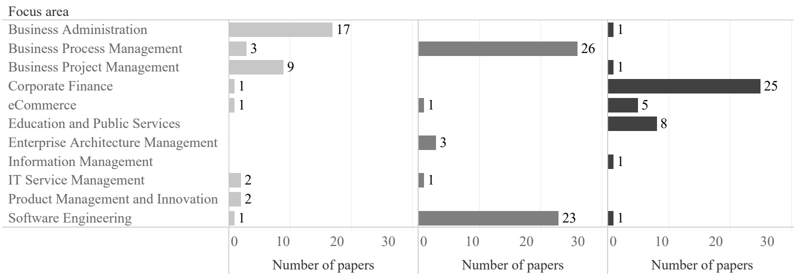


Figure 9. Focus areas per complexity type.

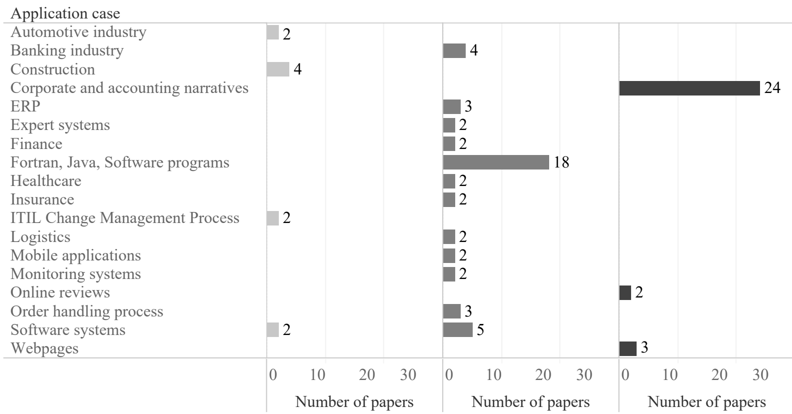


Figure 10. Application cases per complexity type.

4.5. Complexity Research Future Research Directions

In this subsection, we share our findings regarding the future research directions we identified in the reviewed papers.

In Figure 11, we summarize the absolute distributions of the future research directions over time. The most significant future research directions are new validation studies, approaches, and metrics extensions. Such directions as guidelines development, metrics comparison, complexity factors, and complexity effects are rarely mentioned.

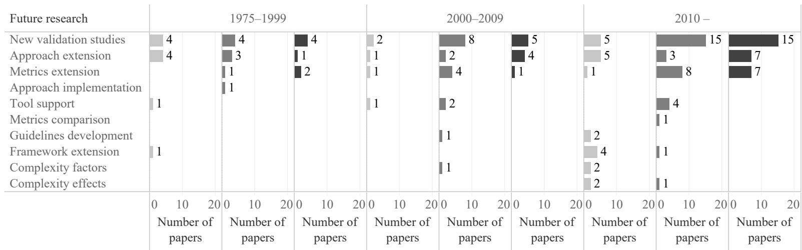


Figure 11. Future research directions over time per complexity type.

5. Literature Classification

In this section, we first classify the SLR results explained in the previous section and develop a morphological box. Second, we exploit a multi-dimensional framework to synthesize the SLR findings.

5.1. Morphological Box

To build the morphological box, the three aspects and ten features described in the research approach section are employed. Initially, each paper in the final set comprised of 130 papers is distributed to the three aspects based on the values in its ten features. Then, thematically related feature values are grouped, and the relative number of papers in each group is calculated. For example, in the novelty feature of generic aspects, “new approach”, “new framework”, and “new metrics” novelties are grouped. As they are mentioned in 30 out of the 130 papers, the relative frequency for that group is 23%. Using the frequency distribution in each group in the three aspects, the morphological box shown in Table 4 is formed. As one paper may discuss multiple topics (that is, have multiple values in a single feature), the percentage value for a group in a feature does not necessarily add up to 100.

In the aspects with an observable amount of values, like validation or tool support, no groups were formed. Hence, the relative number of papers containing the certain value can be directly derived from the table. For example, an empirical validation has been performed in 66% of papers. The three aspects in the created morphological box are explained below.

5.1.1. Generic Aspects

The motivations for performing complexity research can be that multiplex as the term complexity is. As our review shows, they vary from measurement, reviews of complexity studies, and development of new complexity metrics to the investigation of complexity factors and effects. Motivations are highly beneficial to make researchers familiar with the complexity research. Furthermore, for the researchers already into the topic, it would be advantageous to explore the studies related to complexity metrics development as well as practical studies. Practitioners can gain insights into complexity factors, effects, as well as reasons for applying complexity metrics.

The overview of the most frequent and also uncommon focus areas and application cases will help researchers to propose fruitful directions for case study-oriented complexity research and beyond. Business Process and Project Management, Corporate Finance, Software Engineering, and IT Service Management are some of the examples of focus areas. We use Statistical Classification of Economic Activities in the European Community [98] (RAMON) to provide a consistent overview of application areas, like Healthcare, Governmental Services, and Automotive.

Diverse novelty aspects (for example, a new approach, framework, new metrics or adaption of existing ones, findings on empirical studies, application area, and comparative evaluations) may serve as a source of inspiration for both complexity novices and expert researchers. Similarly, the future research aspect can point to some significant limitations of the studies and potential further research directions. Additionally, gaps or problems mentioned in future research can be further investigated to contribute to the existing body of knowledge.

5.1.2. Complexity Metrics and Analysis Aspects

Metrics are common instruments for measurement. They are often used as a communication tool among stakeholders to control and assess the quality and status of artifacts. Moreover, various complexity metrics are applied to measure complexity broadly and in distinct areas. Such complexity metrics and analyses are characterized by certain input and output data. For example, depending on the type of complexity and application case, textual documentation, event logs, software programs, expert interviews, or even image data can serve as an input for the complexity analysis and measurements. The output is directly related to the research motivation in general and associated input in particular.

In the metrics development process, another important aspect is their theoretical base or origin. As a rule, theories are used to name observed concepts and explain relationships between them. The theory is a well-known tool that helps to identify a problem and plan a solution. Hence, metrics origin is included as a feature in the proposed morphological box. Accordingly, the disciplines that are widely used for metrics development are summarized to give an overview and support the metrics development process. Some frequent examples are Organizational Sciences, Cognitive Sciences, Cognitive Informatics, Human Sciences, Software Engineering, Process Mining, and Mathematics.

As a matter of fact, metrics can be of different quality, depending on how precisely they describe an attribute of an entity. Hereby, metrics validity is among the most critical quality characteristics [86]. Validation is essential to guarantee that the outcomes of the metric application are legitimate. There are two types of validation methods: theoretical and empirical [99,100]. Theoretical validation is conducted using one of the following typical ways: (i) metrics development exclusively based on a theory, (ii) metrics development solely based on existing studies, (iii) checking compliance with a standard framework (for example, Briand's framework properties [101] and Weyuker's properties [102]), or (iv) illustration of the metrics application with the help of an example. In general, empirical validation of metrics complements theoretical validation. For empirical validation, different strategies are used, for example, case studies, surveys, or experiments. The objective of empirical validation is to find out whether the given metric measures what it is supposed to. Thus, for a metric to be structurally sound and useful, both theoretical and empirical validation are required [103].

5.1.3. Implementation Aspects

As can be seen in Table 4, approximately a third of the papers (34%) provide information about tool support. In particular, 23% of the papers mention that they use existing tools. In the remaining 11%, a tool was developed to support complexity measurement. We observed that in organizational complexity, there is a high demand for tool development, whereas, in technological complexity, tool support is more common. Notably, the studies providing tool support are prevailingly motivated by complexity metrics development.

Table 4. Morphological box describing three complexity aspects.

Generic Aspects					**Complexity Metrics and Analysis Aspects**				**Implementation Aspects**
motivation	novelty	focus area	application case	future research	input	output	metrics origin	validation	tool support
Complexity metrics development 38%	Specific application area 31%	Business Administration 27%	Information and Communication 29%	Approach extension; Metrics extension; Framework extension 46%	Business information 32%	Complexity measurements 62%	Cognitive Informatics; Cognitive Sciences; Human Sciences; Organizational Sciences 63%	Empirical 66%	No 35%
Complexity factors 19%	New approach; New framework; New metrics 23%	Business Process Management 22%	Professional, scientific and technical activities 28%	New validation studies 22%	Software and architectures 17%	Complexity factors 15%	Mathematics; Process Mining; Software Engineering 39%	Theoretical 17%	Using existing tools 23%
Reasons for complexity metrics 16%	Complexity studies review 16%	Corporate Finance 20%	Manufacturing 9%	Tool support 6%	Business process models 11%	Complexity effects 14%	Linguistics 32%		Yes 11%
Complexity effects 12%	Empirical study findings 9%	Software Engineering 19%	Financial and insurance activities 6%	Complexity effects; Complexity factors 3%	Event logs; Workflows 10%	Complexity studies review findings 6%	Decision making 17%		
Complexity measurements 9%	Complexity effect analysis; Complexity factor analysis 11%	eCommerce 5%	Wholesale and retail trade 5%	Approach implementation; Guidelines development; Metrics comparison 3%	Case studies and interviews 5%	Metrics selection; Metrics evaluation 5%	Graph Theory; System Theory 13%		
Complexity studies review 8%	Metrics evaluation; Metrics adaptation; Tool support 5%	Enterprise Architecture Management; IT Service Management 5%	Education 4%		Others 12%	Complexity reduction method 3%	Information Theory 5%		
Complexity metrics analysis 5%	Specific research artifacts 5%	Information and Innovation Management 2%	Others 23%				Others (Psychology; Complexity Theory; Complexity Sciences) 7%		

5.2. Integrated Multi-Dimensional Complexity Framework

While performing the comprehensive literature analysis, we identified multiple specific approaches focusing on complexity and, using them, developed the mind map of complexity concepts shown in Figure 2. Although these approaches somewhat reflect a generic commonly accepted complexity definition, that is, quantity and variety of the elements and their relationships, they are loosely coupled and focus on narrow areas. While addressing only one particular area, they fall short to embrace the whole variety of elements in the environment and support complexity management on large scales. Hence, motivated by the identified gaps IG1 and IG2, we synthesize the SLR results in the multi-dimensional complexity framework explained in the research approach section. In this subsection, we elaborate on that synthesis.

We consolidate the most common complexity concepts with the four dimensions of the framework. Specifically, we integrate the identified concepts to the complexity notions in the framework. Then, we calculate the relative distributions of the complexity notions per complexity type. Similar to the comprehensive literature analysis, the synthesis of the SLR results was performed using the qualitative data analysis tool NVivo. Hereby, each paper was analyzed and assigned to the complexity dimensions and respective notions by two researchers independently. Further, the discrepancies were discussed, and a common decision was taken. The resulting synthesis is depicted in Table 5 and explained below using an illustrative example.

The business process model is one of the complexity concepts we identified in technological complexity. As can be seen in the first row of technological complexity in Table 5, 91 % of the technological complexity papers focus on the objective observer perspective, whereas 36 % of the papers focus on the subjective observer perspective. In the remainder of this subsection, we elaborate on the synthesis from the three complexity perspectives, namely organizational, technological, and textual.

In technological complexity, the indicated complexity concepts are the business process model, event logs and workflows, and software. Though technological complexity subtypes are expected to be similar in their nature and, hence, complexity, we observed the diversity of combinations, especially if compared to organizational and textual complexities (see Table 5). Owing to its technological nature, we expected technological complexity to be objective (D1) and quantitatively measured (D3). Interestingly, both subjective (D1) and qualitative (D3) complexity approaches gained popularity in this type of complexity. Accordingly, subjective approaches to analyze complexity are exemplified below:

- Business process models, event logs and workflows: understandability, cognitive load perspectives [104] and quality measure of a process based on the number of generated process logs [105];
- IT services: complexity theory-based conceptualization [106];
- Enterprise systems: defining case study-based complexity factors [107];
- Rule-based systems: difficulty of problems that can be solved [108];
- User interfaces: case study-based evaluation [109];
- Software programs: dependency on the programmer's skills [110].

In the time perspective (D2) for technological complexity, all but one concepts in the analyzed papers are either purely or substantially structural. As can be seen in Table 5, IT services is the only concept considered completely dynamic (100 %). Similarly, in D4, it is the single disorganized technological complexity concept. Additionally, rule-based systems is the concept with a considerable disorganized relative distribution value.

Table 5. Relative distribution of the complexity notions per complexity type.

	Complexity Concept	D1: Observer	D2: Time	D3: Measures	D4: Dynamics Predictability
	see Figure 2	OBJ*ective*, SUBJ*ective*	STR*uctural*, DYN*amic*	QUAN*titative*, QUAL*itative*	ORG*anized*, DisORG*anized*
	Technological complexity (%)				
	Business process model	Obj (91) Subj (36)	Str (100) Dyn (18)	Quan (77) Qual (50)	Org (95) Dorg (27)
	Event logs and workflows	Obj (100)	Str (100)	Quan (80) Qual (20)	Org (100)
Software	Enterprise systems	Obj (100) Subj (50)	Str (100)	Quan (100) Qual (50)	Org (100)
Software	IT architectures	Obj (100)	Str (100)	Quan (100)	Org (100)
Software	IT services	Obj (100)	Dyn (100)	Qual (100)	Dorg (100)
Software	Rule-based systems	Obj (50) Subj (50)	Str (100) Dyn (50)	Qual (100)	Org (100) Dorg (50)
Software	User interfaces	Obj (100) Subj (57)	Str (100)	Quan (100) Qual (43)	Org (100) Dorg (29)
Software	Programs	Obj (92) Subj (17)	Str (100) Dyn (8)	Quan (75) Qual (42)	Org (100)
	Organizational complexity (%)				
	Organization as a whole	Obj (20) Subj (100)	Str (100) Dyn (40)	Quan (40) Qual (80)	Org (80) Dorg (40)
	Task	Obj (86) Subj (79)	Str (100) Dyn (7)	Quan (64) Qual (71)	Org (100) Dorg (14)
	Project	Obj (70) Subj (70)	Str (100) Dyn (20)	Quan (50) Qual (90)	Org (100) Dorg (10)
	Process	Obj (100) Subj (20)	Str (100) Dyn (20)	Quan (80) Qual (40)	Org (100) Dorg (20)
	Product(-ion)	Obj (50) Subj (50)	Str (100)	Quan (50) Qual (50)	Org (100)
	Textual complexity (%)				
	Legislative documentation	Obj (83) Subj (41)	Str (100)	Quan (83) Qual (41)	Org (100)
	News articles	Obj (100)	Str (100)	Quan (100)	Org (100)
	Webpages	Obj (100)	Str (100)	Quan (100)	Org (100)
	Online reviews	Obj (100)	Str (100)	Quan (100)	Org (100)
	Textbooks and other teaching materials	Obj (33) Subj (100)	Str (100)	Quan (33) Qual (100)	Org (100)

Due to its intrinsic diversity, that is, the diversity of tasks, projects, and processes as well as approaches to analyze their complexity, organizational complexity mostly includes all complexity notions of four dimensions, except for product (-tion) complexity. The latter, as a rule, reflects a defined number of elements (structural notion, D2) that interact in a specifically designed way (organized notion, D4) [85].

Regarding textual complexity, we observed the same complexity notions in the most textual complexity concepts, that is, legislative documentation, news articles, webpages, online reviews, and textbooks. It is noteworthy that they are considered purely structural (D2) and organized (D4). Moreover, except for legislative documentation and textbooks and other teaching materials, textual complexity concepts are highlighted as objective (D1) and quantitative (D3).

6. Method to Address Complexity

In this section, we present our method to address complexity in organizations with an illustrative example taken from a real-life setting.

For developing the method, GQM is taken as the basis and extended using the complexity analysis approaches obtained from the SLR, the developed morphological box, and the integrated multi-dimensional complexity framework. Aligned with the hierarchy in GQM, the starting point of our method is a goal definition. In other words, the method exploits a top-down mechanism in tackling complexity. This is also necessary to provide a staged complexity guidance, starting with a goal statement and reaching a specific solution. In our case, the specific solution is a collection data structure, each element of which is a map with the following key-value pairs: a question derived from a given goal and a set of complexity measurement and analysis approaches relevant for the problem origin of the question.

At the question level, the morphological box is used to refine an expressed goal into several questions. In particular, except for metrics origin, all features in the morphological box and the groups in these features are employed to facilitate deriving specific questions. For example, the groups in the input feature can be used to break down a given goal and define specific questions about the available inputs regarding complexity in the organization.

Next, at the measurement level, each question is analyzed with the help of the integrated multi-dimensional complexity framework. In the analysis, complexity types, complexity concepts, complexity dimensions, and notions are investigated to determine which of the complexity approaches obtained from the SLR can be beneficial for answering each question. With this, the aim is to distill the comprehensive analysis and provide a relevant subset of approaches mapping to each question. To build such a subset, initially, related complexity types are identified for each question. Then, per complexity type, matching complexity concepts, complexity dimensions, and notions are discovered. Based on these obtained inputs, the complexity approaches retrieved in the SLR are filtered, and a subset is created. The created subsets are merged into a final subset per question.

In the final level, that is, the data level, each subset is further filtered considering the available data related to each question. Importantly, in the case of data unavailability, data transformation possibilities can be investigated based on the input groups in the morphological box and the input attributes of the approaches in the created subsets. With this, our method enables organizations to consider and evaluate alternatives that may be beneficial depending on their context.

Based on the explanation of each level above, we summarize the steps of the proposed method:

1. Define a goal;
2. Formulate specific questions from the goal using the *morphological box;*
   For each question:
3. Analyze the question using the *integrated multi-dimensional complexity framework;*
4. Identify *complexity types* related to the question;
   For each complexity type:
   4.1. Using the *integrated multi-dimensional complexity framework*, find matching
      i complexity concepts;
      ii complexity dimensions;
      iii notions;
   4.2. Create complexity approaches subset based on *complexity type*, *i*, *ii*, and *iii*;
5. Merge complexity approaches subsets and form a final subset;
6. Filter the final subset considering the available data;
7. Check the need and possibilities for data transformation using the *morphological box.*

In the remainder of this section, we provide an example of addressing complexity (see Table 6). The setting of the example is taken from a company. The company operates in the healthcare sector and provides Business Intelligence (BI) solutions to various healthcare institutions. To gather requirements and change requests, business analysts of the company visit its clients. However, due to the Covid-19 pandemic and its consequent restrictions, visits were not possible. Moreover, healthcare practitioners were deeply in need of adequate dashboards to better understand the pandemic and were sending multiple immediate requests to the company via email. Hence, the business analysis department of the company has to tackle mostly unstructured and ambiguous texts in the received emails. Regarding the situation, the company is looking for a way to deal with the complexity that emerged due to the changes in its work environment.

Following our illustrative example in Table 6, in the growing popularity of remote working, the employees receive their tasks prevailingly in a textual form. The goal is to analyze the workload of employees. This goal is broken down into four questions.

Table 6. An example illustrates the use of the proposed method.

Goal: Currently, in the prevailingly remote way of working, managers are giving the tasks to the employees very often in a textual form, for example, per email. Analyze the workload of employees.			
Question 1: How difficult is it for employees to read and comprehend the manager's emails?			
a. Data: textual email data	**b. Complexity type:** textual	**c. Complexity dimensions and notions:** D1: objective, D2: structural, D3: quantitative, D4: organized	**d. Complexity analysis approach:** Flesh Reading Ease score, Gunning Fog Index, etc. [111]
Question 2: How many activities does an email contain?			
a. Data: textual email data	**b. Complexity type:** textual	**c. Complexity dimensions and notions:** D1: objective, D2: structural, D3: quantitative, D4: organized	**d. Complexity analysis approach:** count of verbs, specific approaches suggested in the recent research [69,112]
Question 3: What is the task complexity from the employee's point of view?			
a. Data: textual email data, specific information from employees regarding task complexity	**b. Complexity type:** organizational, tasks	**c. Complexity dimensions and notions:** D1: objective, subjective, D2: structural, dynamic, D3: quantitative, qualitative, D4: organized, disorganized	**d. Complexity analysis approach:** amount and clarity of inputs, processing, and output [11], objective (size, distance functions) and subjective (experience, motivation, etc.) [25]
Question 4: How often does an employee receive such emails per day?			
a. Data: email event log data	**b. Complexity type:** technological, event logs and workflows	**c. Complexity dimensions and notions:** D1: objective, D2: structural, D3: quantitative, D4: organized	**d. Complexity analysis approach:** count of specific case event IDs per day

In *Question 1*, the required data are email texts. Hence, the complexity type in focus is a textual one. Based on Table 5, we can conclude about the most typical complexity dimensions analyzed within this complexity type. Afterward, one can make use of the literature analysis results (for the details of this study, please check out our project on GitHub) we obtained with the help of Nvivo. Based on the complexity type (and subtype if applicable) and selected suitable complexity dimensions, one can determine relevant papers and, this way, derive necessary complexity analysis approaches and metrics. In the case of *Question 1* of the illustrative example, these are standard readability formulae, such as Flesh Reading Ease score or Gunning Fog Index [111].

In *Question 2*, the specific recent complexity analysis approaches, such as [112], are more relevant to answer the question and extract the exact activities from the textual data.

In *Question 3* and *Question 4*, different data and complexity types are required. In *Question 3*, organizational complexity type and task subtype are to be identified. Due to the inherent variability of organizational complexity, all possible complexity dimensions and notions are observed in this case (see Table 5), which complicates the choice. Thus, individual managerial decisions need to be taken to estimate the efforts and further actions. Additionally, gathering necessary data, that is, specific information from employees regarding task complexity, can be problematic and time-consuming. However, one can rely on solid research works for guidance, for example, [11]. At the same time, in *Question 4*, the required data, derived complexity type, dimensions, and notions are rather straightforward, and complexity analysis is easy to implement.

To summarize, GQM provides an opportunity to demonstrate the direct practical value of the study findings, that is, the comprehensive literature analysis, obtained complexity types, complexity concepts mind map, the morphological box, and the integrated multi-dimensional complexity framework. We extended the standard GQM process with additional points based on our study findings: data, complexity type, complexity dimensions, notions, and complexity analysis approaches.

7. Discussion

In this section, we discuss the findings of our study in accordance with the demands we have inferred from the gaps listed in the related work section and which served as a motivation for our study. These are:

- Demand for an extensive literature analysis embracing different types of complexity in organizations (IG1, IG4);
- Demand for a structured overview and integration of findings based on existing standard approaches, frameworks, and vocabularies to promote the common acceptance of complexity research in organizations (IG1);
- Demand for practical guidance supporting managers in addressing the complexity and planning comprehensive complexity management initiatives in companies (IG2, IG3).

7.1. Demand for an Extensive Literature Analysis

In the related literature, we have noted a high fragmentation in the research on complexity in general and in organizations in particular. This tendency has been observed both in regular and review papers. For example, the literature review works consider different types of complexity implicitly related to organizations, such as task complexity [10], project complexity [113], or interorganizational complexity [23]. However, each of such works focuses on a specific aspect of an organization. Our study sets itself apart by addressing multiple types of complexity in an extensive SLR of the complexity studies from the People Process Technology (the PPT framework [14]) perspective. Moreover, our SLR has shown that the peak popularity of complexity research in organizations has been in 2011–2015, which highlights the demand for an up-to-date review.

7.2. Demand for a Structured Overview

The related work shows that the mentioned problem of high fragmentation applies not only to the term complexity but also to the complexity drivers [25], measurements, and approaches to address complexity [26]. We address this limitation while classifying the SLR findings using the morphological box and the integrated multi-dimensional complexity framework. The developed morphological box (see Table 4) provides a logically structured outcome highlighting the most interesting and important points with the help of the three groups of aspects: generic, complexity metrics and analysis, and implementation aspects.

The generic aspects represent a summary on research motivations, novelty, focus areas, application cases, and future research on complexity in organizations, the information typical for any research study. This way, practitioners can get a quick overview on how the complexity research is structured and approached in the academic community. The researchers can also draw valuable insights using the information on motivations or future research directions. For example, some complexity studies are driven by research on the reasons for performing complexity analysis (16%) and complexity effects (12%), that is, hidden benefits of working on the complexity, interesting insight for both researchers and practitioners.

Accordingly, our study aims to facilitate and promote the efforts in this direction while streamlining the complexity research projects. Thus, the less frequent motivations, for example, complexity metrics analysis (5%) and complexity studies review (7%), may serve as a potential gap for further complexity studies. Likewise, the information about novelty areas can be useful for better positioning and justifying research projects in the sense of envisioned contributions. The summary on focus areas and application cases is valuable for researchers and practitioners in two ways. First, the information on the popularity of focus areas can help them to identify relevant areas requiring more research. Second, similar application cases can be closely studied for comparison-related purposes.

In addition, our findings can serve to prove the reusability of existing application cases. In the context of future research directions, the interested researchers can either follow the strategy of the trendy topics or pick up rare cases on which a few research activities have been conducted so far. For example, one can clearly see that the suggested complexity analysis approaches, metrics, and frameworks lack a thorough validation in all three complexity types.

The complexity metrics and analysis aspects include the inputs, resulting outputs, origins of complexity metrics, and their validation. Whereas metrics origin provides a theoretical background, input and output types as well as validation approaches provide practical insights into complexity measurement for both researchers and practitioners.

Aside from that, tool support, that is, implementation aspect, is one of the topics that need special attention since most of the complexity metrics (35%, see Table 4) are not using any tool. In other words, there is a clear lack of tool support in complexity measurement, which should be addressed in future research.

7.3. Demand for an Integration of Findings

We integrated our findings into a solid complexity framework [5] whereby four universal dimensions based on time, observer, measures, and dynamics and predictability were applied. Such a generic classification approach allows for independent documentation of diverse complexity concepts. Moreover, it enables more transparency in presenting and understanding the existing work on complexity. In our study, we demonstrated the possibility of integrating seemingly different complexity types into one complexity framework. This also enables other researchers to use their complexity concepts in the context of such a framework without a need to precisely define them [5].

In the classification of our study findings, we could identify various combinations of dimensions' values in the technological complexity indicating approaches not covered in the specific complexity concept. For example, the complexities of event logs and workflows, IT architectures, and IT services evidence the lack of subjective approaches (see Table 5). Though organizational complexity comprises prevailingly all dimensions and values, product(-ion) complexity lacks dynamic and disorganized approaches. Being structural and organized in its nature, textual complexity demonstrates the absence of subjective approaches in the case of the analysis of news articles, webpages, and reviews. All these open points offer potential for further exploration and, hence, future work. Furthermore, though the classification is not always straightforward, we highlight the importance of formalization in the context of complexity measurement.

7.4. Demand for a Practical Guidance

As mentioned in Section 2, [36,37] state the need for frameworks and practices assisting managers in the development of broad scope complexity management strategies. In our work, we address this shortcoming by proposing a method for complexity management by adopting GQM, a well-known approach for deriving and selecting metrics for a specific task in a goal-oriented manner [16]. In GQM, goals and metrics are adjusted for a specific setting. The upfront problem and, hence, goal statements allow for the selection of the metrics relevant for achieving these goals, what reduces the data collection effort considerably. The interpretation of the measurements becomes also straightforward due to the traceability between data and metrics. Hence, the wrong interpretations can be prevented [114]. Despite such advantages, GQM reveals certain limitations, such as high flexibility and generation of a large number of solutions [60], which were discussed in Section 3. The attempts to deal with this limitation end up in other shortcomings, like dismissing important perspectives [63]. Therefore, we have observed multiple extensions of GQM in line with a particular study purpose. In our work, we adopt and extend GQM with complexity types, dimensions, and notions making it more specific to our objectives and study setting. This way, we address the high flexibility of GQM. Further, in the metric selection stage of GQM, we consider not only metrics but also comprehensive complexity analyses approaches we found in the literature. In doing so, we deal with another shortcoming, that is, the risk of missing relevant solutions.

Although addressing the shortcomings of GQM [60,63], we are aware that our method also reveals some limitations. For example, it uses the SLR results as an input. Hence, it naturally inherits typical SLR limitations, such as authors' bias while building the search strings, exclusion and inclusion criteria definition, coding, and synthesizing the results [115]. Further, though our SLR contains significant ground work, there is a need for repeating such reviews to have an up-to-date list of complexity metrics and analysis approaches. Moreover, in its current state, our method represents a conceptually designed artifact that is manually applied in a real-world example to illustrate its relevance. Hence, automation solutions, such as [116,117], should be considered as a part of future work.

8. Conclusion and Future Work

In this paper, we proposed a method to address complexity in organizations. With the method, our main goal was to develop practical guidelines for the selection of complexity analysis approaches for a particular problem about complexity management in organizations. To achieve this goal, we extended the Goal Question Metric approach [16] using the SLR on complexity and its results, which we comparatively analyzed and synthesized.

In particular, to retrieve the body of knowledge on complexity research and practice, we conducted an SLR on the complexity research considering the three main complexity types, that is, organizational, technological, and textual, as the starting point. We analyzed 130 papers that discuss complexity. Hereby, we took the PPT framework [14] as the basis. Then, to provide structured outcomes and address the problem of high fragmentation, we designed and implemented two classification approaches: a morphological box and an integrated multi-dimensional framework. The developed morphological box summarizes the three fundamental aspects of complexity research. The three aspects reveal ten features capturing the information extracted from the analyzed papers. To contribute to the standardization efforts in the complexity research field, we synthesized our SLR findings by integrating them in a solid multi-dimensional complexity framework [5].

Next, the obtained knowledge and findings from the classification of the SLR results were used to extend GQM providing the method to address complexity in organizations. With an illustrative example taken from a company, we demonstrated the practical value of the method. Hence, as a practical contribution in the form of the comprehensive guidance, the method can assist organizations in their complexity management initiatives. Thus, our study sets itself apart from the existing work in the way that it serves as a guideline both for complexity researchers and practitioners willing to perform the complexity analysis in an organization. It is important to note that the structuring and conceptualization of complexity presented in our study is the first attempt to align diverse types of complexity.

For future work, two prominent avenues considering the limitations of the method can be highlighted. First, to have an up-to-date set of complexity analysis approaches, the SLR on complexity needs to be automated, as it is currently manual. For this, we aim to use automated systematic review solutions that employ machine learning technologies, for example, ASReview [117]. Such solutions may help to enrich results serving as a reusable basis to better deal with the researchers' bias. Second, we plan to conduct case studies and investigate the usefulness of the method in organizations, as it is currently demonstrated using an illustrative example from a real-life setting. Moreover, with case studies, we would like to collect the opinions of practitioners in terms of the fit of the method in their daily routine.

Author Contributions: Conceptualization, Aleksandra Revina; methodology, Aleksandra Revina; software, Ünal Aksu; formal analysis, Aleksandra Revina and Ünal Aksu; investigation, Aleksandra Revina and Ünal Aksu; resources, Vera G. Meister; data curation, Ünal Aksu; writing—original draft preparation, Aleksandra Revina and Ünal Aksu; writing—review and editing, Aleksandra Revina; visualization, Ünal Aksu; supervision, Vera G. Meister; project administration, Aleksandra Revina. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on https://github.com/complexityreview/comprehensiveOverview (accessed on 15 October 2021).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. List of Papers Analyzed in this Study

Below, in Table A1, we list the papers analyzed in this study. For the sake of simplicity and due to limited space, we only provide the most relevant features of the papers in addition to their title and authors. In particular, complexity type, motivation, input, and future research features are kept in the list. For the detailed list with all features and encoding, we refer to the project page on check out our project on GitHub.

Table A1. Analyzed papers in this study.

Reference	Complexity Type	Title	Motivation	Input	Future Research
[72]	Organizational	Task complexity and contingent processing in decision making: An information search and protocol analysis	Reasons for applying complexity metrics	Other; Case studies and interviews	Tool support
[46]	Organizational	Task complexity: Definition of the construct	Complexity metrics development	Other	Framework extension
[21]	Organizational	Task complexity: A review and analysis	Complexity factors; Complexity metrics development	Not Applicable	New validation studies; Approach extension
[118]	Organizational	Review of concepts and approaches to complexity	Complexity studies review	Not Applicable	Not Specified
[11]	Organizational	A model of the effects of audit task complexity	Complexity effects	Not Applicable	Approach extension
[119]	Organizational	Task complexity affects information seeking and use	Complexity effects	Other; Case studies and interviews	New validation studies
[120]	Organizational	The impact of knowledge and technology complexity on information systems development	Complexity metrics development	Software and Architectures	Approach extension; New validation studies
[68]	Organizational	Task and technology interaction (TTI): a theory of technological support for group tasks	Complexity metrics development	Other	New validation studies
[121]	Organizational	Perspective: Complexity theory and organization science	Complexity metrics development	Not Applicable	Approach extension
[84]	Organizational	Sources and assessment of complexity in NPD projects	Complexity factors	Business textual information	New validation studies
[81]	Organizational	Quantifying the Complexity of IT Service Management Processes	Complexity metrics development	Business textual information	Metrics extension; New validation studies
[122]	Organizational	Complexity of megaprojects	Complexity factors	Not Applicable	Not Specified
[80]	Organizational	Estimating business value of IT services through process complexity analysis	Reasons for applying complexity metrics	Business textual information	Approach extension; Tool support
[123]	Organizational	The inherent complexity of large scale engineering projects	Complexity metrics development	Not Applicable	Not Specified
[70]	Organizational	Complexity of Proceduralized Tasks	Complexity metrics development; Reasons for applying complexity metrics	Not Applicable	Not Specified
[66]	Organizational	Finding and reducing needless complexity	Complexity factors	Other	Not Specified
[79]	Organizational	Revisiting project complexity: Towards a comprehensive model of project complexity	Complexity metrics development	Not Applicable	Approach extension
[71]	Organizational	Model-based identification and use of task complexity factors of human integrated systems	Complexity factors	Other	Framework extension; Guidelines development
[10]	Organizational	Task complexity: A review and conceptualization framework	Complexity studies review; Complexity metrics development	Not Applicable	New validation studies; Framework extension
[85]	Organizational	Testing complexity index-a method for measuring perceived production complexity	Complexity effects	Business textual information	Not Specified
[83]	Organizational	The impact of business process complexity on business process standardization	Complexity factors	Other	Not Specified

Table A1. *Cont.*

Reference	Complexity Type	Title	Motivation	Input	Future Research
[124]	Organizational	Relationships between project complexity and communication	Complexity effects	Other; Case studies and interviews	Approach extension
[78]	Organizational	Building up a project complexity framework using an international Delphi study	Complexity metrics development	Other; Case studies and interviews	New validation studies
[76]	Organizational	An extended literature review of organizational factors impacting project management complexity	Complexity factors	Not Applicable	Not Specified
[125]	Organizational	Revisiting complexity in the digital age	Reasons for applying complexity metrics	Not Applicable	Not Specified
[77]	Organizational	Complexity in the Context of Systems Approach to Project Management	Complexity effects	Not Applicable	New validation studies; Framework extension
[126]	Organizational	Review of complexity drivers in enterprise	Complexity factors	Not Applicable	Not Specified
[45]	Organizational	Measurement model of project complexity for large-scale projects from task and organization perspective	Complexity metrics development	Business textual information	Metrics extension
[75]	Organizational	Work Autonomy and Workplace Creativity: Moderating Role of Task Complexity	Reasons for applying complexity metrics	Other; Case studies and interviews	New validation studies
[127]	Organizational	Managing complexity in service processes. The case of large business organizations	Complexity metrics development	Business textual information	New validation studies; Approach extension
[69]	Organizational	Modeling task complexity in crowdsourcing	Complexity factors	Business textual information	Approach extension
[128]	Organizational	Construction project complexity: research trends and implications	Complexity studies review	Not Applicable	Complexity factors; Complexity effects
[129]	Organizational	Revisiting complexity theory to achieve strategic intelligence	Reasons for applying complexity metrics	Other	Not Specified
[25]	Organizational	Revisiting task complexity: A comprehensive framework	Complexity metrics development	Other	Framework extension
[130]	Organizational	Complexity drivers in digitalized work systems: implications for cooperative forms of work	Complexity factors	Other; Case studies and interviews	Approach extension
[74]	Organizational	Robotic Process Automation Contemporary themes and challenges	Reasons for applying complexity metrics	Not Applicable	Complexity factors; Complexity effects; Guidelines development
[7]	Technological	A Complexity Measure	Complexity metrics development	Software and Architectures	Not Specified
[131]	Technological	A measure of control flow complexity in program text	Complexity metrics development	Software and Architectures	Not Specified
[132]	Technological	Software structure metrics based on information flow	Reasons for applying complexity metrics; Complexity metrics development	Software and Architectures	Approach extension
[133]	Technological	Measuring the quality of structured designs	Reasons for applying complexity metrics; Complexity metrics development	Software and Architectures	Metrics extension; New validation studies

Table A1. *Cont.*

Reference	Complexity Type	Title	Motivation	Input	Future Research
[110]	Technological	An empirical study of a syntactic complexity family	Complexity metrics development	Software and Architectures	Not Specified
[134]	Technological	System structure and software maintenance performance	Complexity effects	Software and Architectures	Approach extension; New validation studies
[108]	Technological	Verifying, validating, and measuring the performance of expert systems	Complexity effects	Software and Architectures	Approach extension; Approach implementation
[135]	Technological	Software complexity and maintenance costs	Complexity effects	Software and Architectures	Not Specified
[136]	Technological	Complexity metrics for rule-based expert systems	Complexity metrics analysis	Software and Architectures	New validation studies
[137]	Technological	An information theory-based approach for quantitative evaluation of user interface complexity	Complexity metrics development	Software and Architectures	New validation studies
[138]	Technological	Software metrics by architectural pattern mining	Reasons for applying complexity metrics; Complexity metrics development	Software and Architectures	Tool support
[86]	Technological	Finding a complexity measure for business process models	Complexity metrics development	Not Applicable	Metrics extension
[139]	Technological	A new measure of software complexity based on cognitive weights	Complexity metrics development	Software and Architectures	Not Specified
[140]	Technological	Measures of information complexity and the implications for automation design	Complexity metrics development	Not Applicable	New validation studies
[141]	Technological	Complexity and Automation Displays of Air Traffic Control: Literature Review and Analysis	Complexity metrics development	Not Applicable	Metrics extension
[47]	Technological	A discourse on complexity of process models	Complexity metrics analysis; Complexity metrics development	Not Applicable	New validation studies
[142]	Technological	Business process quality metrics: Log-based complexity of workflow patterns	Complexity metrics development	Event log; workflows	Not Specified
[105]	Technological	Complexity analysis of BPEL web processes	Complexity factors	Event log; workflows	New validation studies
[143]	Technological	Approaches for business process model complexity metrics	Complexity metrics development	Software and Architectures	Metrics extension
[144]	Technological	Error Metrics for Business Process Models.	Reasons for applying complexity metrics	Business process models	New validation studies
[145]	Technological	A weighted coupling metric for business process models.	Complexity metrics development	Business process models	New validation studies
[90]	Technological	A metric for ERP complexity	Complexity metrics analysis	Not Applicable	Tool support
[146]	Technological	Business process control-flow complexity: Metric, evaluation, and validation	Complexity metrics development	Event log; workflows	Not Specified
[147]	Technological	Evaluating workflow process designs using cohesion and coupling metrics	Complexity metrics development; Reasons for applying complexity metrics	Business process models	Approach extension; New validation studies

Table A1. *Cont.*

Reference	Complexity Type	Title	Motivation	Input	Future Research
[148]	Technological	On a quest for good process models: the cross-connectivity metric	Complexity metrics development	Business process models	New validation studies; Complexity factors; Guidelines development
[149]	Technological	Complex network model for software system and complexity measurement	Complexity metrics development	Software and Architectures	Metrics extension; Approach extension
[150]	Technological	Complexity metrics for Workflow nets	Complexity metrics development	Business process models	New validation studies
[103]	Technological	A Survey of Business Process Complexity Metrics	Complexity studies review	Not Applicable	Metrics extension; New validation studies; Tool support
[151]	Technological	Prediction of business process model quality based on structural metrics	Complexity metrics analysis	Business process models	New validation studies
[107]	Technological	Enterprise systems complexity and its antecedents: a grounded-theory approach	Complexity factors	Software and Architectures	Complexity effects; New validation studies
[152]	Technological	Optimizing the trade-off between complexity and conformance in process reduction	Complexity effects	Business process models	New validation studies
[153]	Technological	A simpler model of software readability	Complexity metrics development	Software and Architectures; Case studies and interviews	Not Specified
[154]	Technological	Integrated framework for business process complexity analysis	Complexity metrics development	Business textual information	Metrics extension; New validation studies
[155]	Technological	Complexity in Enterprise Architectures - Conceptualization and Introduction of a Measure from a System Theoretic Perspective	Complexity metrics development	Software and Architectures	Approach extension; New validation studies
[109]	Technological	GUIEvaluator: A Metric-tool for Evaluating the Complexity of Graphical User Interfaces.	Complexity metrics development	Not Applicable	Metrics extension; Metrics comparison
[156]	Technological	Examining case management demand using event log complexity metrics	Complexity metrics development	Event log; workflows	Metrics extension; New validation studies; Tool support
[106]	Technological	A complexity theory approach to IT-enabled services (IESs) and service innovation: Business analytics as an illustration of IES	Complexity metrics development	Not Applicable	New validation studies
[157]	Technological	Square complexity metrics for business process models	Complexity metrics development	Business process models	New validation studies
[158]	Technological	Quantification of interface visual complexity	Complexity metrics development	Other	Metrics extension; Tool support
[5]	Technological	Adopting Notions of Complexity for Enterprise Architecture Management	Complexity metrics development	Not Applicable	Not Specified
[159]	Technological	An exploratory study on the relation between user interface complexity and the perceived quality	Complexity effects	Software and Architectures	New validation studies
[160]	Technological	A systematic literature review of studies on business process modeling quality	Complexity studies review	Not Applicable	Framework extension

Table A1. *Cont.*

Reference	Complexity Type	Title	Motivation	Input	Future Research
[161]	Technological	Metrics and performance indicators to evaluate workflow processes on the cloud	Complexity measurements	Event log; workflows	Approach extension; New validation studies
[162]	Technological	Measuring complexity of business process models integrated with rules	Complexity metrics development	Business process models	Not Specified
[163]	Technological	Metrics for the case management modeling and notation (CMMN) specification	Complexity metrics development	Business process models	New validation studies
[88]	Technological	UI-CAT: calculating user interface complexity metrics for mobile applications	Complexity measurements	Event log; Software and Architectures	Not Specified
[164]	Technological	Complexity-aware generation of workflows by process-oriented case-based reasoning	Reasons for applying complexity metrics	Event log; workflows	Metrics extension; New validation studies
[165]	Technological	How visual cognition influences process model comprehension	Complexity metrics development	Business process models	New validation studies
[29]	Technological	Complexity metrics for process models-A systematic literature review	Complexity studies review	Not Applicable	Metrics extension
[166]	Technological	Decision support for reducing unnecessary IT complexity of application architectures	Complexity factors	Software and Architectures	Tool support
[104]	Technological	Dealing with Process Complexity: A Multiperspective Approach	Complexity factors	Business process models	New validation studies
[167]	Technological	Towards understanding code readability and its impact on design quality	Complexity effects	Software and Architectures	Metrics extension
[168]	Technological	Integrating Business Process Models with Rules	Reasons for applying complexity metrics	Business process models	Not Specified
[169]	Technological	Complexity metrics for DMN decision models	Complexity metrics analysis	Business process models	Approach extension
[170]	Textual	Accounting narratives: A review of empirical studies of content and readability	Complexity studies review	Not Applicable	New validation studies; Approach extension
[91]	Textual	Readability of annual reports: Western versus Asian evidence	Complexity factors	Business textual information	Not Specified
[171]	Textual	Readability of annual reports: Western versus Asian evidence-a comment to contexualize	Complexity studies review	Not Applicable	New validation studies
[95]	Textual	The application of the marketing concept in textbook selection: Using the Cloze procedure	Reasons for applying complexity metrics	Business textual information	Metrics extension
[172]	Textual	Annual report readability variability: tests of the obfuscation hypothesis	Complexity factors	Business textual information	Metrics extension; New validation studies
[173]	Textual	Communication in auditors' reports: Variations in readability and the effect of audit firm structure	Complexity factors	Business textual information	Not Specified
[174]	Textual	A texture index for evaluating accounting narratives	Complexity metrics development	Business textual information	New validation studies
[175]	Textual	The effect of thematic structure on the variability of annual report readability	Complexity factors	Business textual information	New validation studies

Table A1. *Cont.*

Reference	Complexity Type	Title	Motivation	Input	Future Research
[176]	Textual	An approach to evaluating accounting narratives: a corporate social responsibility perspective	Reasons for applying complexity metrics	Business textual information	Metrics extension
[177]	Textual	E-comprehension: Evaluating B2B websites using readability formulae	Complexity measurements	Business textual information	Approach extension
[178]	Textual	Obfuscation, textual complexity and the role of regulated narrative accounting disclosure in corporate governance	Reasons for applying complexity metrics	Business textual information	Not Specified
[179]	Textual	Evaluating a measure of content quality for accounting narratives (with an empirical application to narratives from Australia, Hong Kong, and the United States)	Complexity factors	Business textual information	Approach extension
[180]	Textual	Readability of corporate annual reports of top 100 Malaysian companies	Complexity factors	Business textual information	New validation studies
[181]	Textual	Readability of financial statement footnotes of Kuwaiti corporations	Complexity measurements	Business textual information	New validation studies
[182]	Textual	Voluntary narrative disclosures by local governments: A comparative analysis of the textual complexity of mayoral and chairpersons' letters in annual reports	Complexity factors	Business textual information	Approach extension; New validation studies
[183]	Textual	The textual complexity of annual report narratives: A comparison of high-and low-performance companies	Complexity effects	Business textual information	New validation studies; Approach extension
[92]	Textual	Are business school mission statements readable?: Evidence from the top 100	Complexity measurements	Business textual information	Not Specified
[184]	Textual	Enhancing compliance through improved readability: Evidence from New Zealand's rewrite "experiment"	Reasons for applying complexity metrics	Business textual information	New validation studies
[185]	Textual	How readable are mission statements? An exploratory study	Complexity measurements	Business textual information	Approach extension; New validation studies
[186]	Textual	Readability of accountants' communications with small business—Some Australian evidence	Complexity measurements	Business textual information	Metrics extension; New validation studies
[111]	Textual	The readability of managerial accounting and financial management textbooks	Reasons for applying complexity metrics	Business textual information	New validation studies
[187]	Textual	Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content	Complexity effects	Business textual information	Metrics extension; New validation studies

Table A1. *Cont.*

Reference	Complexity Type	Title	Motivation	Input	Future Research
[93]	Textual	Reading between the vines: analyzing the readability of consumer brand wine web sites	Complexity measurements	Business textual information	Approach extension; New validation studies
[188]	Textual	Essays on the issues of readability in business disciplines	Complexity studies review	Not Applicable	Approach extension
[96]	Textual	Revisiting the role of linguistic complexity in ESL reading comprehension	Complexity factors	Business textual information	Not Specified
[189]	Textual	Textual complexity of standard conditions used in the construction industry	Complexity factors	Business textual information	Not Specified
[190]	Textual	Tourism websites in the Middle East: readable or not?	Complexity measurements	Business textual information	New validation studies; Approach extension
[191]	Textual	Developing the Flesch reading ease formula for the contemporary accounting communications landscape	Complexity metrics analysis; Complexity metrics development	Not Applicable	Metrics extension
[192]	Textual	Text complexity: State of the art and the conundrums it raises	Complexity studies review	Not Applicable	Approach extension
[193]	Textual	Traditional and alternative methods of measuring the understandability of accounting narratives	Complexity metrics analysis	Business textual information	New validation studies
[97]	Textual	When complexity becomes interesting	Complexity effects	Business textual information	New validation studies
[194]	Textual	Readability and Thematic Manipulation in Corporate Communications: A Multi-Disclosure Investigation	Reasons for applying complexity metrics	Business textual information	New validation studies
[195]	Textual	Guiding through the Fog: Does annual report readability reveal earnings management?	Complexity effects	Business textual information	Metrics extension
[196]	Textual	From Accountability to Readability in the Public Sector: Evidence from Italian Universities	Complexity factors	Business textual information	Approach extension
[197]	Textual	The readability of integrated reports	Complexity measurements	Business textual information	Metrics extension; New validation studies
[198]	Textual	Readability of Mission Statements: A Look at Fortune 500	Complexity measurements	Business textual information	Metrics extension; New validation studies
[199]	Textual	Assessing social and environmental performance through narrative complexity in CSR reports	Reasons for applying complexity metrics	Business textual information	New validation studies
[200]	Textual	A conceptual model for measuring the complexity of spreadsheets	Complexity metrics development	Other	Metrics extension
[201]	Textual	The influence of business strategy on annual report readability	Complexity factors	Business textual information	New validation studies
[94]	Textual	Roles of review numerical and textual characteristics on review helpfulness across three different types of reviews	Complexity effects	Business textual information	Approach extension; New validation studies

References

1. Benbya, H.; Nan, N.; Tanriverdi, H.; Yoo, Y. Complexity and Information Systems Research in the Emerging Digital World. *MIS Q.* **2020**, *44*, 1–17.
2. Törnblom, O. Managing Complexity in Organizations: Analyzing and Discussing a Managerial Perspective on the Nature of Organizational Leadership. *Behav. Dev.* **2018**, *23*, 51–62.
3. Legner, C.; Eymann, T.; Hess, T.; Matt, C.; Böhmann, T.; Drews, P.; Maedche, A.; Urbach, N.; Ahlemann, F. Digitalization: Opportunity and Challenge for the Business and Information Systems Engineering Community. *Bus. Inf. Syst. Eng.* **2017**, *59*, 301–308.
4. Essaides, N. The Impact of Organizational Complexity on Finance Performance. Available online: thehackettgroup.com/blog/the-impact-of-organizational-complexity-on-finance-performance/ (accessed on 15 August 2021).
5. Schneider, A.; Zec, M.; Matthes, F. Adopting Notions of Complexity for Enterprise Architecture Management. In Proceedings of the 20th Americas Conference on Information Systems, Savannah, Georgia, USA, 7–9 August 2014; Association for Information Systems: Atlanta, GA, USA, 2014; pp. 1–10.
6. Bergman, T.J.; Beehner, J.C. Measuring social complexity. *Anim. Behav.* **2015**, *103*, 203–209.
7. McCabe, T.J. A Complexity Measure. *IEEE Trans. Softw. Eng.* **1976**, *2*, 308–320.
8. Chronéer, D.; Bergquist, B. Managerial Complexity in Process Industrial R&D Projects: A Swedish Study. *Proj. Manag. J.* **2012**, *43*, 21–36.
9. Daryani, S.M.; Amini, A. Management and Organizational Complexity. *Procedia—Soc. Behav. Sci.* **2016**, *230*, 359–366.
10. Liu, P.; Li, Z. Task complexity: A review and conceptualization framework. *Int. J. Ind. Ergon.* **2012**, *42*, 553–568.
11. Bonner, S.E. A model of the effects of audit task complexity. *Account. Organ. Soc.* **1994**, *19*, 213–234.
12. Kohr, D.; Budde, L.; Friedli, T. Identifying Complexity Drivers in Discrete Manufacturing and Process Industry. *Procedia CIRP* **2017**, *63*, 52–57.
13. Leavitt, H. Applied organizational change in industry: Structural, technological, and humanistic approaches. In *Handbook of Organizations*; Routledge: London, UK, 1965; Volume 264.
14. Prodan, M.; Prodan, A.; Purcarea, A. Three new dimensions to people, process, technology improvement model. In *New Contributions in Information Systems and Technologies*; Springer: Cham, Switzerland, 2015; Volume 353, pp. 481–490.
15. Zwicky, F. *Discovery, Invention, Research through the Morphological Approach*; The Macmillian Company: Toronto, ON, Canada, 1969.
16. van Solingen, R.; Basili, V.; Caldiera, G.; Rombach, D. Goal question metric (GQM) approach. In *Encyclopedia of Software Engineering*; Wiley: New York, USA, 2002.
17. Arora, S.; Barak, B. *Computational Complexity—A Modern Approach*; Cambridge University Press: Cambridge, UK, 2009.
18. Grobman, G. Complexity theory: A new way to look at organizational change. *Public Adm. Q.* **2005**, *29*, 350–382.
19. Miestamo, M.; Sinnemäki, K.; Karlsson, F. In *Language Complexity: Typology, Contact, Change*; John Benjamins Publishing: Amsterdam, The Netherlands, 2008.Volume 94.
20. Edmonds, B. What is complexity? The philosophy of complexity per se with application to some examples in evolution. In *The Evolution of Complexity*; Kluwer: Dordrecht, The Netherlands, 1995.
21. Campbell, D. Task complexity: A review and analysis. *Acad. Manag. Rev.* **1988**, *13*, 40–52.
22. Poutanen, P.; Soliman, W.; Ståhle, P. The complexity of innovation: An assessment and review of the complexity perspective. *Eur. J. Innov. Manag.* **2016**, *19*, 189–213.
23. Milch, V.; Laumann, K. Interorganizational complexity and organizational accident risk: A literature review. *Saf. Sci.* **2016**, *82*, 9–17.
24. Vogel, W.; Lasch, R. Complexity drivers in manufacturing companies: A literature review. *Logist. Res.* **2016**, *9*, 25:1–25:66.
25. Efatmaneshnik, M.; Handley, H. Revisiting task complexity: A comprehensive framework. In Proceedings of the 2018 Annual IEEE International Systems Conference, Vancouver, BC, Canada, 23–26 April 2018; pp. 1–4.
26. Hvam, L.; Hansen, C.L.; Forza, C.; Mortensen, N.H.; Haug, A. The reduction of product and process complexity based on the quantification of product complexity costs. *Int. J. Prod. Res.* **2020**, *58*, 350–366.
27. Augusto, A.; Mendling, J.; Vidgof, M.; Wurm, B. The Connection between Process Complexity of Event Sequences and Models discovered by Process Mining. *arXiv* **2021**, arXiv:2106.07990.
28. Fernández-Cerero, D.; Varela-Vaca, Á.J.; Fernández-Montes, A.; López, M.T.G.; Álvarez-Bermejo, J.A. Measuring data-centre workflows complexity through process mining: the Google cluster case. *J. Supercomput.* **2020**, *76*, 2449–2478.
29. Polancic, G.; Cegnar, B. Complexity metrics for process models—A systematic literature review. *Comput. Stand. Interfaces* **2017**, *51*, 104–117.
30. Rizun, N.; Revina, A.; Meister, V.G. Assessing business process complexity based on textual data: Evidence from ITIL IT ticket processing. *Bus. Process. Manag. J.* **2021**, doi:10.1108/BPMJ-04-2021-0217.
31. Damasiotis, V.; Fitsilis, P.; O'Kane, J.F. Modeling Software Development Process Complexity. *Int. J. Inf. Technol. Proj. Manag.* **2018**, *9*, 17–40.
32. Zuse, H. *Software Complexity: Measures and Methods*; W. de Gruyter: New York, NY, USA, 1991.

33. Antinyan, V.; Staron, M.; Sandberg, A. Evaluating code complexity triggers, use of complexity measures and the influence of code complexity on maintenance time. *Empir. Softw. Eng.* **2017**, *22*, 3057–3087.
34. Bosch-Rekveldt, M.; Bakker, H.; Hertogh, M. Comparing Project Complexity across Different Industry Sectors. *Complexity* **2018**, *2018*, 3246508.
35. Morcov, S.; Pintelon, L.; Kusters, R. Definitions, characteristics and measures of IT Project Complexity—A Systematic Literature Review. *Int. J. Inf. Syst. Proj. Manag.* **2020**, *8*, 5–21.
36. Camposa, P.F.; Truccoa, P.; Huatuco, L.H. Managing structural and dynamic complexity in supply chains: Insights from four case studies. *Prod. Plan. Control* **2019**, *30*, 611–623.
37. Aitken, J.; Bozarth, C.; Garn, W. To eliminate or absorb supply chain complexity: A conceptual model and case study. *Prod. Plan. Control* **2016**, *21*, 759–774.
38. Reeves, M.; Levin, S.; Fink, T.; Levina, A. Taming Complexity. Make Sure the Benefits of Any Addition to an Organization's Systems Outweigh Its Costs. Published January-February 2020. Available online: https://hbr.org/2020/01/taming-complexity (accessed on 9 October 2021).
39. Uhl-Bien, M.; Arena, M. Complexity leadership: Enabling people and organizations for adaptability. *Organ. Dyn.* **2017**, *46*, 4–20.
40. Watkins, D.; Earnhardt, M.; Pittenger, L.; Roberts, R.; Rietsema, K.; Cosman-Ross, J. Thriving in Complexity: A Framework for Leadership Education. *J. Leadersh. Educ.* **2017**, *16*, 148.
41. Afsar, B.; Umrani, W.A. Transformational leadership and innovative work behavior: The role of motivation to learn, task complexity and innovation climate. *Eur. J. Innov. Manag.* **2020**, *23*, 402–428.
42. Axelos. *Information Technology Infrastructure Library*; Technical Report; Axelos: The Stationery Office: London, UK, 2019.
43. Mintzberg, H. The Structuring of Organizations. In *Readings in Strategic Management*; Macmillan Education UK: London, UK, 1989; pp. 322–352.
44. Lunenburg, F.C. Organizational Structure: Mintzberg's Framework. *Int. J. Sch. Acad. Intellect. Divers.* **2012**, *14*, 1–7.
45. Lu, Y.; Luo, L.; Wang, H.; Le, Y.; Shi, Q. Measurement model of project complexity for large-scale projects from task and organization perspective. *Int. J. Proj. Manag.* **2015**, *33*, 610–622.
46. Wood, R. Task complexity: Definition of the construct. *Organ. Behav. Hum. Decis. Process.* **1986**, *37*, 60–82.
47. Cardoso, J.; Mendling, J.; Neumann, G.; Reijers, H. A discourse on complexity of process models. In Proceedings of the International Conference on Business Process Management, Vienna, Austria, 4–7 September 2006; Springer: Berlin/Heidelberg, Germany, 2006, pp. 117–128.
48. Rizkallah, J. The Big (Unstructured) Data Problem. Available online: www.forbes.com/sites/forbestechcouncil/2017/06/05/the-big-unstructured-data-problem/#1ccea420493a (accessed on 15 August 2021).
49. Revina, A. Business Process Management: Integrated Data Perspective. A Framework and Research Agenda. In Proceedings of the Information Systems Development: Crossing Boundaries between Development and Operations (DevOps) in Information Systems (ISD2021 Proceedings), Valencia, Spain, 8–10 September 2021; Association for Information Systems: Atlanta, GA, USA, 2021.
50. Revina, A. Considering Business Process Complexity through the Lens of Textual Data. In Proceedings of the Sixteenth International Multi-Conference on Computing in the Global Information Technology (ICCGI 2021) Proceedings), Nice, France, 18–22 July 2021; ThinkMind(TM), 2021.
51. Kitchenham, B. *Procedures for Performing Systematic Reviews*; Technical Report; Keele University: Keele, UK, 2004; #TR/SE-0401.
52. Templier, M.; Paré, G. A framework for guiding and evaluating literature reviews. *Commun. Assoc. Inf. Syst.* **2015**, *37*, 113–137.
53. Bandara, W.; Furtmueller, E.; Gorbacheva, E.; Miskon, S.; Beekhuyzen, J. Achieving rigor in literature reviews: Insights from qualitative data analysis and tool-support. *Commun. Assoc. Inf. Syst.* **2015**, *34*, 155–204.
54. Gusenbauer, M. Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics* **2019**, *118*, 177–214.
55. Schoknecht, A.; Thaler, T.; Fettke, P.; Oberweis, A.; Laue, R. Similarity of business process models—A state-of-the-art analysis. *ACM Comput. Surv. (CSUR)* **2017**, *50*, 1–33.
56. Thaler, T.; Ternis, S.F.; Fettke, P.; Loos, P. A Comparative Analysis of Process Instance Cluster Techniques. *Wirtschaftsinformatik* **2015**, *2015*, 423–437.
57. Kaplan, R.; Norton, D. The Balanced Scorecard: Measures That Drive Performance. Available online: hbr.org/1992/01/the-balanced-scorecard-measures-that-drive-performance-2 (accessed on 15 August 2021).
58. Buglione, L.; Abran, A. Balanced Scorecards and GQM: What are the differences. In Proceedings of the Third European Software Measurement Conference (FESMA-AEMES 2000), Madrid, Spain, 18–20 October 2000; pp. 18–20.
59. Martinsons, M.; Davison, R.; Tse, D. The balanced scorecard: A foundation for the strategic management of information systems. *Decis. Support Syst.* **1999**, *25*, 71–88.
60. Ghani, A.A.; Wei, K.T.; Muketha, G.M.; Wen, W.P. Complexity metrics for measuring the understandability and maintainability of business process models using goal-question-metric (GQM). *Int. J. Comput. Sci. Netw. Secur.* **2008**, *8*, 219–225.
61. Cyra, Ł.; Górski, J. Extending GQM by argument structures. In Proceedings of the IFIP Central and East European Conference on Software Engineering Techniques, Poznan, Poland, 10–12 October 2007; Springer: Berlin/Heidelberg, Germany, 2007; pp. 26–39.
62. Fenton, N.; Pfleeger, S. *Software Metrics: A Rigorous and Practical Approach*; PWS Publishing Co.: Boston, MA, USA, 1997.

63. Berander, P.; Jönsson, P. A Goal Question Metric Based Approach for Efficient Measurement Framework Definition. In Proceedings of the 2006 ACM/IEEE International Symposium on Empirical Software Engineering, Rio de Janeiro, Brazil, 21–22 September 2006; Association for Computing Machinery: New York, USA, 2006; pp. 316–325.
64. Sanaa, H.; Afifi, W.; Darwish, N.R. The goal questions metrics for agile business intelligence. *Egypt. Comput. Sci. J.* **2016**, *40*, 24–42.
65. Jarke, M.; Jeusfeld, M.A.; Quix, C.; Vassiliadis, P. Architecture and Quality in Data Warehouses: An Extended Repository Approach. *Inf. Syst.* **1999**, *24*, 229–253.
66. Olson, E.G.; Reger, S.J.M.; Singer, D.D. Finding and reducing needless complexity. *Horizon* **2010**, *18*, 53–61.
67. Murthy, U.; Kerr, D. Comparing audit team effectiveness via alternative modes of computer-mediated communication. *Audit. J. Pract. Theory* **2004**, *23*, 141–152.
68. Rana, A.; Turoff, M.; Hiltz, S.R. Task and technology interaction (TTI): A theory of technological support for group tasks. In Proceedings of the Thirtieth Hawaii International Conference on System Sciences, Wailea, HI, USA, 7–10 January 1997; Volume 2, pp. 66–75.
69. Yang, J.; Redi, J.; Demartini, G.; Bozzon, A. Modeling task complexity in crowdsourcing. In Proceedings of the Fourth AAAI Conference on Human Computation and Crowdsourcing, Austin, TX, USA, 30 October–3 November 2016; AAAI Press: Menlo Park, CA, USA, 2016, pp. 249–258.
70. Park, J. Complexity of Proceduralized Tasks. In *Series in Reliability Engineering*; Springer: London, UK, 2009; pp. 13–21.
71. Ham, D.; Park, J.; Jung, W. Model-based identification and use of task complexity factors of human integrated systems. *Reliab. Eng. Syst. Saf.* **2012**, *100*, 33–47.
72. Payne, J. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organ. Behav. Hum. Perform.* **1976**, *16*, 366–387.
73. Saastamoinen, M.; Kumpulainen, S.; Vakkari, P.; Järvelin, K. Task complexity affects information use: A questionnaire study in city administration. *Inf. Res.* **2013**, *19*, 592.
74. Syed, R.; Suriadi, S.; Adams, M.; Bandara, W.; Leemans, S.; Ouyang, C.; ter Hofstede, A.; van de Weerd, I.; Wynn, M.T.; Reijers, H. Robotic Process Automation: Contemporary themes and challenges. *Comput. Ind.* **2020**, *115*, 103162.
75. Sia, S.K.; Appu, A.V. Work Autonomy and Workplace Creativity: Moderating Role of Task Complexity. *Glob. Bus. Rev.* **2015**, *16*, 772–784.
76. Gutierrez, C.; Hussein, B. An extended literature review of organizational factors impacting project management complexity. In Proceedings of the 28th IPMA World Congress: Innovation through Dialogue (IPMA), Rotterdam, Netherlands, 29 September–1 October 2014; IPMA: Amsterdam, Netherlands, 2014; pp. 1–10.
77. Botchkarev, A.; Finnigan, P. Complexity in the Context of Systems Approach to Project Management. *Organ. Proj. Manag.* **2015**, *2*, 15–34.
78. Vidal, L.A.; Marle, F.; Bocquet, J.C. Building up a project complexity framework using an international Delphi study. *Int. J. Technol. Manag.* **2013**, *62*, 251–283.
79. Gul, S.; Khan, S. Revisiting project complexity: Towards a comprehensive model of project complexity. In Proceedings of the 2nd International Conference on Construction and Project Management, Singapore, 16–18t September 2011; IACSIT Press: Singapore, 2011; pp. 148–155.
80. Diao, Y.; Bhattacharya, K. Estimating business value of IT services through process complexity analysis. In Proceedings of the IEEE Network Operations and Management Symposium, Salvador, Brazil, 7–11 April 2008; IEEE: New York, USA, 2008; pp. 208–215.
81. Diao, Y.; Keller, A. Quantifying the Complexity of IT Service Management Processes. In Proceedings of the 17th IFIP/IEEE International Workshop on Distributed Systems: Operations and Management, Dublin, Ireland, 23–25 October, 2006; Springer: Berlin/Heidelberg, Germany, 2006; Volume 4269, pp. 61–73.
82. Rennung, F.; Luminosu, C.; Draghici, A.; Paschek, D. An Evaluation of Strategic Methods of Complexity Management to Manage Large Outsourcing Projects Successfully. In *Managing Innovation and Diversity in Knowledge Society through Turbulent Time, Proceedings of the MakeLearn and TIIM Joint International Conference, Timisoara, Romania, 25–27 May 2016*; ToKnowPress: Bangkok/Celje/Lublin, 2016; pp. 79–88.
83. Schäfermeyer, M.; Rosenkranz, C.; Holten, R. The impact of business process complexity on business process standardization. *Bus. Inf. Syst. Eng.* **2012**, *4*, 261–270.
84. Kim, J.; Wilemon, D. Sources and assessment of complexity in NPD projects. *R D Manag.* **2003**, *33*, 15–30.
85. Mattsson, S.; Gullander, P.; Harlin, U.; Bäckstrand, G.; Fasth, Å.; Davidsson, A. Testing complexity index—A method for measuring perceived production complexity. *Procedia CIRP* **2012**, *3*, 394–399.
86. Latva-Koivisto, A. *Finding a Complexity Measure for Business Process Models*; Technical Report; Helsinki University of Technology, Systems Analysis Laboratory: Helsinki, 2001.
87. Hinz, G.; Chen, G.; Aafaque, M.; Röhrbein, F.; Conradt, J.; Bing, Z.; Qu, Z.; Stechele, W.; Knoll, A. Online multi-object tracking-by-clustering for intelligent transportation system with neuromorphic vision sensor. In Proceedings of the Joint German/Austrian Conference on Artificial Intelligence (Künstliche Intelligenz), Dortmund, Germany, 25–29 September 2017; Springer: Cham, Switzerland, 2017; pp. 142–154.

88. Riegler, A.; Holzmann, C. UI-CAT: Calculating user interface complexity metrics for mobile applications. In Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia, Linz, Austria, 30 November–2 December 2015; Association for Computing Machinery: New York, USA, 2015; pp. 390–394.
89. Chen, Z.; Suen, C. Measuring the complexity of rule-based expert systems. *Expert Syst. Appl.* **1994**, *7*, 467–481.
90. Bansal, V.; Negi, T. A metric for ERP complexity. In Proceedings of the International Conference on Business Information Systems, Innsbruck, Austria, 5–7 May, 2008; Springer: Berlin/Heidelberg, Germany, 2008; pp. 369–379.
91. Courtis, J.K. Readability of annual reports: Western versus Asian evidence. *Account. Audit. Account. J.* **1995**, *8*, 4–17.
92. Pitt, L.; Sattari, S.; Bevelander, D. Are business school mission statements readable?: Evidence from the top 100. *J. Strateg. Manag. Educ.* **2010**, *6*, 1–16.
93. Mills, A.; Pitt, L.; Sattari, S. Reading between the vines: Analyzing the readability of consumer brand wine web sites. *Int. J. Wine Bus. Res.* **2012**, *24*, 169–182.
94. Zhou, Y.; Yang, S. Roles of review numerical and textual characteristics on review helpfulness across three different types of reviews. *IEEE Access* **2019**, *7*, 27769–27780.
95. Rugimbana, R.; Patel, C. The application of the marketing concept in textbook selection: Using the Cloze procedure. *J. Mark. Educ.* **1996**, *18*, 14–20.
96. Barrot, J.S. Revisiting the role of linguistic complexity in ESL reading comprehension. *3L Southeast Asian J. Engl. Lang. Stud.* **2013**, *19*, 5–18.
97. van der Sluis, F.; van den Broek, E.; Glassey, R.; van Dijk, E.; de Jong, F. When complexity becomes interesting. *J. Assoc. Inf. Sci. Technol.* **2014**, *65*, 1478–1500.
98. Eurostat. *Statistical Classification of Economic Activities in the European Community*; Technical Report; Eurostat: Luxembourg, 2020.
99. Serrano, M.; Calero, C.; Piattini, M. Validating metrics for data warehouses. *IEE Proc.-Softw.* **2002**, *149*, 161–166.
100. Devpriya, S.; Ritu, S.; Kumar, M. A Framework for Validation of Object-Oriented Design Metrics. *Int. J. Comput. Sci. Inf. Secur.* **2009**, *6*, 46–52.
101. Briand, L.; Differding, C.; Rombach, D. Practical guidelines for measurement-based process improvement. *Softw. Process. Improv. Pract.* **1996**, *2*, 253–280.
102. Weyuker, E. Evaluating software complexity measures. *IEEE Trans. Softw. Eng.* **1988**, *14*, 1357–1365.
103. Muketha, G.; Ghani, A.A.A.; Selamat, M.H.; Atan, R. A Survey of Business Process Complexity Metrics. *Inf. Technol. J.* **2010**, *9*, 1336–1344.
104. Krenn, F. Dealing with Process Complexity: A Multiperspective Approach. In Proceedings of the 10th International Conference on Subject-Oriented Business Process Management, Linz, Austria, 05–06 April, 2018; Association for Computing Machinery: New York, USA, 2018; pp. 1–10.
105. Cardoso, J. Complexity analysis of BPEL web processes. *Softw. Process. Improv. Pract.* **2007**, *12*, 35–49.
106. Chae, B. A complexity theory approach to IT-enabled services (IESs) and service innovation: Business analytics as an illustration of IES. *Decis. Support Syst.* **2014**, *57*, 1–10.
107. Schoenherr, T.; Hilpert, D.; Soni, A.; Venkataramanan, M.; Mabert, V. Enterprise systems complexity and its antecedents: A grounded-theory approach. *Int. J. Oper. Prod. Manag.* **2010**, *30*, 639–668.
108. Suen, C.Y.; Grogono, P.D.; Shinghal, R.; Coallier, F. Verifying, validating, and measuring the performance of expert systems. *Expert Syst. Appl.* **1990**, *1*, 93–102.
109. Alemerien, K.; Magel, K. GUIEvaluator: A Metric-tool for Evaluating the Complexity of Graphical User Interfaces. In Proceedings of the 26th International Conference on Software Engineering and Knowledge Engineering, Hyatt Regency, Vancouver, Canada, 1–3 July, 2013; KSI Research Inc.: Skokie, IL, USA, 2014; pp. 13–18.
110. Basili, V.; Hutchens, D. An empirical study of a syntactic complexity family. *IEEE Trans. Softw. Eng.* **1983**, *SE-9*, 664–672.
111. Bargate, K. The readability of managerial accounting and financial management textbooks. *Meditari Account. Res.* **2012**, *20*, 4–20.
112. Leopold, H.; van der Aa, H.; Reijers, H. Identifying candidate tasks for robotic process automation in textual process descriptions. In *Enterprise, Business-Process and Information Systems Modeling, Proceedings of the BPMDS 2018, EMMSAD 2018, Tallinn, Estonia, 11–12 June 2018*; Gulden, J., Reinhartz-Berger, I., Schmidt, R., Guerreiro, S., Guedria, W., Bera, P., Eds.; Lecture Notes in Business Information Processing; Springer: Cham, Switzerland, 2018, pp. 67–81.
113. Baccarini, D. The concept of project complexity—A review. *Int. J. Proj. Manag.* **1996**, *14*, 201–204.
114. Koziolek, H. Goal, question, metric. In *Dependability Metrics*; Springer: Cham, Switzerland, 2008; pp. 39–42.
115. Kraus, S.; Breier, M.; Dasí-Rodríguez, S. The art of crafting a systematic literature review in entrepreneurship research. *Int. Entrep. Manag. J.* **2020**, *16*, 1023–1042.
116. Kilpi, T. Implementing a software metrics program at Nokia. *IEEE Softw.* **2001**, *18*, 72–77.
117. van de Schoot, R.; de Bruin, J.; Schram, R.; Zahedi, P.; de Boer, J.; Weijdema, F.; Kramer, B.; Huijts, M.; Hoogerwerf, M.; Ferdinands, G.; et al. An open source machine learning framework for efficient and transparent systematic reviews. *Nat. Mach. Intell.* **2021**, *3*, 125–133.
118. Modrick, J. Review of concepts and approaches to complexity. In Proceedings of the Human Factors Society Annual Meeting, Atlanta, Georgia, USA, 12–16 October, 1992; SAGE Publications Sage CA: Los Angeles, CA, 1992; Volume 36, pp. 1166–1170.
119. Byström, K.; Järvelin, K. Task complexity affects information seeking and use. *Inf. Process. Manag.* **1995**, *31*, 191–213.

120. Meyer, M.; Curley, K.F. The impact of knowledge and technology complexity on information systems development. *Expert Syst. Appl.* **1995**, *8*, 111–134.
121. Anderson, P. Perspective: Complexity theory and organization science. *Organ. Sci.* **1999**, *10*, 216–232.
122. Brockmann, C.; Girmscheid, G. Complexity of megaprojects. In Proceedings of the CIB World Building Congress: Construction for Development, Cape Town, South Africa, 14–17 May 2007; CIB: Ottawa, ON, Canada, 2007; pp. 219–230.
123. Girmscheid, G.; Brockmann, C. The inherent complexity of large scale engineering projects. *Proj. Perspect.* **2008**, *29*, 22–26.
124. Senescu, R.R.; Guillermo, A.M.; Haymaker, J.R. Relationships between project complexity and communication. *J. Manag. Eng.* **2013**, *29*, 183–197.
125. Mocker, M.; Weill, P.; Woerner, S. Revisiting complexity in the digital age. *MIT Sloan Manag. Rev.* **2014**, *55*, 73–81.
126. Gorzeń-Mitka, I.; Okręglicka, M. Review of complexity drivers in enterprise. In Proceedings of the Liberec Economic Forum, Liberec, Czech Republic, 16–17 September 2015; Technical University of Liberec: Liberec, Czech Republic, 2015; pp. 253–260.
127. Rennung, F.M. Managing Complexity in Service Processes. The Case of Large Business Organizations. Ph.D. Thesis, Politehnica Unviersty of Timisoara, Timisoara, Romania, 2016.
128. Luo, L.; He, Q.; Jaselskis, E.; Xie, J. Construction project complexity: Research trends and implications. *J. Constr. Eng. Manag.* **2017**, *143*, 04017019.
129. Basile, G.; Kaufmann, H.R.; Savastano, M. Revisiting complexity theory to achieve strategic intelligence. *Int. J. Foresight Innov. Policy* **2018**, *13*, 57–70.
130. Latos, B.A.; Harlacher, M.; Burgert, F.; Nitsch, V.; Przybysz, P.; Mütze-Niewöhner, S. Complexity drivers in digitalized work systems: Implications for cooperative forms of work. *Adv. Sci. Technol. Eng. Syst.* **2018**, *3*, 171–185.
131. Woodward, M.R.; Hennell, M.A.; Hedley, D. A measure of control flow complexity in program text. *IEEE Trans. Softw. Eng.* **1979**, *SE-5*, 45–50.
132. Henry, S.; Kafura, D. Software structure metrics based on information flow. *IEEE Trans. Softw. Eng.* **1981**, *SE-7*, 510–518.
133. Troy, D.; Zweben, S. Measuring the quality of structured designs. *J. Syst. Softw.* **1981**, *2*, 113–120.
134. Gibson, V.; Senn, J. System structure and software maintenance performance. *Commun. ACM* **1989**, *32*, 347–358.
135. Banker, R.; Datar, S.; Kemerer, C.; Zweig, D. Software complexity and maintenance costs. *Commun. ACM* **1993**, *36*, 81–95.
136. Chen, Z.; Suen, C.Y. Complexity metrics for rule-based expert systems. In Proceedings of the 1994 International Conference on Software Maintenance, Victoria, BC, Canada, 19–23 September 1994; IEEE: New York, USA, 1994; pp. 382–391.
137. Kang, H.; Seong, P. An information theory-based approach for quantitative evaluation of user interface complexity. *IEEE Trans. Nucl. Sci.* **1998**, *45*, 3165–3174.
138. Paakki, J.; Karhinen, A.; Gustafsson, J.; Nenonen, L.; Verkamo, I. Software metrics by architectural pattern mining. In Proceedings of the International Conference on Software: Theory and Practice (16th IFIP World Computer Congress), Beijing, China, 21–24 August 2000; pp. 325–332.
139. Shao, J.; Wang, Y. A new measure of software complexity based on cognitive weights. *Can. J. Electr. Comput. Eng.* **2003**, *28*, 69–74.
140. Xing, J. *Measures of Information Complexity and the Implications for Automation Design*; Technical Report; Federal Aviation Administration: Washington, DC, USA, 2004; #FAA/AM-04/17.
141. Xing, J.; Manning, C.A. *Complexity and Automation Displays of Air Traffic Control: Literature Review and Analysis*; Technical Report; Federal Aviation Administration: Washington, DC, USA 2005; # DOT/FAA/AM05-4.
142. Cardoso, J. Business process quality metrics: Log-based complexity of workflow patterns. In Proceedings of the OTM Confederated International Conferences "On the Move to Meaningful Internet Systems", Vilamoura, Portugal, 25–30 November, 2007; Springer: Heidelberg, Berlin, 2007; pp. 427–434.
143. Gruhn, V.; Laue, R. Approaches for Business Process Model Complexity Metrics. In *Technologies for Business Information Systems*; Springer: Dordrecht, The Netherlands, 2007; pp. 13–24.
144. Mendling, J.; Neumann, G. Error Metrics for Business Process Models. In Proceedings of the International Conference on Advanced Information Systems Engineering Forum, Trondheim, Norway, 11–15 June 2007; Springer: Heidelberg, Berlin, 2007; Volume 247, pp. 53–56.
145. Vanderfeesten, I.; Cardoso, J.; Reijers, H. A weighted coupling metric for business process models. In Proceedings of the International Conference on Advanced Information Systems Engineering Forum, Trondheim, Norway, 11–15 June 2007; Springer: Heidelberg, Berlin, 2007; Volume 247, pp. 41–44.
146. Cardoso, J. Business process control-flow complexity: Metric, evaluation, and validation. *Int. J. Web Serv. Res.* **2008**, *5*, 49–76.
147. Vanderfeesten, I.; Reijers, H.; van der Aalst, W. Evaluating workflow process designs using cohesion and coupling metrics. *Comput. Ind.* **2008**, *59*, 420–437.
148. Vanderfeesten, I.; Reijers, H.; Mendling, J.; van der Aalst, W.; Cardoso, J. On a quest for good process models: The cross-connectivity metric. In Proceedings of the International Conference on Advanced Information Systems Engineering, Montpellier, France, June 18–20 2008; Springer: Heidelberg, Berlin, 2008; pp. 480–494.
149. Gao, S.; Li, C. Complex network model for software system and complexity measurement. In Proceedings of the WRI World Congress on Computer Science and Information Engineering, Los Angeles, California, USA, 31 March – 2 April 2009; Volume 7, pp. 624–628.
150. Lassen, K.B.; van der Aalst, W. Complexity metrics for workflow nets. *Inf. Softw. Technol.* **2009**, *51*, 610–626.

151. Sánchez-González, L.; García, F.; Mendling, J.; Ruiz, F.; Piattini, M. Prediction of business process model quality based on structural metrics. In Proceedings of the International Conference on Conceptual Modeling, Vancouver, BC, Canada, 1–4 November 2010; Springer: Heidelberg, Berlin, 2010; pp. 458–463.
152. Marchetto, A.; Francescomarino, C.D.; Tonella, P. Optimizing the trade-off between complexity and conformance in process reduction. In Proceedings of the International Symposium on Search Based Software Engineering, Szeged, Hungary, 10–12 September 2011; Springer: Heidelberg, Berlin,2011; pp. 158–172.
153. Posnett, D.; Hindle, A.; Devanbu, P. A simpler model of software readability. In Proceedings of the 8th Working Conference on Mining Software Repositories, Waikiki, Honolulu, USA, 21–22 May 2011; Association for Computing Machinery: New York, USA, 2011; pp. 73–82.
154. Setiawan, M.A.; Sadiq, S. Integrated framework for business process complexity analysis. In Proceedings of the 21st European Conference on Information Systems, Utrecht, The Netherlands, 5–8 June 2013; Association for Information Systems: Atlanta, GA, USA, 2013; pp. 1–12.
155. Schütz, A.; Widjaja, T.; Kaiser, J. Complexity in Enterprise Architectures - Conceptualization and Introduction of a Measure from a System Theoretic Perspective. In Proceedings of the 21th European Conference on Information Systems, Utrecht, The Netherlands, 5–8 June 2013; Association for Information Systems: Atlanta, GA, USA, 2013; pp. 1–12.
156. Benner-Wickner, M.; Book, M.; Brückmann, T.; Gruhn, V. Examining case management demand using event log complexity metrics. In Proceedings of the IEEE 18th International Enterprise Distributed Object Computing Conference Workshops and Demonstrations, Ulm, Germany, 1–2 September 2014; pp. 108–115.
157. Kluza, K.; Nalepa, G.; Lisiecki, J. Square complexity metrics for business process models. In *Advances in Business ICT*; Mach-Król, M., Pełech-Pilichowski, T., Eds.; Advances in Intelligent Systems and Computing; Springer: Cham, Switzerland, 2014; Volume 257, pp. 89–107.
158. Miniukovich, A.; Angeli, A.D. Quantification of interface visual complexity. In Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces, Como, Italy, 27–29 May 2014; Association for Computing Machinery: New York, USA, 2014; pp. 153–160.
159. Taba, S.E.S.; Keivanloo, I.; Zou, Y.; Ng, J.; Ng, T. An exploratory study on the relation between user interface complexity and the perceived quality. In Proceedings of the International Conference on Web Engineering, Toulouse, France, 1–4 July 2014; Springer: Cham, Switzerland, 2014; pp. 370–379.
160. de Oca, I.M.M.; Snoeck, M.; Reijers, H.; Rodríguez-Morffi, A. A systematic literature review of studies on business process modeling quality. *Inf. Softw. Technol.* **2015**, *58*, 187–205.
161. Debnath, N.; Peralta, M.; Salgado, C.; Baigorria, L.; Montejano, G.; Riesco, D. Metrics and performance indicators to evaluate workflow processes on the cloud. In Proceedings of the IEEE/ACS 12th International Conference of Computer Systems and Applications, Marrakech, Morocco, 17–20 November 2015; pp. 1–6.
162. Kluza, K. Measuring complexity of business process models integrated with rules. In Proceedings of the International Conference on Artificial Intelligence and Soft Computing, Zakopane, Poland, 14–18 June 2015; Springer: Cham, Switzerland, 2015; pp. 649–659.
163. Marin, M.; Lotriet, H.; van der Poll, J. Metrics for the case management modeling and notation (CMMN) specification. In Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists, Stellenbosch, South Africa, 28–30 September 2015; Association for Computing Machinery: New York, USA, 2015; pp. 1–10.
164. Müller, G.; Bergmann, R. Complexity-aware generation of workflows by process-oriented case-based reasoning. In Proceedings of the Joint German/Austrian Conference on Artificial Intelligence (Künstliche Intelligenz), Dortmund, Germany, 25-–29 September 2017; Springer: Cham, Switzerland, 2017; pp. 207–221.
165. Petrusel, R.; Mendling, J.; Reijers, H. How visual cognition influences process model comprehension. *Decis. Support Syst.* **2017**, *96*, 1–16.
166. Wehling, K.; Wille, D.; Seidl, C.; Schaefer, I. Decision support for reducing unnecessary IT complexity of application architectures. In Proceedings of the 2017 IEEE International Conference on Software Architecture Workshops, Gothenburg, Sweden, 5–7 April 2017; pp. 161–168.
167. Mannan, U.A.; Ahmed, I.; Sarma, A. Towards understanding code readability and its impact on design quality. In Proceedings of the 4th ACM SIGSOFT International Workshop on NLP for Software Engineering, Lake Buena Vista, Florida, USA, 4 November 2018; Association for Computing Machinery: New York, USA, 2018; pp. 18–21.
168. Nalepa, G. Integrating Business Process Models with Rules. In *Modeling with Rules Using Semantic Knowledge Engineering*; Springer: Cham, Switzerland, 2018; pp. 313–337.
169. Hasić, F.; Vanthienen, J. Complexity metrics for DMN decision models. *Comput. Stand. Interfaces* **2019**, *65*, 15–37.
170. Jones, M.J.; Shoemaker, P.A. Accounting narratives: A review of empirical studies of content and readability. *J. Account. Lit.* **1994**, *13*, 142–184.
171. Jones, M.J. Readability of annual reports: Western versus Asian evidence—A comment to contexualize. *Account. Audit. Account. J.* **1996**, *9*, 86–91.
172. Courtis, J.K. Annual report readability variability: Tests of the obfuscation hypothesis. *Account. Audit. Account. J.* **1998**, *11*, 459–472.

173. Hay, D. Communication in auditors' reports: Variations in readability and the effect of audit firm structure. *Asia-Pac. J. Account.* **1998**, *5*, 179–197.
174. Sydserff, R.; Weetman, P. A texture index for evaluating accounting narratives: An alternative to readability formulas. *Account. Audit. Account. J.* **1999**, *12*, 459–488.
175. Clatworthy, M.; Jones, M.J. The effect of thematic structure on the variability of annual report readability. *Account. Audit. Account. J.* **2001**, *14*, 311–326.
176. Iu, J.; Clowes, C. Approaches to evaluating accounting narratives: A corporate social responsibility perspective. In Proceedings of the Governance and Social Responsibility Conference, Burwood, Victoria, Australia, 26-27 November 2001; Deakin University: Burwood, Victoria, Australia, 2001; pp. 111–123.
177. Leong, E.; Ewing, M.; Pitt, L. E-comprehension: Evaluating B2B websites using readability formulae. *Ind. Mark. Manag.* **2002**, *31*, 125–131.
178. Rutherford, B. Obfuscation, textual complexity and the role of regulated narrative accounting disclosure in corporate governance. *J. Manag. Gov.* **2003**, *7*, 187–210.
179. Iu, J.; Clowes, C. Evaluating a measure of content quality for accounting narratives (with an empirical application to narratives from Australia, Hong Kong, and the United States). In Proceedings of the Fourth Asia Pacific Interdisciplinary Research in Accounting Conference APIRA 2004 Proceedings, Singapore, 4-6 July 2004; Nanyang Technological University: Singapore, 2004; pp. 1–21.
180. Mohamad, R.; Rahman, A.A. Readability of corporate annual reports of top 100 Malaysian companies. *Malays. Manag. J.* **2006**, *10*, 33–47.
181. Hewaidy, A. Readability of financial statement footnotes of Kuwaiti corporations. *Eur. J. Econ. Financ. Adm. Sci.* **2007**, *8*, 18–28.
182. Clarke, D.; Hrasky, S.; Tan, C. Voluntary narrative disclosures by local governments: A comparative analysis of the textual complexity of mayoral and chairpersons' letters in annual reports. *Aust. J. Public Adm.* **2009**, *68*, 194–207.
183. Hrasky, S.; Mason, C.; Wills, D. The textual complexity of annual report narratives: A comparison of high-and low-performance companies. *N. Z. J. Appl. Bus. Res.* **2009**, *7*, 31–45.
184. Sawyer, A. Enhancing compliance through improved readability: Evidence from New Zealand's rewrite "experiment". In *Recent Research on Tax Administration and Compliance*; Gangi, M.E., Plumley, A., Eds.; IRS Research Bulletin: Washington D.C., USA, 2010; pp. 31–56.
185. Sattari, S.; Pitt, L.; Caruana, A. How readable are mission statements? An exploratory study. *Corp. Commun. Int. J.* **2011**, *16*, 282–292.
186. Stone, G.W. Readability of accountants' communications with small business—Some Australian evidence. *Accounting Forum* **2011**, *35*, 247–261.
187. Korfiatis, N.; GarcíA-Bariocanal, E.; SáNchez-Alonso, S. Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. *Electron. Commer. Res. Appl.* **2012**, *11*, 205–217.
188. Sattari, S. Essays on the Issues of Readability in Business Disciplines. Ph.D. Thesis, Luleå University of Technology, Luleå, Sweden, 2012.
189. Rameezdeen, R.; Rodrigo, A. Textual complexity of standard conditions used in the construction industry. *Constr. Econ. Build.* **2013**, *13*, 1–12.
190. Sattari, S.; Wallström, Å. Tourism websites in the Middle East: Readable or not? *Int. J. Leis. Tour. Mark.* **2013**, *3*, 201–214.
191. Stone, G.; Parker, L. Developing the Flesch reading ease formula for the contemporary accounting communications landscape. *Qual. Res. Account. Manag.* **2013**, *10*, 31–59.
192. Goldman, S.R.; Lee, C.D. Text complexity: State of the art and the conundrums it raises. *Elem. Sch. J.* **2014**, *115*, 290–300.
193. Jones, M.; Smith, M. Traditional and alternative methods of measuring the understandability of accounting narratives. *Account. Audit. Account. J.* **2014**, *27*, 183–208.
194. Richards, G.; Fisher, R.; van Staden, C. Readability and Thematic Manipulation in Corporate Communications: A Multi-Disclosure Investigation. In Proceedings of the 2015 Conference Accounting and Finance Association of Australia and New Zealand, Hobart, Australia, 5–7 July 2015; Accounting and Finance Association of Australia and New Zealand: Melbourne, Australia 2015; pp. 1–38.
195. Ajina, A.; Laouiti, M.; Msolli, B. Guiding through the Fog: Does annual report readability reveal earnings management? *Res. Int. Bus. Financ.* **2016**, *38*, 509–516.
196. Allini, A.; Ferri, L.; Maffei, M.; Zampella, A. From Accountability to Readability in the Public Sector: Evidence from Italian Universities. *Int. J. Bus. Manag.* **2017**, *12*, 27–35.
197. Toit, E.D. The readability of integrated reports. *Meditari Account. Res.* **2017**, *25*, 629–653.
198. Khan, M.S.; Ahmed, I.; Khan, Z.I. Readability of mission statements: A look at fortune 500. *J. Qual. Technol. Manag.* **2017**, *13*, 1–14.
199. Nazari, J.; Hrazdil, K.; Mahmoudian, F. Assessing social and environmental performance through narrative complexity in CSR reports. *J. Contemp. Account. Econ.* **2017**, *13*, 166–178.
200. Reschenhofer, T.; Waltl, B.; Shumaiev, K.; Matthes, F. A conceptual model for measuring the complexity of spreadsheets. In Proceedings of the EuSpRIG 2016 Conference "Spreadsheet Risk Management", England and Wales (ICAEW), Moorgate, London, 7–8 July 2016; pp. 37–48.

201. Lim, E.K.; Chalmers, K.; Hanlon, D. The influence of business strategy on annual report readability. *J. Account. Public Policy* **2018**, *37*, 65–81.

CHAPTER 3

Unveiling the Potential of Textual Data for Business Process Complexity Analysis

Summary

This chapter presents the research work on the textual data-based BP complexity concept. Accordingly, it deals with the conceptual design and extraction of BP complexity-relevant knowledge from textual data using the industrial case study of ITIL Change Management IT ticket processing from a prominent telecommunication provider in Germany. The concept builds on the objective, subjective, and meta-knowledge extracted from the BP textual data and encompasses semantics, syntax, and stylistics.

This chapter is based on the following publication:

Rizun, N., Revina, A. & Meister, V. G., "Assessing business process complexity based on textual data: evidence from ITIL IT ticket processing", *Business Process Management Journal*, vol. 27, no. 7, 2021, pp. 1966–1988, DOI: 10.1108/BPMJ-04-2021-0217.

Title Assessing Business Process Complexity Based on Textual Data: Evidence from ITIL IT Ticket Processing

Authors Nina Rizun[1], Aleksandra Revina[2,3], and Vera G. Meister[3]
[1] Gdansk University of Technology
[2] Technical University of Berlin
[3] Brandenburg University of Applied Sciences

Published in Business Process Management Journal, 2021, Emerald, impact factor 2021: 3.715.

Abstract This study aims to draw the attention of business process management (BPM)research and practice to the textual data generated in the processes and the potential of meaningful insights extraction. The authors apply standard natural language processing (NLP) approaches to gain valuable knowledge in the form of business process (BP) complexity concept suggested in the study. It is built on the objective, subjective and meta-knowledge extracted from the BP textual data and encompassing semantics, syntax and stylistics. As a result, the authors aim to create awareness about cognitive, attention and reading efforts forming the textual data-based BP complexity. The concept serves as a basis for the development of various decision-support solutions for BP workers. The starting point is an investigation of the complexity concept in the BPM literature to develop an understanding of the related complexity research and to put the textual data-based BP complexity in its context. Afterward, utilizing the linguistic foundations and the theory of situation awareness (SA), the concept is empirically developed and evaluated in a real-world application case using qualitative interview-based and quantitative data-based methods. While aiming to draw attention to those valuable insights inherent in BP textual data, the authors propose an unconventional approach to BP complexity definition through the lens of textual data. Hereby, the authors address the challenges specified by BPM researchers, i.e. focus on semantics in the development of vocabularies and organization- and sector-specific adaptation of standard NLP techniques.

Keywords business process management; business process complexity; natural language processing; situation awareness; decision support; ITIL IT tickets

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1463-7154.htm

Assessing business process complexity based on textual data: Evidence from ITIL IT ticket processing

Nina Rizun
Gdansk University of Technology, Gdansk, Poland
Aleksandra Revina
Brandenburg University of Applied Sciences, Brandenburg an der Havel, Germany and Technical University of Berlin, Berlin, Germany, and
Vera G. Meister
Brandenburg University of Applied Sciences, Brandenburg an der Havel, Germany

Received 15 April 2021
Revised 10 July 2021
26 July 2021
Accepted 26 August 2021

Abstract

Purpose – This study aims to draw the attention of business process management (BPM) research and practice to the textual data generated in the processes and the potential of meaningful insights extraction. The authors apply standard natural language processing (NLP) approaches to gain valuable knowledge in the form of business process (BP) complexity concept suggested in the study. It is built on the objective, subjective and meta-knowledge extracted from the BP textual data and encompassing semantics, syntax and stylistics. As a result, the authors aim to create awareness about cognitive, attention and reading efforts forming the textual data-based BP complexity. The concept serves as a basis for the development of various decision-support solutions for BP workers.

Design/methodology/approach – The starting point is an investigation of the complexity concept in the BPM literature to develop an understanding of the related complexity research and to put the textual data-based BP complexity in its context. Afterward, utilizing the linguistic foundations and the theory of situation awareness (SA), the concept is empirically developed and evaluated in a real-world application case using qualitative interview-based and quantitative data-based methods.

Findings – In the practical, real-world application, the authors confirmed that BP textual data could be used to predict BP complexity from the semantic, syntactic and stylistic viewpoints. The authors were able to prove the value of this knowledge about the BP complexity formed based on the (1) professional contextual experience of the BP worker enriched by the awareness of cognitive efforts required for BP execution (objective knowledge), (2) business emotions enriched by attention efforts (subjective knowledge) and (3) quality of the text, i.e. professionalism, expertise and stress level of the text author, enriched by reading efforts (meta-knowledge). In particular, the BP complexity concept has been applied to an industrial example of Information Technology Infrastructure Library (ITIL) change management (CHM) Information Technology (IT) ticket processing. The authors used IT ticket texts from two samples of 28,157 and 4,625 tickets as the basis for the analysis. The authors evaluated the concept with the help of manually labeled tickets and a rule-based approach using historical ticket execution data. Having a recommendation character, the results showed to be useful in creating awareness regarding cognitive, attention and reading efforts for ITIL CHM BP workers coordinating the IT ticket processing.

Originality/value – While aiming to draw attention to those valuable insights inherent in BP textual data, the authors propose an unconventional approach to BP complexity definition through the lens of textual data. Hereby, the authors address the challenges specified by BPM researchers, i.e. focus on semantics in the development of vocabularies and organization- and sector-specific adaptation of standard NLP techniques.

Keywords Business process management, Business process complexity, Natural language processing, Situation awareness, Decision support, ITIL IT tickets

Paper type Research paper

Business Process Management Journal
© Emerald Publishing Limited
1463-7154
DOI 10.1108/BPMJ-04-2021-0217

1. Introduction

The significance of natural language in human work and private life cannot be overestimated. It is a means of sharing thoughts and feelings and storing knowledge.

Over the last decade, the maturity of natural language processing (NLP) techniques, along with the proliferation of big data, has shifted the focus to new opportunities in a range of applications. In these applications, documents and textual data are extensively used to manage customer service, legal issues, logistics or accounting (van der Aa *et al.*, 2018a). Unstructured text is commonly believed to account for more than 80% of data in companies (Kobayashi *et al.*, 2018). Yet, as also stated by (Kobayashi *et al.*, 2018), few researchers have applied NLP to tackle organizational challenges despite the abundance of textual data. In business process management (BPM), recent research demonstrates the capabilities of NLP-based analysis techniques to support various tasks in a scalable manner (Mendling *et al.*, 2017). However, there are still many challenges of NLP-supported BPM, especially related to its enhancement in the sense of semantics and developing domain or even organization-specific adaptations (van der Aa *et al.*, 2018a).

At the same time, due to the fast development and penetration of digital technologies into BPM, the overworked term of process complexity and solutions addressing this complexity gain new attention. In this respect, the mainstream BPM research has been inspired by the software complexity metrics and is directed towards estimating the complexity of technical artifacts (Cardoso *et al.*, 2006). Hence, it does not consider the textual data. This observation explains that while being a popular subject area in business (Müller *et al.*, 2016), text analytics and NLP have not been used to study the business process (BP) complexity so far.

The demand for complexity research is especially evident in the most impacted IT service management (ITSM) domain (Lei *et al.*, 2021). Practitioners state a dramatic increase in software errors and a lack of experts to deal with them. Software maintenance and its costs, constituting up to 90% of total software development (Goyal and Sardana, 2021), remain in the research and practice focus (Jang and Kim, 2021; Peimbert-García *et al.*, 2021). This complexity and the dynamic nature of processes make the problem of providing business process (BP) workers with structured knowledge to enable informed decision-making especially significant (Lee *et al.*, 2020).

Based on the above motivation, we aim to create a textual data-based instrument for increasing the awareness of BP workers regarding process complexity. Accordingly, we set to develop a BP complexity concept as a basis for various decision-support solutions for BP workers. The concept development involves solving some important issues, which make up the *specific objectives* of this study:

(1) Extending an understanding and conceptualizing the BP complexity based on the textual data generated in BPs using a theoretical background.

(2) Developing a set of BP complexity measures based on the textual data using a linguistic justification.

(3) Exploring, adapting and illustrating the benefits of the BP complexity concept application using an industrial example.

To achieve these objectives and ensure the comprehensiveness of the examined phenomena, we employ a *triangulation* approach based on the following five steps. *First*, we analyze the related work to develop and extend an understanding of BP complexity and closely associated research in Section 2. *Second*, expert knowledge is used to build the theoretical background for the concept model development and adapt the NLP and linguistic considerations to the BPM context. It forms a solid foundation for designing BP complexity measures based on the textual data in Section 3. *Third*, a real-world application case is used to develop a set of BP complexity measures while adapting standard NLP techniques to the BPM context and BP complexity resulting in the BP complexity concept in Section 4. *Fourth*, the BP complexity concept is applied to a real-world scenario to demonstrate its practical value and relevance in Section 5. *Fifth*, expert knowledge collected in onsite workshops, interviews and as a part of remote feedback (qualitative evaluation) as well as statistical methods (quantitative evaluation) are iteratively used to evaluate the results in Section 6.

Hence, the BP complexity concept has been developed in relation to a specific BP from the ITSM domain, which is ITIL change management (CHM) IT ticket processing of an international telecommunication provider. The concept aims to create awareness about certain efforts needed to process a ticket [1]: (1) awareness of *cognitive efforts* obtained with the help of domain-specific taxonomy and necessary for the process/task execution, i.e. a comprehensive understanding of the current situation, including the professional contextual experience of the BP worker, (2) awareness of *attention efforts* to be paid to individual process elements or the entire process, i.e. business emotions contained in the BP text, extracted with the help of domain-specific business sentiment (BS) lexicon and (3) awareness of *reading efforts* obtained with the help of stylistic features contextually related to the text quality, i.e. indicating professionalism, expertise and stress level of the text author. Such awareness can serve as prioritization support, necessary expert identification, selection of process automation candidates and aggregated analyses of BP textual data over specific periods.

Thus, our work contributes to BPM by (1) proposing a BP complexity concept based on the three knowledge types, (2) addressing semantics, syntax and stylistics and (3) creating awareness about certain efforts necessary for BP execution. To the best of our knowledge, this is the first time in the literature that BP textual data are analyzed from these three perspectives to predict BP complexity. Using qualitative and quantitative research methods, we illustratively apply and evaluate our concept in the ITIL CHM IT ticket processing case. Hence, we adapt common NLP techniques to the domain specificity of ITIL CHM to increase their performance on the semantic level.

2. Related work

According to the research artifact, our study naturally lies at the intersection of (1) BPM, (2) BP complexity and (3) NLP techniques for extracting the knowledge about the latter. This section provides an understanding of BP complexity and complexity-related research and gives an overview of the NLP application in BPM while outlining the research gaps. Thus, Section 2.1 reviews the BP complexity approaches in BPM. Section 2.2 presents the research closely related to but not directly addressing BP complexity. Finally, Section 2.3 introduces the status quo of the NLP research in BPM.

2.1 Business process complexity

As organizations develop and expand their businesses, interdependencies between their processes and information systems increase rapidly. To address this problem, organizations modify the technology supporting their businesses. As a result of such developments, organizations face substantial problems. One of the first and most significant problems is complexity, which impedes decision-making and leads to excessively high and often hidden costs. The term complexity has received much attention in different fields. For example, organizational sciences adapt concepts from complexity theory and define an organization as a complex dynamic system consisting of elements interacting with each other and their environment (Grobman, 2005). In computer sciences, as a rule, the term complexity determines the complexity of an algorithm, i.e. the number of resources required to execute the algorithm (Arora and Barak, 2009).

In this study, we limit our scope to the BPM discipline. BPs are sequences of well-defined actions that must be modeled and redesigned as needed (van der Aalst, 2013). Hence, BPM focuses on modeling whereby processes are recorded, evaluated, planned and redesigned. This is also a dominant research direction in BPM (Leno *et al.*, 2020), demonstrating its closeness to computer sciences. Fundamental concepts and approaches of complexity measures applied to BPs have attracted researchers' attention since the 1970s. The necessity to measure complexity became apparent in software development projects with the purpose of management and control. One of the first essential measures, graph theory-based McCabe complexity (McCabe, 1976), or cyclomatic complexity, was designed to identify software modules that are difficult to test or maintain. Later on, it was applied to different subject areas, including BPs, whereby it is known as control-flow complexity (Cardoso *et al.*, 2006). Another popular measurement applied to BPs is Halstead software complexity (Halstead, 1977), calculated based on program operands (variables and constants) and operators (arithmetic operators and keywords influencing the program control flow) (Cardoso *et al.*, 2006). Accordingly, various software complexity approaches have been adapted to BPs. The cited (Cardoso *et al.*, 2006) can be reasonably considered one of the pioneers of software complexity adaption in BPM. Other works, such as (Henry and Kafura, 1981; Jingqiu and Wang, 2003; Jukka *et al.*, 2000; Woodward *et al.*, 1979) and (Conte *et al.*, 1986; Troy and Zweben, 1981), were studied in detail by (Laue and Gruhn, 2006) and (Vanderfeesten *et al.*, 2008a; Vanderfeesten *et al.*, 2007).

At the same time, some research work breaks away from the software complexity and explores other subject fields. (Vanderfeesten *et al.*, 2008) draw inspiration from cognitive sciences, and (Kluza *et al.*, 2014; Sánchez-González *et al.*, 2010) link their research to mathematics. Other researchers experiment with visual cognition of BP models (Petrusel *et al.*, 2017) in a broader context of decision sciences and test various perspectives to BP model complexity, such as errors and rules (Kluza, 2015; Mendling and Neumann, 2007). A number of studies on BP complexity use the widely deployed Business Process Model and Notation (BPMN) (OMG, 2013) modeling framework (Pozzi *et al.*, 2011; Rolón *et al.*, 2009). With the BPMN counterparts' adoption in the BPM field, i.e. Case Management Model and Notation (CMMN) for the case and Decision Model and Notation (DMN) for decision modeling, the corresponding work on their complexity has started to appear. The complexity approaches are similar to the BP model complexity (Hasić and Vanthienen, 2019; Marin *et al.*, 2015). It is important to note that whereas complexity considerations for BPMN and DMN are comparable, the complexity in CMMN can get incomparably high. Two other fields worth mentioning are expert systems (Chen and Suen, 1994; Kaisler, 1986; Suen *et al.*, 1990) and IT architectures (Kinnunen, 2006; Solic *et al.*, 2011; Wehling *et al.*, 2016, 2017). To sum up, BPs consist of many different elements (splits, joins, resources, diverse data types, activities, etc.). Therefore, there can be no universal measure of process complexity addressing all BP elements.

As we can conclude from the summary in Table 1, most of the existing BP complexity approaches come from the software subject area and consider a BP from the angle of programming language, i.e. as a technical artifact. Similar to the software complexity, in the sense of the practical contributions, BP complexity research mainly aims to achieve more transparency, understandability, reducing errors, defects and exceptions of BPs. The observation also proves the intense focus on technical artifacts dominant in the BPM community.

2.2 Research closely related to business process complexity

As can be observed in Section 2.1, software complexity originated in the 1970s has paved the way to the major complexity approaches in BPM, i.e. complexities of process models, event

Business process complexity

Complexity	Studies	Approach	Pursued goals/practical contributions
Software	McCabe (1976)	Graph-theoretic complexity measures	Management and control of software program complexity
	Halstead (1977)	Program operands and operators-based measures	
	Gao and Li (2009)	Complex network theory-based measures	
	Henry and Kafura (1981)	Information flow-based measures (fan-in and fan-out)	Evaluating the structure of large-scale systems
	Woodward *et al.* (1979)	Knots as a measure of control-flow complexity in program texts	Structuring programs
	Jingqiu Shao and Yingxu Wang (2003)	A measure of the cognitive and psychological complexity of software as a human intelligence artifact	Analysis and prediction of software complexity
	Jukka *et al.* (2000)	Discovery of architectural and design patterns	Analysis of the quality of architecture
	Conte *et al.* (1986), Troy and Zweben (1981)	Five design quality measures – coupling, cohesion, complexity, modularity and size	Evaluation of software designs
	Banker *et al.* (1989, 1993), Basili and Hutchens (1983), Gibson and Senn (1989)	The average size of module's procedures, application's modules and the density of goto statements	Understanding and managing computer software complexity in terms of the maintenance costs
BP model	Cardoso *et al.* (2006)	Number of activities, control flows, joins and splits in general and unique (not repeating), interface complexity and graph theory-oriented metrics measuring the complexity of a graphic	Understandability, fewer errors, defects, and exceptions, more robust processes requiring less time to be developed, tested and maintained
	Laue and Gruhn (2006)	Cognitive weights for BP models, information flow, max/mean nesting depth, number of handles and (anti) patterns	
	Vanderfeesten *et al.* (2008a, b), Vanderfeesten *et al.* (2007)	Adapted cohesion and coupling metrics and cross-connectivity (strength of the links between BP model elements)	
	Mendling and Neumann (2007)	Graph theory-based metrics incl. size, separability, sequentially, structuredness, cyclicity and parallelism	
	Sánchez-González *et al.* (2010)	Structural metrics incl. diameter, nodes, density, gateway degrees and mismatch, and the coefficient of connectivity	
	Kluza (2015)	BP model metrics integrated with rules	
	Petrusel *et al.* (2017)	Visual comprehension of a BP model with an eye-tracking experiment	
	Kluza *et al.* (2014)	Durfee and perfect square	
Work- and control-flow	Cardoso (2006, 2008), Lassen and Aalst (2009)	Compound control-flow complexity of all split constructs	
Event log	Cardoso (2007)	Number of process logs that are generated when workflows are executed	
	Benner-Wickner *et al.* (2014)	Average trace length, size, event density and trace diversity	Metrics which can measure the degree of event log quality that is needed so that discovery algorithms can be applied

(*continued*)

Table 1.
Related work review of BP complexity

Complexity	Studies	Approach	Pursued goals/practical contributions
DMN	Hasić and Vanthienen (2019)	Number of decisions, elements, information requirements, density, data objects, durfee and perfect square metric, sequentially, diameter, longest path, vertex degree, knot count, network complexity, decision nesting depth, cyclomatic complexity and interface complexity	Complexity metrics for DMN models
CMMN	Marin *et al.* (2015)	Size, length and complexity	Complexity metrics for CMMN models
Expert systems and rule bases	Chen and Suen (1994), Kaisler (1986), Suen *et al.* (1990)	Number of rules, decision components, breadth of the search path, depth of search space, number of antecedents and consequents of a rule, content, connectivity and size complexity and entropy-based rule base complexity	Systematic and reliable techniques for evaluating expert systems
Enterprise IT architectures	Wehling *et al.* (2016, 2017)	Variability mining	Decision support to determine and remove redundant architectural artifacts
	Kinnunen (2006)	Interface complexity multiplier	Complexity measures for object-process models, compensating the hidden information at interfaces
	Solic *et al.* (2011)	Roger sessions' methodology	Reduce complexity to enhance security, increase functionality and reduce costs of maintenance of the IT system

Source(s): Own elaboration

Table 1.

logs, work and control flows. However, along with this mainstream, we could identify other standalone BPM research directions closely related to BP complexity and most relevant for our research. These are (1) task complexity and cognition and (2) process knowledge intensiveness, approaches that we also partially use, extend and adapt in our BP complexity concept development.

In fact, BPs represent a sequence of steps in the form of activities or tasks. Although not well recognized in the BPM community, task complexity should be reviewed as closely related to the BP complexity. Along with the software complexity, task complexity research going back to the 1980s (Campbell, 1988; Wood, 1986) can be reasonably considered one of the oldest and most extensively studied in the organizational context (Efatmaneshnik and Handley, 2018). In the literature, task complexity is often used in respect to the discussion on task routineness vs cognition caused by the effect of technological change on labor demand (Fernández-Macías and Bisello, 2016). Thus, (Autor *et al.*, 2003) pose the question "how computerization affects skill demands" and differentiate between routine (manual, cognitive) and non-routine (manual, analytic, interactive) tasks. With the growing importance of automation, the attention of the BPM community also shifts towards new extended perspectives on task classifications. For example, (Koorn *et al.*, 2018) suggest differentiating between such task dimensions as creative, adaptive, interactive (routine), analytical (evaluation, standardization), system supervision, routine cognitive, information processing and information exchange (data stream). Further, in the light of the automation trend and gaining popularity robotic process automation (RPA), (Leopold *et al.*, 2018) propose analyzing textual BP descriptions to derive tasks best fitting the RPA application.

In their very essence, BPs are closely intertwined with knowledge. Variations or deviations from a standard BP, insufficient or unrealistic process rules, or the absence of a well-structured process model may only be overcome by the employee's knowledge to keep the process flowing (Gronau and Weber, 2004). To address this challenge, so-called knowledge-intensive processes (KIPs) have been introduced in BPM. KIPs are concerned with the dynamic knowledge conversion among the individuals engaged in the BP execution and often include tacit and continuously changing pieces of knowledge (França *et al.*, 2012). As fairly stated by (Van Leijen and Baets, 2003), almost every process needs some level of knowledge intensiveness to recover from errors, handle unusual cases and improve or adapt the process itself. As a rule, knowledge intensiveness is considered one of the process complexity characteristics. (Eppler *et al.*, 2008) define knowledge intensiveness by the following conditions: highly dependent outcomes, reliance on random events, many options, creativity in problem-solving, performance dependency on the skills and extended learning time. To deal with this complexity, recommendations and requirements to Knowledge Management (KM) tools (Eppler *et al.*, 2008), design of specific KM systems for KIPs (Sarnikar and Deokar, 2010), knowledge modeler languages (Gronau and Weber, 2004), frameworks (Van Leijen and Baets, 2003) and lately data mining (Khanbabaei *et al.*, 2019) are suggested.

Considering this work, we also observe the dominant technical perspective in BPM and BP complexity. Although the discussed approaches do not address textual data generated by BP workers in the BP execution, we base our concept on the notion of knowledge intensiveness of the processes and its characteristics. Focusing on the most typical data type in organizations, we aim to deal with the KIPs characterized by large decision scope, long learning time, demand for much contextual knowledge, skills and complex problem-solving. Thus, knowledge intensity is typically considered in tandem with BP complexity. As also follows from above, another notion closely related to BP complexity is cognition, which gained the attention of BPM research in the context of BP automation potential assessment. Accordingly, we consider BP cognition one of the BP complexity determinants when developing the taxonomy-based approach to assess cognitive efforts.

So far, we have reviewed BPM approaches mainly concerned with technical artifacts. Hereby, the question regarding the role of textual data and NLP in BPM remains discarded. Hence, in Section 2.3, we revise those aspects in which BPM research considers textual data and applies NLP.

2.3 Business process management and natural language processing

Thanks to publicly available frameworks and maturity, NLP has become popular in many application areas. As approximately 80% of enterprise data are textual (Kobayashi *et al.*, 2018; Rizkallah, 2017), the business applications based on textual data are rather broad. They range from accounting, production and logistics to legal office, marketing and customer service and support such tasks as sentiment analysis, automatic text classification, summarization and extraction of topics from a large document corpus (Pröllochs and Feuerriegel, 2020; Zamuda and Lloret, 2020). The complete list of analyzed sources is part of the supplementary material [2].

In the context of BPM, NLP research is streamlined along with the three commonly differentiated layers: multi-process management (identifying the organization's major processes and their prioritization), process model management (managing a single process in a traditional BPM lifecycle) and process instance management (single enactments of a process, i.e. planning, executing and monitoring of a process) (van der Aa *et al.*, 2018a; Mendling *et al.*, 2017). Accordingly, there is a solid prior research in respect to multi-process management dealing with large process model repositories, such as identifying the similarity (Dijkman *et al.*, 2011), matching (Klinkmüller and Weber, 2021; Weidlich *et al.*, 2010) and merging of process models (La Rosa *et al.*, 2013), textual-based (Leopold *et al.*, 2019b) and semantic (Thomas and Fellmann, 2011) search, resolution of lexical ambiguity (Pittke *et al.*, 2015), automatic service derivation (Leopold *et al.*, 2015) and refactoring of large process model repositories (Weber *et al.*, 2011). Next, process model management reveals a significant number of research primarily aimed at BP modeling support, for example, the transformation of textual descriptions into process models (Friedrich *et al.*, 2011) and vice versa (Leopold *et al.*, 2014), text annotations (Stenetorp *et al.*, 2012), multiple languages, semantic quality check (Leopold *et al.*, 2013), checking compliance (van der Aa *et al.*, 2017), correctness and consistency of BP models (Leopold *et al.*, 2019a), BP model discovery (Han *et al.*, 2020), comparing process descriptions with BP models (van der Aa *et al.*, 2018b) and process description autocompletion (Hornung *et al.*, 2007). Finally, process instance management, a primary objective of BPM at the bottom operational level, reveals rather scarce NLP-related research. Whereas there has been a large amount of study on conversational systems, i.e. chatbots, in recent years, such solutions guiding BP workers through possible options and providing BP execution support are at their early stage (Alman *et al.*, 2020; Han, 2019).

A relatively new research direction is integrating process mining with NLP (Fan and Ilk, 2020; Gupta *et al.*, 2020). These approaches aim to include both event logs and textual data into the process analysis. However, such works making BPM discipline more "humanistic" are also at their early stage. Moreover, as noted by (van der Aa *et al.*, 2018a), NLP research has a great potential in the application to BPM to solve several challenges, for example, (1) improvement of the performance, especially at the semantic level, and (2) developing domain- and organization-specific adaptations of common NLP techniques.

To sum up, in the BPM and NLP-related work, we have seen a strong focus on the technical artifacts, i.e. BP models. The actual support of BP workers in the BP execution (process instance management) remains underresearched. Hence, in the present study, by predicting BP complexity based on textual data, we address the demand for direct BP execution support while considering the mentioned challenges. In particular, to develop our BP complexity concept, we use three linguistic levels of text understanding realized through semantics, syntax and stylistics, semantics being in focus. The importance of semantics in NLP is a recognized and attractive research field (Mitra and Jenamani, 2021). However, it remains one of the significant challenges impeding the full exploitation of the NLP benefits in BPM (van der Aa *et al.*, 2018a). Further, motivated by the declared need for domain-specific adaptations, we adapt common NLP techniques to the ITSM subject area.

3. Theoretical and practical background

From the perspective of the theoretical and practical research background, in this section, we briefly present the three levels of the theory of SA to rationalize the knowledge extraction and situation awareness (SA) creation as a basis for the BP complexity concept development. Afterward, we introduce the motivating example to underpin the conceptual model development.

3.1 Conceptual model development based on the theory of situation awareness

In BPM, the term SA is not new. In this context, major research activities have been devoted to studying how to integrate SA into BP models and facilitate the SA of BPs. Similar to complexity and NLP research in BPM, the goal of these research projects has been directed towards BP modeling. In this study, we use the theory of SA (Endsley, 1995) to theoretically justify the extraction of the three knowledge types and illustrate the value of our approach in awareness creation. Hereby, we consider both social and technical aspects enabling the BP workers' decision-making. In any decision-making process, it is important to be able to identify the elements in the environment, understand their relevance for the goal achievement, draw various scenarios of the actions and make informed decisions. These conditions are realized by the theory of SA, which serves as the basis for developing our conceptual model. Our study adapts the SA model suggested by (Endsley, 1995) to the BPM domain, textual data context in general and IT ticket processing as a particular example.

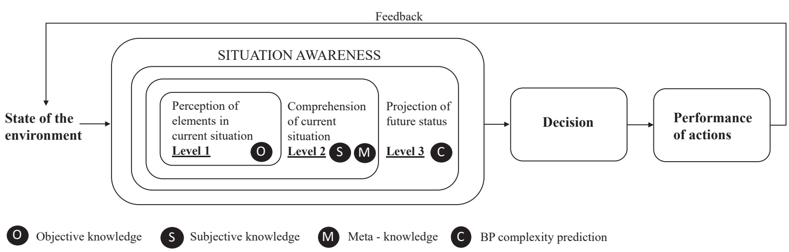
Considering our example, IT tickets are issued following the requests for changes in the IT infrastructure of a big telecommunication company. Implementing these changes means interfering with the organizational environment and its functions, which are often grown historically. The BP worker needs to carefully extract all important information from the textual request he/she receives, put it together and enter into the IT ticket processing system. Hereby, a lot of (often critical) information needs to be filled in, such as a ticket description, its feasible impact, affected systems/items, need for approval by the advisory board, possible service outages, to name only a few. It becomes evident that wrongly entered information can lead to ticket implementation errors with severe service outages. Thus, the domain specificity of our example, including a significant number of BPM cases, can be reasonably considered a high-consequence one, i.e. somewhat similar to those domains common for SA application (Endsley, 1995). However, textual data context needs to be elaborated in detail. According to (Daelemans, 2013), it is commonly distinguished between three levels of text understanding, i.e. types of knowledge that can be extracted from text: (1) *objective* knowledge (answering the who, what, where, when, etc. questions), (2) *subjective* knowledge (sentiment text component) and (3) *meta-knowledge* (further information which can be derived from the text apart from its contents).

Following the SA model (Endsley, 1995), on *SA Level 1*, one perceives the status, attributes and dynamics of relevant elements in the environment and develops an initial understanding of the situation (Endsley, 1995), i.e. *objectively* assesses the situation. In the BPM context, the primary goal of this level is the *perception of basic professional knowledge* about the process/task, namely (1) a deep understanding of its structure enhanced by (2) awareness of the *cognitive efforts* necessary for the process/task execution, which directly follow from the professional context. The first element of such perception is realized by the understanding of the basic elements of the BP text, i.e. ***Resources*** (nouns indicating the specificity of BP elements), ***Techniques*** (verbs of knowledge and information transformation activity affecting ***Resources***), ***Capacities*** (adjectives describing situation specificity of ***Techniques***) and ***Choices*** (adverbs determining the selection of the required set of ***Techniques***), elements of RTCC framework developed in our previous research (Rizun *et al.*, 2019a). The second perception element, i.e. cognitive efforts awareness, is determined by the expected type of activities: (1) simple routine activities happening every day, (2) activities including some nonroutine BP elements or (3) complex activities demanding much cognitive effort. This awareness enables subject matter experts (BP workers) to identify meaningful RTCC elements in the textual descriptions and classify them into the corresponding cognition level.

On *SA Level 2*, the comprehension of *basic elements* in a current situation occurs. It is facilitated by awareness of (1) *attention efforts* needed to be paid to individual BP elements and entire BP and (2) *readability efforts* contextually related to the text quality. This awareness is supported by two other types of knowledge indicated by (Daelemans, 2013), i.e. *subjective* and *meta*-knowledge. One common approach to extracting subjective knowledge in the BPM context is BPM-specific sentiment analysis. Hereby, emotionally loaded keywords, capitalizations and special characters indicate the *attention efforts* needed to address particular BP elements. Additionally, while reading the text, BP workers comprehend meta-knowledge, i.e. text quality, which (1) directly relates to the text author's professionalism, expertise, level of stress and some other important psychological and sociological properties (Daelemans, 2013), (2) influences the understanding of the text and awareness of necessary readability efforts and (3) forms the trust (or doubt) to the written content, i.e. if additional refinements, adjustments and enrichments are needed.

The highest *SA Level 3* is defined by the ability to project the future status of the current situation. The main goal is *to select a mental model* directing the decision strategy necessary for the BP execution. Such ability is enabled by the BP worker's awareness of the *BP complexity*.

Figure 1 provides the extended SA model adapted from (Endsley, 1995). The person's perception and comprehension of the relevant elements in the environment set the foundation for the SA and determine further BP decisions and actions. Hence, the SA is formed based on a comprehensive understanding of the current situation, i.e. the professional contextual experience of the BP worker (SA Level 1), business emotions and quality of the written text, i.e. professionalism, expertise and stress level of the text author (SA Level 2). Further, it allows predicting the required BP worker's efforts (cognitive, attention and reading comprehension) while preparing to execute the process/task at hand and contributes to the BP complexity identification on the SA Level 3.



Source(s): Own elaboration based on (Endsley, 1995)

Figure 1.
Extended situation awareness model of BP complexity

3.2 ITIL change management IT ticket processing motivating example

As an illustrative application, a typical BP scenario of IT ticket processing is used. Processing IT tickets is common for most businesses today. It has a clear start, steps and end. In this process, customer requests, problems and complaints are recorded in the form of IT tickets. Afterward, further steps are taken to process the ticket. As a rule, such a process is carried out by IT service desks. It starts when a customer submits a request and ends with the resolution of that request. It seems to be straightforward with many existing software solutions. However, the recent literature evidences various challenges in service management automation in general and unsolved problems in reporting and processing customer requests in particular (Keller, 2017).

Our study uses ITIL CHM IT ticket texts from a large telecommunication provider, i.e. customer requests to change, improve or resolve a problem regarding IT products and services. CHM is a part of ITIL service transition dealing with the processing of so-called requests for change, IT tickets issued to add, modify or remove anything in the IT infrastructure that could affect IT services (AXELOS, 2011). This dataset was chosen due to the following reasons: (1) as a rule, such requests are written in a free manner, i.e. not following a predefined pattern, (2) in theory, it should contain a clear description of the situation, i.e. the information necessary to process the request, implying the tasks or activities required for an IT ticket resolution and (3) ITIL is a specific but still rather widely used framework (Global Knowledge, 2020). Furthermore, the employees of the application case department have declared the lack of context and related difficulty in processing customer requests during an onsite workshop. Hence, we envision a BP complexity concept-based decision support to assist the ITIL CHM workers creating an awareness of cognitive, attention and reading efforts necessary for understanding and processing the IT ticket. As a motivating example, we use the following anonymized IT ticket received by the ITIL CHM worker per email: *"Dear colleagues, please apply SAP R3 PSU patches on server XXX.YYY.ZZZ for database AAA.BBB.CCC. Attachments - READ RunBook !!! ***************Minimum lead time - 10 h 45 min*************** !!! Otherwise the ticket will be rejected. Disaster recovery tests are prepared by XYZ"*. The motivating example will be used in Section 5 Illustrative application.

The following section describes the knowledge extraction and NLP techniques used on SA Levels 1 and 2 and the resulting textual data-based complexity on SA Level 3, providing decision support to the worker processing the IT ticket.

4. Concept of business process complexity

The previous section introduced an adapted SA model regarding BP complexity, which allows estimating cognitive, attention and reading efforts necessary for BP execution. This section formalizes the BP complexity concept and its components: objective, subjective and meta-knowledge.

4.1 Definition 1: objective knowledge

Core NLP research addresses the extraction of objective knowledge from text, i.e. which concepts, attributes and relations between concepts can be extracted from the text (Daelemans, 2013). Among diverse approaches, taxonomies and ontologies are widely used, also in business, as a necessary resource for many applications (Khadir *et al.*, 2021). Hence, to realize the concept of *SA Level 1* introduced in Section 3.1, which is the perception of basic professional knowledge and *cognitive efforts*, we suggest a specific approach of objective knowledge extraction. It is based on the RTCC framework and decision-making logic (DML) taxonomy (Rizun *et al.*, 2019a). In this approach, the RTCC framework implements the first part – extracting basic professional knowledge and enabling a deep understanding of the BP textual semantic-syntactic structure considering basic BP elements. These are RTCC. DML taxonomy realizes the possibility of the *cognitive effort* awareness necessary for the process/task execution.

To discover the DML, first, we develop an understanding of the three DML levels (Rizun *et al.*, 2019a). Using the systematic literature analysis enhanced with the observation of recent research and market developments, we distinguish three levels, which determine the expected type of activities – *routine, semi-cognitive* and *cognitive* (Rizun *et al.*, 2019a, 2021). The proposed definitions are as follows: (1) *routine* activities are those expressible in rules so that they are easily programmable and can be performed by computers at economically feasible costs (Levy *et al.*, 1996); (2) *semi-cognitive* activities are those where no exact ruleset exists, and there is a clear need of information acquisition and evaluation (Koorn *et al.*, 2018). Here, computer technology cannot substitute but increases employees' productivity (Spitz-Oener, 2006) by partial task processing; (3) *cognitive* activities are the most complex ones where not only information acquisition and evaluation are required but also complex problem-solving. Computers can offer only minimal support.

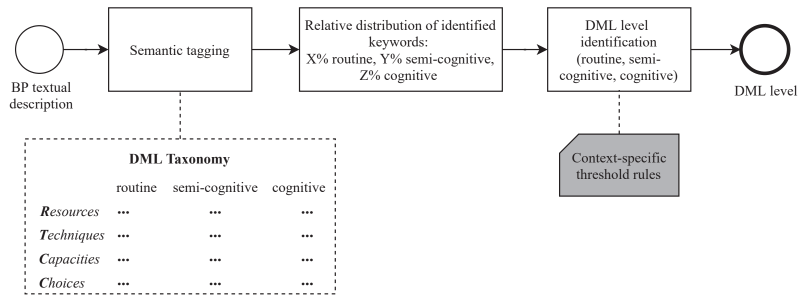
Second, using PoS tagging related to the RTCC elements and a Latent Dirichlet Allocation algorithm (LDA) (Blei, 2012), we identify the most significant keywords in BP texts, in our case, IT tickets. Each of the keywords is associated with an introduced DML level. Performing a systematic literature analysis, we drafted a set of indicators, or contextual variables (Rizun and Taranenko, 2014), based on which subject matter experts (BP workers) categorize words into one of the four RTCC elements and one of the three DML levels. For example, the keywords *interface, tool, client and file* are associated with the IT tickets related to the daily work. In the text, they have an exact, straightforward meaning like *please use file X in the attachment* or *configure interface Y for user Z in the application W*. Hence, these keywords usually denote *routine* ***R****esources*. The keyword *CAB (Change Advisory Board)* belongs to *cognitive* ***R****esources*. The approval of CAB is usually needed when IT tickets are complicated and critical. The process of DML taxonomy development and evaluation is described in (Rizun *et al.*, 2019a).

Third, with the help of the developed domain-specific DML taxonomy, we apply a taxonomy keyword-based pattern matching algorithm to determine the DML level of each ticket. For this, we calculate (1) the total number of routine, semi-cognitive and cognitive keywords extracted from the tickets, (2) the relative occurrence of each category's words in the ticket text and then (3) derive the DML level based on the context-specific threshold rules defined by the subject matter experts.

To sum up, in this study, *objective knowledge* is defined as the one (1) determining the perception of basic professional knowledge about the process/task, namely a deep understanding of its structure enhanced by awareness of the cognitive efforts necessary for its execution, and (2) realized with the BP elements of RTCC organized into one of the three DML levels of routine, semi-cognitive and cognitive. The process of objective knowledge extraction and identification of the DML level is summarized in Figure 2.

4.2 Definition 2: subjective knowledge

To realize the concept of SA Level 2 introduced in Section 3.1, consisting in the comprehension of the elements in a current situation, connected with awareness of the *attention efforts* needed to be paid to individual BP elements, we suggest a specific approach of subjective knowledge extraction (SA Level 2.1.). Subjective knowledge is closely related to sentiment or opinion (Liu, 2012). Hence, we suggest the BS for subjective knowledge extraction in the BPM context. We consider BS as an instrument for measuring awareness of those *attention efforts* needed to be paid to individual BP elements and the entire BP (Rizun and Revina, 2019). We propose extracting this latent information regarding *attention efforts* with the help of a lexicon-based context-specific BS. Hence, in our IT ticket case, we first develop the domain-specific BS lexicon identifying the emotionally loaded keywords and



Source(s): Own elaboration

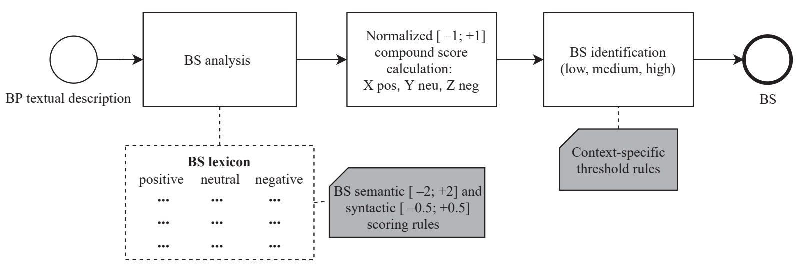
Figure 2.
Objective knowledge extraction and DML level identification

expressions based on two sources: (1) corpus, i.e. IT ticket texts, and (2) CHM descriptions from the ITIL handbook.

Second, based on the subject matter experts' opinion, we (1) refine the developed BS lexicon and (2) assign valence scores to the BS lexicon keywords. Each of the BS lexicon words is associated with positive, negative or neutral BS. Words with valence scores greater than 0 are considered positive, whereas those with less than 0 are considered negative. All other words are considered to have a neutral sentiment. For example, in contrast to such expressions as *no risk, no outage* associated with the positive sentiment, the words *offline* and *downtime* have a negative sentiment. They indicate the need to pay specific attention to the process/task, which will require shutting down the servers or application disconnection. Such activities need to be carefully coordinated with all (possibly) affected parties so that one does not experience any unexpected service outage or other inconvenience.

Third, when the BS lexicon is developed, we apply a lexicon keyword-based pattern matching algorithm to determine the attention efforts in each IT ticket text. For this, (1) we calculate the normalized total score of words with negative, neutral and positive sentiment with the preassigned valence and specific importance markers (syntactic and semantic intensifiers), (2) a set of threshold rules is defined and fine-tuned by subject matter experts and adjusted to the current setting, and (3) using context-specific threshold rules and normalized score, BS is formalized on the ordinal scale of "low", "medium" or "high".

To sum up, *subjective knowledge* is outlined as the one (1) determining the attention efforts to be paid to the BP in general and BP elements in particular, (2) reflecting the emotional component of the BP text and (3) extracted with the help of a domain-specific BS. For an overview of subjective knowledge extraction and BS identification, we refer to Figure 3.



Source(s): Own elaboration

Figure 3.
Subjective knowledge extraction and BS identification

4.3 Definition 3: meta-knowledge

The complete realization of SA Level 2 is introduced in Section 3.1. It also implies the presence of such type of BP worker comprehension as (1) clear understanding of the BP text and (2) his/her awareness of necessary *reading efforts* (SA Level 2.2.). To provide this type of comprehension, we suggest a specific approach of meta-knowledge extraction. In general, meta-knowledge is a conceptually different type of knowledge. As can be seen from the definitions above, objective and subjective knowledge aims to make explicit and structure the knowledge present in the text. Meta-knowledge determines the awareness of the *reading efforts* related to the text quality, i.e. knowledge about the text author or content outside of the text (Daelemans, 2013). Here, as already mentioned, the text quality will likely depend on factors such as the author's professionalism, expertise and stress level. Undoubtedly, a well-written textual description of BPs facilitates successful and fast execution. Vice versa, a poorly written text complicates the work. In our approach, the readability concept is used to extract the meta-knowledge, i.e. measure the text quality (Rizun *et al.*, 2019b). In the BPM context, we suggest the following set of readability measures: (1) *text length*, (2) *parts of speech (PoS)* and *unique PoS distribution* and (3) *wording style*, allowing to formalize the text patterns in terms of a combination of BP text size, its linguistic structure and specificity of BP text presentation correspondingly. To discover the readability, first, we introduce the readability measures' definitions:

(1) *Text length* is based on the principle of least effort by Zipf (1950), i.e. the tendency to communicate efficiently with the least effort, and measures a BP text in the number of words (without stopwords).

(2) *PoS and unique PoS distribution* reflect the same logic from the perspective of the linguistic text structure. The authors tend to use a concise writing style in case of simple processes or tasks. By this, we understand the high usage of unique BP ***R****esources* (nouns) and ***T****echniques* (verbs and verbal nouns) to accurately describe the essence of the problem/process/task.

(3) The *wording style* in our study is based on Zipf's word frequency law (Zipf, 1932) and indicates the information presentation flow as condensed vs disperse. We propose to use the approximation in the equation of Zipf's laws $y = a + b/x$ based on the ordinary least squares method. To describe the wording style concept, we interpret the basic coefficients of Zipf's law as an average frequency of identified keywords (coefficient a) and approximated values of the average speed of appearance of new words in the text (the slope of the hyperbolic function, coefficient b) (Scorobey, 2017). Using these two coefficients, we build the text presentation pattern influencing the specifics of perception and understanding of the text by a BP worker. For example, a *high speed* of new words' appearance can be interpreted as a wording style with a condensed and concrete information presentation pattern. Such a style could testify to transparent and comprehensive readability and required *low reading efforts.* If the *speed* of new words' appearance *slows down*, the wording style becomes more verbose, the same words (RTCC elements) are used more often. The rest of the words are used randomly, depending on the BP text context. The information presentation flow becomes more dispersed and redundant, decreasing readability and increasing required *reading efforts.*

Second, a set of threshold rules is defined and fine-tuned by subject matter experts and adjusted to the current setting. Third, using the obtained measures and threshold rules, readability is identified and formalized on the ordinal scale of "*effortless*", "*involving effort*" and "*telegraphic*". For example, the "telegraphic" readability is assigned if unique BP ***R****esources,* particularly technical specifications, prevail. Thus, in the ticket *"update*

XXX.XXX.XXX, YYY.YYY.YYY, install ZZZ.ZZZ.ZZZ, upgrade AAA.BBB.CCC", the technical names of specific configuration items are in a clear majority. In this case, a BP worker either already knows what needs to be done (for example, the mentioned updates, installation and upgrading are requested every two months), or the (possible) complexity can be captured via objective or subjective knowledge extraction. In this telegraphic example, the BP ***T****echniques update, install* and *upgrade* belong to *routine* DML. Hence, on SA Level 3, the worker will get the decision support based on objective knowledge.

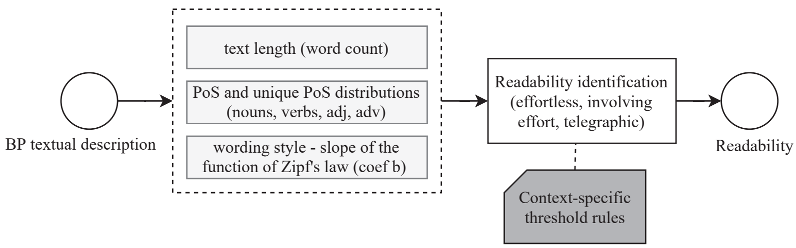
Thus, *meta-knowledge* can be defined as the one (1) containing the information about the text quality, (2) directly influencing the comprehension of the text by its readers and necessary reading efforts and (3) expressed by readability extracted based on text length, PoS and unique PoS distribution and wording style. For an overview of the meta-knowledge extraction and readability identification, we refer to Figure 4.

4.4 Definition 4: business process complexity

To realize the SA Level 3 introduced in Section 3.1, consisting of the BP worker's ability to project the future status of the current situation, we propose a concept of *BP complexity*. It aims to create awareness of BP workers regarding the future BP status, i.e. its execution. The knowledge about BP complexity is formed based on the (1) professional contextual experience of the BP worker enriched by the awareness of cognitive efforts required for BP execution (objective knowledge), (2) business emotions enriched by attention efforts (subjective knowledge) and (3) quality of the text, i.e. professionalism, expertise and stress level of the text author, enriched by reading efforts (meta-knowledge).

Using the expert rule-based approach, the output values of DML level (objective knowledge), BS (subjective knowledge) and readability (meta-knowledge) are aggregated to "low", "medium" or "high" BP complexity. Such a process classification is envisioned to support the BP workers in selecting a mental model directing the strategy necessary for the BP execution. Further, regarding the decision support itself, we refer to the SA-enhanced process design in the context of automation and interpret "low", "medium" and "high" BP complexity as follows: (1) BPs with *low* complexity are those which can be easily automated based on clear rules; (2) BPs of *medium* complexity do not follow exact ruleset and can be only partially automated; (3) in case of highly challenging BPs (*high* complexity), there is no automation expected but minimal assistance in the form of the history of similar BPs. In Figure 5, we formalize the BP complexity concept definition and envisioned decision support.

To sum up, the concept of *BP complexity* is the one (1) estimated using the objective, subjective and meta-knowledge extracted from the BP textual description with the help of NLP techniques, (2) formed based on a comprehensive understanding of the current situation and (3) promoting knowledge creation, transfer and application. Below, we describe the



Source(s): Own elaboration

Figure 4. Meta-knowledge extraction and readability identification

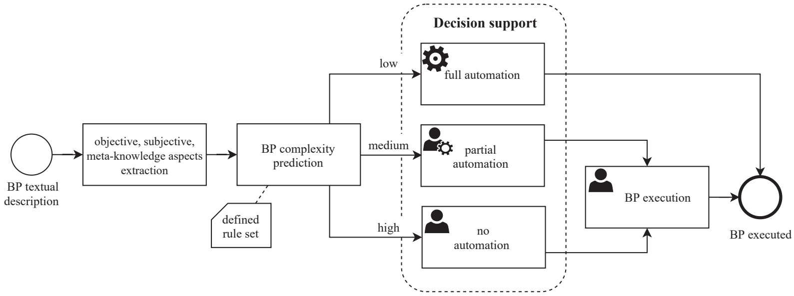


Figure 5.
BP complexity for decision support

Source(s): Own elaboration

illustrative application of the BP complexity concept based on the ITIL CHM dataset of the motivating example.

5. Illustrative application

This section presents the BP complexity concept application on the ITIL CHM dataset. We propose five steps, out of which steps 2–4 can be performed in parallel.

5.1 Step 1: data collection and preprocessing

An important preparation step is collecting BP textual descriptions and converting them into the format required by the program in which the computational analysis will be performed. In our study, two datasets were obtained from the ITIL CHM department according to their availability. After removing duplicates and empty entries, the final datasets comprised 28,157 and 4,625 entries correspondingly. To extract objective and meta-knowledge, classical preprocessing (removal of numbers, special symbols, punctuation, converting to lowercase and stemming) is required. The subjective knowledge is related to the sentiment-relevant information. This information is expressed by the BS lexicon and intensifiers (such as capitalizations, exclamation and question marks, and specific symbols). The correct extraction of mentioned intensifiers requires particular preprocessing, i.e. only removing numbers and stemming. Preprocessing and extraction of the knowledge types were conducted using Python 3.4. The development of handcrafted threshold rules and quantitative evaluation was implemented iteratively using Microsoft Office Excel 2016. Further, collecting other process-related information, such as manuals, handbooks, process descriptions and identification of necessary experts, was performed in this step.

5.2 Step 2: objective knowledge extraction

In the second step (SA Level 1), to extract objective knowledge, we follow Definition 1 in Section 4. Hence, we perform a PoS tagging and assign nouns, verbal nouns, verbs, adjectives and adverbs to the BP elements of the RTCC framework per each IT ticket text. Using the DML taxonomy developed to determine the *cognitive efforts* necessary to process the IT ticket, we identify routine, semi-cognitive and cognitive RTCC BP elements in each IT ticket text. Then, the IT tickets are assigned to one of the three DML levels using the taxonomy keyword-based pattern matching algorithm. Accordingly, based on the total number of detected keywords, we calculate the relative occurrence of the keywords of each category. Using the context-specific threshold rules, we identify the DML level. In our motivating example, all the keywords detected based on the DML taxonomy, i.e. five *routine* ***R****esources* (*PSU, patch, database, server and attachment*) and two *routine* ***T****echniques* (*apply and reject*), belong to the *routine* DML level. Thus, following the threshold rules in Table 2, this IT ticket is associated with the expected 100% routine activity type. On SA Level 3, the CHM worker can expect fast processing of this ticket and plan the time for other effort-intensive or creative tasks in accordance with the expected *cognitive efforts*. Additional information is provided in the repositories [3], and Section 6 describes our iterative evaluation and threshold rule establishment.

5.3 Step 3: subjective knowledge extraction

In the third step (SA Level 2.1.), to extract subjective knowledge, we follow Definition 2 in Section 4. We use the BS lexicon developed based on the process described in (Rizun and Revina, 2019). BS has been introduced as an instrument for measuring the "emotional" component of an IT ticket to determine the *attention efforts* needed to be paid to certain BP elements of the ticket and ticket as a whole. This latent information is extracted from the IT ticket text based on the domain-specific BS lexicon and the lexicon keyword-based pattern matching algorithm. We determine the proportion of words with *negative, neutral* and *positive* valence and the intensifiers (specific punctuation, characters and capitalizations) for each IT ticket text. These values are used to determine the normalized compound score. Finally, the BS is formalized on the ordinal scale of "low", "medium" and "high" based on a set of threshold rules. See Section 6 for a detailed description of our iterative evaluation and threshold rule establishment. In our motivating example, the three BS lexicon keywords (*dear*, *please* and *minimum*), one expression (*disaster recovery*) with 0 valence, one keyword (*rejected*) with −2 valence and time information (*10 h 45 min*) with −0.5 valence have been identified. Furthermore, extensive intensifier usage, i.e. six exclamation marks, 30 stars (total valence is −0.36) and one capitalized word (valence is −0.1). The author of this text used these signs to draw attention to the time. On SA Level 3, the CHM worker will be aware of and consider the necessary time window more carefully while planning and allocating the ticket-related tasks. Hence, according to the threshold rules in Table 3, the processing of such an IT ticket requires much *attention efforts*, making the BS *high*. A detailed description of the lexicon, its scoring semantic and syntactic rules can be found in (Rizun and Revina, 2019) and repositories [4].

5.4 Step 4: meta-knowledge extraction

In the fourth step (SA Level 2.2.), to extract meta-knowledge, we follow Definition 3 in Section 4. Accordingly, we measure the text quality and the required *reading efforts* based on

#	Decision-making logic taxonomy routine (rout), semi-cognitive (semi-cog), cognitive (cog)			DML (cognitive efforts)
1	Rout = 0	Semi-cog = 0	Cog = 1	Cog
2	0 ≤ rout < 0.3	0 ≤ semi-cog<0.5 Semi-	Cog ≥ 0.3	Cog
3	(Rout = 1) and (rout = 0)	cog = 0	Cog = 0	Rout
4	Rout ≥ 0.5	(Semi-cog + cog) ≤ 0.3		Rout
5	Rout = 0	Semi-cog = 1	Cog = 0	Semi-cog
6	Rout = 0	Semi-cog = 0	Cog ≥ 0.3	Semi-cog

Source(s): Revina and Rizun (2019)

Table 2. Threshold rules of objective knowledge extraction (DML)

#	Compound valence positive (pos), neutral (neut), negative (neg)			BS (attention efforts)
1	Pos > 0.2	Neut > 2*abs(neg)	0 < abs(neg) < 0.1	Low
2	Pos ≥ 0	Neut = 0	Neg = 0	Low
3	Pos > 2*neut	Neut > 0	Neg = 0	Low
4	Unrecognized			Low
5	Pos = 0	Neut = 1, neut = 0	Neg = 0	Medium
6	Pos > 0	Neut > 0	Neg = 0	Medium
7	Pos ≥ 0	Neut ≥ 0	0 < abs(neg) < 0.1	Medium
8		Else –		High

Source(s): Revina and Rizun (2019)

Table 3. Threshold rules of subjective knowledge extraction (BS)

the relative number of PoS calculated as related to the whole *ticket length*, i.e. 21 in our motivating example. Afterward, the relative number of *unique PoS* is determined in relation to the *PoS number* in the ticket text. In our example, all PoS are unique (no repetitions) with a substantial prevalence of unique nouns. Next, coefficient b (*wording style*) characterizing the speed of new words' appearance and allowing to identify the pattern of information presentation is computed. In our case, coefficient b has the value 0. This indicates a clearly written text with *effortless* readability. The latter is assigned using the expert-defined threshold rules (see Table 4 and Section 6 for iterative evaluation and threshold rule establishment). Hence, on SA Level 3, the CHM worker will be aware of the high-quality text and low *reading efforts*, which facilitates ticket processing. A comprehensive description of the mentioned readability measures can be found in (Rizun *et al.*, 2019b) and as a part of supplementary material [5].

5.5 Step 5: BP complexity identification

At the end (SA Level 3), the IT ticket complexity is identified using the expert-defined decision rules in Table 5 and obtained DML, BS and readability values in Steps 2–4. The BP worker receives a comprehensive understanding of the situation based on the awareness of the *cognitive, attention* and *reading* efforts needed to process the IT ticket. At this step, the BP worker can also trace back and analyze which BP RTCC, specific sentiment-loaded keywords, punctuation, linguistic text structure and wording style have led to the suggested complexity level. Such an ability is essential for developing trust in the recommendation and creation and transfer of knowledge regarding the process complexity and those factors contributing to this complexity. Hence, after providing the SA, the decision support itself takes place in the form of a recommendation. In our motivating example, the BP complexity is estimated as

#	Text length (L)	PoS and unique PoS σ (n, v, adj, adv)	Wording style (Zipf's coefficient b)	Readability (reading efforts)
1	$L < 25$ words	$\sigma(n) > 0$ and $\sigma(v, \text{adj}, \text{adv}) = 0$	$b = 0$	Telegraphic
2		$\sigma(n, v) > 0$ and $\sigma(n) > \sigma(v, \text{adj}, \text{adv})$ and $\sigma(n) \geq \sigma(\exists! n)$	$b < 3$	Effortless
3		Else –		Involving effort

Source(s): Revina and Rizun (2019)

Table 4. Threshold rules of meta-knowledge extraction (readability)

medium based on the identified *routine* DML, *high* BS and *effortless* readability and using the expert-defined decision rules in Table 4. The decision support should be realized as a recommendation to use a prefilled form from the database and adjust necessary fields. See Table 6 for a summary.

Table 7 summarizes the five-step application scenario described above and specifies required input, processing and necessary outputs for each step, highlighting manual and computer-aided tasks.

#	DML	BS	Readability	BP complexity
1	Routine	Low, medium	Effortless, involving effort	Low
2	Semi-cognitive	Low	Effortless	Low
3	Routine	–	Telegraphic	Low
4	Routine	High	Effortless, involving effort	Medium
5	Cognitive	Low	Effortless	Medium
6	Semi-cognitive, cognitive	Low	Involving effort	Medium
7	Semi-cognitive, cognitive	Medium, high	Effortless	Medium
8	Semi-cognitive	–	Telegraphic	Medium
9	Semi-cognitive, cognitive	Medium, high	Involving effort	High
10	Cognitive	–	Telegraphic	High

Source(s): Revina and Rizun (2019)

Table 5. Expert-defined decision rules for BP complexity identification

"Dear colleagues, please apply SAP R3 PSU patches on server XXX.YYY.ZZZ for database AAA.BBB.CCC. Attachments - READ Runbook !!! ***************minimum lead time - 10 h 45 min*************** !!! otherwise the ticket will be rejected. Disaster recovery tests are prepared by XYZ"				Output
Objective knowledge (DML)				
Routine	*Semi-cognitive*	*Cognitive*		DML/cognitive efforts: "Routine"
1 (7/7)	0	0		
psu, patch, database, server, attachment, apply, reject	–	–		
Subjective knowledge (BS)				
Positive	*Neutral*	*Negative*	*Intensifiers*	BS/attention efforts "High"
0	0.86 (6/7)	0.14 (1/7)	39	
–	Read, dear, please, minimum, disaster recovery	Rejected	Capitalization, time, *, !	
Meta-knowledge (readability)				
Relative # of nouns	*Relative # of verbs*	*Relative # of adjectives and adverbs*	*Wording style (b)*	Readability/ reading efforts: "Effortless"
0.52 (11/21)	0.19 (4/21)	0.05 (1/21)	0	
Colleague, patch, server, database, attachment, lead, time, ticket, disaster, recovery, test	Please, apply, reject, prepare	Dear	–	

Estimated complexity level = "medium"

Decision support recommendation: use a prefilled form as a basis and adjust necessary fields

Source(s): Own elaboration

Table 6. BP complexity concept application on the motivating example

BPMJ

Table 7. Generalized summary of the five-step application scenario of BP complexity prediction

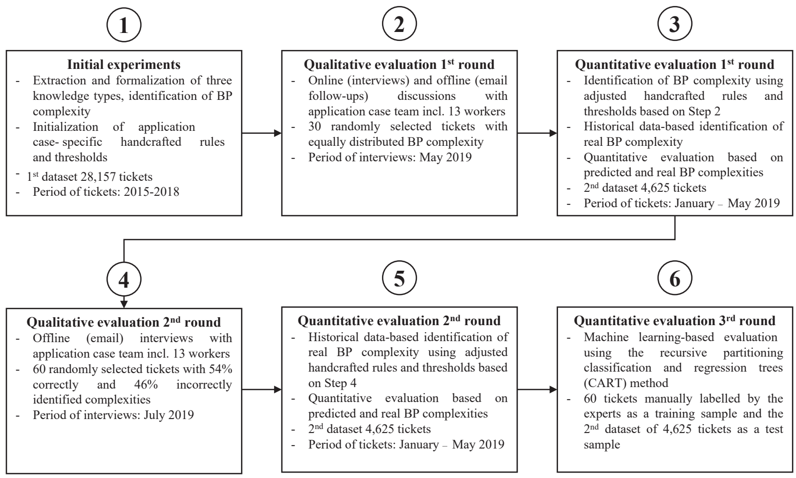
Input	Processing	Output
Step 1. Data collection and preprocessing		
BP textual descriptions Tools: Standard NLP processing software, e.g. python and NLTK library[1]	(1) Standard preprocessing: removal of numbers, special symbols, punctuation, converting to lowercase, stemming	Preprocessed BP texts for the extraction of objective and meta-knowledge
	(2) Special preprocessing: only removal of numbers and stemming	Preprocessed BP texts for the extraction of subjective knowledge
Step 2. Objective knowledge extraction		
*.Csv file with preprocessed BP texts from step 1.1 DML taxonomy Threshold rules Tools: Python, NLTK, MS excel	(1) Computational analysis: PoS tagging (2) Computational analysis: identification of DML keywords; calculation of the relative occurrence of the keywords of each DML level Manual implementation of threshold rules to identify DML in MS excel	PoS assigned to the RTCC elements for each BP text DML keywords and assigned DML for each BP text
Step 3. Subjective knowledge extraction		
*.Csv file with preprocessed BP texts from Step 1.2 BS lexicon Threshold rules Tools: Python, NLTK, MS excel	Computational analysis: identification of BS keywords and their valence and intensifiers; calculation of the normalized compound score Manual implementation of threshold rules to identify BS in MS excel	BS keywords, normalized compound score, and assigned BS for each BP text
Step 4. Meta-knowledge extraction		
*.Csv file with preprocessed BP texts from Step 1.1 Threshold rules Tools: Python, NLTK, MS excel	Computational analysis: calculation of BP text length (word count), PoS and unique PoS identification, calculation of coefficient *b* Manual implementation of threshold rules to identify readability in MS excel	Text length, PoS and unique PoS distribution, coefficient *b* and assigned readability for each task text
Step 5. BP complexity identification		
Excel *.csv file with BP texts and identified DML, BS and readability for each BP text Decision rules Tools: MS excel	Manual implementation of decision rules to identify BP complexity	BP complexity for each BP text

Note(s): [1]https://www.nltk.org/
Source(s): Own elaboration

6. Evaluation

In the context of practical implications of the research artifact, special attention was paid to the evaluation. We conducted quantitative computer experiments and qualitative interviews and discussions in the experimental and evaluation phase consisting of six main steps (see Figure 6).

In *Step 1*, using the dataset of 28,157 tickets, the initial experiments were carried out. The goal was to extract and formalize the three knowledge types, set up initial case-specific handcrafted rules and thresholds and identify BP complexity based on these values. In *Step 2*, to evaluate the obtained values, 30 randomly selected tickets with



Source(s): Own elaboration

Figure 6.
Evaluation process

equally distributed predicted, i.e. textual data-based, *BP complexity* values were presented to the experts – 13 workers of the application case department. This first qualitative evaluation round, including online discussion (see the evaluation questionnaire as a basis for the interviews [6]) and offline follow-up sessions, was conducted in May 2019. The interview was divided into three parts. First, we introduced the objectives of the interview, research motivations, theoretical and methodological background. Second, the method of knowledge extraction and BP complexity prediction was illustratively presented using the sample of 30 tickets. Afterward, the experts estimated the IT ticket complexity based on the available historical data regarding IT ticket processing.

The discussion of the discrepancies between *predicted BP complexity* and historical data-based complexity, further referred to as *real BP complexity*, was shifted to offline (email). Additional information from the IT ticketing system was needed to estimate *real BP complexity*. In the scope of the research, *real BP complexity* is exclusively used to evaluate the research artifact. Third, a question and answer (Q&A) session was conducted following a so-called funnel model (Runeson and Höst, 2009). We started with open questions and moved towards more specific ones regarding possible practical implications of the complexity prediction. Hereby, providing recommendations in the form of templates or historical ticket data (see illustrative process models [7]), prioritization of an incoming ticket as a dashboard for the correct time and workforce management in the team and automatic filling in of the ticket complexity field in the IT ticketing system were mentioned as possible use cases of BP complexity concept application. The offline discussion of the BP complexity values yielded the results presented in Table 8, Point 1. The table gives an overview of qualitative and quantitative evaluation results of BP complexity prediction. Overall, precision is the relative number of correctly identified *predicted BP complexity* compared to the whole number of identified *real BP complexity*. Recalls are calculated for each of the three possible *predicted BP*

	Low	Medium	High
1. Qualitative evaluation of initial experiment results (28,157 tickets) based on 30 tickets – predicted vs expert complexity			
Recall	69%	55%	67%
Overall precision		63%	
2. Quantitative historical data-based evaluation of follow-up experiments results (4,625 tickets) – handcrafted rules			
Recall	51.2%	28.4%	9.1%
Overall precision		45%	
3. Qualitative evaluation of experiment results (4,625 tickets) based on 60 tickets – real vs expert complexity			
Recall	62%	10%	0%
Overall precision		54%	
4. Quantitative historical data-based evaluation of follow-up experiments results (4,625 tickets) – handcrafted rules			
Recall	73.9%	71.9%	40.7%
Overall precision		61.75%	
5. Quantitative historical data-based evaluation of follow-up experiments results (4,625 tickets) – CART – based rules			
Recall	75.6%	61.6%	50.2%
Overall precision		62.27%	

Source(s): Own elaboration

Table 8. Evaluation results of IT ticket complexity prediction

complexity values and represent a fraction of relevant values that have been retrieved over the total amount of relevant values.

In the discussions, we obtained the following findings for qualitative improvements: (1) enrichment of the DML taxonomy and BS lexicon with the one- and bi-grams indicating simple vs complex problem solving, (2) development of the handcrafted threshold rules and (3) identification of necessary historical ticket data allowing to calculate the *real BP complexity*. Hence, we amended the mentioned vocabularies with such one- and bi-grams as "(no, not) affected", "(no) PSO" (projected service outage), "(no) impact", "(no, short, zero) downtime", "test", "(no, not) production", "(no, not) prod". We also added German equivalents of such adverbs as "no", "not", i.e. "kein(e)", "nicht", for the case of English–German ticket texts. Next, the following handcrafted rules and historical data were selected to identify the *real BP complexity*: (1) the presence of the mentioned one- and bi-grams in the IT ticketing system fields "Impact description" and "Brief description" of the ticket-based free text search (RegEx (Prasse *et al.*, 2015)), (2) number of tasks per ticket (count of tasks, integer data type),(3) number of configuration items, specifically applications involved in the ticket (count of applications, integer data type) and (4) risk type of ticket (enumeration, ordinal scale of "low", "medium", "high").

In *Step 3*, we obtained the second dataset of 4,625 tickets with the historical data necessary to identify *real BP complexity,* as discussed with the experts in Step 2. The first quantitative evaluation results were not satisfying, revealing the overall precision of approximately 45% with the following recalls of *predicted BP complexity* – low 51.2%, medium 28.4% and high 9.1% (see Table 8, point 2). Therefore, in *Step 4*, we conducted the second qualitative evaluation round in the form of an interview in an offline (email) mode in July 2019. For this purpose, 60 randomly selected tickets with *predicted* and *real BP complexity values* were presented to the experts. The sample contained 54% correctly and 46% incorrectly identified complexities with the random structure of low, medium and high values. The goal was to adjust the rules and thresholds to identify *real BP complexity* based on the historical data. In the offline discussions, the cases of discrepancies between the *real BP complexity* and the one assigned by the experts were reviewed in detail (for the evaluation results, see Table 8, point 3). Finally, in *Step 5*, we conducted the second quantitative evaluation round. Using adjusted rules regarding the keywords and thresholds for the historical ticket data, such as number of applications and tasks, we achieved an improvement resulting in a better prediction (see Table 8, point 4).

Additionally, in *Step 6*, to compare the evaluation results, we applied a machine learning (ML)-based approach, i.e. the recursive partitioning classification and regression tree (CART) method (Podgorelec *et al.*, 2002), with complexity parameter cp = 0.056 and measures of the error in classification xerror = 0.39. For this purpose, we used the mentioned set of 60 tickets manually evaluated by the experts as a training sample and a dataset of 4,625 tickets as a test sample. The results can be seen in Table 8, Point 5. Comparing the evaluation results of Points 4 and 5 in Table 8, we observe relatively consistent results and can conclude that the performance of our method is acceptable. Looking into ML-based ticket classification approaches in the literature, sophisticated ML classification pipelines report accuracy in a rather broad range from 30 to 90% (Banerjee *et al.*, 2012; Mandal *et al.*, 2019).

The dataset structure obtained at the end of the experiments and evaluation is presented in Table 9. Considering both datasets, we could identify some clear trends. Hence, in the DML distribution, the predominant values are *routine* and *semi-cognitive*, with only a few *cognitive* values. This trend follows a general understanding and expectation of the distribution of daily tasks. In the BS distribution, there is an evident discrepancy between the two datasets. In the first case, the prevalent BS is *medium* (68.5%). Generally, CHM workers tended to use the BS intensifiers (capitalizations, special characters and punctuation) to highlight certain text parts. The reason was that the IT ticket processing software did not support standard text highlighting functions, like bold or cursive letters, underlining and colors. Thus, we observe most tickets of *medium* BS in the first dataset. In the second dataset, the majority of tickets evidence *low* BS (63.2%). Such a discrepancy can be explained by the different sizes of the datasets and their imbalance. The *high* BS is distributed almost equally in both datasets. The distribution values of readability demonstrate a trend similar to that of DML. The most common values are *effortless* and *involving effort*, with relatively few *telegraphic* values. The most frequent value in the first dataset is *effortless* and in the second *involving effort*. The IT ticket complexity values of both datasets reveal comparable distributions, i.e. prevailing *low* complexity tickets followed by *medium* and *high*.

Tables 2–4 in Section 5 provided the final values for the case study-specific handcrafted rules and thresholds of the knowledge aspects extraction. These were obtained after the

DML (objective knowledge)	BS (subjective knowledge)	Readability (meta-knowledge)	IT ticket complexity
1. Dataset of 28,157 tickets			
Routine – 60%	Low – 8.7%	Effortless – 53.1%	Low – 56.3%
Semi-cognitive – 39%	Medium – 68.5%	Involving effort – 43.6%	Medium – 26.8%
Cognitive – 1%	High – 22.8%	Telegraphic – 3.3%	High – 16.9%
2. Dataset of 4,625 tickets			
Routine – 48.6%	Low – 63.2%	Effortless – 35.5%	Low – 52.4%
Semi-cognitive – 49.5%	Medium – 11.3%	Involving effort – 52.8%	Medium – 31.7%
Cognitive – 1.9%	High – 25.5%	Telegraphic – 11.7%	High – 15.9%

Source(s): Own elaboration

Table 9. Distribution statistics of DML, BS, readability and IT ticket complexity

experiments and evaluation rounds described in this section. The final case study-specific rules for the *predicted BP complexity* are illustratively presented in Table 5.

7. Discussion, contribution and limitations

This study develops the BP complexity concept based on textual data using the theory of SA and addresses the awareness of BP workers regarding *cognitive, attention and reading efforts* needed to perform a BP task or activity. Such awareness is realized with common NLP techniques including domain-specific semantics, syntax and stylistics. Hereby, we rely on the three levels of text understanding well-known in linguistics (Daelemans, 2013), i.e. objective, subjective and meta-knowledge. Finally, we illustrate our research findings using a real-world IT ticket processing case.

Hence, the main *theoretical* contributions of our work are as follows:

(1) We propose a novel textual data-based BP complexity based on the three levels of text understanding following the theoretical foundations of computational linguistics by Daelemans (2013).

(2) Our BP complexity concept includes three linguistic perspectives: semantics, syntax and stylistics, the semantics being in the most focus. The latter is a declared challenge impeding the full exploitation of the NLP benefits in BPM (van der Aa *et al.*, 2018a).

(3) Motivated by the recent studies (Karami *et al.*, 2020) and a declared need for domain-specific adaptations (Endsley, 2015), we use the theory of SA to adapt the well-known linguistic foundations to the BPM context.

(4) Hereby, the major contribution to the literature is the confirmation that BP textual data can be used to predict BP complexity from the semantic, syntactic and stylistic viewpoints.

The *methodological* contribution of our research is the combination of common NLP techniques to operationalize the knowledge extraction on the three levels of text understanding differentiated by linguists. In particular, the three levels of text understanding, i.e. objective, subjective and meta-knowledge, are realized by (1) domain-specific taxonomy, (2) sentiment lexicon and (3) stylistic features such as text length, PoS and unique PoS distribution and wording style calculated based on the Zipf's law. These are widely used NLP techniques, which can be relatively easily implemented. The difficulty consists in the preparatory work of vocabularies' compilation and threshold rule establishment. However, one can think about implementing ML approaches using our linguistic features as text representation.

The mentioned *theoretical* and *methodological* contributions enable the realization of the incremental *practical* research value:

(1) The three levels of text understanding aim to provide awareness regarding the *cognitive, attention and reading efforts* required to perform the BP task or activity, hence estimating the BP complexity.

(2) The knowledge about the BP complexity is formed based on the (1) professional contextual experience of the BP worker enriched by the awareness of cognitive efforts required for BP execution (objective knowledge), (2) business emotions enriched by attention efforts (subjective knowledge) and (3) quality of the text, i.e. professionalism, expertise and stress level of the text author, enriched by reading efforts (meta-knowledge).

(3) This work uses a real-world industrial dataset of IT ticket processing to receive expert feedback regarding the BP, i.e. IT ticket processing, complexity.

(4) Further, our BP complexity concept allows a granular perspective on the analyzed data. BP workers can trace back the suggested level of complexity. This is especially important in the context of erroneous classifications and the explainable artificial intelligence (XAI) paradigm.

Lastly, we are aware that our contributions have generalization problems and limitations. The concept of BP complexity reveals certain constraints resulting in additional efforts of the different degree and limitations to adjust the approach while applying in different areas, in particular:

(1) The threshold rules should be adjusted with the subject matter experts for each application case.

(2) The assumptions underlying the meta-knowledge extraction are made based on the mentioned Zipf's law, the principle of least effort and observations and interviews with the subject matter experts. There is a need to prove, extend or refine these assumptions with the subject matter experts in other cases.

(3) Current vocabularies of DML taxonomy and BS lexicon are developed for the ITIL CHM ticket processing. The efforts to adjust these vocabularies for ITIL-related ticket processing cases, such as incident or problem management, should be minimal. In other IT ticket processing cases, the efforts are estimated to be moderate, i.e. some parts of the mentioned vocabularies can be reused. It is worth mentioning that ITIL is widely used, having been ranked in the ten top-paying IT certifications for 2020 based on the survey conducted in the USA (Global Knowledge, 2020). Moreover, managing IT tickets in general remains a crucial concern for the IT service industry (Paramesh and Shreedhara, 2019).

(4) In entirely different cases, like other customer service areas, marketing, software development or strategy, all the vocabularies need to be developed from scratch following the processes described in this paper.

(5) If textual descriptions are written in a language other than English, all the vocabularies also need to be compiled from the beginning.

8. Conclusion and future work

This research work aimed to propose a concept of BP complexity and a set of measures based on the unstructured textual data generated in BPs. The BP complexity can be used to prioritize the BP tasks and activities correctly, estimate necessary effort and provide adequate decision support. The theoretical background of computational linguistics and SA was used to develop, structure and justify a set of the three knowledge types. Diverse common NLP techniques were implemented to extract the knowledge. Afterward, data analysis based on an industrial example from the ITIL CHM department of a telecommunication provider illustrated the concept application.

Our study evidences certain limitations opening the opportunities for future work. For example, manual adjustment of threshold rules is a rather tedious and time-consuming process demanding a constant involvement of subject matter experts. Further, as the goal of our study states, we use textual data to determine the process complexity. Hereby, the event logs known to contain important process insights remain out of scope. Hence, as a part of future work, we plan to further exploit the potential applications of our complexity concept and experiment with (1) ML approaches in combination with the three knowledge types as the text representation to avoid manual adjustments of threshold rules, (2) process mining-, i.e. event logs, based complexity prediction and (3) combining textual and process mining-based complexities into one framework. Bringing these two perspectives together represents a promising but understudied research area (Fan and Ilk, 2020). At the same time, as a demonstration of the practical value of the research, the following business cases of the concept can be developed: (1) dashboard for prioritization of an incoming ticket for a correct time and workforce management in team and (2) prototype of a recommender system for BP workers (Revina and Rizun, 2019) that automatically extracts the knowledge types of the BP complexity concept from the incoming textual requests and adapts the type and the way of recommendation according to the identified BP complexity.

Notes

1. By ticket processing, we understand opening a ticket in an IT ticketing system, filling in necessary fields, identifying and performing preparatory and follow-up work required for successful ticket resolution.
2. See the overview and references of text analytics applications in business on our Github project page.
3. See our Github project page repository Decision-Making-Logic-Taxonomy with the DML taxonomy vocabulary, python file for extracting DML keywords (as an input for python files serve ticket textual descriptions and DML taxonomy), excel file with the calculation of DML based on the motivating example, threshold rules (as an input for excel file serve threshold rules).
4. See our Github project page repository Business-Sentiment with the BS Lexicon, python file for extracting BS (as an input for python file serve ticket textual descriptions and BS Lexicon), excel file with the BS calculation based on the motivating example, threshold rules and scoring, semantic and syntactic rules (as an input for excel file serve threshold rules).
5. See our Github project page repository Stylistic-Patterns-and-Readability with python file for extracting readability measures (as an input for python files serve ticket textual descriptions), excel file with the calculation of readability based on the motivating example and threshold rules (as an input for excel file serve threshold rules) and illustrative application of Zipf's law on tickets (wording style).
6. See the evaluation questionnaire used as a preparation for the interviews.
7. See the process models illustrating the recommendations.

References

Alman, A., Balder, K.J., Maggi, F. and Aa, H.V.D. (2020), "Declo: a chatbot for user-friendly specification of declarative process models", *Proceedings of the Best Dissertation Award, Doctoral Consortium, and Demonstration and Resources Track at BPM 2020 Co-Located with the 18th International Conference on Business Process Management (BPM)*, CEUR Workshop Proceedings, Sevilla, pp. 122-126.

Arora, S. and Barak, B. (2009), *Computational Complexity: A Modern Approach*, Cambridge University Press, New York, NY, available at: http://theory.cs.princeton.edu/complexity/ (accessed 18 July 2020).

Autor, D.H., Levy, F. and Murnane, R.J. (2003), "The skill content of recent technological change: an empirical exploration", *The Quarterly Journal of Economics*, Oxford Academic, Vol. 118 No. 4, pp. 1279-1333.

AXELOS (2011), *ITIL® Service Transition*, TSO, London.

Banerjee, S., Cukic, B. and Adjeroh, D. (2012), "Automated duplicate bug report classification using subsequence matching", *Proceedings of IEEE International Symposium on High Assurance Systems Engineering*, pp. 74-81.

Banker, R.D., Datar, S.M. and Zweig, D. (1989), "Software complexity and maintainability", *Proceedings of the Tenth International Conference on Information Systems - ICIS '89*, ACM Press, New York, NY, pp. 247-255.

Banker, R.D., Datar, S.M., Kemerer, C.F. and Zweig, D. (1993), "Software complexity and maintenance costs", *Communications of the ACM*, ACM, Vol. 36 No. 11, pp. 81-94.

Basili, V.R. and Hutchens, D.H. (1983), "An empirical study of a syntactic complexity family", *IEEE Transactions on Software Engineering*, Vols SE-9 No. 6, pp. 664-672.

Benner-Wickner, M., Book, M., Bruckmann, T. and Gruhn, V. (2014), "Examining case management demand using event log complexity metrics", *2014 IEEE 18th International Enterprise Distributed Object Computing Conference Workshops and Demonstrations*, IEEE, Ulm, pp. 108-115.

Blei, D. (2012), "Probabilistic topic models", *Communications of the ACM*, ACM, Vol. 55 No. 4, pp. 77-84.

Campbell, D.J. (1988), "Task complexity: a review and analysis", *The Academy of Management Review*, Vol. 13 No. 1, pp. 40-52.

Cardoso, J. (2006), "Complexity analysis of BPEL web processes", *Software Process: Improvement and Practice*, John Wiley and Sons, Vol. 12 No. 1, pp. 35-49.

Cardoso, J. (2007), "Business process quality metrics: log-based complexity of workflow patterns", *On the Move to Meaningful Internet Systems 2007: CoopIS, DOA, ODBASE, GADA, and IS*, Springer, Vilamoura, pp. 427-434.

Cardoso, J. (2008), "Business process control-flow complexity: metric, evaluation, validation", *International Journal of Web Services Research*, IGI Global, Vol. 5 No. 2, pp. 49-76.

Cardoso, J., Mendling, J., Neumann, G. and Reijers, H. (2006), "A discourse on complexity of process models", *Business Process Management Workshops*, Springer, Vienna, pp. 117-128.

Chen, Z. and Suen, C.Y. (1994), "Complexity metrics for rule-based expert systems", *Proceedings International Conference on Software Maintenance ICSM-94*, IEEE, Victoria, pp. 382-391.

Conte, S.D., Dunsmore, H.E. and Shen, V. (1986), *Software Engineering Metrics and Models*, Benjamin/Cummings Publishing, Redwood City, CA.

Daelemans, W. (2013), "Explanation in computational stylometry", *International Conference on Intelligent Text Processing and Computational Linguistics*, Springer, Samos, pp. 451-462.

Dijkman, R., Dumas, M., Van Dongen, B., Krik, R. and Mendling, J. (2011), "Similarity of business process models: metrics and evaluation", *Information Systems*, Pergamon, Vol. 36 No. 2, pp. 498-516.

Efatmaneshnik, M. and Handley, H.A. (2018), "Revisiting task complexity: a comprehensive framework", *12th Annual IEEE International Systems Conference, SysCon 2018 - Proceedings* April, pp. 1-4.

Endsley, M.R. (1995), "Toward a theory of situation awareness in dynamic systems", *Human Factors: The Journal of the Human Factors and Ergonomics Society*, Sage, Vol. 37 No. 1, pp. 32-64.

Endsley, M.R. (2015), "Situation awareness misconceptions and misunderstandings", *Journal of Cognitive Engineering and Decision Making*, Sage Publications, Vol. 9 No. 1, pp. 4-32.

Eppler, M.J., Seifried, P. and Röpnack, A. (2008), "Improving knowledge intensive processes through an enterprise knowledge medium (1999)", *Kommunikationsmanagement Im Wandel*, Gabler, pp. 371-389.

Fan, S. and Ilk, N. (2020), "A text analytics framework for automated communication pattern analysis", *Information and Management*, Elsevier, Vol. 57 No. 4, 103219.

Fernández-Macías, E. and Bisello, M. (2016), *Measuring the Content and Methods of Work: A Comprehensive Task Framework*, European Foundation for the Improvement of Living and Working Conditions, Dublin.

França, J.B.S., Baião, F.A. and Santoro, F.M. (2012), "Towards characterizing knowledge intensive processes", *Proceedings of the 2012 IEEE 16th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2012*, pp. 497-504.

Friedrich, F., Mendling, J. and Puhlmann, F. (2011), "Process model generation from natural language text", in Mouratidis, H. and Rolland, C. (Eds), *Advanced Information Systems Engineering. CAiSE 2011. Lecture Notes in Computer Science*, Springer, Berlin, Vol. 6741, pp. 482-496.

Gao, S. and Li, C. (2009), "Complex network model for software system and complexity measurement", *2009 WRI World Congress on Computer Science and Information Engineering*, IEEE, Los Angeles, pp. 624-628.

Gibson, V.R. and Senn, J.A. (1989), "System structure and software maintenance performance", *Communications of the ACM*, ACM, Vol. 32 No. 3, pp. 347-358.

Global Knowledge (2020), "15 highest-paying IT Certifications in 2020 | global knowledge", available at: https://www.globalknowledge.com/us-en/resources/resource-library/articles/top-paying-certifications/ (accessed 16 December 2020).

Goyal, A. and Sardana, N. (2021), "Bug handling in service sector software", *Research Anthology on Recent Trends, Tools, and Implications of Computer Programming*, Information Resources Management Association, pp. 1941-1960.

Grobman, G.M. (2005), "Complexity theory: a new way to look at organizational change", *Public Administration Quarterly*, Vol. 29 Nos 3/4, pp. 350-382.

Gronau, N. and Weber, E. (2004), "Management of knowledge intensive business processes", *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, Vol. 3080, pp. 163-178.

Gupta, M., Agarwal, P., Tater, T., Dechu, S. and Serebrenik, A. (2020), "Analyzing comments in ticket resolution to capture underlying process interactions", *Business Process Management Workshops. BPM 2020*, Springer, Berlin, pp. 219-231.

Halstead, M.H. (1977), *Elements of Software Science*, Elsevier, New York, NY.

Han, S. (2019), *Business Process Automation through ChatBots Implementation: A Case Study of an IT Service Process at Philips*, Delft University of Technology, Delft.

Han, X., Hu, L., Dang, Y., Agarwal, S., Mei, L., Li, S. and Zhou, X. (2020), "Automatic business process structure discovery using ordered neurons LSTM: a preliminary study", ArXiv, available at: http://arxiv.org/abs/2001.01243 (accessed 20 February 2021).

Hasić, F. and Vanthienen, J. (2019), "Complexity metrics for DMN decision models", *Computer Standards and Interfaces*, Elsevier, Vol. 65, pp. 15-37.

Henry, S. and Kafura, D. (1981), "Software structure metrics based on information flow", *IEEE Transactions on Software Engineering*, Vols SE-7 No. 5, pp. 510-518.

Hornung, T., Koschmider, A. and Oberweis, A. (2007), "Rule-based autocompletion of business process models", *Proceedings of the CAiSE'07 Forum at the 19th International Conference on Advanced Information Systems Engineering*, Trondheim, Norway.

Jang, K.A. and Kim, W.J. (2021), "A method of activity-based software maintenance cost estimation for package software", *Journal of Supercomputing*, Springer, Vol. 22 No. 8, pp. 1-21.

Jingqiu, S. and Wang, Y. (2003), "A new measure of software complexity based on cognitive weights", *Canadian Journal of Electrical and Computer Engineering*, Vol. 28 No. 2, pp. 69-74.

Jukka, P., Karhinen, A., Gustafsson, J., Nenonen, L. and Inkeri Verkamo, A. (2000), "Software metrics by architectural pattern mining", *Proceedings of the International Conference on Software: Theory and Practice (16th IFIP World Computer Congress)*, Beijing, pp. 325-332.

Kaisler, S.H. (1986), "Expert system metrics", *Proceedings of the 1986 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 114-120.

Karami, A., Shah, V., Vaezi, R. and Bansal, A. (2020), "Twitter speaks: a case of national disaster situational awareness", *Journal of Information Science*, SAGE Publications, Vol. 46 No. 3, pp. 313-324.

Keller, A. (2017), "Challenges and directions in service management automation", *Journal of Network and Systems Management*, Springer, Vol. 25 No. 4, pp. 884-901.

Khadir, A.C., Aliane, H. and Guessoum, A. (2021), "Ontology learning: grand tour and challenges", *Computer Science Review*, Elsevier, Vol. 39, p. 100339.

Khanbabaei, M., Alborzi, M., Sobhani, F.M. and Radfar, R. (2019), "Applying clustering and classification data mining techniques for competitive and knowledge-intensive processes improvement", *Knowledge and Process Management*, John Wiley and Sons, Vol. 26 No. 2, pp. 123-139.

Kinnunen, M.J. (2006), *Complexity Measures for System Architecture Models*, Massachusetts Institute of Technology, Cambridge, MA.

Klinkmüller, C. and Weber, I. (2021), "Every apprentice needs a master: feedback-based effectiveness improvements for process model matching", *Information Systems*, Elsevier, Vol. 95, p. 101612.

Kluza, K. (2015), "Measuring complexity of business process models integrated with rules", *International Conference on Artificial Intelligence and Soft Computing*, Springer, Zakopane, pp. 649-659.

Kluza, K., Nalepa, G.J. and Lisiecki, J. (2014), "Square complexity metrics for business process models", *Advances in Business ICT*, Springer, pp. 89-107.

Kobayashi, V.B., Mol, S.T., Berkers, H.A., Kismihók, G. and Den Hartog, D.N. (2018), "Text mining in organizational research", *Organizational Research Methods*, SAGE Publications, Vol. 21 No. 3, pp. 733-765.

Koorn, J.J., Leopold, H. and Reijers, H.A. (2018), "A task framework for predicting the effects of automation", *26th European Conference on Information Systems: Beyond Digitization - Facets of Socio-Technical Change, ECIS 2018*, Portsmouth, pp. 1-14.

La Rosa, M., Dumas, M., Uba, R. and Dijkman, R. (2013), "Business process model merging: an approach to business process consolidation", *ACM Transactions on Software Engineering and Methodology*, ACM PUB27 New York, NY, Vol. 22 No. 2, pp. 1-42.

Lassen, K.B. and Aalst, W.M.P.V.D. (2009), "Complexity metrics for Workflow nets", *Information and Software Technology*, Elsevier, Vol. 51 No. 3, pp. 610-626.

Laue, R. and Gruhn, V. (2006), "Complexity metrics for business process models", *9th International Conference on Business Information Systems*, Springer, Klagenfurt, pp. 1-12.

Lee, O.K.D., Choi, B. and Lee, H. (2020), "How do knowledge management resources and capabilities pay off in short term and long term?", *Information and Management*, Elsevier, Vol. 57 No. 2, p. 103166.

Lei, Y., Guo, Y., Zhang, Y. and Cheung, W. (2021), "Information technology and service diversification: a cross-level study in different innovation environments", *Information and Management*, Elsevier, Vol. 58 No. 6, p. 103432.

Leno, V., Dumas, M., Maggi, F.M., La Rosa, M. and Polyvyanyy, A. (2020), "Automated discovery of declarative process models with correlated data conditions", *Information Systems*, Elsevier, Vol. 89, p. 101482.

Leopold, H., Eid-Sabbagh, R.H., Mendling, J., Azevedo, L.G. and Baião, F.A. (2013), "Detection of naming convention violations in process models for different languages", *Decision Support Systems*, Vol. 56 No. 1, pp. 310-325.

Leopold, H., Mendling, J. and Polyvyanyy, A. (2014), "Supporting process model validation through natural language generation", *IEEE Transactions on Software Engineering*, Institute of Electrical and Electronics Engineers, Vol. 40 No. 8, pp. 818-840.

Leopold, H., Pittke, F. and Mendling, J. (2015), "Automatic service derivation from business process model repositories via semantic technology", *Journal of Systems and Software*, Elsevier, Vol. 108, pp. 134-147.

Leopold, H., van der Aa, H. and Reijers, H.A. (2018), "Identifying candidate tasks for robotic process automation in textual process descriptions", in Gulden, J., Reinhartz-Berger, I., Schmidt, R., Guerreiro, S., Guédria, W. and Bera, P. (Eds), *19th International Conference, BPMDS 2018, 23rd International Conference, EMMSAD 2018, Proceedings*, Springer, Tallin, pp. 67-81.

Leopold, H., van der Aa, H., Offenberg, J. and Reijers, H.A. (2019a), "Using Hidden Markov Models for the accurate linguistic analysis of process model activity labels", *Information Systems*, Elsevier, Vol. 83, pp. 30-39.

Leopold, H., van der Aa, H., Pittke, F., Raffel, M., Mendling, J. and Reijers, H.A. (2019b), "Searching textual and model-based process descriptions based on a unified data format", *Software and Systems Modeling*, Springer Verlag, Vol. 18 No. 2, pp. 1179-1194.

Levy, F., Murnane, R.J., Levy, F. and Murnane, R. (1996), "With what skills are computers a complement?", *American Economic Review*, American Economic Association, Vol. 86 No. 2, pp. 258-262.

Liu, B. (2012), "Sentiment analysis and opinion mining", *Synthesis Lectures on Human Language Technologies*, Vol. 5 No. 1, pp. 1-184.

Mandal, A., Malhotra, N., Agarwal, S., Ray, A. and Sridhara, G. (2019), "Automated dispatch of helpdesk email tickets: pushing the limits with AI", *Proceedings of the AAAI Conference on Artificial Intelligence*, Association for the Advancement of Artificial Intelligence (AAAI), Vol. 33, pp. 9381-9388.

Marin, M.A., Lotriet, H. and Van Der Poll, J.A. (2015), "Metrics for the case management modeling and notation (CMMN) specification", *Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists - SAICSIT '15*, ACM Press, Stellenbosch, pp. 1-10.

McCabe, T.J. (1976), "A complexity measure", *IEEE Transactions on Software Engineering*, Vols SE-2 No. 4, pp. 308-320.

Mendling, J. and Neumann, G. (2007), "Error metrics for business process models", *Proceedings of the CAiSE'07 Forum at the 19th International Conference on Advanced Information Systems Engineering*, Trondheim, pp. 53-56.

Mendling, J., Baesens, B., Bernstein, A. and Fellmann, M. (2017), "Challenges of smart business process management: an introduction to the special issue", *Decision Support Systems*, Elsevier, Vol. 100, pp. 1-5.

Mitra, S. and Jenamani, M. (2021), "Helpfulness of online consumer reviews: a multi-perspective approach", *Information Processing and Management*, Pergamon, Vol. 58 No. 3, p. 102538.

Müller, O., Junglas, I., Debortoli, S. and vom Brocke, J. (2016), "Using text analytics to derive customer service management benefits from unstructured data", *MIS Quarterly Executive*, Vol. 15 No. 4, pp. 243-258.

OMG (2013), *Business Process Model and Notation (BPMN) - Version 2.0.2*, Object Management Group, Needham.

Paramesh, S.P. and Shreedhara, K.S. (2019), "Automated IT service desk systems using machine learning techniques", *Lecture Notes in Networks and Systems*, Springer, Vol. 43, pp. 331-346.

Peimbert-García, R.E., Vázquez-Serrano, J.I. and Limón-Robles, J. (2021), "The impact of early failures on maintenance costs: an empirical study in Latin America", *Journal of Quality in Maintenance Engineering*, Emerald Publishing, Vol. ahead-of-print No. ahead-of-print, doi: 10.1108/JQME-08-2020-0086.

Petrusel, R., Mendling, J. and Reijers, H.A. (2017), "How visual cognition influences process model comprehension", *Decision Support Systems*, Elsevier, Vol. 96, pp. 1-16.

Pittke, F., Leopold, H. and Mendling, J. (2015), "Automatic detection and resolution of lexical ambiguity in process models", *IEEE Transactions on Software Engineering*, Vol. 41 No. 6, pp. 526-544.

Podgorelec, V., Kokol, P., Stiglic, B. and Rozman, I. (2002), "Decision trees: an overview and their use in medicine - semantic scholar", *Journal of Medical Systems*, Vol. 26 No. 5, pp. 445-463.

Pozzi, G., Morasca, S., Antonini, A. and Ferreira, A.M. (2011), "Software measures for business processes", *Proceedings II of the 15th East-European Conference on Advances in Databases and Information Systems*, CEUR, Vienna.

Prasse, P., Sawade, C., Landwehr, N. and Scheffer, T. (2015), "Learning to identify concise regular expressions that describe email campaigns", *The Journal of Machine Learning Research*, Vol. 16 No. 1, pp. 3687-3720.

Pröllochs, N. and Feuerriegel, S. (2020), "Business analytics for strategic management: identifying and assessing corporate challenges via topic modeling", *Information and Management*, Elsevier, Vol. 57 No. 1, p. 103070.

Revina, A. and Rizun, N. (2019), "Multi-criteria knowledge-based recommender system for decision support in complex business processes", *ComplexRec Workshop Proceedings, Recommender Systems Conference RecSys*, Copenhagen, CEUR Workshop Proceedings, pp. 16-22.

Rizkallah, J. (2017), "Forbes technology council: the big (unstructured) data problem", available at: https://www.forbes.com/sites/forbestechcouncil/2017/06/05/the-big-unstructured-data-problem/#782a3256493a (accessed 17 September 2019).

Rizun, N. and Revina, A. (2019), "Business sentiment analysis. Concept and method for perceived anticipated effort identification", in Siarheyeva, A., Barry, C., Lang, M., Linger, H. and Schneider, C. (Eds), *Information Systems Development: Information Systems beyond 2020 (ISD2019 Proceedings)*, AIS eLibrary, pp. 1-12.

Rizun, N. and Taranenko, Y. (2014), "Simulation models of human decision-making processes", *Management Dynamics in the Knowledge Economy*, College of Management (NUPSPA) and Tritonic Books, Vol. 2 No. 2, pp. 241-264.

Rizun, N., Revina, A. and Meister, V. (2019a), "Method of decision-making logic discovery in the business process textual data", in Abramowicz, W. and Corchuelo, R. (Eds), *Proceedings of 22nd International Conference on Business Information Systems*, Springer, Vol. 353, pp. 70-84.

Rizun, N., Revina, A. and Meister, V. (2019b), "Discovery of stylistic patterns in business process textual descriptions: IT ticket case", in Soliman, K.S. (Ed.), *Innovation Management and Education Excellence through Vision 2020, International Business Information Management Association (IBIMA-2019)*, Granada, pp. 1-10.

Rizun, N., Revina, A. and Meister, V.G. (2021), "Analyzing content of tasks in Business Process Management. Blending task execution and organization perspectives", *Computers in Industry*, Elsevier, Vol. 130, p. 103463.

Rolón, E., Cardoso, J., García, F., Ruiz, F. and Piattini, M. (2009), *Analysis and Validation of Control-Flow Complexity Measures with BPMN Process Models*, Springer, Berlin, Heidelberg, pp. 58-70.

Runeson, P. and Höst, M. (2009), "Guidelines for conducting and reporting case study research in software engineering", *Empirical Software Engineering*, Springer US, Vol. 14 No. 2, pp. 131-164.

Sánchez-González, L., García, F., Mendling, J., Ruiz, F. and Piattini, M. (2010), "Prediction of business process model quality based on structural metrics", *International Conference on Conceptual Modeling*, Springer, pp. 458-463.

Sarnikar, S. and Deokar, A. (2010), "Knowledge management systems for knowledge-intensive processes: design approach and an illustrative example", *Proceedings of the Annual Hawaii International Conference on System Sciences*. doi: 10.1109/HICSS.2010.248.

Scorobey (2017), "Простая программа на Python для гиперболической аппроксимации статистических данных", 1 March, available at: https://m.habr.com/ru/post/322954/(accessed 21 February 2021).

Solic, K., Sebo, D. and Jovic, F. (2011), "Methodology for complexity reduction of IT system (adjustment of the sessions' methodology)", *MIPRO, 2011 Proceedings of the 34th International Convention*, Opatija, pp. 1-4.

Spitz-Oener, A. (2006), "Technical change, job tasks, and rising educational demands: looking outside the wage structure", *Journal of Labor Economics*, Vol. 24 No. 2, pp. 235-270.

Stenetorp, P., Pyysalo, S., Topić, G., Ohta, T., Ananiadou, S. and Tsujii, J. (2012), "BRAT: a web-based tool for NLP-assisted text annotation", *Proceedings of the Demonstrations Session at EACL 2012*, Association for Computational Linguistics, Avignon, pp. 102-107.

Suen, C.Y., Grogono, P.D., Shinghal, R. and Coallier, F. (1990), "Verifying, validating, and measuring the performance of expert systems", *Expert Systems with Applications*, Elsevier, Vol. 1 No. 2, pp. 93-102.

Thomas, O. and Fellmann, M. (2011), "Semantic process modeling – design and implementation of an ontology-based representation of business processes", *Business and Information Systems Engineering*, Vol. 1 No. 6, pp. 438-451.

Troy, D.A. and Zweben, S.H. (1981), "Measuring the quality of structured designs", *Journal of Systems and Software*, Elsevier, Vol. 2 No. 2, pp. 113-120.

van der Aa, H., Leopold, H. and Reijers, H.A. (2017), "Comparing textual descriptions to process models – the automatic detection of inconsistencies", *Information Systems*, Elsevier, Vol. 64, pp. 447-460.

van der Aa, H., Carmona, J., Leopold, H., Mendling, J. and Padró, L. (2018a), "Challenges and opportunities of applying natural language processing in business process management", *The 27th International Conference on Computational Linguistics*, Association for Computational Linguistics, Santa Fe, pp. 2791-2801.

van der Aa, H., Leopold, H. and Reijers, H.A. (2018b), "Checking process compliance against natural language specifications using behavioral spaces", *Information Systems*, Elsevier, Vol. 78, pp. 83-95.

van der Aalst, W.M.P. (2013), "Business process management: a comprehensive survey", *ISRN Software Engineering*, Vol. 2013, 507984, doi: 10.1155/2013/507984.

Van Leijen, H. and Baets, W.R.J. (2003), "A cognitive framework for reengineering knowledge-intensive processes", *Proceedings of the 36th Annual Hawaii International Conference on System Sciences, HICSS 2003*, Institute of Electrical and Electronics Engineers. doi: 10.1109/HICSS.2003.1174232.

Vanderfeesten, I.T.P., Cardoso, J. and Reijers, H.A. (2007), "A weighted coupling metric for business process models", *Workshop Proceedings of the 19th International Conference on Advanced Information Systems Engineering (CAiSE 2007)*, Springer, Trondheim, pp. 41-44.

Vanderfeesten, I., Reijers, H.A. and van der Aalst, W.M.P. (2008a), "Evaluating workflow process designs using cohesion and coupling metrics", *Computers in Industry*, Elsevier, Vol. 59 No. 5, pp. 420-437.

Vanderfeesten, I., Reijers, H.A., Mendling, J., van der Aalst, W.M.P. and Cardoso, J. (2008b), "On a quest for good process models: the cross-connectivity metric", *20th International Conference CAiSE'08*, Springer, Montpellier, pp. 480-494.

Weber, B., Reichert, M., Mendling, J. and Reijers, H.A. (2011), "Refactoring large process model repositories", *Computers in Industry*, Elsevier, 1 June, Vol. 62 No. 5, pp. 467-486.

Wehling, K., Wille, D., Pluchator, M. and Schaefer, I. (2016), "Towards reducing the complexity of enterprise architectures by identifying standard variants using variability mining", in Kubica, S., Ringshausen, H., Reiff-Stephan and Jörg Schlingelhof, M. (Eds), *Automobil Symposium Wildau: Tagungsband*, Technische Hochschule Wildau, Wildau, pp. 37-43.

Wehling, K., Wille, D., Seidl, C. and Schaefer, I. (2017), "Decision support for reducing unnecessary IT complexity of application architectures", *2017 IEEE International Conference on Software Architecture Workshops (ICSAW)*, Gothenburg, IEEE, pp. 161-168.

Weidlich, M., Dijkman, R. and Mendling, J. (2010), "The ICoP framework: identification of correspondences between process models", in Pernici, B. (Ed.), *Advanced Information Systems Engineering. CAiSE 2010. Lecture Notes in Computer Science*, Springer, Berlin, Vol. 6051, pp. 483-498.

Wood, R.E. (1986), "Task complexity: definition", *Organizational Behavior and Human Decision Processes*, Vol. 37, pp. 60-82.

Woodward, M.R., Hennell, M.A. and Hedley, D. (1979), "A measure of control flow complexity in program text", *IEEE Transactions on Software Engineering*, Vols SE-5 No. 1, pp. 45-50.

Zamuda, A. and Lloret, E. (2020), "Optimizing data-driven models for summarization as parallel tasks", *Journal of Computational Science*, Elsevier, Vol. 42, p. 101101.

Zipf, G.K. (1932), *Selected Studies of the Principle of Relative Frequency in Language*, Harvard University Press, Cambridge, MA.

Zipf, G.K. (1950), "Human behavior and the principle of least effort", *Journal of Clinical Psychology*, John Wiley and Sons, Vol. 6 No. 3, doi: 10.1002/1097-4679(195007)6:3<306::AID-JCLP2270060331>3.0.CO;2-7.

Corresponding author
Aleksandra Revina can be contacted at: revina@th-brandenburg.de

CHAPTER 4

Predicting Business Process Complexity Using Textual Data

Summary

This chapter presents a machine learning approach for BP complexity prediction. A list of linguistic features is derived based on the BP complexity-relevant knowledge aspects explained in Chapter 3. These linguistic features as well as the TF-IDF technique serve as a text representation to experiment with various machine learning algorithms in the ITIL Change Management IT ticket processing case study setting. Further, in terms of a feature selection process, the features most beneficial for BP complexity prediction are identified.

This chapter is based on the following publication:

Revina, A., Buza, K. & Meister, V. G., "IT ticket classification: the simpler, the better", *IEEE Access*, vol. 8, 2020, pp. 193380–193395, DOI: 10.1109/ACCESS.2020.3032840.

Title IT Ticket Classification: The Simpler, the Better

Authors Aleksandra Revina[1,2], Krisztian Buza[3], and Vera G. Meister[2]
[1] Technical University of Berlin
[2] Brandenburg University of Applied Sciences
[3] Eötvös Loránd University

Published in IEEE Access, 2020, Open Access, impact factor 2021: 3.476.

Abstract Recently, automatic classification of IT tickets has gained notable attention due to the increasing complexity of IT services deployed in enterprises. There are multiple discussions and no general opinion in the research and practitioners' community on the design of IT ticket classification tasks, specifically the choice of ticket text representation techniques and classification algorithms. Our study aims to investigate the core design elements of a typical IT ticket text classification pipeline. In particular, we compare the performance of TF-IDF and linguistic features-based text representations designed for ticket complexity prediction. We apply various classifiers, including kNN, its enhanced versions, decision trees, naïve Bayes, logistic regression, support vector machines, as well as semi-supervised techniques to predict the ticket class label of low, medium, or high complexity. Finally, we discuss the evaluation results and their practical implications. As our study shows, linguistic representation not only proves to be highly explainable but also demonstrates a substantial prediction quality increase over TF-IDF. Furthermore, our experiments evidence the importance of feature selection. We indicate that even simple algorithms can deliver high-quality prediction when using appropriate linguistic features.

Keywords IT tickets; linguistics; machine learning; text classification; TF-IDF; process complexity

IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received September 15, 2020, accepted October 13, 2020, date of publication October 21, 2020, date of current version November 3, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3032840

IT Ticket Classification: The Simpler, the Better

ALEKSANDRA REVINA[1,2], KRISZTIAN BUZA[3], AND VERA G. MEISTER[2]
[1]Chair of Information and Communication Management, Faculty of Economics and Management, Technical University of Berlin, 10623 Berlin, Germany
[2]Faculty of Economics, Brandenburg University of Applied Sciences, 14770 Brandenburg an der Havel, Germany
[3]Faculty of Informatics, Eötvös Loránd University, 1117 Budapest, Hungary

Corresponding author: Aleksandra Revina (revina@tu-berlin.de)

This work was supported in part by the Department of Data Science and Engineering, Faculty of Informatics, Eötvös Loránd University, and in part by the Thematic Excellence Programme, Industry and Digitization Subprogramme, National Research, Development, and Innovation (NRDI) Office, 2019 (Hungary) under Project no. ED_18-1-2019-0030.

ABSTRACT Recently, automatic classification of IT tickets has gained notable attention due to the increasing complexity of IT services deployed in enterprises. There are multiple discussions and no general opinion in the research and practitioners' community on the design of IT ticket classification tasks, specifically the choice of ticket text representation techniques and classification algorithms. Our study aims to investigate the core design elements of a typical IT ticket text classification pipeline. In particular, we compare the performance of TF-IDF and linguistic features-based text representations designed for ticket complexity prediction. We apply various classifiers, including kNN, its enhanced versions, decision trees, naïve Bayes, logistic regression, support vector machines, as well as semi-supervised techniques to predict the ticket class label of low, medium, or high complexity. Finally, we discuss the evaluation results and their practical implications. As our study shows, linguistic representation not only proves to be highly explainable but also demonstrates a substantial prediction quality increase over TF-IDF. Furthermore, our experiments evidence the importance of feature selection. We indicate that even simple algorithms can deliver high-quality prediction when using appropriate linguistic features.

INDEX TERMS IT tickets, linguistics, machine learning, text classification, TF-IDF, process complexity.

I. INTRODUCTION

With today's increased digitization, any enterprise maintains a broad application portfolio, which is often grown historically. Such a portfolio must be supported by large scale complex IT service environments [1]. These developments reveal a fundamental role of IT support systems in any organization's support operations. Two essential steps of any IT ticket processing which are their correct prioritization and assignment keep getting the attention of practitioners and the scientific community, especially in the context of the ever-increasing amount of tickets, errors, long lead times, and lack of human resources [2]–[7]. While small companies still tend to perform these steps manually, large organizations dedicate large budgets to the implementation of various commercial text classification solutions. Usually, these approaches are sophisticated monolithic software focused on accuracy at the cost of explainability and understandability. In our study, we are guided by this important issue in the choice of text representation and classification techniques.

Being one of the fastest-growing sectors of the service economy, recently, enterprise IT services have gained importance both in research and practice. One popular stream of this research is IT Service Management (ITSM) [8], [9], which focuses on the servitization of IT function, organization and management of IT service provision [10]. For a successful establishment of organizational infrastructure for IT services, IT practitioners implement IT service delivery with a reference process – the IT Infrastructure Library (ITIL) [11], [12]. Separate areas of ITIL, such as Incident, Problem or Change Management, deal with a large amount of unstructured text data, i.e., tickets issued in relation to the incidents, problems, or changes in the IT infrastructure products and services.

In the context of text classification, remarkable research has been done on the advantages and disadvantages of different text classification algorithms [13] and text representation techniques. In a specific context of IT tickets, a considerable amount of studies have been performed on the analysis of the ticket text data addressing such problems as correct

The associate editor coordinating the review of this manuscript and approving it for publication was Najah Abuali.

assignment and prioritization, identification of duplicates, and poorly described tickets [14], [15]. Prevailingly, TF-IDF (Term Frequency-Inverse Document Frequency) ticket text representation technique is used [16], [17]. Various ticket text classification [18], [19] and clustering approaches [20], [21] are studied.

In our previous research [22]–[24], we performed an in-depth linguistic analysis of the ITIL Change Management (CHM) ticket texts originating from the IT CHM ticket processing department of a big enterprise with more than 200,000 employees worldwide. Using the elaborated linguistic representations of the ticket texts, we implemented a rule-based approach to predict the process complexity [25]. The precision of this approach reached approximately 62%. Apart from such a precision quality, the process of rule set establishment demanded a lot of analysis and testing effort on our and process experts' side. Another substantial disadvantage of such an approach, particularly in complex scenarios, is difficulty in maintenance, inability to learn and scale, as adding new rules requires revising the existing ones.

In this study, our objective is to address the challenges above (choice of text representation technique, choice of a classifier, a growing amount of IT tickets, and the need for explainable text classification solutions to support the IT helpdesk workers). Furthermore, we aim to develop an understanding which factors are relevant for the prediction quality with focus on text representation techniques, feature selection, and text classification algorithms.

In the context of representation of ticket text data, we compare the performance of a commonly accepted and widely used TF-IDF [26] with the linguistic approach described in our previous work [22]–[24]. Despite massive research efforts on IT ticket classification, the representation based on the linguistic features considered in this study has not been systematically compared with TF-IDF representation in the ticket classification context yet.

In the classification methods, we focus on various classifiers widely used for text [13] and ticket [14] classification, including kNN, its enhanced versions, so-called hubness-aware classifiers [27], [28], decision trees [29], naïve Bayes [30, pp. 253–286], logistic regression [31], support vector machines [32], as well as semi-supervised techniques. The latter includes kNN with self-training [33], Semi-sUpervised ClassifiCation of TimE SerieS algorithm (SUCCESS) [34], and its improved version QuickSUCCESS [35]. Although state-of-the-art technology allows us to capture a considerable amount of instances in many applications (e.g., millions of images may be observed), in our case, it is costly and difficult to obtain class labels for a large number of tickets. On the one hand, to label the tickets, expert knowledge is required. On the other hand, as various aspects have to be taken into account, even experts need much time for this task. Therefore, due to the limited amount of labeled tickets, we also experimented with semi-supervised approaches for ticket classification.

The rest of the paper is organized as follows: we give an overview of related works in Section II. Section III presents our methods, followed by the experiments in Section IV. Subsequently, we discuss the implications of the findings and conclude with limitations and future research directions.

II. RELATED WORK

In this paper, as we focus on the problem of CHM ticket classification, the related work is structured as follows: (i) text representation, (ii) text and ticket classification.

A. TEXT REPRESENTATION

Text representation is one of the essential building blocks of approaches for text mining and information retrieval. It aims to transform a text document into a numeric feature vector [36], [37] to allow a computational analysis of textual data. We studied different techniques of text representation and feature extraction and structured them into three main categories: weighted word techniques, word embedding [13], and linguistic features.

1) WEIGHTED WORDS

Weighted words approaches are based on counting the words in the document and computing the document similarity directly from the word-count space [13]. Due to their simplicity, Bag-of-Words (BoW) [38] and TF-IDF [39] can be considered as the most common weighted words approaches. The employed techniques usually rely on the aforementioned text representations rather than in-depth linguistic analysis or parsing [40]. As these models represent every word in the corpus as a one-hot-encoded vector, they are incapable of capturing the word semantic similarity and can become very large and technically challenging [13]. To address these limitations, we use linguistic features in the proposed approach. As described in Section III, the number of our linguistic features is independent of the size of the corpus, and they are capable of capturing relevant aspects of semantics.

2) WORD EMBEDDING

With the progress of research, new methods, such as word embedding, have come up to deal with the limitations of weighted words approaches. Word embedding techniques learn from sequences of words by considering their occurrence and co-occurrence information. Each word or phrase from the corpus is mapped to a high dimensional vector of real numbers, and the mapping from words to vectors is learned based on the surrounding words in the corpus. Various methods have been proposed, such as Word2Vec [41], Doc2Vec [42], GloVe [43], FastText [44], [45], contextualized word representations [46], [47]. These methods consider the semantic similarity of the words. However, they need a large corpus of text datasets for training. To solve this issue, pre-trained word embedding models have been published [48]. Unfortunately, the models do not work for the words outside of the corpus. Another limitation of the word embedding models is related to the fact that they are trained based on the words appearing in a pre-defined window. This is an inherent limitation for considering different meanings of the same word occurring in different contexts. While addressing this limitation, the contextualized word representation techniques based on the context of the word in a document [46], [47] were developed. Nonetheless, in real-world applications, new words may appear (in the description of a new ticket, the customer may use phrases and specific words that have not been used before). These new words are not included in the corpus at training time. Therefore, these word embedding techniques will fail to produce a correct mapping for these new words. While this problem may be alleviated by retraining the model, this requires a lot of data and computational resources (CPU, RAM). In contrast, as it is discussed next, it is straightforward to use our linguistic features text representation in case of new words, as it is not based on word embedding.

3) LINGUISTIC FEATURES

Text representations based on linguistic features have been introduced to address the mentioned limitations of weighted words and word embeddings. Below, we list some examples.

Lexicon-based sentiment analysis is a well-established text classification technique [49]. Synonymy and hypernymy are known approaches to increase prediction quality [36]. Extensive research on the linguistic analysis of ticket texts has been performed in [50], [51]. The authors use parts-of-speech (PoS) count and specific term extractions to define the reported problem's severity. Coussement and Van den Poel extract the following linguistic features: word count, question marks, unique words, the ratio of words longer than six letters, pronouns, negations, assents, articles, prepositions, numbers, time indication [52]. The study results indicated a profoundly beneficial impact of combining traditional features, like TF-IDF and singular value decomposition, with linguistic style into one text classification model. However, the authors declare a demand for more experiments and research. This demand is addressed in our work.

There is no unified opinion in the research community, whether the linguistic approaches are good enough to substitute or enhance the traditional text representations. They are more complex to implement and do not compensate for this complexity with the expected performance increase. Both proponents [53]–[58] and opponents [59] of the linguistic approach provide convincing experimental results. In our work, we amend these research discussions with case study-based findings while comparing the performance of linguistic features with TF-IDF. Word embedding techniques are not applicable in our study due to the following reasons: (i) pre-trained models would not perform well considering the domain-specific vocabulary which contains many new words compared to the training corpus, (ii) at the same time, a limited text corpus would not be enough for training a new model (or retraining an existing one), (iii) industrial applications prevailingly demand explainable models to be able to understand and correct classification mistakes.

TABLE 1. Text representation techniques.

Technique	Strengths	Weaknesses
	Weighted words	
BoW, TF-IDF	• simple to implement • work well with new words • established approaches to extract the most descriptive terms in a document • do not need data to train the mapping	• do not consider syntax and semantics
	Word embedding	
Word2Vec, Doc2Vec, GloVe, contextualized word representations	• consider syntax and semantics • contextualized word representations: consider polysemy	• need much data for training • consider only words that appear in the training data • computationally expensive to train (CPU, RAM)
	Linguistic features	
word count, special characters, parts of speech, unique words, long words, context-specific taxonomies, lexicons, sentiment, etc.	• highly explainable and understandable • wide choice of features • do not necessarily depend on capturing semantics and context • depending on the selected features, can capture both syntax and semantics • do not need data to train the mapping	• expert knowledge is required to define an appropriate set of features

Table 1 summarizes the strengths and weaknesses of the discussed techniques.

B. TEXT AND TICKET CLASSIFICATION

Text classification, also referred to as text categorization or text tagging, is the task of assigning a text to a set of pre-defined categories [60]. Traditionally, this task has been done by human experts. Expert text classification, for example, remains widely used in qualitative research in the form of coding or indexing [61], [62]. Nevertheless, with the growing amount of text data, the attention of the research community and practitioners shifted to automatic methods. One can differentiate three main groups of automatic text classification approaches: rule-based, approaches based on machine learning (ML), and a combination of both in a hybrid system. Ticket classification in software development and maintenance has been studied to tackle challenges such as correct ticket assignment and prioritization, predicting time and number of potential tickets, and avoiding duplicate tickets.

1) RULE-BASED APPROACHES

As the name suggests, this kind of text classification system is based on a set of rules determining classification into pre-defined categories. In general, rule-based classifiers are a popular data mining method applicable to diverse data types. It became popular due to its transparency, explainability, and relative simplicity [63]. Various types of rule development can be distinguished: ML-based such as decision trees [64] or association rules [65], [66], handcrafted rules created by the experts, or hybrid approaches [67], [68]. It is essential to mention that rule-based systems work well with small rule sets. However, they become difficult to build, manage, and change with the growing number of rules.

In our previous work, we conceptualized a recommender system for IT ticket complexity prediction using rule-based classification, i.e., handcrafted rules and rules based on decision trees [25]. Nonetheless, due to the complexity of the domain, the process of rule development consumed much time and effort. The system was difficult to manage and change.

2) MACHINE LEARNING APPROACHES

In the context of ML, text classification can be defined using the following formalization approach. If there is a set of classification labels S and a set of training instances F, each labeled using class labels in S, the model must use F to learn to predict the class labels of unseen instances of the same type [69]. In text classification, F is a set of labeled documents from a corpus. The labels can be extracted topics, themes, writing styles, judgments of the documents' relevance [36]. Regarding the prediction itself, various techniques such as kNN, its enhanced versions (hubness-aware classifiers), decision trees, naïve Bayes, logistic regression, support vector machines, neural networks have been introduced [13]. ML approaches have been shown to be more accurate and easier to maintain compared to rule-based systems [70]. At the same time, it is challenging to select the best ML technique for a particular application [13]. Table 2 summarizes the strengths and weaknesses of the main approaches for text classification.

Most ML techniques, including all the approaches above, require a large amount of training data. However, as described previously, in our application, it is challenging to obtain labeled training data. In contrast, semi-supervised learning (SSL) allows inducing a model from a large amount of unlabeled data combined with a small set of labeled data. For an overview of major SSL methods, their advantages and disadvantages, we refer to [71]. For the above reasons, we experimented with semi-supervised ML approaches. In particular, we used kNN with self-training and SUCCESS [34]. Although SUCCESS showed promising results, it does not scale well to large datasets. We addressed this limitation by suggesting a scaling technique for SUCCESS, and we called the resulting approach QuickSUCCESS [35].

III. METHODS

In the choice of the methods, we put a special emphasis on the explainability and understandability of the classification results for a process worker.

TABLE 2. Text classification techniques.

Technique	Strengths	Weaknesses
Rule-based classification		
ML-based rule development with decision trees, association rules, handcrafted rules	• able to handle a variety of input data • explainable	• with the growing number of rules – difficult to build, manage, maintain, and change
ML-based classification		
kNN, hubness-aware classifiers	• non-parametric • adapts easily to various feature spaces	• finding an appropriate distance function for text data is challenging • prediction might become computationally expensive
SUCCESS/ QuickSUCCESS	• learns both from labeled and unlabeled instances	• finding an appropriate distance function for text data is challenging • computationally expensive training
decision trees	• fast in learning and prediction	• overfitting
	• explainable	• instability even to small variations in the data
naïve Bayes	• showed promising results on text data [72]	• a strong assumption about the data independence
logistic regression	• computationally inexpensive	• a strong assumption about the data independence • is not appropriate for non-linear problems
support vector machines	• robust against overfitting • able to solve non-linear problems	• lack of transparency in results • the problem of choosing an efficient kernel function
neural networks	• flexible with the design of features • can achieve rather accurate predictions	• requires a large amount of data for training • may be extremely computationally expensive • finding an efficient architecture and structure is difficult • not explainability

A. FEATURE EXTRACTION

To allow for automated analysis, texts are usually transformed into a vector space. Unnecessary characters and words (stop words) are removed. Afterward, diverse feature extraction techniques can be applied. In our study, we compare two types of features: TF-IDF and linguistic features.

Next, we describe linguistic features in detail. These features are specifically designed for the task of IT ticket complexity prediction. Three levels of text understanding are commonly distinguished: (1) objective (answering the questions who, what, where, when) measured by semantic technologies, (2) subjective (an emotional component of a text) – by sentiment analysis, and (3) meta-knowledge (information about the author outside of the text) measured by stylometry or stylistic analysis [73]. Accordingly, we develop a set of features which are aggregated by the respective measures indicating the IT ticket complexity. We proposed these features in our initial works [22]–[24].[1] A detailed explanation of linguistic features is provided below using an anonymized IT ticket example (see also Table 3).

TABLE 3. Overview of linguistic features illustrated by a ticket example.

Ticket example: *"Refresh service registry on the XYZ-ZZ YYY server. See attachment for details."*		
Aspects	**Description**	**Linguistic feature**
objective knowledge aspect [24]	relative occurrence of words according to the taxonomy of routine, semi-cognitive and cognitive terms	routine = 0.8
		semi-cognitive = 0.2
		cognitive = 0
subjective knowledge aspect [23]	relative occurrence of words with positive, neutral, and negative sentiment	negative = 0
		neutral = 1
		positive = 0
meta-knowledge aspect [22]	word count	12
	occurrence of nouns in all words	0.5
	occurrence of unique nouns in all nouns	1
	occurrence of verbs in all words	0.17
	occurrence of unique verbs in all verbs	1
	occurrence of adjectives in all words	0.07
	occurrence of unique adjectives in all adjectives	1
	occurrence of adverbs in all words	0
	occurrence of unique adverbs in all adverbs	0
	wording style [22]	0 (no repeating words)

1) OBJECTIVE KNOWLEDGE ASPECT

Core research in Natural Language Processing (NLP) addresses the extraction of objective knowledge from text, i.e., which concepts, attributes, and relations between concepts can be extracted from text, including specific relations such as causal, spatial, and temporal ones [73]. Among diverse approaches, specifically, taxonomies and ontologies, are widely used in the business context [74], [75]. Thus, we suggest a specific approach of objective knowledge extraction using the Decision-Making Logic (DML) taxonomy [24] illustratively presented in Appendix I. Herewith, it is aimed to discover the decision-making nature of activities, called DML level. We use the following DML levels: *routine, semi-cognitive,* and *cognitive* (corresponding to the columns of the table in Appendix I). Using a Latent Dirichlet Allocation Algorithm (LDA) [76], we identify the most important keywords, see [24] for details. Each of the keywords is associated with a DML level. For example, the keywords *user*, *test, request, etc.* are associated with *routine*, whereas the keyword *management* belongs to *cognitive*. We detect these keywords in IT ticket texts. Based on the total number of detected keywords, we calculate the relative occurrence of the words of each category in the ticket text. In the example shown in Table 3, the total number of detected words equals five out of which four words belong to *routine* (*server, attach, detail, see*), one to *semi-cognitive* (*service*), and no word to *cognitive*. Thus, the corresponding features are calculated as follows: *routine* 4/5 = 0.8 and *semi-cognitive* 1/5 = 0.2.

[1] https://github.com/IT-Tickets-Text-Analytics

2) SUBJECTIVE KNOWLEDGE ASPECT

To extract a subjective knowledge aspect, we perform a sentiment analysis [77]. In [23], we suggest a specific business sentiment approach as an instrument for measuring the emotional component of an IT ticket. This latent information is extracted from the unstructured IT ticket texts with the help of a lexicon-based approach. As standard lexicons do not work well in our context of IT ticket classification, using the state-of-the-art VADER [78] and LDA, we developed a domain-specific lexicon, see [23] for details. Our lexicon can also be found in Appendix II.

Each of the words is associated with positive, negative, or neutral sentiment. Words with valence scores greater than 0 are considered positive, whereas words with a valence score less than 0 are considered negative. All other words are considered to have a neutral sentiment. We determine the proportion of words with negative, neutral, and positive sentiment for each IT ticket text and use these values as features. In our example, there are no words with positive or negative sentiment. Therefore, the ticket is assigned to be entirely neutral.

3) META-KNOWLEDGE ASPECT

In our case, meta-knowledge is the knowledge about the author of an IT ticket. The quality of the written text will likely depend on such factors as the author's professionalism and expertise, level of stress, and different psychological and sociological properties [73]. To extract the meta-knowledge aspect, we use the following features [22]: (1) IT ticket text length, (2) PoS features, (3) wording style [22] calculated with the Zipf's law of word frequencies [79].

By length, we mean the number of words in the IT ticket text. This feature is motivated by the following observation: in most cases, IT ticket texts containing a few words, such as *update firewalls,* refer to simple daily tasks. Therefore, short ticket texts may be an indication of the low complexity of the ticket. In the example shown in Table 3, the length of the ticket is 12 words.

As for PoS features, we consider the following PoS tags: nouns, verbs, adjectives, and adverbs. For each of them, we calculate their absolute occurrence, i.e., the *total* number of words having that PoS tag (for example, the total number of nouns). Subsequently, we calculate their relative occurrence, called *occurrence* for simplicity, i.e., the ratio of nouns, verbs, adjectives, and adverbs relative to the length of the text. We use these occurrences as features. In the example shown in Table 3, the occurrence of nouns (*registry, XYZ-ZZ, YYY, server, attachment, details*) in all words is $6/12 = 0.5$.

We also calculate the number of *unique* words having the considered PoS tags (for example, number of unique nouns). Then, we calculate the occurrence of *unique* nouns, verbs, adjectives, and adverbs relative to the number of *all* nouns, verbs, adjectives, and adverbs, respectively. We use these occurrences as features as well. In Table 3, the uniqueness of nouns is $6/6 = 1$ (no repeating words).

According to Zipf's word frequency law, the distribution of word occurrences is not uniform, i.e., some words occur very frequently. In contrast, others appear with a low frequency, such as only once. Our wording style feature describes how extreme this phenomenon is in the IT ticket text, i.e., whether the distribution of occurrences is close to being uniform or not. For details, we refer to [22].

B. SUCCESS

After the features have been extracted and the texts are represented in the form of numerical vectors, they can be fed into ML classifiers. To make sure that our paper is self-contained, below, we shortly review SUCCESS and its scaled version.

We define the semi-supervised classification problem as follows: given a set of labeled instances $L = \{(x_i, y_i)\}_{i=1}^{l}$ and a set of unlabeled instances $U = \{x_i\}_{i=l+1}^{n}$, the task is to train a classifier using both L and U. We use the phrase *set of labeled training instances* to refer to L, x_i is the i-th instance, y_i is its label, whereas we say that U is the *set of unlabeled training instances*. The labeled instances (elements of L) are called seeds. We wish to construct a classifier that can accurately classify any instance, i.e., not only elements of U. For this problem, we proposed the SUCCESS approach that has the following phases:

1. The labeled and unlabeled training instances (i.e., instances of U and L) are clustered with constrained single-linkage hierarchical agglomerative clustering [34]. While doing so, we include cannot-link constraints for each pair of labeled seeds, even if both seeds have the same class labels.
2. The resulting top-level clusters are labeled by their corresponding seeds.
3. The final classifier is 1-nearest neighbor trained on the labeled data resulting at the end of the 2nd phase. This classifier can be applied to unseen test data.

As we noticed the relatively slow training speed of semi-supervised classifiers, we implemented resampling [80], a technique similar to bagging, to accelerate the classification and possibly improve classification results [81].

In particular, we select a random subset of the data and train the model on the selected instances. This process is repeated several times. When making predictions for new instances, the predictions of all the models above are aggregated by majority voting. As the sample size is much smaller than the size of the original dataset and the training has a superlinear complexity, the training of several models on small samples is computationally cheaper than the training of the model on the entire dataset. While kNN with resampling has been established in the research community since a long time ago [82], to the best of our knowledge, this is the first attempt to speed up the SUCCESS algorithm using resampling. We call the resulting approach QuickSUCCESS [35]. Next, we explain QuickSUCCESS in detail.

Algorithm 1 Training QuickSUCCESS

Require: labeled training data L, unlabeled training data U, sample size r, number of classifiers m
Ensure: trained classifiers $\{C^{(i)}\}_{i=1}^{m}$, predicted labels for U

for i in $1 \ldots m$ **do**
  $U^{(i)} \leftarrow$ random_sample (U, r)
  $C^{(i)}, \hat{y^{(i)}} \leftarrow$ t rain SUCCESS $(L, U^{(i)})$
  for x in $U^{(i)}$ **do**
    vote for the label $\hat{y^{(i)}}[x]$ for unlabeled instance x
  end for
end for
return $\{C^{(i)}\}_{i=1}^{m}$, class labels predicted based on the majority vote

The core component of SUCCESS is constrained hierarchical agglomerative clustering. Although various implementations are possible, we may assume that it is necessary to calculate the distance between each pair of unlabeled instances as well as between each unlabeled instance and each labeled instance. This results in a theoretical complexity of $O(l(n-l) + (n-l)^2) = O(n^2 - l^2 - nl)$ at least,[2] where l denotes the number of labeled instances ($l = |L|$), while n indicates the number of all instances (labeled and unlabeled) that are available at training time ($n = |L| + |U|$). Under the assumption that l is a small constant, the complexity of distance computations is $O(n^2)$.

Considering only the aforementioned distance computations required for SUCCESS, the computational costs in the case of a dataset containing n/c instances is c^2-times lower than in the case of a dataset containing n instances. Therefore, even if the computations have to be performed on several "small" datasets, the overall computational costs may be an order of magnitude lower. In particular, computing SUCCESS m-times on a dataset containing r instances has a total computational cost of $O(m \times r^2)$. Under the assumption that $r = n/c$ and $m \approx O(c)$, the resulting complexity is $O(n^2/c)$. Based on the analysis, we propose to speed up SUCCESS by repeated sampling. In particular, we sample the set of unlabeled instances m-times. From now on, we will use $U^{(j)}$ to denote the j-th sample, $1 \leq j \leq m$.

[2] We use the $O(\ldots)$ notation of Complexity Theory

Each $U^{(j)}$ is a random subset of U containing r instances ($|U^{(j)}| = r$). For simplicity, each instance has the same probability of being selected. To train the j-th classifier, we use all the labeled instances in L together with $U^{(j)}$. When sampling the data, the size r of the sample should be chosen carefully so that the sampled data is representative in the sense that the structure of the classes can be learned. This is illustrated in Figure 1.

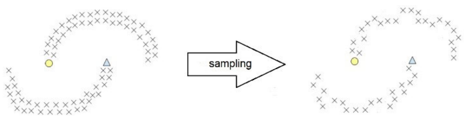


FIGURE 1. When sampling the data, the size r of the sample should be chosen carefully so that the sampled data is representative.

As the sampling is repeated m-times, we induce m classifiers denoted as $C^{(1)}, \ldots, C^{(m)}$. Each classifier $C^{(j)}$ predicts the label for the unlabeled instances in the corresponding sample of the data, i.e., for the instances of $U^{(j)}$. More importantly, each of these classifiers can be used to predict the label of new instances, i.e., instances that are neither in L nor in U. Labels for the instances in U are predicted as follows. For each $x_i \in U$, we consider all those sets $U^{(j)}$ for which $x_i \in U^j$ and the classifiers that were trained using these datasets. The majority vote of these classifiers is the predicted label of x_i. Our approach is summarized in Algorithm 1.

The label of a new instance $x \notin L \cup U$ is predicted as the majority vote of all the classifiers $C^{(1)}, \ldots, C^{(m)}$. We note that the computations related to the classifiers mentioned above $C^{(1)}, \ldots, C^{(m)}$ can be executed in parallel, which may result in additional speed-up in the case of systems where multiple CPUs are available, such as high-performance comput-ing (HPC) systems.

The same resampling technique can be applied with other classifiers as well. In particular, in our experiments, we used it with a semi-supervised variant of kNN, called *kNN with self-training and resampling*.

C. OTHER CLASSIFIERS

We experimented with various other classifiers such as kNN and its enhanced versions, kNN with self-training [83], and kNN with self-training and resampling. As the emergence of bad hubs was shown to characterize textual data [84], we included hubness aware variants of kNN in our study, in particular: kNN with **E**rror **C**orrection (ECkNN) [85], [86], **H**ubness-**W**eighted kNN (HWkNN) [87], **H**ubness-**F**uzzy kNN (HFNN) [88], **N**aive **H**ubness-**B**ayesian kNN (NHBNN) [89]. Additionally, we used decision trees, naïve Bayes, logistic regression, and support vector machines.

Self-learning is a semi-supervised technique known to improve the learning process in case of a large number of unlabeled and a small number of labeled instances [83]. Therefore, we also use kNN with self-training. Hereby, kNN is iteratively retrained with its own most confident predictions [33].

IV. EXPERIMENTAL EVALUATION

The goal of our experiments is to analyze the contribution of the proposed text representation, study the effect of feature selection and the performance of various classifiers.

Our experimental design is summarized in Figure 2 and includes the following steps: (1) collect the case study data; (2) pre-process the data; (3) label part of the data; (4) split the data into training and test sets; (5) extract two sets of features: TF-IDF and linguistic features; (6) apply the classifier; (7) evaluate the results with standard metrics. All the experiments were conducted using Python 3.6. In the case of naïve Bayes, logistic regression, decision trees, and support vector machines, we used the publicly available implementation from scikit-learn,[3] whereas in the case of hubness-aware classifiers and SUCCESS, we used PyHubs.[4] We implemented kNN with and without self-training as well as the proposed speed-up technique based on resampling on our own. We performed all the experiments on an Intel®Core™i7 16 GB RAM machine.

A. DATASETS

The datasets (see Table 4) in the form of IT ticket texts originate from an ITIL CHM department of a big enterprise. The datasets were received (step 1) according to their availability. They covered the whole period obtainable at the moment of this study. The first dataset (Data1) comprised 28,243 tickets created from 2015 to 2018.

TABLE 4. Datasets.

#	time period	# of ticket texts	# of acquired labels
Data1	2015 – 2018	28,243	30
Data2	January – May 2019	4,684	60

The data was pre-processed (step 2) by removing stop words, punctuation, turning to lowercase, stemming and converted into a CSV-formatted corpus of ticket texts. The ticket texts contained prevailingly English texts (more than 80% of English words, a small portion of German words was present). The second dataset (Data2) comprised 4,684 entries in prevailingly English language from the period January – May 2019. The length of IT ticket texts varies from 5-10 words to 120-150.

As labeling of the data is a time-consuming and tedious process and our data comes from an operational unit fully overbooked in the daily work, which is often the case in the industry, experts labeled 30 and 60 tickets of Data1 and Data2 (step 3). To provide a correct label, the case study workers needed to analyze all the factors influencing the IT

[3] https://scikit-learn.org/stable/
[4] http://www.biointelligence.hu/pyhubs/

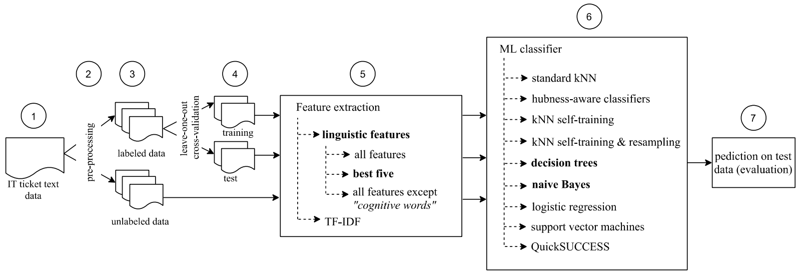


FIGURE 2. Overview of the experimental pipeline.

ticket complexity, such as the number of tasks; the number of Configuration Items (CIs), specifically applications; if the ticket had to be executed online or offline (planning of downtime and considering affected CIs); involvement of Change Advisory Board (CAB), etc. Therefore, labeling a single ticket could take up to 30 minutes. For complexity class labels, a qualitative scale of low, medium, and high has been selected to simplify the classification task and as a well-known scale of priority ratings [90]. Although two datasets are coming from the same case, a distinct period of time and a different number of labels allowed us to test our approaches in two independent settings.

B. EXPERIMENTAL SETTINGS

To evaluate classifiers, we use leave-one-out cross-validation [91]. We select one of the labeled instances as a test instance and use all the remaining instances as training data (step 4). The process of selecting a test instance and training the classifier using all the other instances is repeated as many times as the number of labeled instances. Hereby, in each round, a different instance is selected as the test instance. For example, in the case of Data2, we select one labeled instance out of the available 60 labeled instances as a test instance and use the remaining 59 as labeled training data L. The SSL classifiers also used the 4,684-60 = 4,624 unlabeled instances as unlabeled training data U. We repeat the process 60 times.

As evaluation metrics, we use accuracy, average precision, average recall, and F-score (step 7). To compare the differences in the classifiers' performance, we use the binomial test by Salzberg [92] at the significance threshold of 0.05.

When applying semi-supervised classifiers, we consider all the unlabeled instances available at the training time. Whenever the opposite is not stated, we use kNN with k=1 and Euclidian distance, which is also justified by theoretical and empirical studies [93], [94]. In the case of kNN with self-training, we set the number of self-training iterations to 100.

In respect to support vector machines (SVMs), we used internal leave-one-out cross-validation on the training data to identify appropriate settings of hyper-parameters. In particular, we considered polynomial and RBF kernels, and we used a grid search to determine the best values of the complexity constant C and the kernel coefficient *gamma* in the range 10^{-5}, 10^{-4}, …, 10^{4}, 10^{5}. As for the degree d of the polynomial kernel, we considered all integer values between 1 and 10.

Regarding linguistic features, we include feature selection tests to identify which features play an important role in prediction quality. Similarly to [95], [96], we use logistic regression to determine the most predictive features based on their weights. We also compare classifiers with a statistical method, i.e., an expert-defined decision rule based on the presence of cognitive words.

C. COMPARISON OF SUCCESS AND QUICKSUCCESS

As an initial experiment, we measured the execution time of SUCCESS and QuickSUCCESS. We used the linguistic features representation. In the case of QuickSUCCESS approach, we selected $r = 100$ instances and trained $m = 100$ models. Table 5 shows our results: the execution time of one round of cross-validation (in seconds) and the accuracy. As one can see, the proposed technique leads to several orders of magnitude speed-up, both in the case of Data1 and Data2, with negligible loss of prediction quality (the difference corresponds to the misclassification of one additional instance). Hence, in further experiments, we used the QuickSUCCESS algorithm.

TABLE 5. Execution time (in seconds) and accuracy of SUCCESS and QuickSUCCESS.

	time (seconds)	accuracy
Data1 SUCCESS	75,831	0.567
Data1 QuickSUCCESS	2	0.533
Data2 SUCCESS	452	0.767
Data2 QuickSUCCESS	2	0.750

D. RESULTS

Table 6 provides the obtained values of accuracy, an average of precision and recall calculated over the three classes of low, medium, and high complexity. F-score is calculated based on the average precision and recall.

We compare the classifiers' performance using linguistic features with that of classifiers using TF-IDF on Data1 and Data2. We point out the systematic improvement in the prediction quality of algorithms when using linguistic features.

TABLE 6. Evaluation results using linguistic features and TF-IDF.

Algorithm	Accuracy	Average precision	Average recall	F-score
Data1: linguistic features				
standard kNN	0.533	0.526	0.498	0.511
kNN self-training	0.533	0.526	0.498	0.511
kNN self-training & resampling	0.533	0.526	0.498	0.511
ECkNN	0.633	0.599	0.580	0.589
HFNN	0.600	0.411	0.490	0.447
HWkNN	0.533	0.526	0.498	0.511
NHBNN	0.433	0.144	0.333	0.202
decision trees	1.000	1.000	1.000	1.000
naïve Bayes	0.967	0.972	0.944	0.958
logistic regression	0.500	0.475	0.463	0.469
SVM	0.600	0.575	0.579	0.577
QuickSUCCESS	0.533	0.542	0.523	0.532
Data1: TF-IDF				
standard kNN	0.200	0.067	0.333	0.111
kNN self-training	0.200	0.067	0.333	0.111
kNN self-training & resampling	0.200	0.067	0.333	0.111
ECkNN	0.433	0.144	0.333	0.202
HFNN	0.433	0.144	0.333	0.202
HWkNN	0.200	0.067	0.333	0.111
NHBNN	0.433	0.144	0.333	0.202
decision trees	0.433	0.144	0.333	0.202
naïve Bayes	0.267	0.738	0.389	0.510
logistic regression	0.433	0.144	0.333	0.202
SVM	0.400	0.138	0.308	0.190
QuickSUCCESS	0.200	0.067	0.333	0.111
Data2: linguistic features				
standard kNN	0.750	0.593	0.576	0.584
kNN self-training	0.750	0.593	0.576	0.584
kNN self-training & resampling	0.750	0.593	0.576	0.584
ECkNN	0.733	0.586	0.568	0.577
HFNN	0.733	0.539	0.528	0.534
HWkNN	0.750	0.593	0.576	0.584
NHBNN	0.700	0.233	0.333	0.275
decision trees	1.000	1.000	1.000	1.000
naïve Bayes	1.000	1.000	1.000	1.000
logistic regression	0.800	0.552	0.582	0.567
SVM	0.767	0.531	0.544	0.537
QuickSUCCESS	0.750	0.593	0.576	0.584
Data2: TF-IDF				
standard kNN	0.700	0.233	0.333	0.275
kNN self-training	0.700	0.233	0.333	0.275
kNN self-training & resampling	0.700	0.233	0.333	0.275
ECkNN	0.700	0.233	0.333	0.275
HFNN	0.700	0.233	0.333	0.275
HWkNN	0.700	0.233	0.333	0.275
NHBNN	0.700	0.233	0.333	0.275
decision trees	0.700	0.233	0.333	0.275
naïve Bayes	0.467	0.419	0.468	0.442
logistic regression	0.700	0.233	0.333	0.275
SVM	0.700	0.233	0.333	0.275
QuickSUCCESS	0.700	0.233	0.333	0.275

Namely, we observed that the classifiers' performance with linguistic features was almost always higher than that one with TF-IDF under comparable conditions (the difference between the two experiments is only text representation technique). In the case of TF-IDF features, most classifiers predicted the dominant class for the vast majority of the instances, resulting in relatively low values for precision and recall, i.e., all classifiers had difficulty predicting the correct class labels.

Due to the higher number of labeled tickets, Data2 revealed an expected systematic classification quality increase compared to Data1, independently of applied algorithm or text representation technique.

When enhancing kNN with a self-training, we expected a noticeable increase in performance quality. Nonetheless, the evaluation results evidenced no improvement.

As stated in the Experimental settings subsection, using logistic regression, we identified the most predictive features for Data1 and Data2. According to the weights of logistic regression, in the case of Data1, the five most important features are (1) relative occurrence of cognitive words, (2) occurrence of unique adjectives in all adjectives, (3) wording style, (4) occurrence of unique verbs in all verbs, (5) relative occurrence of words with positive sentiment. In the case of Data2, the five most important features are (1) relative occurrence of cognitive words, (2) occurrence of unique adjectives in all adjectives, (3) occurrence of unique verbs in all verbs, (4) relative occurrence of words with negative and (5) positive sentiments. As can be concluded, the most essential feature appeared to be a relative occurrence of cognitive words, further referred to as the *"cognitive words"* features. After that, we trained the classifiers with the selected linguistic features. The results are summarized in Table 7.

As can be seen in Tables 6 and 7, the usage of our linguistic features delivers excellent performance with simple algorithms, such as decision trees and naïve Bayes. Both of them are statistically significantly better than other classifiers. Additionally, we have shown that selecting the best set of features improves the performance of classifiers consistently. Hence, using a smaller set of features also simplifies the process of extraction and reduces computational costs.

As mentioned above and according to the weights of logistic regression, the most important feature in the ticket text appears to be the *"cognitive words"* feature. To confirm this observation, we tested the algorithms' performance without this feature (see Table 8). In the case of exclusion of the most predictive feature, we see a statistically significant decrease in terms of prediction quality both for decision trees and naïve Bayes. In the case of classifiers that treat all features equally important (e.g., kNN and its variants), the accuracy was generally low, and we did not observe differences when excluding the *"cognitive words"* feature.

Furthermore, we experimented with a statistical method, i.e., the expert decision rule based on the occurrence of the most predictive "*cognitive words*" feature in a ticket text denoted as *COG*. The prediction ($\hat{y}$) is based on our already mentioned previous research [25], and it is determined as follows:

TABLE 7. Evaluation results using the five best performing linguistic features.

Algorithm	Accuracy	Average precision	Average recall	F-score
Data1: five best linguistic features				
standard kNN	0.667	0.589	0.585	0.587
kNN self-training	0.667	0.589	0.585	0.587
kNN self-training & resampling	0.667	0.589	0.585	0.587
ECkNN	0.667	0.578	0.580	0.579
HFNN	0.633	0.429	0.515	0.468
HWkNN	0.667	0.589	0.585	0.587
NHBNN	0.433	0.144	0.333	0.202
decision trees	1.000	1.000	1.000	1.000
naïve Bayes	1.000	1.000	1.000	1.000
logistic regression	0.633	0.611	0.580	0.595
SVM	0.967	0.976	0.970	0.973
QuickSUCCESS	0.733	0.814	0.636	0.714
Data2: five best linguistic features				
standard kNN	0.817	0.711	0.670	0.690
kNN self-training	0.817	0.711	0.670	0.690
kNN self-training & resampling	0.700	0.233	0.333	0.275
ECkNN	0.750	0.523	0.558	0.539
HFNN	0.800	0.526	0.582	0.553
HWkNN	0.817	0.711	0.670	0.690
NHBNN	0.700	0.233	0.333	0.275
decision trees	1.000	1.000	1.000	1.000
naïve Bayes	1.000	1.000	1.000	1.000
logistic regression	0.850	0.554	0.606	0.579
SVM	0.983	0.958	0.970	0.962
QuickSUCCESS	0.867	0.791	0.693	0.739

$$\hat{y} = \begin{cases} low, & \text{if } COG = 0 \\ medium, & \text{if } 0 < COG < 0.3 \\ high, & \text{otherwise} \end{cases}$$

This expert rule is competitive with some of the examined classifiers. However, its performance is significantly worse than that of decision trees and naïve Bayes (see Table 9).

In the case of kNN, we also tested other k values, in particular, all odd numbers between 1 and 10 and different distance functions, such as Euclidian, Manhattan, and cosine. As a result, we did not observe any substantial differences (see Figure 3).

V. DISCUSSION

In this study, our focus was to gain a better understanding of the factors influencing the quality of prediction in text classification tasks. At the same time, we addressed (i) the need for further experiments in industrial settings, especially with recent classifiers that have not been used for ticket classification before, and (ii) limitations of our rule-based approach.

TABLE 8. Evaluation results using linguistic features without "*cognitive words*" feature.

Algorithm	Accuracy	Average precision	Average recall	F-score
Data1: linguistic features except for the "cognitive words" feature				
standard kNN	0.533	0.526	0.498	0.511
kNN self-training	0.533	0.526	0.498	0.511
kNN self-training & resampling	0.533	0.526	0.498	0.511
ECkNN	0.633	0.599	0.580	0.589
HFNN	0.600	0.411	0.490	0.447
HWkNN	0.533	0.526	0.498	0.511
NHBNN	0.433	0.144	0.333	0.202
decision trees	0.233	0.188	0.189	0.188
naïve Bayes	0.467	0.409	0.483	0.443
logistic regression	0.367	0.274	0.296	0.284
SVM	0.633	0.610	0.610	0.610
QuickSUCCESS	0.523	0.523	0.523	0.523
Data2: linguistic features except for the "cognitive words" feature				
standard kNN	0.750	0.593	0.576	0.584
kNN self-training	0.750	0.593	0.576	0.584
kNN self-training & resampling	0.750	0.593	0.576	0.584
ECkNN	0.733	0.586	0.568	0.577
HFNN	0.733	0.539	0.528	0.534
HWkNN	0.750	0.593	0.576	0.584
NHBNN	0.700	0.233	0.333	0.275
decision trees	0.717	0.538	0.560	0.548
naïve Bayes	0.400	0.379	0.409	0.393
logistic regression	0.767	0.504	0.544	0.523
SVM	0.750	0.529	0.536	0.533
QuickSUCCESS	0.767	0.603	0.584	0.593

TABLE 9. Evaluation results using expert decision rule.

Data	Accuracy	Average precision	Average recall	F-score
Data1	0.633	0.451	0.667	0.538
Data2	0.883	0.537	0.667	0.595

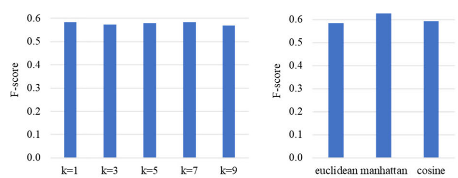


FIGURE 3. kNN performance with various k values (in the left, using Euclidean distance) and distance functions (in the right, k=1) on Data2.

As stated at the beginning of the study, since Artificial Intelligence (AI) becomes an increasing part of our business and daily lives, we recognize the paramount importance of the explainability of AI-based solutions. The need to understand the decisions made by an AI system is crucial in the context of trust and primarily to efficiently address the errors made by such a system. In our work, we closely study the ITIL Change Management-based IT ticketing process of a big telecommunication company. Implementing the changes in the IT infrastructure means intervening in the organization's IT environment, its functions, and experience every time the change is performed. Below, we review the study findings in light of the experimental results.

Our study findings show that feature selection is an important component of ticket classification pipelines not only to reduce the dimensionality of the data and computational costs but also to increase the performance of classifiers. We demonstrate that five linguistic features we identified based on the weights of logistic regression are enough to deliver accurate predictions. While comparing the algorithms' performance with and without the most important feature *relative occurrence of cognitive words*, we evidence that this feature undoubtedly influences the prediction results. In the case of kNN classifier when experimenting with different values of k and distance functions, we also show that their choice does not significantly impact the prediction, especially when comparing with the choice of features (TF-IDF vs. linguistic features).

As there is a general discussion on the advantages and disadvantages of various text representation techniques and classification algorithms, in our paper, we offer a new analysis of both design decisions (text representation and classifier) and their interaction. First, we systematically compared the efficiency of linguistic features and TF-IDF representations in the context of IT ticket complexity prediction. To the best of our knowledge, this is the first attempt of such a research setting. Second, using different datasets and text representation techniques, we consistently tested twelve ML classifiers and showed that simple algorithms, such as decision trees and naïve Bayes, achieved accurate prediction with the linguistic features representation. Hereby, the five best performing features delivered results of the same quality. Hence, building explainable text classification pipelines, i.e., using linguistic features with simple algorithms, can have great potential when applied in real-world tasks.

Our study offers valuable insights for managers and ML experts by enabling them to understand the interdependencies or their absence between selected text representation techniques, classification algorithms, and prediction quality.

The most remarkable practical implication of this study is improving our rule-based approach. We used the described linguistic features approach and both handcrafted and decision trees-based rules to predict the IT ticket complexity value of low, medium, or high [25]. Using the ML classification pipeline discussed in the paper, we managed to improve prediction quality significantly without the need to define and constantly update the rules with the experts, which is a big hurdle in the management, maintenance, and scalability of such systems.

In real-life scenarios, IT ticket processing teams continuously undergo changes and fluctuations, causing the resolver groups to be split, merged, or reorganized. Due to the fast technological developments, the problems they have to solve also change. Hence, the trained machine learning model can become outdated very fast and will demand retraining [97]. This again justifies our message – the simpler, the better, i.e., we recommend using those classifiers which do not demand complex parameter search and are simple and cheap to implement if their prediction quality is acceptable.

To sum up, systematic testing of the discussed representation techniques and classification algorithms and their application for the IT ticket classification task of complexity prediction can be of interest for text data scientists and managers who seek to understand the role of factors influencing the prediction.

VI. CONCLUSION

Our work aimed to provide a comparative analysis of text representation techniques and classifiers while developing an IT ticket classification pipeline. Our observations can be useful for the design of decision support systems in enterprise applications, specifically in the IT ticket area.

The contributions of our work can be summarized as follows: (i) the comprehensive comparative analysis of linguistic features with TF-IDF and various ML algorithms confirms the positive influence of linguistic style predictors [52] on the prediction quality; (ii) our observation that simple algorithms work well if using appropriate linguistic features contributes to the general discussion on the advantages and disadvantages of various text classification algorithms [13], [14]; (iii) we showed that ML-based IT ticket classification outperforms our rule-based approach.

As a part of future work, one can: (i) consider further information regarding IT ticket complexity prediction, such as the number of tasks and configuration items per ticket; (ii) test other application cases in the IT ticket area and beyond, i.e., further explore the potential of linguistic features; (iii) as we showed that selecting an appropriate subset of linguistic features can considerably improve the performance of classifiers, one may conduct further experiments with more advanced feature selection techniques [98].

APPENDIX

List of attached Appendices:

Appendix I. Taxonomy of Decision-Making Logic Levels

Appendix II. Business Sentiment Lexicon with assigned valences

APPENDIX I. **Taxonomy of decision-making logic levels. following [24], we consider diverse semantic concepts: *Resources, Techniques, Capacities*, and *Choices*, elements of RTCC framework. We designed contextual variables [99], based on which experts categorized words into one of the three DML levels and one of the four semantic concepts.**

CONTEXTUAL VARIABLES	DECISION-MAKING LOGIC LEVELS		
	routine	*semi-cognitive*	*cognitive*
	CONCEPTUAL ASPECTS		
	RESOURCES		
Problem Processing Level	user, user request, task, test, check, target, release, contact role, access, interface, cluster, tool, client, file system, partner, node	team, leader, project, colleague, property	management, system, CAB, measure, approval
Accuracy	time, application, product, configuration item, CI, right, instance, machine, minute, hour, day, week, detail, description	description, environment, requirement, validity, reason, solution, method, problem, rule, modification	
Situational Awareness	name, password, group, directory, number, email, package, phone, ID, IP, attachment	request for change, RfC, customer, rollout, backout	server farm
Information	server, file, location, dataset, network, data, patch, port, information, type, root, certificate, account, device, cable, parameter, agent, folder, disk, fallback, database, db, backup, version, tool, firewall, system, hotfix, supervisor, reference, instruction, format	Requestor, software, downtime, production, power-supply, outage, service, case	risk, freeze, impact
	TECHNIQUES		
Experience	need, see, deploy, document, monitor, use, follow, note, provide, test, contain, accompany, inform, consist, describe	implement, create, support, require, classify	approve, delegate, propose
Action Choice	start, finish, monitor, import, export, run, stop, step, end, put, send, switch, install, reject, update, upgrade, include, replace, remove, move, begin, make, get, migrate, open, initialize, revoke	deploy, migrate, process, modify, forget, increase, miss	freeze
Effort	cancel, rundown, decommission, restart, delete, set, add, activate, reboot, specify, agree upgrade, mount, execute, transfer, write, find	perform, modify, assign, check, need, expect, verify	define
	CAPACITIES		
Specificity	additional, preapproved, affected, initial, attached, internal, external, reachable, regular, active, scheduled, next, whole, formal, virtual, wrong, individual, administrative, local	secure, separate, specific, technical, urgent, separate, corrected, minor, normal	related, multiple, multi-solution, major, high, small, big
Decisions Formulation	new, old, preinstalled, fixed, ready, following, current, valid, primary, necessary	available, necessary, important, significant, successful, appropriate, relevant, main, further, responsible	possible, many, desired, different, various
Predictability	actual, full, online, standard, responsible, administrative, existing, minimum, same, visible	strong, temporary, offline, previous, last, other, more, much, similar, standard	random, strong randomized, encrypted, expected
	CHOICES		
Precision	automatically, instead, manually, there, where, here, separately, additionally, internally	normally, newly, shortly, urgently, temporarily	maybe, randomly, likely
Scale	permanently, currently, still, now, often, never, already, just, always, yet, anymore, firstly, before, together, daily, meanwhile, really, furthermore, afterwards, therefore	again, later, however, usually, previously, recently	soon
Ambiguity	correctly, therefore, accordingly, actually, consequently, completely, simultaneously, anyway, necessarily	well, enough, immediately, easily, simply	approximately, properly

APPENDIX II. **Business sentiment lexicon with assigned valences.**

Tickets	ITIL	Valence
Expressions		
no risk, no outage		+2
be so kind, would be nice		0.5
disaster recovery, set alarms warnings, poison attack vulnerability, critical security leaks, fan, outstanding windows updates, thank you, kind regards, would like, best regards	request for change, RfC	0
big measure	projected service outage, change advisory board, high impact, major change	-0.5
Single keywords		
kind, success, correct, like, nice	well, successful, happy	0.5
disaster, recovery, affected, stop, disable, dump, alarm, warning, poison, attack, vulnerability, error, prevent, drop, cancel, delete, exclude, problem, problems, faulty, failed, destroy, defective, obsolete, lack, security, leak, crash, please, support, optimize, grant, privilege, create, dear, acceptance, clarity, restore, increase, danger, balance, right, deny, wrong, retire, missing, weak, invalid, see, follow, yes, allow, approve, approval, confirm	problem, failed, information, operational, identify, order, include, adequately, procedure, necessary, assess, criteria, clear, provide, potentially, identification, adequate, initiate, value, KPI, standard, schedule, align, properly, release, accurate, report, organization, continuous, ensure, service, beneficial, stakeholder, requirement, correct, record, essential, clearly, RfC, support, tool, relevant, attempt, subsequently, configuration, different, follow, directly, CI, potential, request, individual, plan, work, evaluate, author, organizational, manage, number, financial, status, low, chronological, recommend, responsible, model, accountable, handle, timescale, business, normal, submit, update, create, manual, consider, backout, accept, item, project, deliver, formal, data, iterative, produce, local, describe, test, improve, result, deployment, deploy, technical, management, repeatable, determine, minimum, develop, appropriate, activate, implement, require, process, evaluation, customer, contractual, authorize, share, acceptable	0
blocked, critical	cost, PSO, CAB, important, unauthorized, major, significant, undesirable, incomplete, delegate, avoid, coordinate, immediately, significantly	-0.5
offline, risk, outage, emergency, downtime	impact, risk, emergency, incident, outage, downtime	-1
rejected	unacceptable	-2

REFERENCES

[1] Y. Diao and K. Bhattacharya, "Estimating business value of IT services through process complexity analysis," in *Proc. IEEE Netw. Oper. Manage. Symp. (NOMS)*. Salvador, Brazil, 2008.

[2] S. P. Paramesh and K. S. Shreedhara, *Automated IT Service Desk Systems Using Machine Learning Techniques* (Lecture Notes in Networks and Systems), vol. 43. Singapore: Springer, 2019, pp. 331–346.

[3] S. P. Paramesh, C. Ramya, and K. S. Shreedhara, "Classifying the unstructured IT service Des. tickets using ensemble of classifiers," in *Proc. 3rd Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solutions (CSITSS)*, Dec. 2018, pp. 221–227.

[4] S. Roy, D. P. Muni, J. J. Y. T. Yan, N. Budhiraja, and F. Ceiler, *Clustering and Labeling IT Maintenance Tickets*. Berlin, Germany: Springer, 2016.

[5] G. B. Dasgupta, T. K. Nayak, A. R. Akula, S. Agarwal, and S. J. Nadgowda, *Towards Auto-Remediation in Services Delivery: Context-Based Classification of Noisy and Unstructured Tickets*. Berlin, Germany: Springer, 2014.

[6] S. Agarwal, R. Sindhgatta, and B. Sengupta, "SmartDispatch: Enabling efficient ticket dispatch in an IT service environment," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, vol. 2012, pp. 1393–1401.

[7] S. Agarwal, V. Aggarwal, A. R. Akula, G. B. Dasgupta, and G. Sridhara, "Automatic problem extraction and analysis from unstructured text in IT tickets," *IBM J. Res. Develop.*, vol. 61, no. 1, pp. 41–52, Jan. 2017.

[8] E. Orta, M. Ruiz, N. Hurtado, and D. Gawn, "Decision-making in IT service management: A simulation based approach," *Decis. Support Syst.*, vol. 66, pp. 36–51, Oct. 2014.

[9] M. Ruiz, J. Moreno, B. Dorronsoro, and D. Rodriguez, "Using simulation-based optimization in the context of IT service management change process," *Decis. Support Syst.*, vol. 112, pp. 35–47, Aug. 2018.

[10] E. Fielt, T. Böhmann, A. Korthaus, S. Conger, and G. Gable, "Service management and engineering in information systems research," *J. Strategic Inf. Syst.*, vol. 22, no. 1, pp. 46–50, Mar. 2013.

[11] T. R. Eikebrokk and J. Iden, "Strategising IT service management through ITIL implementation: Model and empirical test," *Total Qual. Manage. Bus. Excellence*, vol. 28, nos. 3–4, pp. 238–265, Feb. 2017.

[12] R. D. Galliers, "Towards a flexible information architecture: Integrating business strategies, information systems strategies and business process redesign," *Inf. Syst. J.*, vol. 3, no. 3, pp. 199–213, Jul. 1993.

[13] Kowsari, J. Meimandi, Heidarysafa, Mendu, Barnes, and Brown, "Text classification algorithms: A survey," *Information*, vol. 10, no. 4, p. 150, Apr. 2019.

[14] Y. C. Cavalcanti, P. A. da Mota Silveira Neto, I. D. C. Machado, T. F. Vale, E. S. de Almeida, and S. R. D. L. Meira, "Challenges and opportunities for software change request repositories: A systematic mapping study," *J. Softw., Evol. Process*, vol. 26, no. 7, pp. 620–653, Jul. 2014.

[15] A. Lazarov and P. Shoval, "A rule-based system for automatic assignment of technicians to service faults," *Decis. Support Syst.*, vol. 32, no. 4, pp. 343–360, Mar. 2002.

[16] J. Kanwal and O. Maqbool, "Bug prioritization to facilitate bug report triage," *J. Comput. Sci. Technol.*, vol. 27, no. 2, pp. 397–412, Mar. 2012.

[17] T. Menzies and A. Marcus, "Automated severity assessment of software defect reports," in *Proc. IEEE Int. Conf. Softw. Maintenance*, Sep. 2008, pp. 346–355.

[18] S. N. Ahsan and F. Wotawa, "Impact analysis of SCRs using single and multi-label machine learning classification," in *Proc. ACM-IEEE Int. Symp. Empirical Softw. Eng. Meas. (ESEM)*, 2010, pp. 1–4.

[19] S. N. Ahsan, J. Ferzund, and F. Wotawa, "Automatic classification of software change request using multi-label machine learning methods," in *Proc. 33rd Annu. IEEE Softw. Eng. Workshop*, Oct. 2009, pp. 79–86.

[20] V. Rus, X. Nan, S. Shiva, and Y. Chen, "Clustering of defect reports using graph partitioning algorithms," in *Proc. 21st Int. Conf. Softw. Eng. Knowl. Eng. (SEKE)*, 2009, pp. 442–445.

[21] A. Santana, J. Silva, P. Muniz, F. Araújo, and R. M. C. R. de Souza, *Comparative Analysis of Clustering Algorithms Applied to the Classification of Bugs*. Berlin, Germany: Springer, 2012.

[22] N. Rizun, A. Revina, and V. Meister, "Discovery of stylistic patterns in business process textual descriptions: IT ticket case," in *Proc. Conf., 33rd Int. Bus. Inf. Manage. Assoc. Conf. (IBIMA)*, 2019, pp. 2103–2113.

[23] N. Rizun and A. Revina, "Business sentiment analysis. Concept and method for perceived anticipated effort identification," in *Proc. Inf. Syst. Develop., Inf. Syst. Beyond (ISD)*, A. Siarheyeva, C. Barry, M. Lang, H. Linger, and C. Schneider, Eds. Toulon, France: ISEN Yncréa Méditerranée, 2019. [Online]. Available: https://aisel.aisnet.org/isd2014/proceedings2019/ManagingISD/3/

[24] N. Rizun, A. Revina, and V. Meister, "Method of decision-making logic discovery in the business process textual data," in *Proc. Int. Conf. Bus. Inf. Syst. (BIS)*, Seville, Spain. Cham, Switzerland: Springer, 2019, pp. 70–84.

[25] A. Revina and N. Rizun, "Multi-criteria knowledge-based recommender system for decision support in complex business processes," in *Proc. Workshop Workshop Recommendation Complex Scenarios Co-Located 13th ACM Conf. Recommender Syst. (RecSys)*, 2019, pp. 16–22.

[26] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Inf. Process. Manage.*, vol. 24, no. 5, pp. 513–523, Jan. 1988.

[27] N. Tomašev and K. Buza, "Hubness-aware kNN classification of high-dimensional data in presence of label noise," *Neurocomputing*, vol. 160, pp. 157–172, Jul. 2015.

[28] N. Tomašev, K. Buza, K. Marussy, and P. B. Kis, "Hubness-aware classification, instance selection and feature construction: Survey and extensions to time-series," in *Feature Selection for Data and Pattern Recognition*, U. Stańczyk and L. C. Jain, Eds. Berlin, Germany: Springer, 2015, pp. 231–262.

[29] J. R. Quinlan, "Induction of decision trees," *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, Mar. 1986.

[30] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge, U.K.: Cambridge Univ. Press, 2009, pp. 253–286.

[31] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, vol. 398. Hoboken, NJ, USA: Wiley, 2013.

[32] T. Joachims, *Text Categorization With Support Vector Machines: Learning With Many Relevant Features*. Berlin, Germany: Springer, 1998, pp. 137–142.

[33] I. Triguero, S. García, and F. Herrera, "Self-labeled techniques for semi-supervised learning: Taxonomy, software and empirical study," *Knowl. Inf. Syst.*, vol. 42, no. 2, pp. 245–284, Feb. 2015.

[34] K. Marussy and K. Buza, "SUCCESS: A new approach for semi-supervised classification of time-series," in *Proc. Int. Conf. Artif. Intell. Soft Comput.*, 2013, pp. 437–447.

[35] K. Buza and A. Revina, "Speeding up the SUCCESS approach for massive industrial datasets," in *Proc. Int. Conf. Innov. Intell. Syst. Appl. (INISTA)*, Aug. 2020, pp. 1–6.

[36] S. Scott and S. Matwin, "Text classification using WordNet Hypernyms," in *Proc. Usage WordNet Natural Language Process. Syst.*, 1998, pp. 45–51.

[37] J. Yan, "Text representation," in *Encyclopedia of Database Systems*. Boston, MA, USA: Springer, 2009, pp. 3069–3072.

[38] Y. Zhang, R. Jin, and Z.-H. Zhou, "Understanding bag-of-words model: A statistical framework," *Int. J. Mach. Learn. Cybern.*, vol. 1, nos. 1–4, pp. 43–52, Dec. 2010.

[39] K. J. Sparck, "A statistical interpretation of term specificity and its application in retrieval," *J. Document.*, vol. 28, no. 1, pp. 11–21, Jan. 1972.

[40] M. Radovanović and M. Ivanović, "Text mining: Approaches and applications," *Novi Sad J. Math*, vol. 38, no. 3, pp. 227–234, 2008.

[41] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," 2013, *arXiv:1301.3781*. [Online]. Available: http://arxiv.org/abs/1301.3781

[42] Q. V. Le and T. Mikolov, "Distributed representations of sentences and documents," 2014, *arXiv:1405.4053*. [Online]. Available: http://arxiv.org/abs/1405.4053

[43] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2014, pp. 1532–1543.

[44] A. Joulin, E. Grave, P. Bojanowski, M. Douze, H. Jégou, and T. Mikolov, "FastText.Zip: Compressing text classification models," 2016, *arXiv:1612.03651*. [Online]. Available: http://arxiv.org/abs/1612.03651

[45] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," 2016, *arXiv:1607.04606*. [Online]. Available: http://arxiv.org/abs/1607.04606

[46] O. Melamud, J. Goldberger, and I. Dagan, "Context2vec: Learning generic context embedding with bidirectional LSTM," in *Proc. 20th SIGNLL Conf. Comput. Natural Lang. Learn.*, 2016, pp. 51–61.

[47] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep contextualized word representations," 2018, *arXiv:1802.05365*. [Online]. Available: http://arxiv.org/abs/1802.05365

[48] S. Mahdi Rezaeinia, A. Ghodsi, and R. Rahmani, "Improving the accuracy of pre-trained word embeddings for sentiment analysis," 2017, *arXiv:1711.08609*. [Online]. Available: http://arxiv.org/abs/1711.08609

[49] F. Yin, Y. Wang, J. Liu, and L. Lin, "The construction of sentiment lexicon based on context-dependent part-of-speech chunks for semantic disambiguation," *IEEE Access*, vol. 8, pp. 63359–63367, 2020.

[50] A. Sureka and K. V. Indukuri, "Linguistic analysis of bug report titles with respect to the dimension of bug importance," in *Proc. 3rd Annu. ACM Bengaluru Conf. (COMPUTE)*, 2010, pp. 1–6.

[51] A. J. Ko, B. A. Myers, and D. Horng Chau, "A linguistic analysis of how people describe software problems," in *Proc. Vis. Lang. Human-Centric Comput. (VL/HCC)*, 2006, pp. 127–134.

[52] K. Coussement and D. Van den Poel, "Improving customer complaint management by automatic email classification using linguistic style features as predictors," *Decis. Support Syst.*, vol. 44, no. 4, pp. 870–882, Mar. 2008.

[53] J. Fürnkranz, T. Mitchell, and E. Riloff, "Case study in using linguistic phrases for text categorization on the WWW," Assoc. Advancement Artif. Intell., Palo Alto, CA, USA, Workshop Ser. Tech. Rep. WS-98-05, 1998.

[54] D. Mladenic and M. Grobelnik, "Word sequences as features in text-learning," in *Proc. 17th Electrotech. Comput. Sci. Conf. (ERK)*, 1998, pp. 145–148.

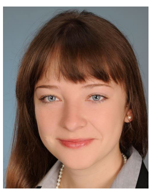
[55] B. Raskutti, H. Ferrá, and A. Kowalczyk, "Second order features for maximising text classification performance," in *Proc. Eur. Conf. Mach. Learn.*, 2001, pp. 419–430.

[56] R. Bekkerman, R. El-Yaniv, N. Tishby, and Y. Winter, "On feature distributional clustering for text categorization," in *Proc. 24th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR)*, 2001, pp. 146–153.

[57] C.-M. Tan, Y.-F. Wang, and C.-D. Lee, "The use of bigrams to enhance text categorization," *Inf. Process. Manage.*, vol. 38, no. 4, pp. 529–546, Jul. 2002.

[58] S. Scott and S. Matwin, "Feature engineering for text classification," in *Proc. ICML*, 1999, pp. 379–388.

[59] A. Moschitti and R. Basili, "Complex linguistic features for text classification: A comprehensive study," in *Proc. Eur. Conf. Inf. Retr.*, 2004, pp. 181–196.
[60] M. Sasaki and K. Kita, "Rule-based text categorization using hierarchical categories," in *Proc. Conf. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 1998, pp. 2827–2830.
[61] J. Mason, Qualitative researching. Sage Publications Ltd, 2002.
[62] J. Seidel and U. Kelle, *Computer-Aided Qualitative Data Analysis: Theory, Methods and Practice*. London, U.K.: SAGE, 1995.
[63] S. Chua, F. Coenen, and G. Malcolm, "Classification inductive rule learning with negated features," in *Proc. 6th Int. Conf. Adv. Data Mining Appl. (ADMA)*, 2010, pp. 125–136.
[64] L. Rokach, "Decision forest: Twenty years of research," *Inf. Fusion*, vol. 27, pp. 111–125, Jan. 2016.
[65] E. García, C. Romero, S. Ventura, and T. Calders, "Drawbacks and solutions of applying association rule mining in learning management systems" in *Proc. Int. Workshop Applying Data Mining e-Learning*, 2007, pp. 13–22.
[66] G. Kaur, "Association rule mining: A survey," *Int. J. Comput. Sci. Inf. Technol.*, vol. 5, no. 2, pp. 2320–2324, 2014.
[67] H. Hu and J. Li, "Using association rules to make rule-based classifiers robust," in *Proc. 16th Austral. Database Conf.*, 2005, pp. 47–54.
[68] V. Chakravarthy, S. Joshi, G. Ramakrishnan, S. Godbole, and S. Balakrishnan, "Learning decision lists with known rules for text mining," in *Proc. 3rd Int. Joint Conf. Natural Lang. Process.*, 2008, pp. 835–840.
[69] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.
[70] O. Uzuner, X. Zhang, and T. Sibanda, "Machine learning and rule-based approaches to assertion classification," *J. Amer. Med. Inform. Assoc.*, vol. 16, no. 1, pp. 109–115, Jan. 2009.
[71] X. Zhu, *Semi-Supervised Learning Literature Survey*. New York, NY, USA: Wisconsin, Jul. 2008.
[72] E. Frank and R. R. Bouckaert, "Naive Bayes for text classification with unbalanced classes," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (LNAI), vol. 4213. 2006, pp. 503–510.
[73] W. Daelemans, "Explanation in computational stylometry," in *Proc. Int. Conf. Intell. Text Process. Comput. Linguistics*, 2013, pp. 451–462.
[74] *Taxonomy Strategies | Bibliography–Taxonomy Strategies*. Accessed: Sep. 2, 2019. [Online]. Available: https://taxonomystrategies.com/library/bibliography/
[75] A. Blumauer. *Taxonomies and Ontologies | LinkedIn Blog Article*. Accessed: Sep. 2, 2019. [Online]. Available: https://www.linkedin.com/pulse/taxonomies-ontologies-andreas-blumauer/
[76] D. M. Blei, "Probabilistic topic models," *Commun. ACM*, vol. 55, no. 4, pp. 77–84, Apr. 2012.
[77] B. Liu, "Sentiment analysis and opinion mining," *Synth. Lectures Hum. Lang. Technol.*, vol. 5, no. 1, pp. 1–184, 2012.
[78] C. J. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proc. 8th Int. Conf. Weblogs Social Media (ICWSM)*, 2014, pp. 1–10.
[79] G. K. Zipf, *Selected Studies of the Principle of Relative Frequency in Language*. Cambridge, MA, USA: Harvard Univ. Press, 1932.
[80] J. Shao and D. Tu, *The Jackknife and Bootstrap*. New York, NY, USA: Springer, 1995.
[81] E. Bauer and R. Kohavi, "An empirical comparison of voting classification algorithms: Bagging, boosting, and variants," *Mach. Learn.*, vol. 36, nos. 1–2, pp. 105–139, 1999.
[82] L. Firte, C. Lemnaru, and R. Potolea, "Spam detection filter using KNN algorithm and resampling," in *Proc. IEEE 6th Int. Conf. Intell. Comput. Commun. Process.*, Aug. 2010, pp. 27–33.
[83] V. Ng and C. Cardie, "Weakly supervised natural language learning without redundant views," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics Human Lang. Technol. (NAACL)*, 2003, pp. 173–180.
[84] K. Hara, I. Suzuki, M. Shimbo, K. Kobayashi, K. Fukumizu, and M. Radovanovic, "Localized centering: Reducing hubness in large-sample data," in *Proc. 29th AAAI Conf. Artif. Intell.*, B. Bonet and S. Koenig, Eds. Austin, TX, USA: AAAI Press, Jan. 2015, pp. 2645–2651. [Online]. Available: http://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9898
[85] K. Buza and D. Neubrandt, "A new proposal for person identification based on the dynamics of typing: Preliminary results," *Theor. Appl. Informat.*, vol. 28, nos. 1–2, pp. 1–12, Feb. 2017.
[86] K. Buza, A. Nanopoulos, and G. Nagy, "Nearest neighbor regression in the presence of bad hubs," *Knowl.-Based Syst.*, vol. 86, pp. 250–260, Sep. 2015.
[87] M. Radovanović, A. Nanopoulos, and M. Ivanović, "Nearest neighbors in high-dimensional data," in *Proc. 26th Annu. Int. Conf. Mach. Learn. (ICML)*, 2009, pp. 1–8.
[88] N. Tomašev, M. Radovanović, D. Mladenić, and M. Ivanović, "Hubness-based fuzzy measures for high-dimensional k-nearest neighbor classification," *Int. J. Mach. Learn. Cybern.*, vol. 5, no. 3, pp. 16–30, 2011.
[89] N. Tomasev, M. Radovanović, D. Mladenić, and M. Ivanović, "A probabilistic approach to nearest-neighbor classification: Naive Hubness Bayesian kNN," in *Proc. 20th ACM Int. Conf. Inf. Knowl. Manage. (CIKM)*, 2011, pp. 2173–2176.
[90] T. L. Saaty, "The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making," in *Multiple Criteria Decision Analysis: State of the Art Surveys*. New York, NY, USA: Springer, 2005, pp. 345–405.
[91] G. I. Webb, "Leave-one-out cross-validation," in *Encyclopedia of Machine Learning*. Boston, MA, USA: Springer, 2011, pp. 600–601.
[92] S. L. Salzberg, "On comparing classifiers: Pitfalls to avoid and a recommended approach," *Data Mining Knowl. Discovery*, vol. 1, no. 3, pp. 317–328, 1997.
[93] X. Xi, E. Keogh, C. Shelton, L. Wei, and C. A. Ratanamahatana, "Fast time series classification using numerosity reduction," in *Proc. 23rd Int. Conf. Mach. Learn. (ICML)*, 2006, pp. 1033–1040.
[94] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, Jan. 1967.
[95] A. Y. Ng, "Feature selection, l1 vs. l2 regularization, and rotational invariance," in *Proc. 21st Int. Conf. Mach. Learn. (ICML)*, 2004, pp. 615–622.
[96] Q. Cheng, P. K. Varshney, and M. K. Arora, "Logistic regression for feature selection and soft classification of remote sensing data," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 4, pp. 491–494, Oct. 2006.
[97] A. Mandal, N. Malhotra, S. Agarwal, A. Ray, and G. Sridhara, "Automated Dispatch of Helpdesk Email Tickets: Pushing the Limits with AI," *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 01, pp. 9381–9388, Jul. 2019.
[98] A. Szenkovits, R. Meszlényi, K. Buza, N. Gaskó, R. Ioana Lung, and M. Suciu, "Feature selection with a genetic algorithm for classification of brain imaging data," in *Advances in Feature Selection for Data and Pattern Recognition*, vol. 138, U. Stanczyk, B. Zielosko, and L. C. Jain, Eds. Cham, Switzerland: Springer, 2018, pp. 185–202.
[99] N. Rizun and Y. Taranenko, "Simulation models of human decision-making processes," *Manage. Dyn. Knowl. Economy*, vol. 2, no. 2, pp. 241–264, 2014.



ALEKSANDRA REVINA received the M.Sc. degree in business administration with a focus on business process and knowledge management and engineering from the Brandenburg University of Applied Sciences, in 2015. Afterward, she worked as a Research Scientist and a Project Manager with the Deutsche Telekom Innovation Laboratories, Berlin. In early 2018, she started her industrial Ph.D. at the Technical University of Berlin and the Brandenburg University of Sciences (cooperative procedure). Her research interests include diverse methods and tools for business process analysis and automation with the goal of developing efficient decision support systems for process workers.

KRISZTIAN BUZA received the diploma degree in computer science from the Budapest University of Technology and Economics, in 2007, and the Ph.D. degree from the University of Hildesheim, in 2011. He is the coauthor of more than 50 publications, including the best paper of the IEEE Conference on Computational Science and Engineering in 2010. He regularly serves as a Reviewer for renowned journals, such as *Neurocomputing*, and *Knowledge-Based Systems* or *Knowledge and Information Systems*. He is also a member of the program committee of international conferences, such as the Pacific-Asia Conference on Knowledge Discovery and Data Mining and the International Conference on Computational Collective Intelligence. His research interests include applied machine learning and data mining.

VERA G. MEISTER received the Ph.D. degree in mathematics from St. Petersburg State University, in 1988. She worked as a Researcher, a Lecturer, and a Project Manager in vocational training for more than 20 years. She joined the Brandenburg University of Applied Sciences in 2012 as a Full Professor for business information systems in the Department of Economics. Her research interests include knowledge engineering and information retrieval.

CHAPTER 5

Enriching Textual Data-Driven Business Process Complexity Analysis with Event Log-Driven Analysis

Summary

This chapter aims to enrich textual data-based BP complexity analysis with the insights obtained with the help of process mining. Whereas the research presented in Chapter 3 and Chapter 4 is taken as a basis for textual data-based BP complexity, for event log-based complexity, a set of metrics is identified. Using the second case study of Service Request IT ticket processing from an academic institution in the Netherlands, the developed textual data-based as well as suggested event log data-based metrics are evaluated, and their relation is investigated. As a contribution, the benefits of using two data types for BP complexity identification are explored and illustrated.

This chapter is based on the following publication:

Revina, A. & Aksu, Ü., "An approach for analyzing business process execution complexity based on textual data and event log", *Information Systems*, vol. 114, 2023, p. 102184, DOI: 10.1016/j.is.2023.102184.

Title An Approach for Analyzing Business Process Execution Complexity Based on Textual Data and Event Log

Authors Aleksandra Revina[1,2] and Ünal Aksu[3]
[1] Technical University of Berlin
[2] Brandenburg University of Applied Sciences
[3] Utrecht University

Published in Information Systems, 2023, Elsevier, impact factor 2021: 3.18.

Abstract With the advent of digital transformation, organizations increasingly rely on various information systems to support their business processes (BPs). Recorded data, including textual data and event log, expand exponentially, complicating decision-making and posing new challenges for BP complexity analysis in Business Process Management (BPM). Herein, Process Mining (PM) serves to derive insights based on historic BP execution data called event log. However, in PM, textual data is often neglected or limited to BP descriptions. Hence, in this study, we propose a novel approach for analyzing BP execution complexity by combining textual data serving as an input at the BP start and event log. The approach is aimed at studying the connection between complexities obtained from these two data types. For textual data-based complexity, the approach employs a set of linguistic features. In our previous work, we have explored the design of linguistic features favorable for BP execution complexity prediction. Accordingly, we adapt and incorporate them into the proposed approach. Using these features, various machine learning techniques are applied to predict textual data-based complexity. Moreover, in this prediction, we show the adequacy of our linguistic features, which outperformed the linguistic features of a widely-used text analysis technique. To calculate event log-based complexity, the event log and relevant complexity metrics are used. Afterward, a correlation analysis of two complexities and an analysis of the significant differences in correlations are performed. The results serve to derive recommendations and insights for BP improvement. We apply the approach in the IT ticket handling process of the IT department of an academic institution. Our findings show that the suggested approach enables a comprehensive identification of BP redesign and improvement opportunities.

Keywords business process execution complexity; event log; IT service management; linguistic features; machine learning; process mining; textual data

Information Systems 114 (2023) 102184

Contents lists available at ScienceDirect

Information Systems

journal homepage: www.elsevier.com/locate/is

An approach for analyzing business process execution complexity based on textual data and event log

Aleksandra Revina [a,b], Ünal Aksu [c,*]

[a] Chair of Information and Communication Management, Faculty of Economics and Management, Technical University of Berlin, 10623 Berlin, Germany
[b] Faculty of Economics, Brandenburg University of Applied Sciences, 14770, Brandenburg an der Havel, Germany
[c] Department of Information and Computing Sciences, Utrecht University, 3584 CC Utrecht, The Netherlands

ARTICLE INFO

Article history:
Received 1 October 2021
Received in revised form 5 July 2022
Accepted 30 January 2023
Available online 3 February 2023
Recommended by Dennish Sasha

Keywords:
Business process execution complexity
Event log
IT service management
Linguistic features
Machine learning
Process mining
Textual data

ABSTRACT

With the advent of digital transformation, organizations increasingly rely on various information systems to support their business processes (BPs). Recorded data, including textual data and event log, expand exponentially, complicating decision-making and posing new challenges for BP complexity analysis in Business Process Management (BPM). Herein, Process Mining (PM) serves to derive insights based on historic BP execution data, called event log. However, in PM, textual data is often neglected or limited to BP descriptions. Therefore, in this study, we propose a novel approach for analyzing BP execution complexity by combining textual data serving as an input at the BP start and event log. The approach is aimed at studying the connection between complexities obtained from these two data types. For textual data-based complexity, the approach employs a set of linguistic features. In our previous work, we have explored the design of linguistic features favorable for BP execution complexity prediction. Accordingly, we adapt and incorporate them into the proposed approach. Using these features, various machine learning techniques are applied to predict textual data-based complexity. Moreover, in this prediction, we show the adequacy of our linguistic features, which outperformed the linguistic features of a widely-used text analysis technique. To calculate event log-based complexity, the event log and relevant complexity metrics are used. Afterward, a correlation analysis of two complexities and an analysis of the significant differences in correlations are performed. The results serve to derive recommendations and insights for BP improvement. We apply the approach in the IT ticket handling process of the IT department of an academic institution. Our findings show that the suggested approach enables a comprehensive identification of BP redesign and improvement opportunities.

1. Introduction

Today's organizations use various information systems like Enterprise Resource Planning (ERP) systems or Information Technology (IT) ticketing systems to support their business processes (BPs) and operations [1]. As such, they highly rely on IT. Since digital transformation engages organizations in rapidly changing environments [2], it continually demands them to have a thorough understanding of their BPs and operations to remain resilient [3].

Accordingly, Business Process Management (BPM) has become popular as a well-known way to enable efficient business operations and improvements in quality and productivity in organizations. To accomplish these goals, BPM research and practice have established various approaches. Process Mining (PM) is one of the commonly used techniques to derive insights for process analysis based on BP execution data extracted as event logs. In this context, a large number of studies in BPM and PM are devoted to BP execution complexity analysis and complexity metrics [4]. However, in these studies focusing on BP executions, the analysis of textual data remains limited [5], despite the fact that textual data make up more than 80% of data in companies [6]. The relevant studies in the literature related to BP executions mainly consider BP descriptions, documentation, and texts in BP models, such as labels [5]. In addition, many unsolved challenges in applying Natural Language Processing (NLP) in BPM are highlighted, such as semantic enhancements and domain or organization-specific adaptations of NLP solutions [5]. Thus, a more rigorous relation between these two areas discloses an untapped potential to substantially improve the BPM toolset.

In fact, textual data serving as an input to a BP at its very start, i.e., triggering a BP, highly influence its execution. For example, in

* Corresponding author.
E-mail addresses: revina@th-brandenburg.de (A. Revina), u.aksu@uu.nl (Ü. Aksu).

https://doi.org/10.1016/j.is.2023.102184
0306-4379/

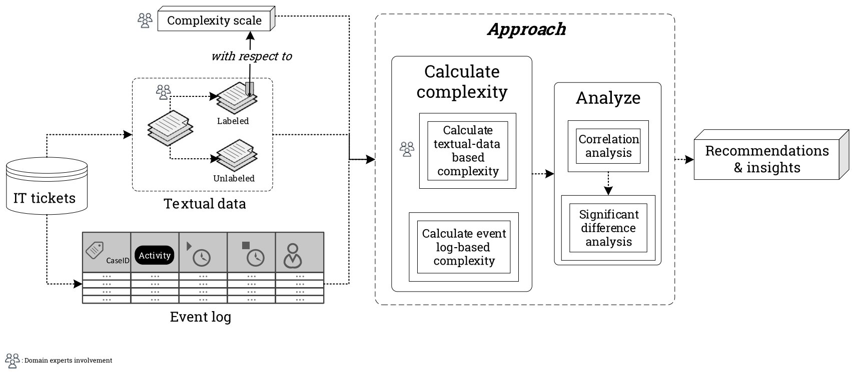


Fig. 1. Our approach for analyzing business process execution complexity based on textual data and event log.

the IT Service Management (ITSM) area, activities performed in a Change Management (CHM) process strongly depend on the textual descriptions of Requests for Changes (RfCs) communicated by a customer. Specifically, the urgency of a given RfC and whether it needs to be analyzed and approved for implementation are determined primarily based on its textual description. Similarly, the involvement of roles, such as a Change Advisory Board (CAB), also depends on the RfC texts. For example, urgent RfCs may not require CAB involvement. Hence, the same input, i.e., textual data, determines important decision points of any RfC processing, such as: (i) which activities in the CHM process will be skipped, (ii) from which activity RfC processing will start, and (iii) which roles will be involved. In many other areas, one can also observe similar influences of textual data on the execution of BPs. For example, in healthcare, the complaints expressed by a patient typically determine the required diagnostics and related BP activities. Generally speaking, in BPs, textual data can influence the decision points, activities, and their order. Thus, BP complexity is affected by textual data. In the related literature, it is shown that there is a connection between BP complexity and BP performance and management [7,8]. For this reason, process redesign and improvement initiatives are often motivated by BP complexity analysis [9,10].

The potential of textual data serving as an input to BP execution in the context of complexity has been extensively studied in our previous work [11–15]. Within that work, we have explored and developed a set of linguistic features, including semantic, syntactic, and stylistic ones, i.e., taxonomy-based [11], sentiment-based [15], and stylistic features [12], which potentially influence BP execution complexity [13]. In an industrial case study of a CHM IT ticket handling process [14], we have investigated the linguistic features favorable for BP execution complexity prediction.

Overall, in this study, we propose a novel BP execution complexity analysis approach in which we combine textual data and event log. For the development of the approach, we set the following *specific objectives*:

- Enriching an understanding of event log-based (EL) complexity common in BPM with textual data-based (TD) complexity,
- Identifying a set of metrics for TD and EL complexities taking existing works as a basis,
- Studying the relation between TD and EL complexities and investigating how textual data can contribute to EL complexity prediction,
- Exploring, adapting, and illustrating the benefits of our approach by applying it in a real-world setting.

To achieve these objectives and ensure the comprehensiveness of our approach, we build our study on the following steps. In Section 2, we analyze the related work on the application of NLP in BPM and BP execution complexity highlighting the unsolved issues regarding the use of textual data. Section 3 summarizes the aspects from our previous work that we use for TD complexity calculation and explains the state-of-the-art event log complexity metrics serving as the basis for EL complexity calculation. Afterward, in Section 4, using a running example, we adapt and incorporate our previous work on TD complexity and well-established studies on EL complexity and introduce the BP execution complexity analysis approach. As can be seen in Fig. 1, in the first block, TD and EL complexities are calculated. For TD complexity calculation, linguistic features extracted from textual data are used. In this regard, we take the designed linguistic features (taxonomy-based and stylistic features) from our previous work [11,12] and identify relevant features for TD complexity. Using these linguistic features, the TD complexity of BPs is predicted with respect to an agreed-upon complexity scale. Further, we assess the adequacy of our linguistic features in predicting TD complexity. This is done by comparing their prediction performance with the prediction performance of the linguistic features from a well-accepted text analysis technique. To calculate EL complexity, the state-of-the-art event log complexity metrics are analyzed, and suitable ones are applied to the given event log. In the second block, two analyses are performed, namely correlation analysis and significant difference analysis. More specifically, how calculated complexities correlate is determined in the correlation analysis. Following that, significant differences within the correlations are analyzed. Using the obtained results, recommendations and insights for process improvement are derived. To illustrate the value of our approach, a Service Request Management process case study from an IT Service Management (ITSM) of an academic institution is conducted and explained in Section 5. We discuss the implications of our findings in Section 6. Finally, in Section 7, we present our conclusions and directions for future work.

Thus, our work contributes to BPM by proposing a new approach to analyze BP execution complexity, considering textual data serving as an input to BP execution and event log. Although one of the dominant research directions in BPM regarding BP analysis is BP complexity [4], to the best of our knowledge, no other works combine textual data and event log to analyze BP execution complexity. Using qualitative (interviews) and quantitative (computational analysis) research methods, we demonstrate the value of our approach by means of a case study. As a practical contribution, our study findings show a comprehensive way of identifying process redesign and improvement opportunities.

2. Related work

This section lists the studies associated with the approach we propose in this paper. In particular, we highlight the increasing relevance of textual data in organizations and review the state-of-the-art NLP applications in various BPM lifecycle phases. Afterward, we present prominent complexity research in the BPM-related literature.

Organizations increasingly focus on insights into understanding and improving their BPs. In this regard, BPM investigates the potentials of NLP to benefit from its maturity and availability in multiple BP applications providing support to different BPM lifecycle phases [5]. In the following, relevant research is reviewed according to BPM lifecycle phases.

In the **BP discovery phase**, considerable research effort has been made to develop BP model discovery approaches from data. Whereas BP discovery from event log is a well-established and matured subject area that already has tangible practical applications, BP model discovery from textual data is still a promising research topic lacking the ability to scale [16]. Below, we review the most prominent and recent developments:

BP model and event log generation from textual data: Creation of BP models makes up to 60% of the time spent in BPM projects [5]. Further, due to the current dynamics of work environments, BP modeling has become a time-consuming and costly activity requiring constant updates of BP models, which might lead to BPM project failures [17,18]. Thus, an automatic generation of BP models from available textual data becomes an attractive application paving the way for multiple research projects. For example, recent research by [19] proposes an automatic BP model discovery from textual BP descriptions based on neural networks. Further, [20] extend existing NLP techniques to extract activities and their relations defining BP constraints from textual descriptions. [21] present a method to generate an event log from textual data using action and topic analysis. Thereafter, BP models are mined based on common techniques. [22] use natural language inference to construct event log from customer service conversations. [23] deal with the problem of multi-grained text classification by introducing a hierarchical neural network to extract multi-grained information from BP descriptions. In [24], the early developments of a tool to extract BP models from text and then maintain their alignment using Dynamic Condition Response (DCR) Graphs are presented.

Enrichment of event log with textual data: Process Mining (PM) represents the most typical approach to automatically create BP models from event log [25]. Hereby, textual data massively generated by BP participants in the BP execution, such as comments or email communication, are not considered. To address this shortcoming, several research projects started to appear. For example, [26] extract key phrases denoting activities from comments related to IT ticket processing enriching the event log with this information. Subsequently, a more comprehensive BP model can be derived. Further, [27] enhance the event log analysis with the analysis of textual attributes contained in it using a novel attribute classification technique.

Automatic discovery of decisions in BP models: Decisions make up an important effort-intensive part of BP modeling. Accordingly, [28] propose a deep learning approach to obtain decision constraints and conditional clauses from text. [29] provide an NLP pipeline to automatically extract the decisions and their dependencies to build the decision requirements diagram making part of a decision model. The study by [30] describes a method for generating entire decision models from textual inputs. The suggested technique based on NLP and customized syntactic patterns enables the extraction of both decision requirements and decision logic from a document.

Text annotation: The efforts related to BP model creation can be sufficiently reduced in case the text is well annotated, this way decreasing the noise caused by automatic NLP techniques. Such annotated BP descriptions can be used for both inferring new relations to create more comprehensive BP descriptions and as training data for various NLP analyzers [5]. Hence, in [31], a novel approach using NLP and a query language for tree-based patterns is introduced. It derives annotations representing essential BP elements, i.e., activities, events, actors, roles, and control flow. [32] describe a method based on Semantic Parsing and Graph Convolutional Networks. This method avoids the use of manual rules and outputs much better results than existing neural network-based solutions to derive annotations from BP descriptions.

Automatic BP modeling recommendations and semantic autocompletion: Considerable research has been devoted to automatic activity recommendations to support BP modeling task [33,34]. Hence, grounding on a similar technique as [35,36] exploit label semantics for rule-based activity recommendation. Additionally, [37] propose to use semantic similarities between BPs to enable design-time autocompletion by relying on pre-trained NLP models. The method converts BP sequences into text paragraphs and encodes them as sentence embeddings, i.e., learned text representations that include semantics as real-number vectors [37].

The next phase of the BPM lifecycle, i.e., **BP analysis**, aims to identify flaws and bottlenecks in the discovered BP models. Hereby, NLP techniques can also be of support in specific applications. We present some up-to-date developments below:

BP model semantic correctness and completeness verification: The semantic quality of BP models is critical for understanding BPs correctly. A number of NLP research projects are naturally aimed at automating the verification of this characteristic. Accordingly, many BP model analysis strategies rely on a thorough examination of the natural language information included in the activity labels of the models. Standard NLP is not adequate for analyzing these labels since they are often short and heterogeneous in terms of grammatical style. Dealing with this challenge, [38] propose a Hidden Markov Models-based approach for a linguistic analysis of BP model activity labels. Additionally, research by [39] addresses the problem of ambiguity of BP textual descriptions and suggests a compliance checking technique using behavioral spaces.

BP model and text consistency check: As mentioned above, maintaining various BP-related data allows for improving the knowledge of BPs in organizations. However, as BPs change over time, it is important to constantly identify inconsistencies among various BP descriptions so that expectations for BP outputs remain the same for every stakeholder [40]. The latter research [40] proposes an approach to detect conflicts between textual and model-based descriptions using NLP. Further, [41] design a technique to align BP models and textual descriptions, mapping the knowledge derived from these two representations into a unified, comparable format.

BP-related data querying: Having a variety of BP-related data allows organizations to better analyze their BPs. In such an analysis, a common task is querying these data to get insights into specific BP parts. In the case of event log data, to be able to use common BP querying techniques, end-users must be familiar with the query language and database schema. Addressing this challenge, [42] introduce a natural language interface. Hereby, questions can be asked in a normal language, and the interface will automatically translate them into a structured query to be run in a database. [43] also address this problem by suggesting a technique to search both textual and model-based BP descriptions.

Sentiment analysis: Sentiment analysis has already shown its high value for e-business and e-commerce providing insights based on the textual data collected in social media and social networks [44]. [45] explore its potential in BPM and develop a BP modeling tool considering stakeholders' comments and feedback. Applying sentiment analysis, the tool identifies positive and negative feedback to support BP analysts in designing the to-be process.

In the following **BP redesign phase**, the BP model needs to be modified to address the concerns discovered in the previous phase. Contemporary NLP techniques can also be used to achieve this goal:

BP redesign using textual data: As a rule, experts dealing with BP redesign focus on proposing to-be BP models and redesign patterns with little or no consideration of end-user feedback. To address this shortcoming, recent research suggests an NLP approach based on a novel set of annotation guidelines to identify redesign suggestions directly from end-user feedback [46].

Comparing BP models: To produce a sound to-be BP model in the BP redesign phase, a BP analyst might need to examine large sets of BP models that are often organized hierarchically based on the level of abstraction. Hereby, one of the most difficult tasks is ensuring that BP models at the same level in the hierarchy have the same level of abstraction [5]. To solve this challenge, various BP model matching algorithms can be used. For example, [47] present a semantic multi-phase matching algorithm based on a vector space model and NLP to match the models. [48] provide a technique for discovering sets of related activities based on constrained k-means clustering considering both BP semantics and control flow order.

BP model refactoring: The quality of BP models may significantly vary since BP modeling is time-consuming and error-prone. Moreover, the competence of various modelers differs. Hereby, refactoring can be used to improve the quality of BP models. Refactoring is a popular approach in software engineering to restructure the code without changing its external behavior. As BP modeling and coding are similar to a certain extent, existing refactoring technologies from software engineering have been adapted for BP workflows [49]. In the context of NLP, such an approach as linguistic refactoring has appeared. Accordingly, [50] elaborates NLP techniques for syntactic, semantic, and pragmatic refactoring in the dissertation.

In addition to typical monitoring techniques for assessing performance and conformity requirements, in the **BP controlling phase**, NLP can be used to make available different forms of BP descriptions [5]:

Transformation of BP model to text: In the controlling phase, it is important that all stakeholders are able to understand BP-related data. However, event log, workflows, and BP models are not straightforward and require certain expertise for comprehension. On the contrary, a written BP textual description can be understood by any stakeholder. Hence, it is highly recommended to support event log and BP models with the latter [5]. To solve this problem, BP model-based natural language generation has been researched [51]. Further, [52] introduce a semi-automated approach to transfer knowledge from BP models to natural language requirements documents. [53,54] develop a tool to generate BP textual descriptions from declarative BP models. In this context, another group of researchers [55] deals with the comparison of manually and automatically generated textual descriptions of BP models focusing on the choice of an appropriate matching technique. Additionally, [56] suggest a technique to fix poorly written BP textual descriptions based on BP models.

Multi-lingual support: In international companies as well as in the context of cross-country and cross-organizational learning, it is essential to translate BP models and BP descriptions into multiple languages to enable accessing BP information to various stakeholders. Thus, [57] develop a framework for the automatic generation of multi-language description text using an emergency disposal process example. In line with the latter, [58] enhance the framework with multiple (cross-department) views and operationalize it in a cross-department medical diagnosis process.

As can be concluded from above, several research streams, such as sentiment analysis, BP-related data querying, BP modeling recommendation and autocompletion, might be applied in multiple BPM lifecycle phases, for example, BP discovery, analysis, and redesign. However, the most prominent research stream is aimed at supporting the discovery phase, i.e., the automatic creation of BP models from BP textual descriptions. Accordingly, the most frequently used textual data are related to the BP descriptions, feedback from BP participants, and textual data inherent in BP models, i.e., labels. Hereby, only a limited number of works deal with the comments and emails related to the activities in the event log [26].

In addition, a large number of studies in BPM analyze BP executions from a complexity perspective. For example, [59] describe metrics for measuring BP model complexity based on observations from software complexity. Similarly, in [60], metrics for analyzing BP model complexity are proposed by extending metrics on software complexity. BP model complexity metrics and their theoretical thresholds are studied in [61] to assess BP model complexity and categorize BP models based on their complexity. An overview of the BP model complexity reduction mechanisms is provided in the form of patterns in [62]. Aside from that, in BPM, there is a great interest in analyzing BPs from a complexity perspective using PM. Hence, [63] study the design and applicability of metrics for measuring event log complexity. [64] provide a comprehensive evaluation of state-of-the-art BP discovery techniques considering the complexity of their automatically generated BP models. An approach aimed at the reduction of the complexity of the discovered declarative BP models is proposed in [65]. Moreover, [66] provide an overview of the BP model complexity metrics by conducting a systematic literature review. Lastly, in a recent study [4], the state-of-the-art event log-based complexity metrics are analyzed to determine the relation between the event log and the resulting BP model.

To sum up, despite the ubiquity of textual data in organizations, in the relevant BPM literature, the analysis of textual data related to BP executions is prevailingly limited to BP descriptions and texts in BP models [5]. Moreover, the complexity of textual data has a considerable influence on BP execution, which has been recently studied and proven in our work [13]. To address the shortcoming, in this paper, we propose an approach combining textual data used as an input to BP execution and event log for BP execution complexity analysis.

3. Background

In this section, we provide a background on the linguistic features adapted to calculate TD complexity and event log complexity metrics employed to calculate EL complexity.

3.1. Linguistic features

In our previous work, we studied what characteristics, i.e., features, of a given text have a great potential to influence BP execution complexity [11–13,15]. In particular, we investigated several features and identified that taxonomy-based, so-called Decision-Making Logic (DML) taxonomy, and stylistic features are prevailingly important for the complexity of textual data [14]. In this regard, we summarize how these features were developed.

To design linguistic features that capture TD complexity via cognition and style of textual data, we focus on the distribution of

Table 1
Linguistic features.

Group	Linguistic feature
DML taxonomy-based	Relative occurrence of words based on the DML taxonomy and the DML cognition level *routine*
	Relative occurrence of words based on the DML taxonomy and the DML cognition level *semi-cognitive*
	Relative occurrence of words based on the DML taxonomy and the DML cognition level *cognitive*
Stylistic	Relative occurrence of nouns in all words
	Relative occurrence of unique nouns in all nouns
	Relative occurrence of verbs in all words
	Relative occurrence of unique verbs in all verbs
	Relative occurrence of adjectives in all words
	Relative occurrence of unique adjectives in all adjectives
	Relative occurrence of adverbs in all words
	Relative occurrence of unique adverbs in all adverbs
	Word count
	Wording style

parts-of-speech (PoS) in textual data. In particular, nouns, verbs, adjectives, and adverbs are analyzed since they reveal the most information about decision-making and style in textual data. As a rule, process workers interpret textual data inputs to determine how a BP should be carried out. They map the phrases to the BP elements. For example, a process worker may decide on a BP execution based on the information the customer mentioned in a textual message about previously performed BP activities (nouns) using specific verbs, adjectives, or adverbs indicating the timeline and status of such activities. Hence, extracting this information could assist process workers in handling textual data related to BP execution. Aligned with this, there are naming conventions in BPM [62,67] on the effects of labeling in BPs. More specifically, these conventions provide guidance on those PoS that can be used in labeling and how. As such, we suggest considering nouns as ***R****esources* expressing the specifics of BP elements, verbs as ***T****echniques* of knowledge and information transformation activity impacting ***R****esources*, adjectives as ***C****apacities* revealing contextual specifics of ***T****echniques*, and adverbs as ***C****hoices* defining the selection of the necessary set of ***T****echniques*, elements of RTCC framework developed in our previous research [11].

DML taxonomy: DML taxonomy is a 2-tuple: (1) most important words (nouns, verbs, adjectives, and adverbs) extracted from a given text and (2) decision-making logic levels, i.e., cognition levels, each of which denotes the easiness of the process to understand something for making a decision. In DML taxonomy, each word is associated with a DML cognition level. For detailed information regarding the DML taxonomy development process, we refer to our previous work [11]. In this paper, we present a summary of the most important steps:

Step-1 The first step is collecting BP-relevant textual data from different sources. Whereas in our approach, the focus lies on the textual data provided as input to BP execution, for DML taxonomy, also other textual data, such as official BP descriptions, interview transcriptions, or legal documents, should be considered.

Step-2 The collected data are converted into a machine-readable format, in which the computational analysis will be performed, for example, a CSV file format.

Step-3 Afterward, the data are pre-processed and parsed, building the document-term matrices for the most important parts of speech (nouns, verbs, adjectives, and adverbs).

Step-4 The created document-term matrices are processed using the topic modeling methods, such as a combination of Latent Dirichlet Allocation (LDA) and Latent Semantic Indexing (LSI) [68].

Step-5 In the last step, the extracted topics with descriptive keywords are classified into the decision-making logic levels, i.e., cognition levels, of routine, semi-cognitive, and cognitive. Here, the involvement of experts being familiar with the context is essential for the right keyword classification.

Stylistic patterns: In our previous work [12], we showcased that the style, i.e., stylistic patterns, of a given IT ticket text can reveal information on identifying its BP complexity. More specifically, ticket length, PoS distributions, and wording style are suitable for indicating and understanding how the complexity of handling an IT ticket is affected. To capture such components of a text, we proposed Syntactic Structure (SynS) and Wording Style (WS) as new features. The SynS feature focuses on syntax. The way the words are put together to form phrases influences text comprehension and corresponding BP execution, which uses that text as a primary input. The WS feature takes Zipf's Laws [69] as a basis and focuses on the appearance of new words in a text and the speed of appearance.

In accordance with the considerations explained above, for DML taxonomy-based features, PoS distributions are computed based on a given DML taxonomy. Hereby, all the words are considered as the search space for stylistic features. In Table 1, we list the identified features.

3.2. Event log complexity metrics

Organizations use various information systems (for example, ERP systems or IT ticketing systems) to enact their BPs with the support of such systems. These systems enable organizations to record a large amount of data about BP executions. Such process execution data are then extracted in the form of an event log to analyze and provide insights into improving BPs [70]. An event log consists of events, each of which refers to an activity performed in executing a BP. In Table 2, an exemplary event log of a Service Request Management process is depicted.

As can be seen, each row shows (i) which activity is performed, (ii) when, (iii) for which request, and (iv) other information (for example, resource and priority). The events carried out in the

Table 2
An example of an event log of a Service Request Management process.

Request	Activity	Time stamp	Resource	Priority	Attr.
t001	Register	01-08-2021 10:11:12	Worker1	Low	...
t001	Analyze	01-08-2021 14:10:00	Worker2	Low	...
t002	Register	01-08-2021 16:01:03	Worker3	High	...
t003	Register	01-08-2021 16:05:42	Worker1	Low	...
t002	Analyze	01-08-2021 16:10:42	Worker4	High	...
t002	Resolve	01-08-2021 16:25:51	Worker5	High	...
t003	Analyze	01-08-2021 17:15:02	Worker6	Low	...
t001	Escalate	01-08-2021 17:35:40	Worker7	Low	...
t003	Resolve	02-08-2021 09:01:02	Worker8	Low	...
t004	Register	02-08-2021 09:10:20	Worker3	Medium	...
t004	Analyze	02-08-2021 09:25:33	Worker2	Medium	...
t003	Re-open	02-08-2021 10:01:01	Worker8	Low	...
t004	Reject	02-08-2021 10:06:08	Worker6	Medium	...
t001	De-escalate	02-08-2021 10:24:32	Worker8	Low	...
t003	Resolve	02-08-2021 11:16:10	Worker4	Low	...
t001	Resolve	02-08-2021 11:59:59	Worker4	Low	...

scope of a single process instance execution are called a case. In the example, each request refers to a case that goes through the same Service Request Management process. The sequence of the events in the scope of a particular case is called a trace.

In the literature, there are several studies focusing on quantifying the complexity of such an event log. Within these studies, metrics are proposed in order to assess the complexity of event logs, so that further analysis can be determined considering the characteristics of event logs. In recent research on EL complexity measurement [4], EL complexity metrics are reviewed and studied in detail. Then, a set of entropy-based complexity metrics are proposed to address the issues in the studied EL complexity metrics. We take this research as the basis and analyze both discussed and proposed metrics in it. These metrics are listed in Table 3.

Based on the specifications of the metrics given in the table, one can note that either one or more aspects (size, variation, and distance) of complexity are selected as the focus in each metric. In that sense, some metrics have limitations. For example, metrics measuring the size of event logs would not capture any difference in terms of variation or distance. Despite the fact that some metrics focus on the same aspects of complexity, there is not much dependency among them as they differ from each other in measuring an event log using its various components (for example, traces, event classes, or event relations) [4]. To have a comprehensive view of the aspects of complexity, in our approach, we opt for employing all EL complexity metrics listed in the table. Further, to mitigate the influence of one metric on another, we use majority voting in our approach to obtain a single EL complexity value for a given event log.

In general, processes indicate all the work performed in an organization [1]. Accordingly, the complexity of a process and corresponding ways to measure it can imply a wide range of elements and factors, like those emerging from the process context [73,74], which are often difficult to obtain. Moreover, how a process is reflected in a model affects its perceived complexity. In other words, quality aspects of process models and process modeling notations have a notable impact on perceived process complexity. Hence, we focus on event log-based complexity metrics in this paper and list extending our approach with process complexity metrics as part of our future work.

4. Approach development

As introduced in Section 1, we propose an approach, hereafter *Approach*, aimed at analyzing BP execution complexity based on textual data and event log. More specifically, we investigate the relation between TD complexity and EL complexity. To achieve this, first, for a given BP, attributes of an entity that goes through the BP are identified. For example, a communication channel attribute of a service request, which is an entity handled in a Service Request Management (SRM) process. Then, based on these attributes and time dimension, textual data and event log of the BP are split into subsets. Further, TD and EL complexities are calculated for each subset. Using the computed complexities, correlation analysis is performed to investigate whether textual data may be used for EL complexity prediction. Thereafter, statistically significant differences in the created event log subsets are analyzed to find out the factors affecting EL complexity. For instance, a certain category of service requests may account for the repeated or skipped information collection activities resulting in a considerable increase or decrease of EL complexity. Thus, in terms of such factors, recommendations for process redesign and improvement can be formulated and provided to organizations.

In the separate subsections of this section, we present the inputs required in our *Approach* and introduce a running example. Afterward, we elaborate on how TD and EL complexities are calculated. Finally, correlation analysis and identification of statistically significant differences are explained.

4.1. Inputs

To carry out the tasks mentioned above, three types of input are required in the *Approach*: *textual data*, *event log*, and *complexity scale*. The first two inputs will be described in Sections 4.3 and 4.4. The complexity scale necessary for calculating TD and EL complexities is defined below.

Complexity scale: A complexity scale is a set of ordinal complexity values. They can be numbers or categories that are put in a certain order denoting either increasing or decreasing complexity. A five-point Likert-type scale containing numbers from one to five or a set of category names like low, medium, and high are two examples of a complexity scale [75]. Although a considerable number of metrics for measuring complexity exists (see Table 3), textual complexity metrics are rather generic, i.e., mostly considering language usage in texts [13,76], and have less emphasis on how textual data are perceived by process workers in terms of work instructions. This perception is important because, using a given text, process workers determine which activities will be performed and in what order. Moreover, such generic metrics are not applicable in a given area without considering the jargon, characteristics, and regulations of the area. Therefore, when applying our *Approach*, expert involvement is essential to determine a suitable complexity scale for a particular domain.

Creating textual data and event log subsets: To conduct a well-established analysis of the relation between TD and EL complexities, in the *Approach*, the textual data and event log are split

Table 3
Metrics for event log-based complexity calculation.

Metric	Definition
Number of events (magnitude) [10]	The total number of events an event log contains
Number of event types (variety) [10]	The total number of event classes in an event log
Number of sequences (support) [10]	The total number of traces in an event log
Average sequence length (TL-avg) [25]	The average length of a trace in an event log
Average time difference between consecutive events (time granularity) [10]	The mean duration between two events where the first one is followed by the second one without an interruption
Number of acyclic paths in transition matrix (LOD) [8]	The total number of simple paths (paths without cycles) in the graph network that represents the event connections in an event log
Number of ties in transition matrix (t-comp) [71]	The total number of possible paths in the graph network that represents the event connections in an event log
Lempel–Ziv complexity (LZ) [72]	The minimum number of steps that are required to generate a given trace by either reusing its previous parts or inserting a new symbol
Number and percentage of unique sequences DT(#), DT(%) [25]	The number and percentage of distinct traces in a given event log
Average distinct events per sequence (structure) [10]	The amount of present directly-related event pairs compared to the all possible ones in a given event log
Average affinity (affinity) [10]	The homogeneity of a given event log based on the average overlap of traces in terms of direct following relations (i.e., one event right afterward another)
Deviation from random (dev-random) [72]	The Euclidian distance of the transition matrix that is created using the pairwise associations of events of a given event log
Average edit distance (avg-dist) [72]	The average edit distance of traces to transform one to another using string matching with the lowest cost
Entropy-based metrics (variance and sequence entropy) [4]	The entropy-based metrics that use prefix automation to describe sequences within a given event log to map it to a graph

Table 4
Running example IT tickets.

ID	Channel	Category	Textual data
SR001	IT ticketing system	Application	I would like to get access to *XYZ*. Could you please send me the available document how to install it?
SR002	IT ticketing system	Security	As of this week, I am working in a different building. When I try to login, it says unable to find the trust certificate *CRT-ABC* in the recovery database for this workstation. Would you please activate the broken security configuration?

into subsets. Hereby, we take the following into account: time dimension and a set of attributes of an entity going through a BP, for example, a service request. A time dimension is important to analyze changes in BP execution complexity. Accordingly, relevant time periods can be defined for observing the BP execution complexity over time. Having attributes allows us to perform a drill-down analysis and move from a general to a more detailed view. Further, to enrich the analysis, we combine entity attributes pairwise.

4.2. Running example

To illustrate our *Approach* development, we introduce a running example of two IT tickets (service requests) from an SRM process case study used in this research.As can be seen in Table 4, the two tickets, SR001 and SR002, are entered directly in the IT ticketing system. When tickets are received, their textual contents are analyzed, and they are assigned to a category (i.e., grouping of tickets based on the concerned topics in them) by service desk employees. Using the category of a ticket, how it will be handled, i.e., next activities and involvement of resources, is determined. As shown in the table, the ticket SR001 consists of fewer and more common words compared to SR002. Important to note that the event log for the tickets in the running example is also provided.

In the subsections below, we show how the *Approach* is developed using the illustrative running example, in particular, its textual data and event log.

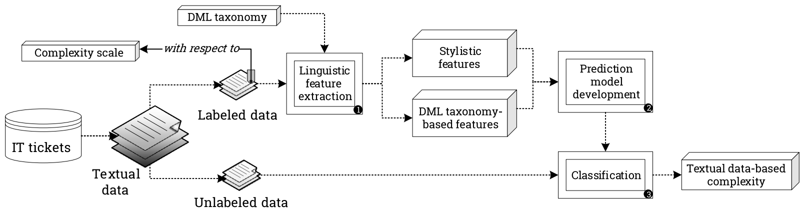


Fig. 2. Textual data-based complexity calculation.

4.3. Calculating textual data-based complexity

As an unstructured data type, one of the important inputs influencing BP execution is textual data. However, due to the dynamic nature and high interdependence of BPs, textual data are often either unlabeled or contain few labeled points. Moreover, labeling is a time-consuming and costly process. Hence, we use machine learning to address this problem. In particular, we start with the textual data that have very few labeled data points, extract linguistic features, and develop prediction models. Afterward, we select the outperforming prediction model. We run that model on the unlabeled data. This flow is depicted in Fig. 2.

To perform a computational analysis of textual data and build prediction models for TD complexity, we extract two sets of linguistic features from textual data: DML taxonomy-based and stylistic features. Specifically, we focus on the two fundamental aspects of textual data that indicate complexity, namely *cognition level* [77] and *style* [78,79]. The decisions made based on textual data depend on the complexity perceived after reading the text, i.e., the cognition of textual data affects decision-making. Further, words and their order in sentences are referred to as style that can contain certain stylistic patterns.

As stated in Section 3, in our previous work, we have extensively analyzed those linguistic features potentially influencing textual complexity [13]. In [14], we have performed a feature selection using a complete set of features for the IT ticket classification task. Whereas the importance of these features is likely dependent on the domain specificity, our analysis demonstrated that the DML taxonomy-based and stylistic features were favorable for complexity prediction in the IT ticket processing case study, which was belonging to the ITSM domain. Accordingly, in our *Approach*, we use the DML taxonomy-based and stylistic features. Using the features in Table 1, prediction modeling techniques are trained on the labeled textual data, i.e., the training set. Since the labeled data comprise very few data points, the problem that the *Approach* deals with is a semi-supervised learning problem. Hence, we use the following commonly applied semi-supervised learning techniques [80] to enrich unlabeled data using labeled data: Label Propagation, Label Spreading, and Self-Training. The first two are very similar: both consider the distance of data points to assign labels using the unlabeled data points by putting all data points in a graph. In Label Spreading, affinity matrix and normalized graph are used. Self-Training assigns labels to unlabeled data points by reinforcing a model as a pseudo-labeler.

While training, in each prediction modeling technique, a set of adequate hyper-parameters is chosen for creating the best prediction model. As soon as all prediction modeling techniques are trained, the test set is used to determine the best-performing technique based on the prediction quality. For prediction quality assessment, we use the F-score metric.

In the prediction model development, to accomplish a better prediction quality, we use three common meta-algorithms [81], namely bagging, boosting, and stacking. In stacking, a single modeling technique aims to learn the best combination of the prediction models of the primary prediction modeling techniques put in a stack. In boosting, to fix the errors in prior prediction models, the prediction modeling techniques are trained in a chain. Bagging involves selecting different sub-samples of a training data set. Predictions for the sub-samples are then aggregated to identify the final and the best prediction model. When the prediction model development is completed, the best-performing model is applied to the unlabeled data. Using the DML taxonomy-based and stylistic features, each data point is classified based on the complexity scale, which is the one used while preparing the labeled data.

In Table 5, linguistic feature value calculations for the running example IT tickets are listed. As can be seen, for each DML taxonomy-based and stylistic feature, a value per ticket is computed. For calculating the DML taxonomy-based feature values, the DML taxonomy of our case study shown in Table A.12 in the appendix is used. These values are fed into the best-performing prediction model to obtain a single TD complexity value for each ticket.

To derive insights into the BP complexity of these tickets handling and analyze how their TD and EL complexities are related, a single value of TD complexity per ticket is necessary. Nevertheless, one can trace back the single feature presence in the text if needed, for example, to understand which features influence the complexity. This becomes notably relevant in the case of inaccurate classifications and the XAI (explainable artificial intelligence) paradigm [82]. For instance, inaccurate classifications can be analyzed by identifying the contribution of each linguistic feature to the complexity prediction. Thus, one can use these contributions to build end-user recommendations to improve the text.

In addition to the overall analysis of TD and EL complexities, i.e., at the ticket level, it is to emphasize that the relation between TD and EL complexities can be further investigated in a higher granularity using ticket attributes. As shown in Table 5, each of the IT tickets in the running example has a different category. Such an attribute of tickets may be beneficial to identify how BP execution complexity varies among subsets of tickets. In this regard, it is essential to calculate an aggregated TD complexity value per subset. To do so, in the *Approach*, we use weight multipliers that are determined based on the same complexity scale. These multipliers are applied to the calculated TD complexities of IT tickets in a given subset, and a weighted average is computed per subset.

4.4. Calculating event log-based complexity

To calculate EL complexity, in addition to an event log, a set of EL complexity metrics and a complexity scale are taken as

Table 5
Textual data-based complexity calculation of the running example IT tickets.

Group	Linguistic feature	Calculated values for tickets			
		Feature value		TD complexity	
		SR001	SR002	SR001	SR002
DML taxonomy-based	Relative occurrence of words based on the DML taxonomy and the DML cognition level *routine*	0.75 (install, send, available)	0.3 (certificate, login, find, active)		
	Relative occurrence of words based on the DML taxonomy and the DML cognition level *semi-cognitive*	0.25 (document)	0.3 (security, configuration, activate, unable)		
	Relative occurrence of words based on the DML taxonomy and the DML cognition level *cognitive*	0	0.4 (database, recovery, workstation, different, broken)		
Stylistic	Relative occurrence of nouns in all words	0.35	0.4	Low	Medium
	Relative occurrence of unique nouns in all nouns	1	0.88		
	Relative occurrence of verbs in all words	0.3	0.2		
	Relative occurrence of unique verbs in all verbs	1	1		
	Relative occurrence of adjectives in all words	0.05	0.05		
	Relative occurrence of unique adjectives in all adjectives	1	1		
	Relative occurrence of adverbs in all words	0.05	0.05		
	Relative occurrence of unique adverbs in all adverbs	1	1		
	Word count	20	40		
	Wording style	0 (min. word repeats)	0 (min. word repeats)		

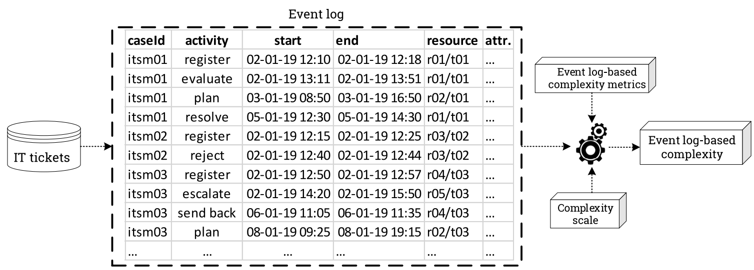


Fig. 3. Event log-based complexity calculation.

inputs. As the complexity scale, we use the same scale as in the TD complexity. Thus, one can perform a correlation analysis between the TD and EL complexities computed in respect to the same complexity scale. As for the EL complexity metrics, in our *Approach*, we use the ones[1] described in Section 3 (see Table 3). Similarly to TD complexity calculation, in this task, we create subsets of a given event log considering the attributes present in the textual data.

Fig. 3 shows how the calculation of the EL complexity is performed in our *Approach*. For each event log subset, a single complexity value is computed using the employed EL complexity metrics. Since these metrics focus on different event log properties (for example, size, variance in executions, distances between sequences and events) and use different measurement units, computed measurements will vary for a single event log subset. For example, in an event log subset, the number of events is a counted value, and measuring a time interval is about calculating an average time difference between consecutive events.

To have a single EL complexity value for each event log subset, the computed complexity values are mapped to the points of the complexity scale using clustering. More specifically, for each metric, calculated values are clustered. Then, for each cluster, a value is determined from the complexity scale, which is the same scale used for the textual data labeling. This flow is depicted in Fig. 4.

In Table 6, each EL-based complexity metric and the resulting EL complexity of our running example tickets are shown. As can be seen, the ticket SR002 was put into the medium cluster for the metric Percentage of Unique Sequences (DT%), whereas the

[1] The Python script provided on this Github page is adopted to calculate those metrics.

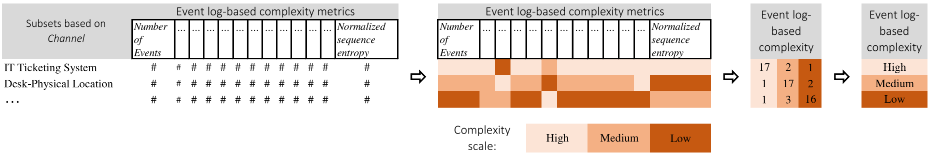


Fig. 4. Determining event log-based complexity of an event log subset.

Table 6
Event log-based complexity calculation of the running example IT tickets.

EL-based complexity metric	Calculated values per ticket			
	Metric value		EL complexity	
	SR001	SR002	SR001	SR002
Number of events (magnitude)	Low	High		
Number of event types (variety)	Low	High		
Number of sequences (support)	Low	High		
Average sequence length (TL-avg)	Low	High		
Average time difference between consecutive events (time granularity)	Low	Medium		
Number of acyclic paths in transition matrix (LOD)	Low	High		
Number of ties in transition matrix (t-comp)	Low	High		
Lempel–Ziv complexity (LZ)	Low	High		
Number of unique sequences DT(#)	Low	High	Low	High
Percentage of unique sequences DT(%)	Low	Medium		
Average distinct events per sequence (structure)	Low	High		
Average affinity (affinity)	Low	High		
Deviation from random (dev-random)	Low	High		
Average edit distance (avg-dist)	Low	High		
Variance entropy	Low	High		
Normalized variance entropy	Low	High		
Sequence entropy	Low	High		
Normalized sequence entropy	Low	High		

ticket SR001 was put into the low cluster. As the result of majority voting, the EL complexity of SR002 is determined as high and, for SR001, it is low. As the TD and EL complexities are now known, one can identify how the TD complexities of these tickets are associated with their EL complexities. Moreover, depending on the direction and strength of the association between the TD and EL complexities, further necessary analyses and inputs can be determined.

4.5. Correlation analysis

As introduced in Section 1, one of the main goals of our research is to determine how textual data can contribute to EL complexity prediction. Thus, in our *Approach*, we conduct correlation analysis [83] to find out how the TD complexity is related to the EL complexity. In the correlation analysis, the strength of association between these two complexities and the direction of their relation are measured. More specifically, we investigate strong-positive, strong-negative, and no correlations. In the case when TD and EL complexities are close to a normal distribution, the Pearson correlation is used. To assess the strength in the Pearson correlation, we follow the general guidelines [83,84] and use 0.1, 0.3, and 0.5 as coefficient thresholds. In the case of non-normality in TD and EL complexities, we choose the Spearman's correlation [83,84]. Using the same general guidelines, 0.2 and 0.8 are taken as the thresholds to detect the strength of correlations. In addition, $p \leq 0.05$ (p denotes probability) is used as the indicator both in the Pearson and Spearman correlations for a significance identification.

When TD and EL complexities are in a strong-positive correlation, we can use textual data to predict the execution complexity of BPs by means of TD complexity. Moreover, organizations can make prior decisions and take actions to mitigate complexity in performing BPs. A negative correlation between TD and EL complexities is a good indicator to identify which data type should be further analyzed. No correlation between these two complexities cannot be directly interpreted. Therefore, more textual data and event log attributes and other BP execution data, such as performance indicator values, should be considered and further analyzed to detect reasons for complexity.

In Table 7, the TD and EL complexities for the running example IT tickets are shown. Solely considering these two tickets, one can notice that the TD complexity has a positive correlation with the EL complexity. Based on such observation, the following interpretation can be deduced: TD complexity affects EL complexity. Hence, textual data of these (and similar) tickets can be used to predict their EL complexity.

As a correlation does not necessarily imply a cause and effect between two complexities, we conduct a significant difference analysis on event logs to identify the factors that may account

Table 7
Textual data- and event log-based complexities of the running example IT tickets.

ID	Channel	Category	TD complexity	EL complexity
SR001	IT ticketing system	Application	Low	Low
SR002	IT ticketing system	Security	Medium	High

for variation in the EL complexity. In the following subsection, we elaborate on that.

4.6. Significant difference analysis of event logs

To understand what process parts or activities affect the EL complexity, statistically significant differences in the event logs need to be analyzed. Aligned with that goal, there are approaches in the related literature. We reviewed the applicability of these approaches for analyzing the event logs within our *Approach*. The approach by Bolt et al. [85] came forward since it provides an extensible basis for the EL complexity metrics. To identify what differences are significant, well-known statistical tests are used within that approach. In particular, the two-tailed Welch's T-test [86] is used in the case of a normal distribution. This basis is beneficial for our *Approach* due to the following reasons: (1) for Welch's T-test, it is not necessary to have equal variance between two groups, and (2) Welch's T-test is less restrictive than the original Student's T-test, which makes it more reliable. For non-normality cases, the non-parametric rank-based Mann–Whitney U-test [87] is applied. To handle outliers and unexpected observations, this test focuses on the median, which is a better measure of the central tendency for skewed data. In this respect, such a non-parametric method is useful for our *Approach*.

Furthermore, this approach is available as a plugin (called Process Comparator) in the Process Mining framework, ProM [88], which offers built-in features for handling event logs. As such, we execute this plugin for each complexity cluster pair and analyze statistically significant differences in the event logs for a particular cluster pair. Complexity cluster pairs are determined using the correlation analysis results. For example, if a strong-positive correlation is observed with respect to an attribute present in textual data, subsets created for that attribute will be used to form cluster pairs.

Taking our running example IT tickets, a cluster pair is created using the category attribute of the tickets. Fig. 5 illustrates how the significant differences between the event logs of this cluster pair are highlighted. In the figure, nodes represent activities, whereas arcs reflect the sequence of these activities in the process. The thickness of the arcs and nodes is determined based on the increasing or decreasing value of the selected process metric, for example, frequency or duration. Distinct colors, i.e., red and blue, are used to indicate significant differences in terms of the selected process metric. In addition, the letters B and R are placed where appropriate to indicate blue and red colors, respectively. As can be seen, the activity "Hand-over" (T10) is significant in the cluster that contains the ticket SR002. Likewise, "Work assignment" (T14) is significant in the cluster that SR001 is put. Considering these insights, one can further analyze the presence or absence of those textual data features which resulted in handovers or work reassignments.

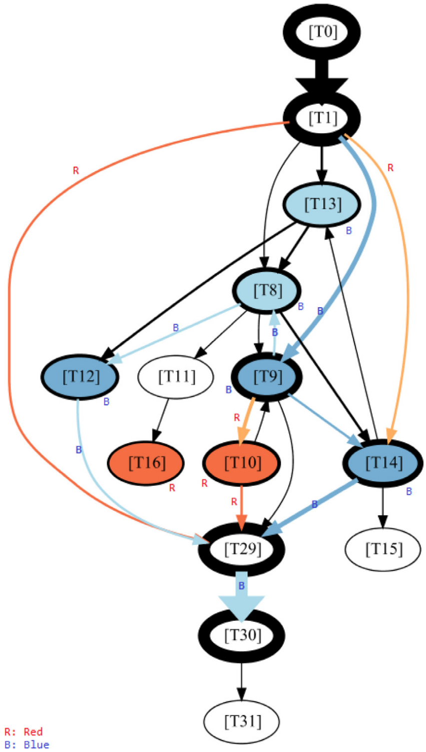


Fig. 5. Statistically significant differences between SR001 (with B) and SR002 (with R). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Case study

In this section, first, we give information about the setting of the case study, in which we apply the proposed *Approach*. Then, we elaborate on the application of our *Approach* and obtained results.

Case study organization: The IT department of an educational institution in the Netherlands initiated a project to learn from the data recorded about its ITSM processes. One of the important processes among them is the SRM process. This process defines the way of handling user requests related to the products and services offered by the educational institution. Upgrading a software product installed on a user device or providing access to a file are two examples of typical service requests. Since the institution offers a wide range of products and services to more than 25K users, multiple resolution teams are involved in the SRM process. Each resolution team handles requests for a particular set of products and services. For example, printing and secure file sharing are two services managed by separate resolution teams.

To accomplish a data-driven service delivery, the IT department, hereafter *Org-IT*, expressed its interest in finding ways to reduce the complexity of incoming requests. As this setting is highly related to the *Approach* introduced in this paper, we apply it in *Org-IT* and explain the findings showing its usefulness.

Table 8
Identified ticket attributes.

Attribute	Description	Values
Channel	The entry source of the ticket	IT ticketing system, email, phone, online chat, desk-physical location
Category	The assigned category on the ticket	A number of categorical values, e.g., printing, IT infrastructure
Organizational unit	The organizational entity that the ticket reporting end-user belongs to	A number of categorical values, e.g., human resources, finance department, or science faculty
Duration	Time spent on the ticket	Less than 1 day, 1–3 days, 3–5 days, 5–10 days, 10–15 days, and more than 15 days
Hold-on	Whether the ticket is held-on, i.e., the ticket processing is paused, and the end-user is asked for more information	Never, once, more than once
Re-open	Whether the ticket is reopened	Never, once, more than once

5.1. Data collection

As explained in the previous section, three types of input (textual data, event log, and complexity scale) are required in the *Approach*. To obtain these, we worked together with five experts who coordinate the resolution teams in *Org-IT*. Importantly, these experts have substantial knowledge of incoming service requests. With the support of the experts, first, we defined a complexity scale. Second, we asked experts to label a set of randomly selected service requests considering their textual data. Lastly, together with the experts, we extracted the SRM process execution data in the form of an event log.

Complexity scale: In a semi-structured discussion meeting, we asked the aforementioned five experts to define a complexity scale based on the service requests handled by the resolution teams they are coordinating. The experts agreed on a three-point scale containing *low*, *medium*, and *high* as complexity values.

Textual data: *Org-IT* provided us with textual data about the service requests handled between Jan 2019 and May 2021. The provided textual data contains 4982 service requests (also called *tickets*). From them, randomly selected 134 (∼2.7% of the total requests) are labeled based on the defined three-point complexity scale with the support of the same five experts. In the process of labeling, the experts make their decision based on the textual data as well as other data recorded about tickets, like priority, category, or attached documents.

Event log: For the aforementioned 4982 service requests, we received an event log consisting of ∼37K events.

Furthermore, to conduct a detailed analysis of the relation between TD and EL complexities, subsets from textual data and event log need to be created. For this purpose, a set of service request attributes is identified together with the aforementioned experts. These attributes are listed in Table 8. In addition, to analyze complexities over time and detect changes, the following time periods are defined: before Covid-19 (Jan 2019–Feb 2020) and during Covid-19 (Mar 2020–May 2021).

5.2. Calculating textual data-based complexity

DML taxonomy is a required input to calculate TD complexity, as explained in Section 4. Since it is not provided in the case study, we develop it. For this purpose, we follow the procedure explained in Section 3, which is taken from our previous work [11].

DML taxonomy for SRM process: As we focus on the SRM process in the case study, we develop a DML taxonomy for that process. To express the cognition level of textual data in terms of linguistic features, PoS (nouns, verbs, adjectives, and adverbs) in textual data are analyzed. First, we clean the textual data in the tickets and create document-term matrices. To clean the textual data, we performed the following activities: (i) we removed signatures in the tickets received via email and (ii) URLs, email addresses, and phone numbers mentioned in the text are replaced with pseudonyms (for example, url1, email1). Second, LDA and LSI topic modeling methods [68] are combined to extract the descriptive keywords of the tickets. Finally, the extracted keywords are grouped into the three DML cognition levels (routine, semi-cognitive, and cognitive). In this grouping, the aforementioned five experts are involved in correctly identify the DML cognition level for each keyword. In particular, these experts are asked to critically evaluate and provide their feedback on the extracted keywords and their corresponding DML cognition levels. Additionally, the Information Technology Infrastructure Library (ITIL) framework [89] is used to enrich the taxonomy. The experts are asked to identify which keywords from the SRM process documentation in ITIL should be included in the taxonomy. The DML taxonomy created for the SRM process is shown in Table A.12 in the appendix. For the implementation purpose, separate text files for each part of speech and DML level are created as presented on our Github page.[2]

Using the developed DML taxonomy, we extract the DML taxonomy-based linguistic features for the given 4982 tickets. Afterward, for the same tickets, we extract the stylistic features as listed in Table 1 in Section 3.

Next, prediction models are developed to classify unlabeled tickets. Specifically, we take the four typical semi-supervised learning techniques as the basis and enrich the unlabeled data using the labeled data. Afterward, labeled and unlabeled data are combined. From the combined data, training and test data sets are created. A number of commonly used prediction modeling techniques are trained on the training data set. The best-performing prediction model is selected using the F-score metric. Then, the best-performing prediction model is run on the unlabeled data to assign TD complexities.

Fig. 6 depicts the flow from cleaning textual data to calculating TD complexities. How textual data is split and processed can be seen in the figure.

As indicated before, developing DML taxonomy-based linguistic features requires additional resources, i.e., domain expert involvement. In fact, one can argue whether such investment would provide more accurate results compared to a feature development technique that does not include user involvement. To address this concern and show the value of our linguistic features

[2] check out the DML taxonomy on our Github page.

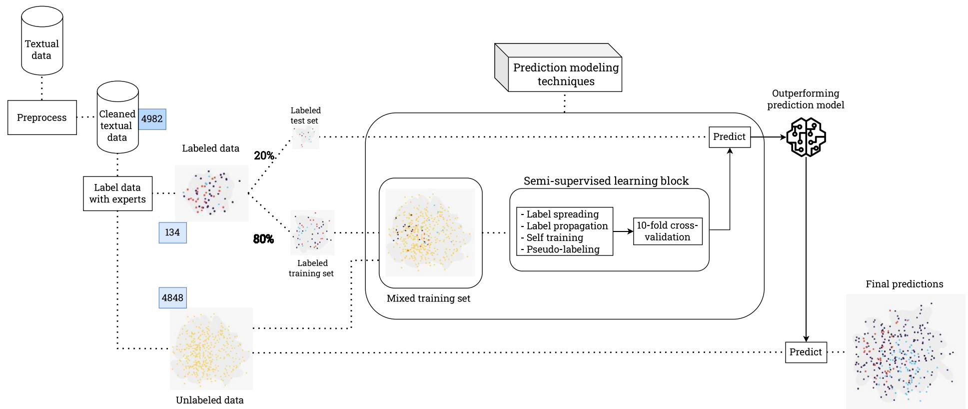


Fig. 6. Textual data-based complexity calculation by means of prediction.

in predicting TD complexity, we take Linguistic Inquiry and Word Count (LIWC), a state-of-the-art text analysis technique, predict TD complexity using LIWC features, and compare the results with our linguistic features. In this regard, first, we elaborate on LIWC features and then assess the evaluation of our approach against it.

LIWC is a dictionary-based text analysis technique that focuses on the connection between important psycho-social constructs and theories with words, phrases, and other linguistic constructions [90]. Words given in a text are analyzed and placed into one or more linguistic, psychological, and topical categories. Each of these categories indicates several aspects of a text, for example, social, cognitive, or affective. From these categories, we select linguistic features to comply with the focus of our *Approach*. The selected features and their definitions with examples (taken from [90]) are given in Table 9. As can be seen, there is considerable overlap between our linguistic features listed in Table 1 and the selected LIWC features. Notably, features derived using parts-of-speech (PoS) in text, for example, nouns, verbs, adverbs, and adjectives, show similarity.

Using the LIWC implementation,[3] we obtain the LIWC feature values (mostly counts and distribution frequencies) for the tickets. Using these values and following the same steps explained above, we develop prediction models and measure the performance of these models in terms of weighted F-score. The prediction models and their performance in the evaluation using both our linguistic features and LIWC features are given in Table 10.

As highlighted in italics in the first row in Table 10, Bagged Decision Trees is the outperforming algorithm with the best weighted F-score value. This performance is achieved with the DML taxonomy-based and stylistic features. In the case of LIWC features, the Bagged Decision Trees algorithm has notably lower performance. Overall, in almost all algorithms, we obtained a better performance with our linguistic features than LIWC features. Only in two cases, with a subtle difference, LIWC features showed a better performance. These are highlighted in the table in dark gray, namely Naïve Bayes with Label Spreading and Stacked SVM-Naïve Bayes with Pseudo-Labeling. Important to note that tree-based algorithms perform significantly well when the base semi-supervised learning technique is Pseudo-Labeling (see the top four rows in the table). However, Pseudo-Labeling is also seen in the least successful performances (see the last row in the table). Apart from that, Label Spreading is another semi-supervised learning technique observed in the majority of the less successful performances.

To perform a drill-down TD complexity calculation, we take Table 8 as the basis and create subsets for the following: per ticket attribute, pairwise combined ticket attributes, and before and during Covid-19 periods. Hereby, the case study experts indicated not only important ticket attributes but also the most relevant pairwise combinations, i.e., by "channel" attribute, as it directly affects the incoming text of a service request. The TD complexity per subset is computed using the weighted TD complexity of the tickets contained in it. For the low, medium, and high points in the complexity scale, 1, 2, and 3 are used as the weight multipliers, respectively. Simply put, the TD complexity per subset is the aggregation of individual TD complexity of tickets in the subset using these weight multipliers.

5.3. Calculating event log-based complexity

The event log provided by the case study organization is filtered, and event log subsets are created for the ticket attributes listed in Table 8, their pairwise combinations, and the defined two time periods. For example, five event log subsets are created for the five values seen in the channel attribute. To obtain a single EL complexity for an event log subset, cluster analysis is conducted. The number of clusters is set to three since the given complexity scale contains three value points. Then, for each EL complexity metric (see in Table 3), a complexity value is determined, resulting in 13 values per subset. With a majority voting, a single complexity value is selected as the EL complexity for each subset.

5.4. Analyzing correlations

In the correlation analysis, we measure to what extent the calculated TD and EL complexities correlate. Aligned with the specification in our *Approach* (see Section 4.5), the Spearman's correlation is chosen for the analysis, as there is non-normality in the data distribution and the complexity scale contains ordinal

3 check out LIWC implementation.

Table 9
LIWC features [90].

Category	Feature	Definition/Example
Summary	Word count	Total word count
	Words per sentence	Average words per sentence
	Big words	Percent words 7 letters or longer
	Dictionary words	Percent words captured by LIWC
	Analytical thinking	Metric of logical, formal thinking
Linguistic	Total function words	the, to, and, I
	Total pronouns	Total amount of pronouns
	Personal pronouns	I, you, my, me
	1st person singular	I, me, my, myself
	1st person plural	we, our, us, lets
	2nd person	you, your, u, yourself
	3rd person singular	he, she, her, his
	3rd person plural	they, their, them,
	Impersonal pronouns	that, it, this, what
	Determiners	the, at, that, my
	Articles	a, an, the, alot
	Numbers	one, two, first, once
	Prepositions	to, of, in, for
	Auxiliary verbs	is, was, be, have
	Adverbs	so, just, about, there
	Conjunctions	and, but, so, as
	Negations	not, no, never, nothing
	Common verbs	is, was, be, have
	Common adjectives	more, very, other, new
	Quantities	all, one, more, some
Psychological-cognition	All-or-none	all, no, never, always
	Cognitive processes	but, not, if, or, know
	Insight	know, how, think, feel
	Causation	how, because, make, why
	Discrepancy	would, can, want, could
	Tentative	if, or, any, something
	Certitude	really, actually, of course, real
	Differentiation	but, not, if, or
	Memory	remember, forget, remind, forgot
Expanded-states	Need	have to, need, had to, must
	Want	want, hope, wanted, wish
	Acquire	get, got, take, getting
	Lack	don't have, didn't have, hungry
	Fulfilled	enough, full, complete, extra
	Fatigue	tired, bored, don't care, boring
Expanded-time	Time	when, now, then, day
	Past focus	was, had, were, been
	Present focus	is, are, I'm, can
	Future focus	will, going to, have to, may

values. We compute the correlations for the subsets created with the help of the case study experts. The computed correlations are displayed in Table 11. The first column *Combination* contains the information regarding the grouping attribute, i.e., its presence ("Channel") or absence ("–"). In the column *Overall*, the overall values of *coef*, i.e., coefficient indicating the strength of the measured relationship between TD and EL complexities, and *p*, i.e., the quantified significance of such a relationship, are shown. They are followed by the columns *Before Covid-19* and *During Covid-19* revealing *coef* and *p* values for the indicated time periods. For instance, for the subset created using the pairwise combination of the ticket attributes "Channel" and "Duration" for the time period before Covid-19, coefficient 0.439 and *p* value 0.101 are computed.

Following the general guidelines explained in the *Approach* (see Section 4.5), 0.8 and 0.2 are chosen as the thresholds for strong and weak correlations. The significance of correlations is identified using the criterion $p \leq 0.05$, which is also explained in the same Section 4.5. Accordingly, in Table 11, strong correlations are highlighted in dark gray cells. Light gray is used for indicating the combinations where *p* values meet the criterion. As can be seen in the first six rows of the table (no grouping attribute "–"), the correlations are strong in the following four ticket attributes: "Channel", "Organizational unit", "Hold-on", and "Re-open". Moreover, in all but one of these attributes ("Organizational unit"), the over-time strong correlations are identified. The correlation strength for "Organizational unit" in one of the time periods (before Covid-19) is only 0.12 less than the defined threshold. In the combinations, i.e., pairs (grouping attribute "Channel"), a strong correlation is identified for the "Channel–Category" combination. The computed correlation during Covid-19 is significant for this pair. However, it is weak, albeit very much above the defined threshold (0.2).

Fig. 7 illustrates some of the correlations. In the figure, we used jittering to reduce overlapping points that hinder getting a sense of density. As can be seen in Table 11, we observe weak correlations for the following two subsets, including over time (no grouping attribute "–"): "Category" and "Duration". For the "Duration" attribute, there is a subtle difference between the obtained correlation and the defined strong threshold (i.e., $0.8 - 0.707 = 0.093$). Apart from that, weak correlations are observed in the three out of five combinations (grouping attribute "Channel" in Table 11): "Channel–Organizational unit", "Channel–Re-open", and "Channel–Duration". The only combination for which no correlation exists is "Channel–Hold-on". Aside from that, no negative correlations are detected.

Table 10
Evaluation of prediction models for textual data-based complexity.

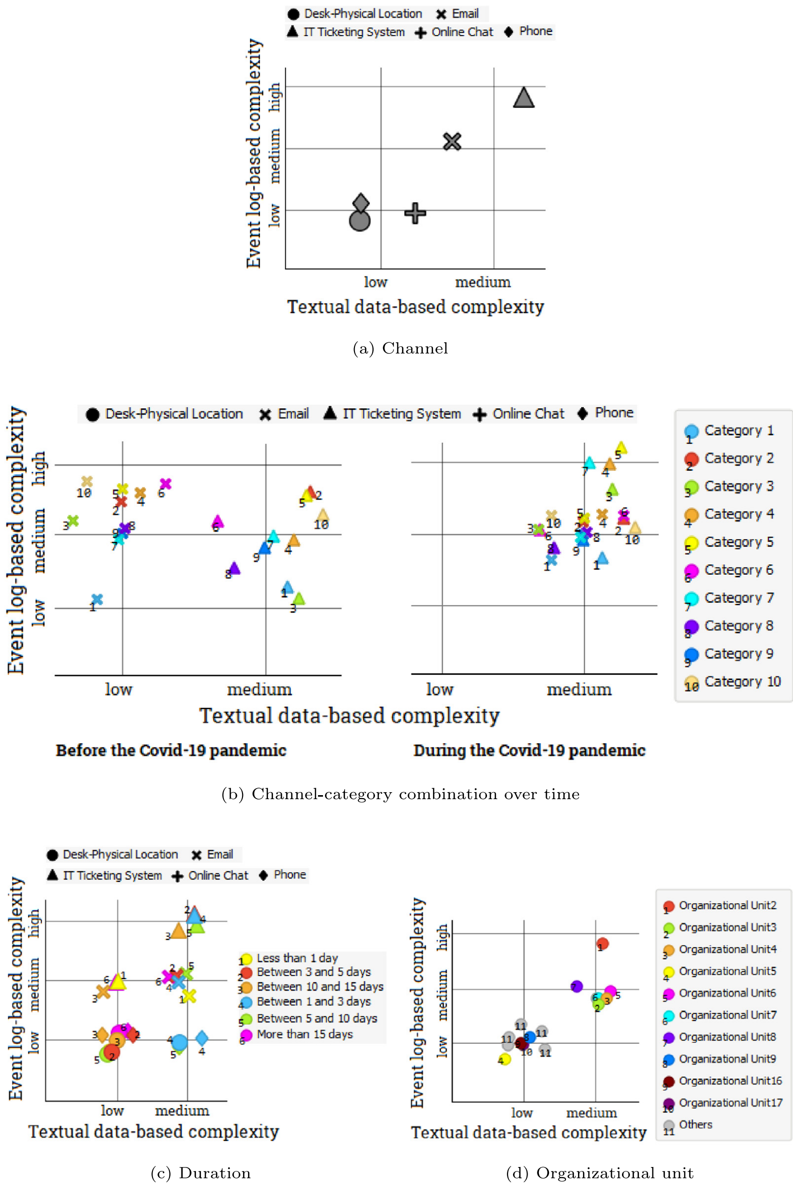
Meta algorithm	Algorithm[a]	Base SSLT[b]	Weighted F-score	
			DML taxonomy-based & stylistic features	LIWC features
Bagging	*Decision Trees*	*PL*	*0.943*	*0.865*
–	Random forest	PL	0.94	0.852
–	Extra trees	PL	0.938	0.881
Boosting	Random forest	PL	0.931	0.872
Boosting	Gradient boosting	PL	0.922	0.832
–	Logistic regression	PL	0.919	0.881
–	K-Nearest neighbors	PL	0.914	0.877
–	Decision trees	PL	0.909	0.85
Boosting	AdaBoost	PL	0.901	0.839
–	Stochastic gradient descent	PL	0.9	0.697
–	Naïve Bayes	PL	0.899	0.904
–	Perceptron	ST	0.897	0.826
–	Support vector machines	ST	0.897	0.818
–	Support vector machines	PL	0.897	0.852
Stacking	Stacked: Support vector machines and Naïve Bayes, finalizer: Logistic regression	PL	0.897	0.896
–	Logistic regression	ST	0.893	0.818
–	Perceptron	PL	0.893	0.837
–	Extra trees	LS	0.89	0.867
–	Support vector machines	LS	0.89	0.826
Stacking	Stacked: Support vector machines and Naïve Bayes, finalizer: Logistic regression	LS	0.89	0.826
–	Random forest	LS	0.89	0.867
–	Stochastic gradient descent	ST	0.889	0.697
–	Decision trees	LS	0.889	0.867
Stacking	Stacked: Support vector machines and Naïve Bayes, finalizer: Decision trees	LS	0.885	0.826
–	K-Nearest neighbors	LS	0.885	0.826
Boosting	Gradient boosting	LS	0.885	0.826
Boosting	Random forest	LS	0.884	0.826
Bagging	Decision trees	LS	0.883	0.826
–	Naïve Bayes	LP	0.871	0.794
Boosting	AdaBoost	LS	0.839	0.818
Stacking	Stacked: Support vector machines and Naïve Bayes, finalizer: Decision trees	PL	0.815	0.828

[a]For algorithms, we refer to the Python scikit-learn library implementation on https://scikit-learn.org/stable/.
[b]Semi-Supervised Learning Technique. PL: Custom Pseudo-Labeling, ST: Self-Training, LS: Label Spreading, and LP: Label Propagation.

Table 11
Correlations between textual data- and event log-based complexities.

Combination		Overall		Before Covid-19		During Covid-19	
Grouping attribute	Attribute	coef.	p	coef.	p	coef.	p
–	Channel	0.968	0.007	1	0	1	0
–	Organizational unit	0.983	0	0.788	0	0.833	0
–	Hold-on	0.866	0.333	0.866	0.333	0.866	0.333
–	Re-open	0.866	0.333	0.866	0.333	0.866	0.333
–	Duration	0.707	0.116	0.707	0.116	0.707	0.116
–	Category	0.577	0.081	0.655	0.056	0.632	0.092
Channel	Category	0.816	0	0.607	0	0.977	0
Channel	Organizational unit	0.577	0	0.364	0.005	0.533	0
Channel	Re-open	0.556	0.048	0.342	0.232	−0.082	0.782
Channel	Duration	0.476	0.019	0.594	0.001	0.536	0.003
Channel	Hold-on	0.037	0.895	0.439	0.101	0.426	0.113

Furthermore, the significance of the correlations can be seen in the *p* column in Table 11. "Channel" and "Organizational unit" are the attributes with significant correlations, including the two time periods. For the pair "Channel–Category", a similar observation can be obtained from the table. In addition, the significance criterion is met when the "Channel" is combined with



(a) Channel

(b) Channel-category combination over time

(c) Duration

(d) Organizational unit

Fig. 7. Correlations between textual data- and event log-based complexities.

"Organizational unit", "Re-open", and "Duration" attributes, albeit their computed correlations are not strong.

The correlations explained above serve as the basis for analyzing statistically significant differences in the execution of BPs based on the event log subsets. In the following subsection, we present the results of that analysis.

5.5. *Analyzing significant differences in business process executions*

Considering Table 11, we focus on strong correlations, weak correlations, and changes in correlations over time. Accordingly, we analyze the statistically significant differences between the event log subsets in the related correlations. The most interesting findings from that analysis are presented in this subsection.

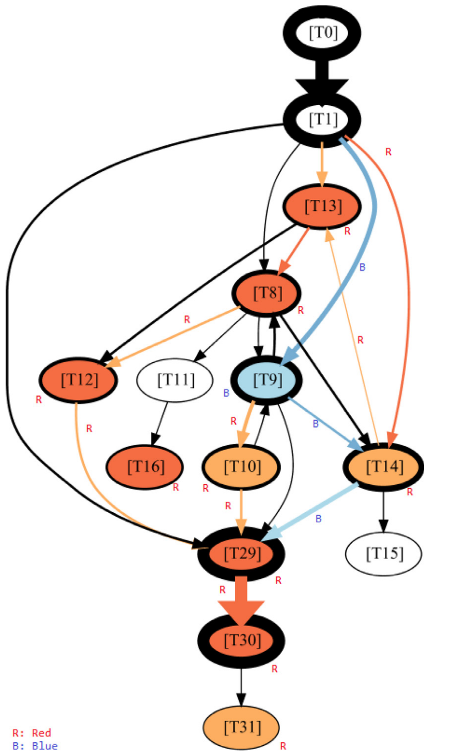


Fig. 8. Significant differences between more interactive channels (with B) and less interactive channels (with R). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

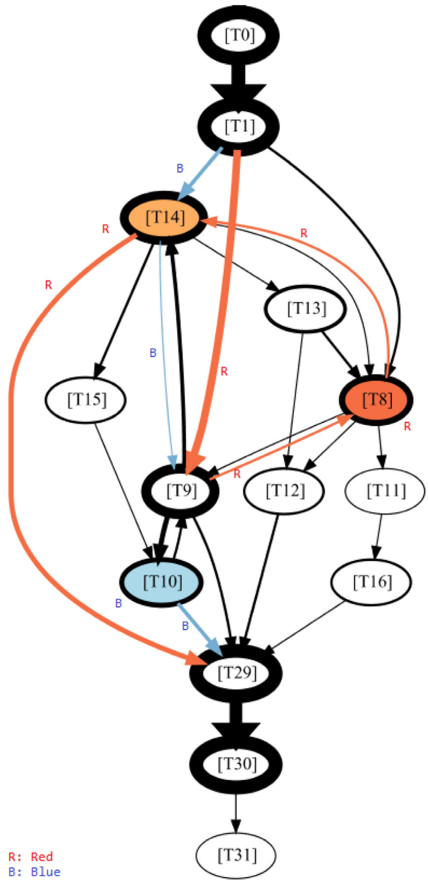


Fig. 9. Significant differences between categories 3, 4, 5, and 7 (with B) and others (with R). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Due to their real-time conversation nature, phone, desk-physical location, and online chat are considered more interactive channels. Unlike, email and IT ticketing system are the two less interactive channels. As shown in Table 11 and in Fig. 7(a), we obtained strong significant correlations in the "Channel" attribute. More specifically, EL complexity increases when the channel is less interactive. Hence, we investigated the differences between the event log subsets of more and less interactive channels.

In Fig. 8, statistically significant differences in the SRM process execution between more and less interactive channels are displayed. Activities and paths that more often reoccur in handling tickets received via email and IT ticketing system are presented in red (with R). Blue (with B) is used for phone, desk-physical location, and online chat. The activity T9 is related to the categorization of tickets and their assignments to the ticket resolution teams. In interactive channels, it is usually performed after the "Registration" activity (T1). However, in less interactive channels, activities "Pause" and "Ask end-user for extra information" (T12 and T13) occur more often. In addition, "Hand-over" (T10) and "Work assignment" (T14) are also frequent activities for these channels. Aside from that, "Reopening" (T31) of the tickets coming via these channels takes place more often compared to the interactive channels. Aligned with that, activities that are related to the resolution and closure of the ticket (T16, T29, and T30) and changing the ticket status (T8) occur more often for the same channels. In a further analysis, the tickets coming via less interactive channels, i.e., email and IT ticketing system, are checked. A notable observation is that the priority of the tickets with a medium TD complexity is frequently changed.

"Channel–Category" is the only pair (grouping attribute "Channel") in which strong significant correlations are observed (see Table 11). Moreover, as depicted in Fig. 7(b), in the four ticket categories (Category 3, 4, 5, and 7), the EL complexity of tickets registered via IT ticketing system increases to high while their TD complexities remain unchanged. Fig. 9 shows the differences in handling the tickets in these four categories (blue color) in comparison to the other categories. Notably, "Changing assigned resolution teams" activity (T10) in the tickets is more common in these categories. In other categories, "Work assignment" (T14) is frequently handled in the same resolution teams.

In addition, as can be seen in Fig. 7(b), the TD complexity of the tickets coming via email channel is low before Covid-19 and increases to medium during Covid-19. However, there is no recognizable change in their EL complexities. Therefore, these tickets are further analyzed. It is found out that the median duration for handling these tickets has been doubled.

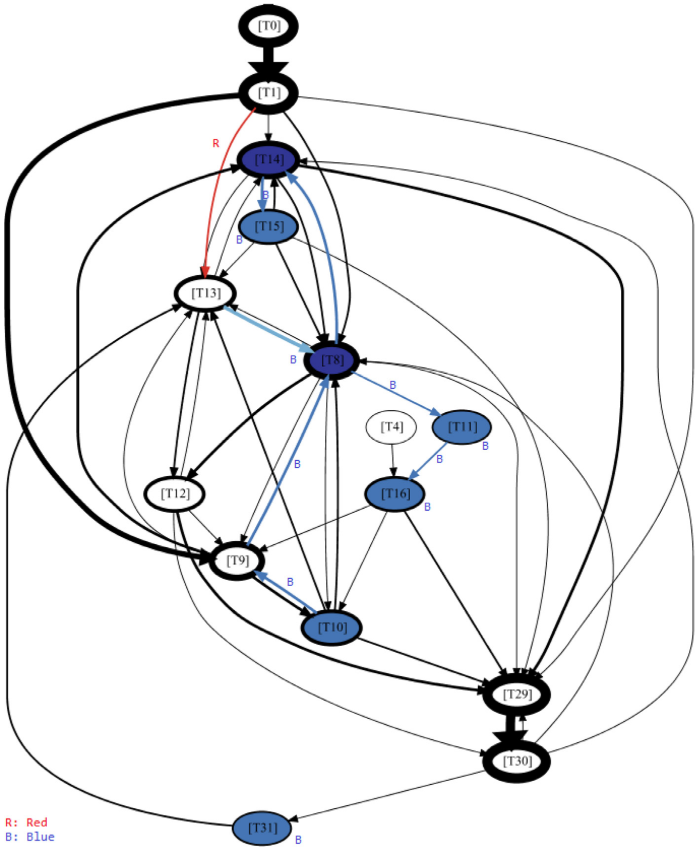


Fig. 10. Significant differences between duration longer than 10 days (with B) and up to 10 days (with R). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 10 shows the difference analysis based on Fig. 7(c), i.e., "Channel–Duration". In particular, we investigate the reasons that may account for the longer execution duration (i.e., more than 10 days) of the tickets with low TD complexity and received via interactive channels (phone and desk-physical location). As highlighted in blue (with B), frequent "Reopening" activity (T31) is observed. Because of frequent reopening, other main activities (for example, T14 and T8) are repeated and highlighted in blue (with B) to indicate their recurrence. Noteworthy, there might be several reasons for a longer execution duration of the tickets with low TD complexity and vice versa, a subject for further investigation. For example, in some cases, end-users trigger the reopening of a prior ticket by appending a follow-up request or requesting minor adjustments.

Next, in Fig. 7(d), we focus on organizational units issuing the requests. Organizational unit2 is the outlier unit in terms of the TD and EL complexities. We analyze the executions of the tickets of that unit and compare them with the tickets coming from the rest of the units. In Fig. 11, the frequently performed activities of handling the tickets coming from that unit are displayed in blue (with B), for example, "Work assignment" (T14) and "Ask end-user for extra information" (T13) activities. As can be seen, "Changing ticket status" (T8) happens more frequently. In addition, in the tickets coming from Organizational unit2, "Work assignment" (T14) is directly followed by "Provide resolution" activity (T29) more often.

Considering the findings explained in the subsections above, we discuss their implications in the following section.

6. Discussion

We discuss our findings, their implications, and the limitations of our *Approach* in two subsections. Section 6.1 is devoted to the interpretations of the case study results and derived relevant observations. Section 6.2 highlights the benefits of the *Approach* while mentioning its limitations.

6.1. Analyzing results and deriving observations

To obtain recommendations for BP redesign and improvement in the proposed *Approach*, we address three important challenges in IT ticket processing: ticket categorization, work assignment, and prioritization. Textual data describing tickets are the basis for identifying ticket categories that are then used to assign resolution teams to the tickets. Moreover, tickets are prioritized based

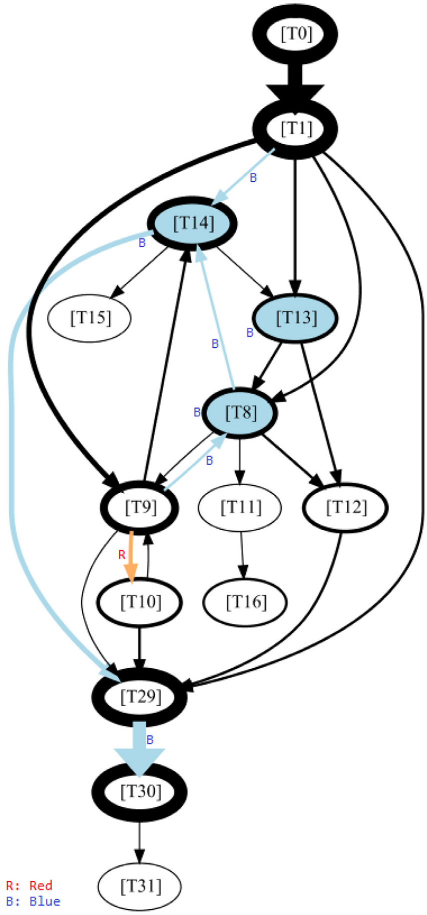


Fig. 11. Significant differences between organizational unit2 (with B) and others (with R). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

on textual data. In IT ticket processing, the priority of tickets has a major influence on the way they are handled. Hence, we discuss our case study findings by focusing on such challenges.

The case study results have shown that the TD and EL complexities increase when the channel is less interactive, i.e., email or IT ticketing system. Specifically, in these channels, tickets are held on ("Pause" and "Ask end-user for extra information" activities) frequently. As a result, constant work reassignments between resolution teams are detected. Moreover, repetitions of such activities show a negative influence on the ticket handling duration, which directly affects Service Level Agreements (SLAs). Considering these, one can infer that less verbal communication when submitting the service request raises the complexity. On the contrary, in interactive channels, textual descriptions of end-user service requests are clear and comprehensive. The real-time verbal communication nature of interactive channels enables the requests to be interpreted and registered correctly. Additionally, in these channels, operators can guide end-users to provide all required information in a single conversation. Accurate ticket categorization, prioritization, and work assignment can be better accomplished using the mentioned advantages of interactive channels.

The analysis using the "Channel–Category" combination shows that resolution teams are changed frequently in some ticket categories. Such changes often happen when textual data are not comprehensive enough to identify the resolution teams correctly. The involved experts in the case study interpret such frequent change of resolution teams as a "ping-pong" behavior. In other words, due to the lack of clarity in textual data, tickets are passing from one resolution team to another until more information is available to detect the correct team.

Next, we note that tickets coming from specific organizational units were more complex. In particular, in one of 17 organizational units, namely Organizational unit2, medium TD and high EL complexities are obtained. To unveil the reasons for high EL complexity, we conducted the significant difference analysis. In particular, the event log of the tickets requested by this organizational unit is compared with the event log of the tickets sent by the remaining 16 organizational units. We found out that "Work assignment" and "Ask end-user for extra information" activities frequently happen and cause high EL complexity in handling tickets coming from Organizational unit2. Moreover, to understand the reasons for TD complexity, we analyzed the attributes of the tickets coming from this organizational unit. Based on the "Category" distribution of the tickets, it is observed that Organizational unit2 often requests services requiring the involvement of multiple resolution teams. Moreover, technical terms are commonly used in the textual data of the tickets. The case study experts mention that end-users belonging to that organizational unit have technical backgrounds and knowledge. In this regard, we have discovered that the stylistic features of these tickets differ from the tickets of other organizational units in the sense of relative occurrences of unique PoS and wording style. This observation indicates that in this case, stylistic features making part of the TD complexity have an important influence on the EL complexity, i.e., actual ticket processing.

In the context of the implications of the observations about organizational units, the following points should be considered by organizations when applying our *Approach*. The compositions of particular departments or teams and regulations within them are likely to have a notable influence on textual data. For example, textual data in the requests of users with rights to install software and change configurations on their computers will differ from the textual data in the requests of users who have no such flexibility in using a computer. Combining the textual data of such organizational units and analyzing separately from the textual data of other grouped units may provide more specific indications of what leads to complexity. Hence, tailored solutions for addressing TD complexity in such organizational units may be more effective than general solutions for the entire organization.

Another observation is that in some tickets coming via interactive channels, the TD complexity and ticket execution duration are inversely related. Noteworthy, the EL complexity of these tickets remains unchanged. In such exceptional cases, further analysis considering other ticket attributes and interactions between teams and end-users is required. For example, due to some changes in the IT infrastructure or outsourced services, a ticket with a low TD complexity may take more time to handle.

Aside from the aforementioned, the case study findings also lead to the following further observations:

- DML taxonomy-based and stylistic features are rather adequate in predicting TD complexity as seen in comparison with LIWC features.

- Simple machine learning algorithms, such as tree-based ones, are powerful and perform well when combined with pseudo-labeling to deal with semi-supervised learning problems.
- Entity attributes (for example, channel, category) play an important role in better understanding the relation between TD and EL complexities by means of a drill-down analysis.

In addition, the DML taxonomy that is developed for the SRM process is of practical value and reusable. Since in most organizations operations are dependent on IT, a similar SRM process exists as part of their ITSM strategy. Hence, it is reasonable that IT ticket texts contain similar concepts in organizations. Accordingly, the developed DML taxonomy can be adjusted and used for calculating TD complexity.

Moreover, the results obtained in the case study serve as a proof of concept for using linguistic features in TD complexity prediction. In other words, the potential of linguistic features as textual data representation discovered in our previous work [13] has been confirmed in a new real-life setting. Notably, DML taxonomy-based and stylistic features are favorable to reveal such potential in analyzing TD complexity.

Despite the fact that domain expert involvement is necessary for developing DML taxonomy-based features, in comparison with LIWC features, we showed that such an effort resulted in a better TD complexity classification performance. More specifically, the average overall increase in the performance is 7.5%, whereas it goes up to 9.5% when only the top 10 best-performing prediction models are considered (see Table 10). Furthermore, the highest increase in the performance, 29%, is observed in Naïve Bayes with Pseudo-Labeling. Likewise, 1.5% is determined as the highest performance loss in the stacking of Naïve Bayes and SVM with Pseudo-Labeling. Though it is difficult to measure precisely, based on the difference in the classification performance, we can infer that such a performance gain outweighs the cost of DML development. Hence, organizations that want to adopt our *Approach* can take the comparison as an initial baseline for a trade-off between costs and performance gain.

6.2. Outlining benefits and limitations

Considering the results discussed above and correlations presented in Section 5, we conclude that TD complexity has connections to EL complexity. We show the benefits of our *Approach* in detecting and analyzing the relation between these two complexities in a real-world setting. Specifically, our observations clearly demonstrate that increasing TD complexity could account for longer BP execution time, less accurate ticket categorization and prioritization, and frequent work reassignments triggered by hold-ons and re-opens.

Overall, the connection between TD complexity and EL complexity indicates that textual data are appropriate and can be used for predicting EL complexity. Organizations operating in various domains often rely on textual data while performing their BPs since textual data are generally the primary input to their BPs. Such organizations can highly benefit from our *Approach*. For example, banks, governmental bodies, universities, infrastructure or supply providers have established BPs for handling various requests coming from their customers in a textual form. Complexity in these texts can considerably influence the execution of their BPs. By studying this complexity, organizations can predict and, therefore, mitigate complexity in performing BPs. Furthermore, using our *Approach*, organizations can have a more comprehensive way of identifying process redesign and improvement opportunities. Such opportunities can be formulated based on the activities in BPs that affect EL complexity. These activities can be detected in the significant difference analysis phase of the *Approach*.

Important to note that our *Approach* reveals a great potential for understanding the implications of each specific linguistic feature for textual complexity. As illustrated in the running example (see Table 5), the contribution of each feature to complexity can be identified separately by tracing back from the aggregated single TD complexity value. Having such information can help organizations in several ways. In real-time, text provided by users to BPs can be annotated in terms of complexity contributions. Hence, based on the indications of those text parts leading to complexity, users can improve text quality. Another way could be assisting process workers in rephrasing text for a more accurate interpretation. For example, depending on the text quality, process workers may apply triage on requests or add an extra explanation in the text to better identify what activities are required in handling requests. Providing support for reducing the use of words or phrases resulting in higher TD complexity and further EL complexity could also be beneficial for organizations.

Further, the *Approach* sets apart from the state-of-the-art approaches on BP complexity analysis by (i) combining textual data and event log, (ii) blending readily available techniques in calculating TD and EL complexities, and (iii) analyzing the relation between them. Our *Approach* is the first effort to analyze the connection between TD and EL complexities. For doing this, publicly accessible and frequently used techniques, such as LDA, LSI, common machine learning algorithms, and existing EL complexity metrics, are employed.

Although the case study of our previous research serving the basis for TD complexity as well as the case study of this paper belong to the ITSM domain, i.e., IT ticket processing, the research value goes far beyond the latter. The *Approach* can be used by organizations relying on textual data as an input to their processes and already executing or interested in initiating BPM projects, for example, healthcare or public administration institutions receiving customer requests in a textual form. The domain-specific adaptations of the *Approach* may require additional efforts of different degrees. Accordingly, among the four inputs of the *Approach*, which are textual data, event log, complexity scale, and DML taxonomy, the latter requires the most manual effort. Hereby, the availability of experts and their willingness to dedicate time and resources as well as top management support can significantly influence the process of DML taxonomy creation. At the same time, DML taxonomy captures the essential semantics enabling context awareness and domain adaptation in the *Approach*.

As also follows from the title, the focus of the paper lies on the BP execution data. However, if we consider the BPM lifecycle [1], the role of textual data goes far beyond BP execution. It can be employed throughout all phases of the lifecycle. For example, in the discovery phase, process analysts might use available BP descriptions, textual data from interviews, legal documents, or ethnography [91–93]. In other BPM lifecycle phases, such as process analysis, process redesign, and process monitoring and controlling, any BP changes must comply with legal requirements, corporate standards, business rules, and service operating procedures which usually exist in textual form. All these documents have a more official character than textual data produced by BP participants (like conversations) in the BP execution revealing different style and, hence, textual complexity. Such textual information is not considered in the present paper. However, this limitation represents a promising direction for future research. Accordingly, our *Approach* can be extended with the analyses of further textual data sources to develop support for process analysts at various phases of the BPM lifecycle.

Our *Approach* reveals several further limitations. Thus, identification of the relevant entity attributes for creating subsets to perform a drill-down analysis is performed manually in the

Approach. In that sense, it is a limitation. However, referred entity attributes in the BPs can be identified and incorporated into the *Approach*. Hence, one can automatically select entity attributes that are relevant for particular BPs. Another limitation of the *Approach* is related to textual data: it only supports the texts written in English. With the advancement of NLP libraries, more languages can be considered in the *Approach*.

7. Conclusion and future work

In this paper, we presented a novel *Approach* aimed at BP execution complexity analysis by combining textual data and event log. In particular, in the *Approach*, the relation between TD and EL complexities is analyzed. To calculate TD complexity, we use two sets of linguistic features aimed at capturing TD complexity in terms of cognition and style: DML taxonomy-based and stylistic features. For calculating EL complexity, the state-of-the-art event log-based complexity metrics are employed. Then, the correlations between these two complexities are measured to check how TD and EL complexities are associated. Based on the computed correlations and analysis of the significant differences, the BP activities affecting EL complexity are determined. Hence, the factors that may account for an increase or decrease in EL complexity are identified.

The proposed *Approach* was applied in the IT department of the academic institution. Specifically, using the textual data (service requests descriptions) and an event log from an SRM process, TD and EL complexities were calculated, and their relation was analyzed. Further, the advantages of DML taxonomy-based and stylistic features in determining TD complexity are assessed by means of comparison against well-accepted linguistic features, LIWC, in predicting TD complexity. Our findings showed that TD and EL complexities highly correlate in most cases. Thus, following our *Approach*, textual data can be used to predict the complexity of BP execution.

The results presented in this paper are, to the best of our knowledge, the first investigation of the connection between TD and EL complexities related to BP execution. The existence of such a connection implies that organizations can benefit from studying the complexity expressed by means of textual data. For instance, organizations that rely on textual data in performing their BPs can approximate the BP execution complexity. Thus, organizations can develop strategies and make prior decisions to deal with the complexity of BP execution. Furthermore, the *Approach* enables organizations to identify the factors that affect the complexity in BP execution. By interpreting these factors, process redesign and improvement directions can be determined.

In future work, we will expand our scope with other ITSM processes and incorporate conversations captured in BP executions into textual data analysis. Moreover, we would like to include decision mining technologies [94] in our *Approach* since the BP decision points are very likely to be associated with BP variants. Apart from that, we aim to develop new linguistic features using sentences and their dependencies to further improve TD complexity prediction. Combining our linguistic features with LIWC features for achieving a better classification performance is another direction we want to pursue. Regarding EL complexity, we want to move one step further, focus on discovered process models, and hence, incorporate process complexity metrics into our *Approach*. In addition, investigating the potential of text summarization for obtaining complexity from textual data is another future work avenue. We will also consider the complexity analysis of other textual data, such as BP descriptions, legal documents, corporate standards, and interview transcriptions, to assist the process analysts at all the stages of the BPM lifecycle. Lastly, we will experiment with other approaches to process complexity, for example, considering process context [73,74], to enhance our *Approach*.

Table A.12
Decision-making logic taxonomy for service request management process.

	Decision-making cognition levels		
	Routine	Semi-cognitive	Cognitive
Resources (Nouns)	account, activation, address, admin, administrator, admission, agenda, appointment, assessment, authentication, authenticator, authorization, booking, capacity, certificate, code, contract, credit, demo, download, guest, host, intranet, licence, license, link, mail, mailaddress, mobile, password, permission, phone, portal, printer, questionnaire, reference, registration, reset, staff, storage, twofactorauthentication, update, url, version	acceptance, app, application, archive, backup, balance, bank, battery, cloud, configuration, document, domain, keychain, mailbox, migration, network, security, software, toner, upgrade, video, vpn	analysis, database, disk, distribution, drive, driver, file, firewall, folder, group, owner, ownergroup, recovery, workstation, server
Techniques (Verbs)	complete, connect, create, download, exist, expand, expire, extend, find, grant, increase, inform, install, login, logon, open, print, reset, restart, run, save, send, share, start, stop, store, write, update	access, activate, add, approve, assign, associate, deactivate, decide, delete, disable, edit, link, make, reactivate, recover, remove, replace, set, unlink	analyse, analyze, change, define, design, lose, migrate, modify
Capacities (Adjectives)	active, additional, available, correct, easier, exact, extra, free, full, higher, important, larger, last, latest, least, new, newest, next, older, online, optional, organizational, original, outdated, private, public, qualitative, ready, relevant, responsible, several, urgent, valid, visible, wide	empty, external, incorrect, international, local, long, multiple, remote, safe, unable, wrong, unknown	broken, different, compatible, offline, temporary, stuck, spatial
Choices (Adverbs)	almost, apparently, beforehand, completely, correctly, directly, easily, efficiently, either, else, ever, exactly, far, fully, last, mainly, much, next, otherwise, properly, quite, rather, since, soon, successfully, together, totally, urgently	accidentally, already, also, always, anymore, anywhere, automatically, constantly, everytime, frequently, indeed, initially, instead, manually, maybe, moreover, mostly, never, often, perhaps, previously, probably, regularly, sometimes, somewhere, suddenly, temporarily, though, twice, usually, wrongfully	hence, however, locally, remotely, somehow, still, therefore, thus, yet

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Appendix A. DML taxonomy

See Table A.12.

References

[1] M. Dumas, M. La Rosa, J. Mendling, H.A. Reijers, Fundamentals of Business Process Management, second ed., Springer, 2018.
[2] C. Matt, T. Hess, A. Benlian, Digital transformation strategies, Bus. Inf. Syst. Eng. 57 (5) (2015) 339–343.
[3] M. Fischer, F. Imgrund, C. Janiesch, A. Winkelmann, Strategy archetypes for digital transformation: Defining meta objectives using business process management, Inf. Manage. 57 (5) (2020) 103262.
[4] A. Augusto, J. Mendling, M. Vidgof, B. Wurm, The connection between process complexity of event sequences and models discovered by process mining, Inform. Sci. 598 (2022) 196–215.
[5] H. van der Aa, J. Carmona, H. Leopold, J. Mendling, L. Padró, Challenges and opportunities of applying natural language processing in business process management, in: Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018, Santa Fe, New Mexico, USA, August 20-26, 2018, Association for Computational Linguistics, 2018, pp. 2791–2801.
[6] V. Kobayashi, S.T. Mol, H. Berkers, G. Kismihók, D.N. den Hartog, Text mining in organizational research, Organ. Res. Methods 21 (2018) 733–765.
[7] M. Schäfermeyer, C. Rosenkranz, R. Holten, The impact of business process complexity on business process standardization, Bus. Inf. Syst. Eng. 4 (5) (2012) 261–270.
[8] B.T. Pentland, P. Liu, W. Kremser, T. Hærem, The dynamics of drift in digitized processes, MIS Q. 44 (2020) 19–47.
[9] B. Münstermann, A. Eckhardt, T. Weitzel, The performance impact of business process standardization: An empirical evaluation of the recruitment process, Bus. Process Manage. J. (2010).
[10] A. Gunasekaran, B. Nath, The role of information technology in business process reengineering, Int. J. Prod. Econ. 50 (2–3) (1997) 91–104.
[11] N. Rizun, A. Revina, V.G. Meister, Method of decision-making logic discovery in the business process textual data, in: International Conference on Business Information Systems, Springer, 2019, pp. 70–84.
[12] N. Rizun, V.G. Meister, A. Revina, Discovery of stylistic patterns in business process textual descriptions: It ticket case, Innov. Manage. Educ. Excell. Through Vis. (2020).
[13] N. Rizun, A. Revina, V.G. Meister, Assessing business process complexity based on textual data: Evidence from ITIL IT ticket processing, Bus. Process Manage. J. (2021).
[14] A. Revina, K. Buza, V.G. Meister, IT ticket classification: The simpler, the better, IEEE Access 8 (2020) 193380–193395.
[15] N. Rizun, A. Revina, Business sentiment analysis. concept and method for perceived anticipated effort identification, in: Proceedings of the 28th International Conference on Information Systems Development: Information Systems Beyond 2020 (ISD 2019), Toulon, France, August 28-30, 2019, 2019.
[16] P. Bellan, M. Dragoni, C. Ghidini, Process extraction from text: state of the art and challenges for the future, 2021, arXiv preprint 2110.03754.
[17] A. Revina, Business process management: Integrated data perspective. A framework and research agenda, in: Proceedings of the 29th International Conference on Information Systems Development: Crossing Boundaries Between Development and Operations (DevOps) in Information Systems (ISD2021), Valencia, Spain, September 8-10, 2021, 2021.
[18] K. Honkisz, K. Kluza, P. Wiśniewski, A concept for generating business process models from natural language description, in: International Conference on Knowledge Science, Engineering and Management, Springer, 2018, pp. 91–103.
[19] X. Han, L. Hu, L. Mei, Y. Dang, S. Agarwal, X. Zhou, P. Hu, A-BPS: Automatic business process discovery service using ordered neurons LSTM, in: 2020 IEEE International Conference on Web Services, ICWS, IEEE, 2020, pp. 428–432.
[20] H. van der Aa, C. Di Ciccio, H. Leopold, H.A. Reijers, Extracting declarative process models from natural language, in: International Conference on Advanced Information Systems Engineering (CAiSE 2019), Springer, 2019, pp. 365–382.
[21] A.J. Chambers, A.M. Stringfellow, B.B. Luo, S.J. Underwood, T.G. Allard, I.A. Johnston, S. Brockman, L. Shing, A. Wollaber, C. VanDam, Automated business process discovery from unstructured natural-language documents, in: International Conference on Business Process Management (BPM 2020), Springer, 2020, pp. 232–243.
[22] C. Kecht, A. Eggert, W. Kratsch, M. Roglinger, Event log construction from customer service conversations using natural language inference, in: 3rd International Conference on Process Mining (ICPM 2021), IEEE, 2021, pp. 144–151.
[23] C. Qian, L. Wen, A. Kumar, L. Lin, L. Lin, Z. Zong, S. Li, J. Wang, An approach for process model extraction by multi-grained text classification, in: International Conference on Advanced Information Systems Engineering (CAiSE 2020), Springer, 2020, pp. 268–282.
[24] H.A. López, S. Debois, T.T. Hildebrandt, M. Marquard, The process highlighter: From texts to declarative processes and back, in: International Conference on Business Process Management (BPM 2018, Dissertation/Demos/Industry), Vol. 2196, 2018, pp. 66–70.
[25] W.M.P. van der Aalst, Process Mining: Data Science in Action, Springer, 2016.
[26] M. Gupta, P. Agarwal, T. Tater, S. Dechu, A. Serebrenik, Analyzing comments in ticket resolution to capture underlying process interactions, in: International Conference on Business Process Management (BPM 2020), Springer, 2020, pp. 219–231.
[27] A. Rebmann, H. van der Aa, Extracting semantic process information from the natural language in event logs, in: International Conference on Advanced Information Systems Engineering (CAiSE 2021), Springer, 2021, pp. 57–74.
[28] A. Goossens, M. Claessens, C. Parthoens, J. Vanthienen, Extracting decision dependencies and decision logic from text using deep learning techniques, in: International Conference on Business Process Management (BPM 2021), Springer, 2021, pp. 349–361.
[29] V. Etikala, Z.V. Veldhoven, J. Vanthienen, Text2Dec: extracting decision dependencies from natural language text for automated DMN decision modelling, in: International Conference on Business Process Management (BPM 2020), Springer, 2020, pp. 367–379.
[30] L. Quishpi, J. Carmona, L. Padró, Extracting decision models from textual descriptions of processes, in: International Conference on Business Process Management (BPM 2021), Springer, 2021, pp. 85–102.
[31] L. Quishpi, J. Carmona, L. Padró, Extracting annotations from textual descriptions of processes, in: International Conference on Business Process Management, Springer, 2020, pp. 184–201.
[32] L. Ackermann, J. Neuberger, S. Jablonski, Data-driven annotation of textual process descriptions based on formal meaning representations, in: International Conference on Advanced Information Systems Engineering (CAiSE 2021), Springer, 2021, pp. 75–90.
[33] H. Wang, L. Wen, L. Lin, J. Wang, Rlrecommender: a representation-learning-based recommendation method for business process modeling, in: International Conference on Service-Oriented Computing, Springer, 2018, pp. 478–486.
[34] S. Deng, D. Wang, Y. Li, B. Cao, J. Yin, Z. Wu, M. Zhou, A recommendation system to facilitate business process modeling, IEEE Trans. Cybern. 47 (6) (2016) 1380–1394.
[35] H. van der Aa, A. Rebmann, H. Leopold, Natural language-based detection of semantic execution anomalies in event logs, Inf. Syst. 102 (2021) 101824.
[36] D. Sola, H. van der Aa, C. Meilicke, H. Stuckenschmidt, Exploiting label semantics for rule-based activity recommendation in business process modeling, Inf. Syst. (2022) 102049.
[37] M. Goldstein, C. González-Álvarez, Augmenting modelers with semantic autocompletion of processes, in: International Conference on Business Process Management (BPM 2021), Springer, 2021, pp. 20–36.
[38] H. Leopold, H. van der Aa, J. Offenberg, H.A. Reijers, Using hidden Markov models for the accurate linguistic analysis of process model activity labels, Inf. Syst. 83 (2019) 30–39.
[39] H. van der Aa, H. Leopold, H.A. Reijers, Checking process compliance against natural language specifications using behavioral spaces, Inf. Syst. 78 (2018) 83–95.
[40] H. van der Aa, H. Leopold, H.A. Reijers, Comparing textual descriptions to process models–the automatic detection of inconsistencies, Inf. Syst. 64 (2017) 447–460.
[41] J. Sànchez-Ferreres, H. van der Aa, J. Carmona, L. Padró, Aligning textual and model-based process descriptions, Data Knowl. Eng. 118 (2018) 25–40.
[42] M. Kobeissi, N. Assy, W. Gaaloul, B. Defude, B. Haidar, An intent-based natural language interface for querying process execution data, in: 3rd International Conference on Process Mining (ICPM 2021), IEEE, 2021, pp. 152–159.

[43] H. Leopold, H. van der Aa, F. Pittke, M. Raffel, J. Mendling, H.A. Reijers, Searching textual and model-based process descriptions based on a unified data format, Softw. Syst. Model. 18 (2) (2019) 1179–1194.
[44] A. Yadav, D.K. Vishwakarma, Sentiment analysis using deep learning architectures: a review, Artif. Intell. Rev. 53 (6) (2020) 4335–4385.
[45] E. Lüftenegger, S. Softic, Sentipromo: a sentiment analysis-enabled social business process modeling tool, in: International Conference on Business Process Management (BPM 2020), Springer, 2020, pp. 83–89.
[46] A. Mustansir, K. Shahzad, M.K. Malik, Towards automatic business process redesign: an NLP based approach to extract redesign suggestions, Autom. Softw. Eng. 29 (1) (2022) 1–24.
[47] T. Niesen, S. Dadashnia, P. Fettke, P. Loos, A vector space approach to process model matching using insights from natural language processing, in: Multikonferenz Wirtschaftsinformatik (MKWI 2016), Universitätsverlag Ilmenau, 2016, pp. 93–104.
[48] N. Wang, S. Sun, D. OuYang, Business process modeling abstraction based on semi-supervised clustering analysis, Bus. Inf. Syst. Eng. 60 (6) (2018) 525–542.
[49] F. Dai, M. Liu, Q. Mo, B. Huang, T. Li, Refactor business process models for efficiency improvement, in: Cloud Computing, Smart Grid and Innovative Frontiers in Telecommunications, Springer, 2019, pp. 454–467.
[50] F. Pittke, Linguistic Refactoring of Business Process Models. Dissertation, Tech. Rep., Vienna University of Economics and Business, 2015.
[51] J. Mendling, H. Leopold, A. Polyvyanyy, Supporting process model validation through natural language generation, Softw. Eng. P-252 (2016) 71–72.
[52] B. Aysolmaz, H. Leopold, H.A. Reijers, O. Demirörs, A semi-automated approach for generating natural language requirements documents based on business process models, Inf. Softw. Technol. 93 (2018) 14–29.
[53] L. Ackermann, Language-centric approaches for improving business process model acceptance., in: International Conference on Business Process Management (BPM 2018, Dissertation/Demos/Industry), 2018, pp. 51–55.
[54] L. Ackermann, Sprachzentrierte Ansätze zur Steigerung der Akzeptanz von Geschäftsprozessmodellen. Dissertation, Tech. Rep., Universitaet Bayreuth (Germany), 2018.
[55] K. Shahzad, S. Zaheer, R.M. Adeel Nawab, F. Aslam, On comparing manual and automatic generated textual descriptions of business process models, J. Softw.: Evol. Process 31 (11) (2019) e2204.
[56] Q. Zeng, X. Tang, W. Ni, H. Duan, C. Li, N. Xie, Missing procedural texts repairing based on process model and activity description templates, IEEE Access 8 (2020) 12999–13010.
[57] G. Yuan, Q. Zeng, H. Duan, W. Guo, W. Ni, N. Xie, Multi-language description text automatic generation of emergency disposal process, in: 2018 IEEE International Conference of Safety Produce Informatization, IICSPI, IEEE, 2018, pp. 31–35.
[58] G. Yuan, Q. Zeng, W. Ni, C. Liu, C. Li, H. Duan, Multi-view and multi-language description generation for cross-department medical diagnosis processes, IEEE Access 6 (2018) 76741–76753.
[59] J. Cardoso, J. Mendling, G. Neumann, H.A. Reijers, A discourse on complexity of process models, in: International Conference on Business Process Management, Springer, 2006, pp. 117–128.
[60] V. Gruhn, R. Laue, Approaches for business process model complexity metrics, in: Technologies for Business Information Systems, Springer, 2007, pp. 13–24.
[61] R.D. Boomsma, I. Vanderfeesten, D. Fahland, H.A. Reijers, S. Cramer, An evaluation of thresholds for business process model metrics, 2017.
[62] M. La Rosa, P. Wohed, J. Mendling, A.H.M. ter Hofstede, H.A. Reijers, W.M.P. van der Aalst, Managing process model complexity via abstract syntax modifications, IEEE Trans. Ind. Inform. 7 (4) (2011) 614–629.
[63] M. Benner-Wickner, M. Book, T. Brückmann, V. Gruhn, Examining case management demand using event log complexity metrics, in: 18th IEEE International Enterprise Distributed Object Computing Conference Workshops and Demonstrations, EDOC Workshops 2014, Ulm, Germany, September 1-2, 2014, IEEE Computer Society, 2014, pp. 108–115.
[64] J. De Weerdt, M. De Backer, J. Vanthienen, B. Baesens, A multi-dimensional quality assessment of state-of-the-art process discovery algorithms using real-life event logs, Inf. Syst. 37 (7) (2012) 654–676.
[65] P.H.P. Richetti, F.A. Baião, F.M. Santoro, Declarative process mining: Reducing discovered models complexity by pre-processing event logs, in: Business Process Management - 12th International Conference, BPM 2014, Haifa, Israel, September 7-11, 2014. Proceedings, Vol. 8659, Springer, 2014, pp. 400–407.
[66] G. Polančič, B. Cegnar, Complexity metrics for process models–A systematic literature review, Comput. Stand. Interfaces 51 (2017) 104–117.
[67] J. Mendling, H.A. Reijers, W.M.P. van der Aalst, Seven process modeling guidelines (7PMG), Inf. Softw. Technol. 52 (2) (2010) 127–136.
[68] D.M. Blei, Probabilistic topic models, Commun. ACM 55 (4) (2012) 77–84.
[69] G.K. Zipf, Human Behaviour and the Principle of Least Effort: An Introduction to Human Ecology, Addison-Wesley, 1949.
[70] W.M.P. van der Aalst, Process Mining - Data Science in Action, second ed., Springer, 2016.
[71] T. Hærem, B.T. Pentland, K.D. Miller, Task complexity: Extending a core concept, Acad. Manag. Rev. 40 (3) (2015) 446–460.
[72] B.T. Pentland, Conceptualizing and measuring variety in the execution of organizational work processes, Manage. Sci. 49 (7) (2003) 857–870.
[73] J. vom Brocke, M.-S. Baier, T. Schmiedel, K. Stelzl, M. Röglinger, C. Wehking, Context-aware business process management, Bus. Inf. Syst. Eng. 63 (5) (2021) 533–550.
[74] M. Rosemann, J. Recker, C. Flender, Contextualisation of business processes, Int. J. Bus. Process Integr. Manage. 3 (1) (2008) 47–60.
[75] T.L. Saaty, The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making, in: Multiple Criteria Decision Analysis, Springer, 2016, pp. 363–419.
[76] A. Revina, Ü. Aksu, V.G. Meister, Method to address complexity in organizations based on a comprehensive overview, Information 12 (10) (2021) 423.
[77] N. Rizun, Y. Taranenko, Simulation models of human decision-making processes, Manage. Dyn. Knowl. Econ. 2 (2) (2014) 241–264.
[78] W. Daelemans, Explanation in computational stylometry, in: International Conference on Intelligent Text Processing and Computational Linguistics, Springer, 2013, pp. 451–462.
[79] G.K. Zipf, Selected studies of the principle of relative frequency in language, 1932.
[80] X. Zhu, A.B. Goldberg, Introduction to semi-supervised learning, Synth. Lect. Artif. Intell. Mach. Learn. 3 (1) (2009) 1–130.
[81] Z.-H. Zhou, Ensemble Methods: Foundations and Algorithms, Chapman and Hall/CRC, 2012.
[82] A. Revina, K. Buza, V.G. Meister, Designing explainable text classification pipelines: Insights from IT ticket complexity prediction case study, in: Interpretable Artificial Intelligence: A Perspective of Granular Computing, Vol. 937, Springer, Cham, 2021, pp. 293–332.
[83] F.J. Gravetter, L.B. Wallnau, Essentials of Statistics for the Behavioral Sciences, tenth ed., Cengage Learning, 2017.
[84] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, second ed., Routledge, 1988.
[85] A. Bolt, M. de Leoni, W.M.P. van der Aalst, Process variant comparison: Using event logs to detect differences in behavior and business rules, Inf. Syst. 74 (2018) 53–66.
[86] B.L. Welch, The generalization of 'student's'problem when several different population variances are involved, Biometrika 34 (1–2) (1947) 28–35.
[87] H.B. Mann, D.R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, Ann. Math. Stat. 18 (1) (1947) 50–60.
[88] W.M.P. van der Aalst, B.F. van Dongen, C.W. Günther, A. Rozinat, H.M.W. Verbeek, T. Weijters, Prom: The process mining toolkit, in: Proceedings of the Business Process Management Demonstration Track (BPMDemos 2009), 2009.
[89] Axelos, ITIL Foundation, Stationery Office Books, 2019, URL https://www.axelos.com/best-practice-solutions/itil.
[90] R.L. Boyd, A. Ashokkumar, S. Seraj, J.W. Pennebaker, The Development and Psychometric Properties of LIWC-22, University of Texas, Austin, TX, 2022, URL https://www.liwc.app/help/psychometrics-manuals.
[91] K. Winter, S. Rinderle-Ma, Detecting constraints and their relations from regulatory documents using nlp techniques, in: OTM Confederated International Conferences "on the Move to Meaningful Internet Systems", Springer, 2018, pp. 261–278.
[92] H.A. López, Challenges in legal process discovery, in: ITBPM@BPM 2021, CEUR, 2021, pp. 68–73.
[93] J.C. de AR Goncalves, F.M. Santoro, F.A. Baião, Business process mining from group stories, in: 2009 13th International Conference on Computer Supported Cooperative Work in Design, IEEE, 2009, pp. 161–166.
[94] E. Bazhenova, F. Zerbato, B. Oliboni, M. Weske, From BPMN process models to DMN decision models, Inf. Syst. 83 (2019) 69–88.

CHAPTER 6

Research Agenda and Future Work

Summary

This chapter sets a goal to put the research of this dissertation into a wider BPM, IS design, and business environment context revealing further possible opportunities and challenges. Three typical BPM logics relating BPs, IT infrastructure, and BP actors are employed to build the research questions for future research beyond the scope of this dissertation and suggest success factors using established theories and methodologies for generating novel research questions. A rigorous interplay of process mining and NLP is regarded as a critical aspect of BPM future research considering the increasing relevance of natural language and developments in NLP.

This chapter is based on the following publication:

Revina, A., "Business process management: integrated data perspective. A framework and research agenda", in: *Proceedings of the 29th International Conference on Information Systems Development: Crossing Boundaries between Development and Operations (DevOps) in Information Systems (ISD2021), Valencia, Spain, September 8-10, 2021*, 2021, ISBN: 978-1-908358-98-1.

Title Business Process Management: Integrated Data Perspective. A Framework and Research Agenda

Authors Aleksandra Revina[1,2]
[1] Technical University of Berlin
[2] Brandenburg University of Applied Sciences

Published in Proceedings of the 29th International Conference on Information Systems Development: Crossing Boundaries between Development and Operations (DevOps) in Information Systems (ISD2021), CORE conference rank 2018: A.

Abstract Business Process Management (BPM) is confronted with rapidly growing data flows of various types. One established way to address the complexity caused by structured log flows produced by Information Systems (IS) is Process Mining (PM). However, in this approach, unstructured natural language data generated by humans remains uncovered. With the significant advances in natural language processing (NLP), we observe the attention of BPM research and practice shifting towards this type of data. In the study, building on the Task Technology Fit Theory and Contingency Theory, we derive a framework that addresses relevant future research questions in the context of integrated process data perspective, including structured logs and unstructured natural language. The proposed framework considers traditional BPM logics and highlights BPM as a socio-technical discipline.

Keywords business process management; natural language; log data; task technology fit; contingency theory

29TH INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS DEVELOPMENT (ISD2021 VALENCIA, SPAIN)

Business Process Management: Integrated Data Perspective. A Framework and Research Agenda

Aleksandra Revina
Brandenburg University of Applied Sciences
Brandenburg an der Havel, Germany *revina@th-brandenburg.de*

Abstract

Business Process Management (BPM) is confronted with rapidly growing data flows of various types. One established way to address the complexity caused by structured log flows produced by Information Systems (IS) is Process Mining (PM). However, in this approach, unstructured natural language data generated by humans remains uncovered. With the significant advances in Natural Language Processing (NLP), we observe the attention of BPM research and practice shifting towards this type of data. In the study, building on the Task Technology Fit Theory and Contingency Theory, we derive a framework that addresses relevant future research questions in the context of integrated process data perspective, including structured logs and unstructured natural language. The proposed framework considers traditional BPM logics and highlights BPM as a socio-technical discipline.

Keywords: Business Process Management, Natural Language, Log Data, Task Technology Fit, Contingency Theory

1. Introduction

Business Process Management (BPM) constantly undergoes various challenges triggered by internal and external factors. The most significant external factors can be digitalization and vigorous environmental changes, such as recent pandemics. Altered work conditions enforce changing communication channels and technologies to support operational and strategic needs. Remote work is becoming more popular, resulting in a substantial increase in data flows of various types. Notably, online communication triggers the massive generation of natural language mainly stored in a textual form.

In the context of BPM, Information Systems (IS) supporting the processes produce the structured data, i.e., logs, which are analyzed with Process Mining (PM) techniques used to discover processes and bottlenecks, check compliance, and propose process improvements [5]. However, unstructured natural language data massively produced during the process execution remains out of scope in this approach. BPM mainly deals with the quantitative analysis of key performance process dimensions of time, cost, or flexibility without considering textual data. While BPM researchers recognize Natural Language Processing (NLP) maturity and the lack of related research, the mainstream efforts are directed to NLP application in business process (BP) modeling support [1].

These efforts are justified by the fact that BP modeling is a rather labor-intensive and tedious task, which becomes even impossible in the current dynamics and leads to BPM project failures [10]. Unquestionably, a BP model remains the primary source for process analysis and improvement. However, recent research shows that a synergetic combination of PM and NLP can provide immediate insights into the actual process execution based on which successful improvement occurs. For example, [27] demonstrate how the event log enrichment with comments resulted in implementing the bots to reduce the delays caused by waiting for information. [19] successfully leverage text analytics and PM to analyze the text from a process perspective, i.e., the flow of activities. This way, the authors minimize human labeling and allow for various process improvements, including training programs. In line with such approaches, emails, comments, and chat communication can be used as an additional information source for process improvement.

In fact, though hard to estimate and verify, it is widely assumed that unstructured text accounts for over 80% of data in organizations [35]. This abundance can open new opportunities also for BPM. Hence, as we observe, some research in the intersection of PM and NLP has started to appear. Yet, it remains unclear how it affects and changes the nature of BPM, especially in a larger context of IS and the organizational environment, i.e., what other improvements or possible challenges can be expected.

In the study, we (i) inquire into the potential of integrated process data perspective, i.e., machine- (logs) and human- (natural language) generated data, and its possible positive effect on the BPM success factors and (ii) pose an overall research question of *how such integrated process data perspective can impact BPM research and practice.* This paper provides a framework and research agenda that can pave the way for innovative future studies and theory building. Moreover, we consider BPM as a socio-technical phenomenon, as we highlight the importance of both machine- and human-generated process data. This represents the growing integration of social and technical artifacts within the processes of today's work environments. Since the socio-technical standpoint implies an adaption of BPM settings, we build a multi-view understanding of BPM. The results provide researchers and managers with a structured overview of the possible research questions and success factors concerning (i) IS design, (ii) BPM as a discipline, and (iii) business environment. As a theoretical framing, we use two theories commonly known in the BPM context: (1) Task Technology Fit (TTF) Theory [25] and (2) Contingency Theory [20].

The rest of the paper is structured as follows. Section 2 provides a critical analysis of central BPM logics of a regular business context. In Section 3, we develop a framework that contributes to a deeper understanding of how the integrated process data perspective can impact BPM discipline, IT/IS, and business environment and highlight possible success factors. In Section 4, we summarize and synthesize our points.

2. Traditional BPM Logics and Their Problematization

The research and practice specify three dominant logics of BPM: (1) business processes (BPs) [13], (2) IT infrastructures to support the BPs [54], and (3) BP actors [12]. These three logics determine how process owners, employees, or other stakeholders analyze and (re-)design BPs.

Whereas a declared necessity for further research into the role of BPM in these contexts exists [50], the BPM dynamics in times of digitalization, rapid technological and economic changes have to be addressed. The questions regarding the suitability of prior logics for BPM theory and practice need to be reconsidered [10]. In the study, we critically review the three logics under the integrated process data perspective. In this context, we also note that any BP is carried out in socio-technical arrangements, i.e., under consideration of both machine- and human-generated data types.

Following Alvesson and Sandberg's problematization methodology for generating novel research questions [8], we identify and problematize assumptions following the mentioned three BPM logics. Using dialectical interrogation, we identify three central problematizations regarding traditional BPM that appear under the socio-technical integrated process data perspective.

2.1. Business Processes

BPs are defined as sequences of clearly understood activities that are to be modeled and remodeled when necessary [3]. The core of BPM consists of modeling work whereby the processes are captured, analyzed, designed, and redesigned, which makes up a prevailing research direction in BPM [37]. It follows that the dominant view in BP logic is modeling. Process-relevant textual data, such as process documentation and descriptions, are also mainly used for modeling support, as in the case of the automatic generation of process models and compliance checks [1].

At the same time, continuous developments bring many difficulties for process owners, consultants, and employees to keep updating BP models for all the processes [10]. Process update requests, in which multiple parties' interests need to be addressed, become difficult due to political, economic, and socio-technical reasons [50]. Hence, according to the new logic forced by fast-paced changes, the processes should be highly adaptive and easily configurable [34]. Although we agree with this perspective, much information and stakeholders make the operationalization of a highly adaptive and configurable process approach challenging. In practice, a high level of modeling abstraction, minimal reliance on process models, and more focus on agile work in terms of so-called "light touch" processes are observed [10]. Considering these two different approaches of rigorous and light touch process modeling, we problematize as follows:

P1: *Under conditions of agility, adaptability, and light touch process modeling, integrated machine- and human-generated data will allow process improvements bypassing constant creation and updates of BP models.*

2.2. IT Infrastructure

IT infrastructure is considered a true enabler of any BP. In the sense of the traditional BPM IT infrastructure logic, infrastructure is to be reengineered and aligned with the goals of the BP it supports [18]. As a rule, such efforts extend to larger IS contexts that connect departments, business units, and various stakeholders [33].

By and large, information processing is more advantageous when data of various types and volumes are gathered from different parts of the organization following BP goals, decreasing redundancy and increasing productivity. It is enabled by the assumption that process capabilities rely on information processing, and IT infrastructure is designed to realize the expected information flow [10]. Hence, the dominant view in the IT infrastructure logic is infrastructural alignment in conformity with BP goals.

While we agree with this view, we cannot neglect the effects of the rapid penetration of new data sources and technologies in the IT infrastructure. According to the traditional IT infrastructure logic, every time the IT infrastructure updates are made, they should be aligned with BP goals and vice versa. However, BP goals, same as models, are rigid to change. One possible way to address this problem found in practice is to build independent middleware solutions responsible for all change requests. For example, in the case of changing information processing requirements, an independent data aggregation platform could be an alternative to the constant updates and realignment of IT infrastructure [10]. This way, somewhat infrastructural flexibility can be provided. Considering these two approaches of infrastructural alignment and flexibility, we problematize as follows:

P2: *Under conditions that integrated machine- and human-generated data are becoming a relevant part of BPs, IT infrastructures should be naturally designed to fulfill these conditions.*

2.3. Process Actors

The first processes in the industrialization and factory automation context were designed under the logic that the work steps follow one another [16]. Accordingly, the traditional BP actors' logic relies on the procedural assumption that the actors should strictly follow the process steps in conformity with the rules. This assumption found mutual acceptance in BPM as it facilitates the creation of understandable models [12], making it easy to demonstrate process improvements [30]. In this respect, the provision of exact guidelines and rules on the BP execution is essential.

However, economic and environmental challenges considerably impact the roles and tasks, demanding changes according to the varying work conditions [66]. Dealing with existing employees, hiring new ones, outsourcing activities, employee turnover – all these factors demand certain flexibility in the process roles definition [14]. Moreover, other factors discussed in Subsection 2.1 that complicate keeping process models up-to-date make the procedural assumption of actors' logic problematic. As a response to these challenges, BPM researchers and practitioners' focus shifts towards so-called "mindful" actors capable of improvisations and adjustments to the changing work conditions and making their own decisions based on the collected information [10]. Considering the assumptions of procedural and mindful actors, we problematize as follows:

P3: *Under conditions of non-availability of up-to-date process models and process instructions, it is important to enhance actors' competencies and abilities with the process execution and decision support based on integrated machine- and human-generated data.*

To sum up, we discussed problematizations of the three BPM logics highlighting the integrated data perspective. Accordingly, both machine- and human-generated data play an important role in successfully adapting to the changing work conditions triggered by dynamic developments.

3. Framework Exploring Integrated Data Perspective in BPM

This section aims to build a framework to deeper understand the importance of the integrated data in the modern BPM, its elements, and associated success factors. In the latter, we aim to outline the expected benefits which the framework potentially enables. As a theoretical framing, we use the TTF and Contingency Theories.

The Contingency Theory affirms that organizational performance depends on the fit between an organization and contingencies such as technology, innovation, environmental change, size, culture [24]. In BPM, the Contingency Theory plays a significant role in shifting from the justification of BPM practices' value to understanding the contextual conditions under which they are effective [61].

The Contingency Theory gave rise to the principle of fit in IS, which has become increasingly important in assessing a technology's performance. As a result, the TTF Theory and model were introduced to emphasize the importance of a fit between technology and the individual task to maximize individual achievement. Some IS studies used both the Contingency and TTF Theories in the IS adoption, implying that a task and technology fit is important for IS performance [26].

Similarly, in BPM, the Contingency and TTF Theories are often considered together to reflect the fit between BPs and (i) business environment and (ii) technology correspondingly [65]. TTF concepts initially reflect that positive individual performance is facilitated by the IT capabilities matched to the user tasks [26]. Extended to organizational context, TTF posits that positive organizational performance is facilitated by the IT capabilities matched to the BPs [33].

These two theories outline the main three elements in our framework, i.e., BPM, IT/IS, and business environment (see Fig. 1). Further, in each of the elements, we identify the relevant constituents. To understand the integrated data perspective in more detail, we propose considering the three framework elements and their constituents in the context of respective research questions.

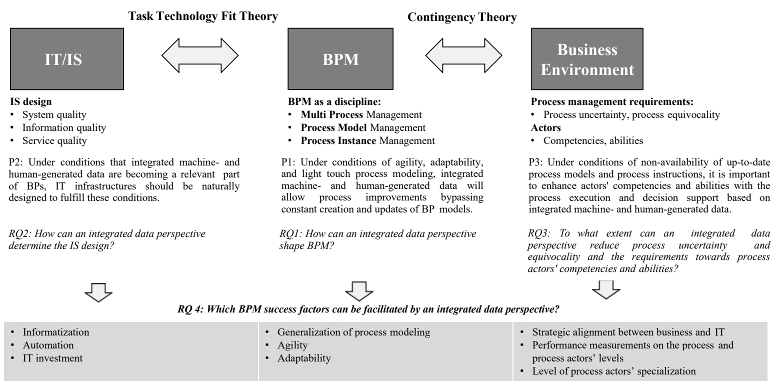


Fig. 1. Framework and research agenda of integrated process data perspective in BPM

First, we define the constituents of the BPM discipline. BPM encompasses a wide range of managerial activities related to BPs. Typically, one differentiates between three layers, i.e., Multi Process Management, Process Model Management, and Process Instance Management [46]. The first layer of Multi Process Management is associated with identifying the organization's major processes and their prioritization. The second layer of Process Model Management deals with managing a single process in a traditional BPM lifecycle, including process discovery, analysis, redesign, implementation, and controlling. The third layer of Process Instance Management is concerned about single enactments of a process, i.e., planning tasks of a process, executing, monitoring, and potentially adapting them to the current setting [1]. All three layers make use of BP models and log data. Hence, this observation proves the gap we focus on in this study: (i) the attention of BPM researchers and practitioners is directed towards structured machine-generated data and process modeling and (ii) the lack of integrated machine- and human-generated data perspective. In this context, we refer to P1 and formulate the first research question (RQ):

RQ1: *How can an integrated data perspective shape BPM?*

Second, we define the most important constituents of IT/IS. Due to the new challenges, experts are working to improve the functionality and quality of solutions [67]. The quality in IS design has been investigated as the function of the second-order structures, such as system, information, and service qualities [32]. System quality is described as the degree to which users believe that systems are enjoyable and straightforward to use, connect, and learn [52]. Information quality refers to the degree to which users view information as reliable, detailed, timely, structured, and up-to-date [28]. Service quality is described with reliability, assurance, tangibility, responsiveness, interactivity, empathy, and functionality [43]. In the study context, two viewpoints are noteworthy: (1) IS design should enable the capture of both machine- and human- generated data types, (2) IS design quality constituents should be partially obtainable (measurable) from the data captured in the IS usage. Hence, we refer to P2 and formulate the second RQ:

RQ2: *How can an integrated data perspective determine the IS design?*

Third, we define the most important constituents concerning a business environment. As a rule, processes reflect the environment where they are performed [44]. Hence, in the study, we summarize the environment's properties with the two process management requirements according to the Organizational Information Processing Theory [23]: (1) process uncertainty determined by the amount of information required to perform a process and (2) process equivocality defined by the ambiguity of interpretations, i.e., information quality [69]. These requirements differentiate one process from another and describe the contingencies typical for a given environment. In addition to process characteristics, people, i.e., process actors, make up another important constituent of a business environment. Hereby, we consider such actors' characteristics as competencies or abilities that differentiate one actor from another and lay the ground for successful performance [11]. Accordingly, we refer to P3 and formulate the third RQ:

RQ3: *To what extent can an integrated data perspective reduce process uncertainty and equivocality and the requirements towards process actors' competencies and abilities?*

As follows from Fig. 1, the fourth research question is devoted to the BPM success factors. In the study, we suggest revisiting this significant yet understudied topic in BPM to highlight and specify the framework's benefits. Today, organizations often use the term BPM as a trendy established concept rather than implement it. Not only is BPM resource-intensive, but it also reveals a large number of failed projects and programs [65]. In this respect, the knowledge on (critical) success factors gains importance. For BPM, somewhat similar and generic success factors are suggested, such as top management support, project management, communication, cooperation, and end-user training [33]. Success factors studied in the IS context, such as leadership and investment, are also applicable to BPM [45]. Here similarly, we detect a strong influence of the IT/IS on the BPM development. Hence, in the study, we urge to reconsider the BPM success factors under an integrated data perspective and formulate the fourth RQ:

RQ4: *Which BPM success factors can be facilitated by an integrated data perspective?*

Using existing established literature, we attempt to suggest and map the selected exemplary success factors according to the three framework elements (see Fig. 1).

In the following subsections addressing each of the RQs, we discuss the theoretical and practical research developments indicating potential research directions, implementation possibilities, and challenges.

3.1. Foundations for RQ1

In times of fast-paced changes of various nature, agility and adaptability become one of the main requirements for successful organizational functioning [31]. Under these conditions, BPM sets out to consider human-generated natural language to support the modeling activities. Referring to the already mentioned three layers of BPM, there is solid prior research in respect to (i) Multi Process Management, such as identifying the similarity [17], matching [68] and merging of process models [60], textual-based [39] and semantic [64] search and (ii) Process Model Management, such as a transformation of textual descriptions into process models [22] and vice versa [42], text annotations [62], multiple languages, semantic quality check [41], and compliance check [2].

However, all these cases are closely related to texts describing BP models or textual data inherent in BP models, such as labels. The research works considering an integrated perspective of structured logs and natural language generated in the processes have recently started to appear, suggesting process analysis and improvement approaches that bypass the laborious and resource-intensive process modeling work, see [19], [27] presented in the introduction section.

From an implementation viewpoint, further exploitation of PM algorithms' enrichment with various text analytics techniques can open new process improvement opportunities. Hereby, word ambiguity and semantics remain the main challenges, like (i) a correct match of the pronouns to the related nouns or several different expressions in text meaning the same thing and (ii) identification and correct match of activities' concurrency, iteration, and decision points described in text [1].

To sum up, it is essential to follow such research endeavors and discover the implications of how and to what extent an integrated data perspective can contribute to agility and adaptability in BPM.

3.2. Foundations for RQ2

The Big Data boom with economic and social transactions going online has generated new demands. The ability to understand the structure and content of the human speech has significantly increased the dimensionality of the available data sets. Such challenges lead to research on better IS design [6].

IS discipline relies on two complementary but different sciences: (1) behavioral science related to human or organizational behavior and (2) design science related to innovative IS artifacts [29]. Regarding the investigation of human behavior with respect to behavioral science, much research exists to analyze the synergies of IS design and natural language, i.e., NLP, in general [47] and in addressing specific challenges, such as large-scale NLP [7], ethical design for NLP [36], building NLP pipelines [55, 56]. In BPM, one of the suggested research directions is the design of conversational systems based on semantic understanding, context resolution, and language generation that would guide process actors through the process [1].

Regarding the IS design and IT artifacts in BPM with respect to design science, one popular approach is the so-called Process-Aware IS (PAIS), executing operational processes based on the process models [4]. Additionally, much research is done on the specific IS design questions based on Process Mining (PM), for example, privacy and system design in the Internet of Things (IoT) [48], PM-enabled decision support systems (DSS) [21], PM-enabled design issues in healthcare [57], to name a few.

As can be concluded from above, IS design implementation technologies and challenges reveal both well-known generic and IT artifact- and domain-specific trends. For example, using natural language as a direct input for conceptual design, finding an appropriate mapping between sentences and objects in the database and operations [47], scalability of text processing and architectural challenges [7], the societal impact of data collection and usage [36], algorithmic and text representation choices in building NLP pipelines [55, 56] are well-known generic challenges related to NLP in the IS context. In the design of specific IT artifacts, we observe such challenges as user-friendly design, seamless navigation and viewing of BP models in PAIS [4], privacy concerns in IoT [48], implementing concrete task-level support in DSS supporting BPM projects [21], adapting the generic approaches to specific domains, like healthcare [57].

Hence, in the unified view of IS artifact- and NLP-related challenges, IS solutions which (i) gather, store, and re-use both data types ensuring their quality, (ii) address IS architectural challenges ensuring system quality and this way (iii) increase service quality are relevant research directions.

3.3. Foundations for RQ3

Due to a rather broad definition of a business environment, we limit the scope to the process perspective and consider the business environment through the lens of the process management requirements, i.e., process uncertainty and equivocality and process actors' competencies and abilities. Increasing the amount of information with an integrated data perspective may help decrease the uncertainty but is ineffective in high equivocality cases. Here, it is essential to assist process actors with various interpretation solutions [15].

Process uncertainty and equivocality have been widely addressed in PM [49], [51]. NLP techniques have also been used to solve these problems. However, the usage is mainly limited to process models, i.e., already mentioned process model semantic quality [41] and compliance check [2].

Concerning process actors, the PM approaches have recently started to explore the "social" aspects of the processes. For example, [63] examine PM from a socio-technical perspective by conceptualizing a PM-based approach to improve work conditions. Furthermore, mentioned in Section 3.1. [19] and [27] demonstrate synergetic usage of PM and NLP in the IT ticket processing scenario and customer service of a financial software company. Another new direction in supporting process actors with PM techniques is Robotic Process Mining aiming to detect those routine tasks that should be automated [38]. At the same time, NLP has been widely implemented in business to enrich process actors' abilities in dealing with large amounts of textual data [9].

The implementation possibilities and challenges of RQ3 are closely related to and can be derived from those of RQ1 and RQ2, for example, semantic considerations in automatic extraction of meaning from textual data [1], user-centricity [4], privacy [48], and ethics [36] in IS design.

To sum up, an integrated data perspective naturally enhances the amount of information helping to reduce process uncertainty. Various PM- and NLP-based solutions processing, interpreting, and making this information useful are essential to deal with process equivocality. Moreover, an integrated data can empower the research on the "social" part of PM, better addressing process actors' needs and helping them adapt to the dynamic transformations in BPs.

3.4. Foundations for RQ4

Following the declared need for the research on success factors in BPM caused by a large number of BPM failures [65], we suggest that an integrated data perspective can be advantageous for realizing success factors. This helps us to outline related benefits. See Fig. 1 for our proposal exemplarily highlighting, but not limiting to, the success factors in each of the three framework elements of BPM, IT/IS, and business environment. Below, we discuss the success factors starting with the BPM discipline.

In BPM, the level of details with which processes should be modeled, or generalization of process modeling, is a well-known problem [53]. An integrated data perspective contributes to this first success factor allowing for light touch BP models. This view underpins the idea that processes can be modeled and organized on a high abstraction level to support agility and adaptability. This approach will reduce the human efforts needed for creating and updating the BP models, enable process analysis and improvements bypassing the laborious modeling activities, and eventually decrease BPM project failures.

In the second framework element, IT/IS, informatization is an established success factor to ensure adequate support from IT/IS [65]. We propose that an integrated data perspective will positively influence informatization while providing various information sources. Informatization is closely related to another success factor, automation, i.e., the use of IT to support or replace process actors in the execution of, as a rule, routine tasks [65]. Here, both data types, logs [38] and natural language [40], [59], appear to be useful in identifying task automation candidates. Finally, we believe that the benefits derived from an integrated process data usage in BPM can provide arguments and justify the third important success factor, i.e., IT investment. Hence, with respect to IT/IS success factors, the consideration of both data types increases the informatization level of the processes providing (more) data sources for adequate IT/IS support, automation projects, and IT investment acquisition.

In the third framework element, business environment, strategic alignment is considered one of the essential factors for reaching the long-term success of BPM programs [65]. Whereas the term strategic alignment includes many constituents, we propose that an integrated data perspective can facilitate the alignment of business and IT. Log data naturally contain information about IT functioning. The business information is likely to be captured in an unstructured textual form, as the most typical data type in organizations [35], [58], and can be used to gain valuable insights, as suggested in [59]. Similarly, an integrated view can be favorable for measuring the performance on the process and process actors' levels [60], [65]. Lastly, we propose that the trade-off between specialists and generalists can be better addressed through enhanced decision-making enabled by integrated data. Thus, the mentioned success factors are beneficial for organizational and employees' performance increase.

For future research, we suggest exploring in-depth the suggested success factors (and beyond) and the role of integrated data, specifically in case study settings.

4. Conclusions

The introduced framework and research agenda demonstrate the multi-view of BPM-related elements structured and justified based on the TTF and Contingency Theories. We focus on the three main elements: BPM as a discipline, IT/IS, and business environment, and their constituents. Next, we problematize the traditional BPM logics regarding BPs, IT infrastructure, and process actors and derive the RQs. Finally, we enrich each framework element with the success factors that an integrated data perspective can facilitate.

We suggest the framework as a guideline for prospective qualitative and quantitative studies based on a structured view of the RQs to investigate the potential of the machine- and human-generated process data in BPM. The framework implies elementary yet future challenges and opportunities. We emphasize the importance of the "social", i.e., "humanistic", aspects of BPM in times of increasing digitalization and convergence of organizational structure and digital infrastructure. Recognizing BPM as a socio-technical phenomenon, we aim to understand its synergetic effects on BPM as a discipline, IT/IS, and business environment expressed by the specified constituents. We consider the interweaving of PM and NLP as an essential determinant of BPM future research under the growing importance of natural language and NLP maturity.

To deeply exploit the potential of the integrated data perspective, we resume the discussion on BPM success factors. We set future research directions to understand how increasing integration of machine- and human-generated data can positively impact the establishment of BPM success factors.

For future research, through focused discussions, we suggest (i) critically reviewing the traditional BPM logics and studying (ii) how an integrated data perspective evolves in BPM in general and in particular, i.e., according to the proposed three framework elements and their constituents, and (iii) which benefits for the realization of success factors can be derived.

References

1. van der Aa, H., Carmona, J., Leopold, H., Mendling, J., Padró, L.: Challenges and Opportunities of Applying Natural Language Processing in Business Process Management. In: The 27th International Conference on Computational Linguistics. pp. 2791–2801. Association for Computational Linguistics, Santa Fe (2018)
2. van der Aa, H., Leopold, H., Reijers, H.A.: Comparing textual descriptions to process models – The automatic detection of inconsistencies. Information Systems. 64 447–460 (2017)
3. van der Aalst, W.M.P.: Business Process Management: A Comprehensive Survey. ISRN Software Engineering. 1–37 (2013)
4. van der Aalst, W.M.P.: Process-aware information systems: Lessons to be learned from process mining. In: LNCS. pp. 1–26. Springer, Berlin (2009)
5. van der Aalst, W.M.P.: Process Mining. Data Science in Action. Springer, Berlin (2016)
6. Agarwal, R., Dhar, V.: Big data, data science, and analytics: The opportunity and challenge for IS research. Information Systems Research. 25 (3), 443–448 (2014)
7. Agerri, R., Artola, X., Beloki, Z., Rigau, G., Soroa, A.: Big data for Natural Language Processing: A streaming approach. Knowledge-Based Systems. 79 36–42 (2015)
8. Alvesson, M., Sandberg, J.: Generating research questions through problematization. Academy of Management Review. 36 (2), 247–271 (2011)
9. Bahja, M.: Natural Language Processing Applications in Business. In: E-Business. IntechOpen (2020)
10. Baiyere, A., Salmela, H., Tapanainen, T.: Digital transformation and the new logics of business process management. European Journal of Information Systems. 29 (3), 238–259 (2020)
11. Boyatzis, R.E.: Managerial and Leadership Competencies. A Behavioral Approach to Emotional, Social and Cognitive Intelligence. Vision: The Journal of Business Perspective. 15 (2), 91–100 (2011)
12. vom Brocke, J., Rosemann, M.: Handbook on Business Process Management 2: Strategic Alignment, Governance, People and Culture. Springer, Berlin (2010)
13. vom Brocke, J., Schmiedel, T., Recker, J., Trkman, P., Mertens, W., Viaene, S.: Ten principles of good business process management. Business Process Management Journal. 20 (4), 530–548 (2014)
14. Cantoni, F., Mangia, G.: Human Resource Management and Digitalization. Routledge (2018)
15. Cooper, R.B., Wolfe, R.A.: Information Processing Model of Information Technology Adaptation: An intra-organizational diffusion perspective. Data Base for Advances in Information Systems. 36 (1), 30–48 (2005)
16. Datta, A.: Automating the Discovery of AS-IS Business Process Models: Probabilistic and Algorithmic Approaches. Information Systems Research. 9 (3), 275–301 (1998)
17. Dijkman, R., Dumas, M., Van Dongen, B., Krik, R., Mendling, J.: Similarity of business process models: Metrics and evaluation. Information Systems. 36 (2), 498–516 (2011)
18. Dumas, M., ter Hofstede, A.H., van der Aalst, W.M.: Process-Aware Information Systems: Bridging People and Software through Process Technology. John Wiley & Sons, New York (2005)
19. Fan, S., Ilk, N.: A text analytics framework for automated communication pattern analysis. Information and Management. 57 (4), 103219 (2020)
20. Fiedler, F.E.: A Contingency Model of Leadership Effectiveness. Advances in Experimental Social Psychology. 1 149–190 (1964)
21. Fleig, C.: Towards the Design of a Process Mining-Enabled Decision Support System for Business Process Transformation. In: Proceedings of the Forum and Doctoral Consortium Papers Presented at the 29th International Conference on Advanced Information Systems Engineering. pp. 170–178. CEUR, Essen (2017)
22. Friedrich, F., Mendling, J., Puhlmann, F.: Process model generation from natural language text. In: Mouratidis, H. and Rolland, C. (eds.) Advanced Information Systems Engineering. CAiSE 2011. Lecture Notes in Computer Science. pp. 482–496. Springer, Berlin (2011)
23. Galbraith, J.R.: Designing complex organizations. Addison-Wesley Pub. Co. (1973)
24. Gangwar, H.: Big Data Analytics Usage and Business Performance: Integrating the Technology Acceptance Model (TAM) and Task Technology Fit (TTF) Model. Electronic Journal of Information Systems Evaluation. 23 (1), 45–64 (2020)

25. Goodhue, D.L.: Development and measurement validity of a task-technology fit instrument for user evaluations of information systems. Decision Sciences. 29 (1), 105–138 (1998)
26. Goodhue, D.L., Thompson, R.L.: Task-technology fit and individual performance. MIS Quarterly: Management Information Systems. 19 (2), 213–233 (1995)
27. Gupta, M., Agarwal, P., Tater, T., Dechu, S., Serebrenik, A.: Analyzing Comments in Ticket Resolution to Capture Underlying Process Interactions. In: Business Process Management Workshops. BPM 2020. pp. 219–231. Springer, Berlin (2020)
28. Halonen, R., Acton, T., Golden, W., Conboy, K.: DeLone & McLean success model as a descriptive tool in evaluating the use of a virtual learning environment. In: International conference on organizational learning, knowledge and capabilities. pp. 1–16. NUI Galway, Amsterdam (2009)
29. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. 28 (1), 75–105 (2004)
30. Hung, R.Y.Y.: Business Process Management as competitive advantage: A review and empirical study. Total Quality Management and Business Excellence. 17 (1), 21–40 (2006)
31. Imgrund, F., Fischer, M., Winkelmann, A.: Approaching Digitalization with Business Process Management. In: Drews, P., Funk, B., Niemeyer, P., and Xie, L. (eds.) Multikonferenz Wirtschaftsinformatik. pp. 1725–1736. Leuphana Universität Lüneburg, Lüneburg (2018)
32. Isaac, O., Abdullah, Z., Ramayah, T., Mutahar, A.M.: Examining the Relationship Between Overall Quality, User Satisfaction and Internet Usage: An Integrated Individual, Technological, Organizational and Social Perspective. Asian Journal of Information Technology. 16 (1), 100–124 (2017)
33. Karimi, J., Somers, T.M., Bhattacherjee, A.: The impact of ERP implementation on business process outcomes: A factor-based study. Journal of Management Information Systems. 24 (1), 101–134 (2007)
34. Kir, H., Erdogan, N.: A knowledge-intensive adaptive business process management framework. Information Systems. 95 101639 (2021)
35. Kobayashi, V.B., Mol, S.T., Berkers, H.A., Kismihók, G., Den Hartog, D.N.: Text Mining in Organizational Research. Organizational research methods. 21 (3), 733–765 (2018)
36. Leidner, J.L., Plachouras, V.: Ethical by Design: Ethics Best Practices for Natural Language Processing. In: Proceedings of the 1st ACL Workshop on Ethics in Natural Language Processing. pp. 30–40. Association for Computational Linguistics, Stroudsburg (2017)
37. Leno, V., Dumas, M., Maggi, F.M., La Rosa, M., Polyvyanyy, A.: Automated discovery of declarative process models with correlated data conditions. Information Systems. 89 101482 (2020)
38. Leno, V., Polyvyanyy, A., Dumas, M., La, M., Fabrizio, R., Maggi, M.: Robotic Process Mining: Vision and Challenges. Business & Information Systems Engineering. (2020)
39. Leopold, H., van der Aa, H., Pittke, F., Raffel, M., Mendling, J., Reijers, H.A.: Searching textual and model-based process descriptions based on a unified data format. Software and Systems Modeling. 18 (2), 1179–1194 (2019)
40. Leopold, H., van der Aa, H., Reijers, H.A.: Identifying Candidate Tasks for Robotic Process Automation in Textual Process Descriptions. In: Gulden, J., Reinhartz-Berger, I., Schmidt, R., Guerreiro, S., Guédria, W., and Bera, P. (eds.) Enterprise, Business-Process and Information Systems Modeling. BPMDS 2018, EMMSAD 2018. Lecture Notes in Business Information Processing. pp. 67–81. Springer, Cham, Tallin (2018)
41. Leopold, H., Eid-Sabbagh, R.H., Mendling, J., Azevedo, L.G., Baião, F.A.: Detection of naming convention violations in process models for different languages. Decision Support Systems. 56 (1), 310–325 (2013)
42. Leopold, H., Mendling, J., Polyvyanyy, A.: Supporting process model validation through natural language generation. IEEE Transactions on Software Engineering. 40 (8), 818–840 (2014)
43. Lin, F., Fofanah, S.S., Liang, D.: Assessing citizen adoption of e-Government initiatives in Gambia: A validation of the technology acceptance model in information systems success. Government Information Quarterly. 28 271–279 (2011)
44. Lindsay, A., Downs, D., Lunn, K.: Business processes-attempts to find a definition. Information and Software Technology. 45 1015–1019 (2003)
45. Lu, X.-H., Huang, L.-H., Heng, M.S.H.: Critical success factors of inter-organizational information systems: a case study of Cisco and Xiao Tong in China. Information and Management. 43 (3), 395–408 (2006)
46. Mendling, J., Baesens, B., Bernstein, A., Fellmann, M.: Challenges of smart business process management: An introduction to the special issue. Decision Support Systems. 100 1–5 (2017)
47. Métais, E.: Enhancing information systems management with natural language processing

techniques. Data and Knowledge Engineering. 41 (2–3), 247–272 (2002)
48. Michael, J., Koschmider, A., Mannhardt, F., Baracaldo, N., Rumpe, B.: User-centered and privacy-driven process mining system design for IoT. In: Cappiello, C. and Ruiz, M. (eds.) Information Systems Engineering in Responsible Information Systems. CAiSE 2019. Lecture Notes in Business Information Processing. pp. 194–206. Springer, Cham (2019)
49. Moreira, C., Haven, E., Sozzo, S., Wichert, A.: Process mining with real world financial loan applications: Improving inference on incomplete event logs. PLOS ONE. 13 (12), e0207806 (2018)
50. Müller, S.D., Mathiassen, L., Saunders, C.S., Kræmmergaard, P.: Political maneuvering during business process transformation: A pluralist approach. Journal of the Association for Information Systems. 18 (3), 173–205 (2017)
51. Pegoraro, M., Van Der Aalst, W.M.P.: Mining uncertain event data in process mining. In: International Conference on Process Mining, ICPM 2019. pp. 89–96. IEEE, Aachen (2019)
52. Petter, S., McLean, E.R.: A meta-analytic assessment of the DeLone and McLean IS success model: An examination of IS success at the individual level. Information and Management. 46 (3), 159–166 (2009)
53. Polyvyanyy, A., Smirnov, S., Weske, M.: Business process model abstraction. In: vom Brocke, J. and Rosemann, M. (eds.) Handbook on Business Process Management 1: Introduction, Methods, and Information Systems. pp. 147–165. Springer, Berlin (2015)
54. Recker, J.: Suggestions for the next wave of BPM research: Strengthening the theoretical core and exploring the protective belt. Journal of Information Technology Theory and Application (JITTA). 15 (2), 5–20 (2014)
55. Revina, A., Buza, K., Meister, V.G.: Designing Explainable Text Classification Pipelines: Insights from IT Ticket Complexity Prediction Case Study. In: Pedrycz, W. and Chen, S.-M. (eds.) Interpretable Artificial Intelligence: A Perspective of Granular Computing. pp. 293–332. Springer, Cham (2021)
56. Revina, A., Buza, K., Meister, V.G.: IT Ticket Classification: The Simpler, the Better. IEEE Access. 8 193380–193395 (2020)
57. Rismanchian, F., Lee, Y.H.: Process Mining–Based Method of Designing and Optimizing the Layouts of Emergency Departments in Hospitals. Health Environments Research and Design Journal. 10 (4), 105–120 (2017)
58. Rizkallah, J.: Council Post: The Big (Unstructured) Data Problem, https://www.forbes.com/sites/forbestechcouncil/2017/06/05/the-big-unstructured-data-problem/#7eb81f84493a, Accessed: September 27, 2020, (2017)
59. Rizun, N., Revina, A., Meister, V.G.: Analyzing content of tasks in Business Process Management. Blending task execution and organization perspectives. Computers in Industry. 130 103463 (2021)
60. La Rosa, M., Dumas, M., Uba, R., Dijkman, R.: Business process model merging: An approach to business process consolidation. ACM Transactions on Software Engineering and Methodology. 22 (2), 1–42 (2013)
61. Sousa, R., Voss, C.A.: Contingency research in operations management practices. Journal of Operations Management. 26 (6), 697–713 (2008)
62. Stenetorp, P., Pyysalo, S., Topić, G., Ohta, T., Ananiadou, S., Tsujii, J.: BRAT: a Web-based Tool for NLP-Assisted Text Annotation. In: Proceedings of the Demonstrations Session at EACL 2012. pp. 102–107. Association for Computational Linguistics, Avignon (2012)
63. Tang, W., Matzner, M.: Creating Humanistic Value with Process Mining for Improving Work Conditions - A Sociotechnical Perspective. In: Forty-First International Conference on Information Systems. ICIS 2020. pp. 1–9. Association for Information Systems, India (2020)
64. Thomas, O., Fellmann, M.: Semantic Process Modeling – Design and Implementation of an Ontology-based Representation of Business Processes. Business & Information Systems Engineering. 1 (6), 438–451 (2011)
65. Trkman, P.: The critical success factors of business process management. International Journal of Information Management. 30 (2), 125–134 (2010)
66. Utesheva, A., Simpson, J.R., Cecez-Kecmanovic, D.: Identity metamorphoses in digital disruption: A relational theory of identity. European Journal of Information Systems. 25 (4), 344–363 (2016)
67. Wang, W.T., Lai, Y.J.: Examining the adoption of KMS in organizations from an integrated perspective of technology, individual, and organization. Computers in Human Behavior. 38 55–67 (2014)
68. Weidlich, M., Dijkman, R., Mendling, J.: The ICoP framework: Identification of correspondences between process models. In: Pernici, B. (ed.) Advanced Information Systems Engineering. CAiSE 2010. Lecture Notes in Computer Science. pp. 483–498. Springer, Berlin

(2010)
69. Zelt, S., Recker, J., Schmiedel, T., vom Brocke, J.: A theory of contingent business process management. Business Process Management Journal. 25 (6), 1291–1316 (2019)

CHAPTER 7

Conclusion

Summary

This chapter finalizes the research work presented in this dissertation, whereby the framework on data-driven BP complexity analysis for decision-making support is developed. The ultimate purpose of the framework is to give insights and draw the attention of the BPM community to the potential of textual data produced in large volumes by BP participants in the BP executions. To this end, the framework combines two data types in one synergistic approach aiming to address the issues introduced in Chapter 1 and enhance the BPM toolset with such instruments for comprehensive analysis. First, the research results and contributions are summarized, and their validation is explained with critical reflections on the research questions introduced in Section 1.2.2. Afterward, implications of the research results and contributions for research and practice are provided, and some of the unresolved challenges and limitations are discussed. Finally, future study directions are suggested.

7.1 Contributions and Validation

The framework on data-driven BP complexity analysis for decision-making support is gradually developed and presented in this dissertation. To determine BP complexity, the framework employs two data types: event log and textual data produced in BP execution. Herewith, it provides a synergistic solution to the three main issues of incomprehensiveness, inefficiency, and bias observed in current BP analysis approaches. With the gradual development of the framework, contributions of this dissertation are derived. The contributions and how they address the research questions specified in the introduction chapter 1 are explained below.

- Framework on data-driven BP complexity analysis for decision-making support comprises the core contribution of the dissertation and provides an answer to the main research question, i.e., **MRQ** (*How can process-related data be combined in a comprehensive, efficient, and objective way to predict BP complexity and provide decision-making support to BP actors?*).
  The framework is developed gradually, each of the stages addressing particular issues of current BP analysis approaches. Specifically, a comprehensive literature review, the framework itself, as well as a research agenda deal with the incomprehensiveness issue. A detailed consideration of textual data and linguistic feature sets addresses the inefficiency and bias issues. Further, the bias issue is also tackled in the framework itself.

- Comprehensive literature analysis on complexity in organizations provides an answer to **RQ1** (*What concepts and measures of complexity are available in the literature?*).
  Before developing the framework, it was important to create a comprehensive overview and analyze existing complexity approaches in organizations. Respectively, a morphological box is introduced and framed with a multi-dimensional complexity basis to integrate the findings of the conducted literature analysis. Further, using the developed overviews of complexity concepts and measures, a method to comprehensively address complexity in organizations is established. Adapting Goal Question Metric, we have introduced step-by-step guidance based on the question, measurement, and data levels leading from a given complexity-related problem statement to an appropriate solution. Herewith, we aim to assist organizations in identifying what is actually needed to deal with complexity.

- Textual data-driven BP complexity analysis sets a goal to answer **RQ2** (*How can a novel BP complexity concept be defined for efficiently measuring BP complexity based on textual data serving as an input to BP?*).
  To gain valuable knowledge in the form of the BP complexity concept, typical NLP techniques are applied to BP textual data to extract objective, subjective, and meta-knowledge aspects encompassing semantics, syntax, and stylistics. Respectively, the knowledge regarding cognitive, attention, and reading efforts is

gained, which forms the textual data-based BP complexity.

- The operationalization of the textual data-driven BP complexity concept answers **RQ3** (*How can the BP complexity concept introduced in the answer to RQ2 be employed objectively and efficiently to predict BP complexity?*).
  To operationalize the BP complexity concept, we have experimented with the linguistic features elaborated as an answer to RQ2 and various machine learning algorithms. Subsequently, a machine learning approach is developed for BP complexity prediction and illustrated in a real-life IT ticket classification setting. Moreover, the benefits of the suggested linguistic features in predicting BP complexity are evaluated by comparison with the widely recognized text analysis technique LIWC.

- Textual and event log data-driven BP complexity analysis gives an answer to **RQ4** (*How can the BP complexity concept introduced in the answer to RQ2 be enhanced with BP execution data, i.e., event log, for a comprehensive and objective BP complexity analysis?*).
  Accordingly, a new approach for measuring BP complexity by combining textual data and event log is suggested. The goal of this approach is to investigate the relationship between the complexity obtained from these two data types. The approach leverages a set of linguistic features for textual data-based complexity developed in **RQ2** and fine-tuned in **RQ3**. Using these features, textual data-based complexity is predicted. To calculate event log-based complexity, state-of-the-art event log complexity metrics are employed. Afterward, a correlation analysis of two complexities is performed, followed by a significant difference analysis in correlations. The findings are used to generate recommendations and insights for redesigning and improving BPs in a synergistic and comprehensive way.

- The research agenda and future work presented in the final part of this dissertation provide an answer to the last research question, i.e., **RQ5** (*How can such an integrated process data perspective addressed in RQ4 impact BPM research and practice?*).
  Building on the Task Technology Fit Theory, Contingency Theory, and traditional BPM logics, we derive the research agenda that addresses relevant future research questions and success factors in the context of integrated process data perspective, including event log and textual data.

When performing any research, validity is an important consideration. As this study contains empirical work, we discuss several threats to validity and our tactics on how to address them. Originating from the software engineering field, threats to validity are defined as those issues limiting the ability to interpret and draw conclusions from research results [87]. Our interpretation of this term is related to the rigor behind the design artifact we introduce and the decision-making support solutions we suggest. We refer to the four criteria for validity outlined by [88], namely construct validity, internal validity, external validity, and reliability.

Construct validity reflects the problem of accordance between measured variables and intended meanings of the theoretical terms. To address this validity, we make use of sound theories and established concepts to underpin the work performed in this dissertation. Thus, in [57] presented in Chapter 2, while creating a comprehensive overview of the literature, we build on the well-known People Process Technology (PPT) framework to streamline our work. Additionally, in the same study, a Goal Question Metric is extended to develop the guidelines for applying the study findings to address complexity in organizations. In Chapter 3 and related publication [12], the linguistic foundations by [54] and the Theory of Situation Awareness are used to empirically develop the textual data-based concept of BP complexity. Further, the concluding study of this dissertation [58] bases on the Task Technology Fit Theory and Contingency Theory to derive a research agenda.

Internal validity is about looking into causal relations. When it is studied if one factor influences another, there is a possibility of a third factor influencing both of them. If a researcher is unaware of this third factor and/or does not know how far it impacts the studied factor, internal validity is put at risk [89]. Hence, special attention should be paid to the study design, particularly if the results are directly emerging from the data. To increase internal validity, we provide a comprehensive description of the data supporting it by anonymized IT ticket examples, see Chapters 3 to 5. Additionally, we present all necessary information on the analysis of the data and discuss the experimental results and evaluation in detail. Further, when experts' involvement was necessary for collecting data or evaluating findings, we looked for experts who have substantial knowledge and expertise on the related topic in organizations. With this, the goal was to identify the experts who could make a valid judgment on findings and their consistency.

External validity is concerned with the rationale of the results' generalizability. In this respect, our framework has certain constraints resulting in additional efforts of various degree while applying it in different areas. As the event log complexity bases on well-researched metrics [10], the main constraints are related to the textual data-based complexity and complexity scale choice:

- *Textual data-based BP complexity*: DML taxonomies serving as an input to linguistic features extraction are developed in two ITIL-aligned case studies coming from ITSM, which are Change Management and Service Request Management IT ticket processing. Since these processes are interconnected to other BPs in the ITSM area and have similarities with, for example, Incident or Problem Management, adjusting DML taxonomies will require less effort. We estimate it as minimal due to the substantial reusability of the taxonomies. It is worth mentioning that ITIL remains widely used, having been ranked in the ten top-paying IT certifications for 2020 based on the survey conducted in the United States [90]. Moreover, managing IT tickets, in general, remains a crucial concern for the IT service industry [91]. In entirely different cases, like other Customer Services areas, Marketing, Software Development, Strategy, and in case of languages other than English, the taxonomy must be developed from scratch following the processes described in Chapter 3 and in [31] particularly.
- *Complexity scale*: a set of ordinal complexity values constitutes a complexity scale. They can be numbers or categories arranged in a certain order to represent increasing or decreasing complexity. Depending on the specifics of the case study, a

complexity scale must be defined by the subject matter experts. A five-point Likert type scale containing numbers from one to five or a set of category names like low, medium, and high are two common examples of a complexity scale [92]. In this dissertation, we use a generic scale with three values of low, medium, and high. This scale is selected for three reasons: (1) preferred by the case study experts, (2) known scale of priority ratings, especially for measuring intangible criteria in the context of decision-making [92], (3) to simplify the method presentation.

Reliability means whether the same results can be obtained when reproduced by other researchers. To address this concern, first, we comprehensively describe the literature search and coding process in Chapter 2. In further chapters, we specify the gradual framework development. Particularly, in Chapter 5, we indicate the inputs, complexity calculation and analysis, as well as outputs and illustrate the application of the framework in the case study setting. We complement the textual data analysis with open source codes and supplementary documentation on Github[1]. We also specify the characteristics of the data used in the case studies. Moreover, we will consider the opportunity of cooperating with other researchers to apply our framework in their (similar) settings, such as ITIL Incident Management of an organization in a different domain.

7.2 Implications

In this section, we list the implications for research and practice of the work presented in this dissertation.

7.2.1 Implications for Practice

The research addressed in this dissertation has a number of implications for organizations relying on textual data as an input to their processes and already executing or interested in initiating BPM projects.

Serving as an input to BPs, textual data influence their execution. As illustrated in the ITIL-aligned ITSM case studies in this dissertation, the processing of IT tickets, i.e., requests of various types, highly depends on their textual descriptions. In the first case study, it is shown that the activities in a Change Management (CHM) process are overly dependent on the textual descriptions of change requests (RfCs) composed by customers. More specifically, the urgency of a specific RfC, its analysis and approval activities, as well as the engagement of roles such as a Change Advisory Board (CAB) are mostly determined by its description. The same input, namely textual data, is also important to determine key decision points in RfC processing, such as (i) which activities will be omitted, (ii) which interactions will take place in the process, and (iii) which roles will be involved.

Similarly, in the second case study, Service Request Management (SRM) process, we handle three problem areas that are typical for any IT ticket processing. These areas are ticket categorization, work assignment, and prioritization. Textual descriptions of service requests are utilized to identify ticket categories, which serve to assign resolution teams to the tickets. Furthermore, textual data are the basis for ticket

[1] https://github.com/IT-Tickets-Text-Analytics

prioritization. The priority of tickets has a significant impact on how they are processed. Despite a vast number of existing vendor solutions, as aforementioned, IT ticket management is still a major challenge for the IT service industry [91], whereby ITIL remains a highly used and ranked framework for ITSM [90]. In this context, decision-making support solutions for IT ticket management gain importance and attraction.

Although the case studies used to address the research questions in this dissertation come from the ITSM area, the generic application aspects of the developed BP complexity analysis framework go far beyond the latter.
First, the textual data complexity analysis presented in Chapter 3 and [12] respectively allows the realization of the following generic incremental practical research values:

- The three levels of text understanding aim to provide awareness regarding the cognitive, attention, and reading efforts required to perform BP activities, hence estimating BP complexity.
- The knowledge about BP complexity is comprised of three parts:
  (i) *objective knowledge*: professional contextual experience of the BP worker enriched by the awareness of cognitive efforts required for BP execution,
  (ii) *subjective knowledge*: business emotions enriched by attention efforts,
  (iii) *meta-knowledge*: quality of the text, i.e., professionalism, expertise, and stress level of the text author, enriched by reading efforts.
- Further, our textual data-based BP complexity concept allows a granular perspective on data analysis. BP workers are able to follow back the suggested level of complexity. This is especially important in the context of misclassifications and explainable artificial intelligence (XAI).

Second, the event log complexity analysis based on the commonly known state-of-the-art metrics [10] reveals several practical implications, such as follows:

- Event log-based BP complexity metrics can be used to predict the quality of BP models discovered with state-of-the-art process discovery algorithms.
- In the context of BP model quality analysis, input data complexity (textual and event log) is valuable to study how it contributes to model quality, such as usefulness perceived by BP analysts.
- Event log-based complexity can be combined with the BP execution performance perspective. This way, insights into identifying the complexity-related root causes can be derived using performance measures, such as long execution, rework, or errors.
- Characteristics of BP execution variants in respect to event log-based complexity can be further analyzed to eliminate and/or improve those variants.

Third, the synergistic consideration of both textual data and event log-based complexities in our BP complexity analysis framework can offer comprehensive decision-making support solutions to multiple external and internal BP stakeholders, i.e., customers triggering the process, BP workers performing activities, and BP analysts working on BP redesign and improvements. We list the potential solutions below:

- *Customers*: real-time text suggestions and chatbots to better collect the input information regarding customer requests. A qualitative description will result in faster request processing.

- *BP analysts*: more comprehensive BP analysis opportunities. Our framework offers an enriched toolset on BP redesign and improvements to identify root causes like bottlenecks, rework, and ping-pong behavior more comprehensively and, afterward, address them while automating, reordering, paralleling, or eliminating certain BP activities. As compared to common process discovery techniques, textual data-based complexity, its linguistic features, and derived patterns provide additional information on the root causes.
- *BP workers*: better work prioritization, categorization, and assignment. This can be achieved by real-time decision-making support, for example, in the form of dashboards indicating the BP complexity and BP complexity-based automation solutions. In the latter, we mean the BP complexity definitions we suggested in [12]: (i) BPs with low complexity are those which can be easily automated based on clear rules, (ii) BPs of medium complexity do not follow exact ruleset and can be only partially automated, (iii) in case of highly challenging BPs (high complexity), there is no automation expected but minimal assistance in the form of the history of similar BPs.

Considering the implications presented above, one can conclude that our framework has great application potential outside ITSM. In fact, in many other areas, one can observe a similar role of textual data in the execution of BPs, i.e., its influence on the decision points, activities, and their order. For example, in healthcare, the complaints expressed by a patient typically determine the required diagnostics and related BP activities. Similarly, in governmental services, the quality and completeness of the described request directly influence the processing speed.

To illustrate how organizations benefit from our framework and research contributions to solve their problems, we apply it in a set of BPM cases and respective processes (see Tables 7.2.1 to 7.2.4). Herewith, we aim to demonstrate how our framework can be used to provide a better understanding of BPs and which complexity-based decision-making support solutions can be considered. Our illustration and understanding of these BPM cases and processes are based on the description presented in [93] and the work experience of the author of this dissertation. The cases are adapted to the present research setting and enriched by our personal experiences. Such illustrations enable us to build justifications for our framework application, reflect on its benefits in various real-life settings, and engage in thought experiments on how BPM practice might be made more efficient [94, 95]. In the illustration of the cases, we provide the following information:

- *Case description* for summarizing the key aspects of the case story,
- *Problem identification* for presenting the initial problem that led to the action taken,
- *Actions taken* for depicting concrete steps to solve the problem, i.e., what measures were undertaken, such as in regard to process optimization,
- *BP complexity framework-based decision-making support* for highlighting the added value that can be achieved by our framework application.

Table 7.2.1: BPM case-1

Insurance Claims Handling [96]	
Case description	Injury-compensation claims are difficult to process as they entail negotiations between multiple parties like claimants, insurers, law firms, and health providers.
Problem identification	In the claims handling process, there are considerable behavioral and performance variations that impact claim costs and duration. The causes for these discrepancies are unknown.
Actions taken	BPM project was launched to employ process mining with a focus on BP identification, discovery, and analysis. To determine where claims handling differed among groups of interest, automated process discovery and comparative performance analysis were conducted. At the same time, a context analysis to determine the context aspects, such as claim amount, claimant's employment, medical, and legal aspects that influence claim costs and duration, was performed.
BP complexity framework-based decision-making support	While additional context aspects are analyzed, such a critical source of information as the claim description by the claimant remained neglected. In this regard, our framework can provide additional insights into this root cause of behavioral and performance variations. Further, the framework adds a complexity perspective to the conducted analysis as another potential root cause for the mentioned variations.

Table 7.2.2: BPM case-2

e-Service Processes in the City of Ghent [97]	
Case description	The City of Ghent, a Belgian public-sector organization, aims to contribute to the larger initiative regarding e-government, e-citizenship, digital identities, and smart cities. Hereby, the goal is to make all services delivered more customer-oriented and driven by the demands of local citizens, organizations, and associations. The long-term plan aims to maximize the city physical and digital services, the latter being in focus.
Problem identification	A large set of heterogeneous services and customers and the historical evolution of working in silos resulted in a large number of applications and forms.
Actions taken	In the digitization of service processes, BPMN was used for the identification, discovery, and analysis of core BPs, albeit with different modeling tools in each department. BP redesign was based on Lean Thinking [98], seeking to minimize waste and maximize customer value. Respectively, the following principles were outlined: (i) fast and personalized services, (ii) easily usable for a heterogeneous audience, and (iii) dynamic forms to collect information.
BP complexity framework-based decision-making support	To be able to fulfill these principles for more than 300 products and services offered by the City of Ghent, one can profit from the complexity analysis of the respective service processes. First, to support BP analysts, our BP complexity framework can be used as a tool for a more comprehensive BP redesign to make the BPs more customer-oriented and lean, reducing the complexity both at the starting point (filling in an application) and in the further processing steps. Second, based on the complexity analysis, one could design chatbots providing the citizens a different level of support to fill in the application. For example, the number of automatic questions vs. human operator support will be different for a simple process, such as a residence application, and a complex one, like a tax declaration. Whereas we used this high-level example to better demonstrate the value of the complexity framework, more interesting and valuable cases could be found within each of the examples. For instance, tax declaration, depending on the specific situation of a submitter, can have various complexity degrees.

Table 7.2.3: BPM case-3

Automate Does Not Always Mean Optimize: Case Study at a Logistics Company [99]	
Case description	A logistics company deals with a large and constantly increasing number of invoices and purchase orders. The workflow automation solution for the purchase order and invoice approval procedures based on the digitized versions of the documents was introduced.
Problem identification	Process automation provided a lot of benefits, but it also necessitated more monitoring, controlling, and the search for further optimization areas. The organization sought to find out how to measure the performance of processes and the involved resources. The major focus was on monitoring and managing the achievement of enterprise-level KPIs and business rules, as well as examining the purchase order process from the perspective of suppliers and quantifying inconsistencies in the delivery of the products. The increasing number of rejected invoices was a source of concern.
Actions taken	Process mining was used to create process models from an event log and identify the main process flows and their deviations, followed by social networks analysis. A problem was discovered in the distribution of work and over-allocation of resources. Further, focusing on the rejected invoices activity, text mining was applied to identify the most common causes. An extension of the event log was made to user comments. Based on a frequency analysis of phrases, the reasons for rejection were identified.
BP complexity framework-based decision-making support	The BP complexity framework can be useful for both textual data and event log analysis. Whereas the presented approach considers user comments, one could include other textual data, such as email texts containing purchase orders and invoices into the textual data analysis. Investigating their complexity as compared to the event log, one can predict possible problems related to purchase order processing and invoice approvals. Hereby, the introduction of new complexity-based business rules can facilitate the work and reduce the over-allocation of resources.

Table 7.2.4: BPM case-4

Document Management System Introduction at the Brandenburg University of Applied Sciences (BUAS) (based on the author's work experience)	
Case description	According to EU Directive 2014/55, every invoice recipient in the public sector must be able to receive and process electronic invoices digitally. To comply with this rule, BUAS launches a project to introduce a Document Management System (DMS). In addition to software acquisition, the processes in the three major areas are captured and modeled using the CWA SmartProcess tool[2]: student enrollment, post office, and procurement.
Problem identification	Whereas the acquired DMS software offers a high degree of automation, there is a number of special cases which should be processed manually. For example, in procurement, there are cases of discounts or partial invoices, which one can derive from email communication.
Actions taken	In the context of the mentioned problem, it is planned to keep manual approval of the invoices.
BP complexity framework-based decision-making support	Already at the beginning of the DMS introduction, one can benefit from the textual data-based analysis. Correspondingly, a taxonomy of keywords in the context of special cases requiring manual processing could be used to automatically forward the respective cases to the BP workers. Further, in addition to the textual data analysis, the event log data from the DMS and its complexity could help to identify bottlenecks, rework, and ping-pong behavior as well as potential root causes.

7.2.2 Implications for Research

The work performed in the scope of this dissertation has some implications for research, particularly in the following subject areas: organizational studies, Business Process Management (BPM), Information Systems (IS) design, and Business Intelligence (BI).

Organizational studies. In Chapter 2, we presented a comprehensive overview of complexity in organizations aimed at supporting organizations in dealing with complexity. As we observe, organizations are under pressure to adapt technological advancements to become more resilient. At the same time, there are new opportunities and challenges caused by emerging technologies, changing organizational structures, and dynamic work environments. Complexity is mostly the result of such developments, which are intensified by the increasing volume and variety of exchanged data. This complexity is typically a major impediment to efficiency, making it difficult to comprehend, regulate, and improve organizational structures, processes, and practices. In research, much work has been done in assessing, reviewing, and studying complexity. However, these attempts are disjointed and lack a common vision. To address these problems, we analyzed the existing works on diverse complexity types and developed structured overviews. Hence, other researchers interested in organizational complexity research can take our findings as a basis and enrich them with new perspectives. Additionally, the most popular and under-researched areas can provide inspiration and serve as guidance for novel research projects. For example, researchers can identify topics to study from the results listed in the future research area dimension. Further, in the results, the proposed complexity metrics and how they were evaluated and implemented can be a starting point of the research aimed at designing adequate complexity metrics.

Business Process Management. Several examples of qualitative and quantitative BP analysis approaches are described in Chapter 1, Section 1.1. Along with that, we highlighted an increasing interest in BP complexity analysis from the BP model and event log perspective. Hereby, we identified the lack of comprehensive solutions considering event log and textual data together. In this context, the suggested framework aims to enrich the BPM toolset with such solutions. Interested researchers could use the framework as a basis to develop process mining tools enabling analysis of complexity by combining event log and textual data. Further, in our textual data-based BP complexity concept, we used three linguistic levels of text understanding (knowledge aspects) realized through semantics, syntax, and stylistics, with semantics being in the most focus. The importance of semantics in NLP, in general, is a recognized and attractive research field [100]. It also remains one of the significant challenges impeding the full exploitation of the NLP benefits in BPM [11]. Hence, our research on textual data-based BP complexity demonstrates another possibility of NLP application in BPM. Additionally, while developing our textual data-based BP complexity concept and studying the BPM and NLP-related work in Chapter 3 and [12], we observed a strong focus on the technical artifacts, i.e., BP models. The actual support of BP workers in the BP execution has shown to be underresearched. As illustrated in the practical implications above, in this dissertation, we met the demand for direct BP execution support, among others.

Information Systems design. The framework we introduced in this dissertation addresses the problem of the negligence of textual data generated in BPs by BP actors and represents a comprehensive solution for how to profit from the latter in synergy with the structured event log. In this regard, it is noteworthy that our data-intensive and knowledge-based economy increases the role of information and its quality significantly [101, 102]. Naturally, one of the prominent research issues in IS design is the quality of information. It is referred to as the degree to which users view information as reliable, detailed, timely, structured, and up-to-date [103]. This issue becomes relevant, especially in the case of unstructured textual data existing in organizations in large volumes. In fact, many organizations do not fully exploit the potential or even completely disregard their information assets being unaware of what information is available, where to get it, or whether it is consistent, up-to-date, or accurate. According to organizations, employees searching for information are frequently confronted with information chaos or overload [104]. Despite considerable investments into technologies to collect, store, and process vast quantities of data, organizations are frequently hampered in their attempts to turn these data into actionable insights that can be used to improve processes, make better decisions, and create strategic advantages. These difficulties are caused by a variety of issues relating to the quality of data and information [102]. Developing the framework on data-driven BP complexity analysis, we considered two data types, unstructured textual data and structured event log, both of them representing a certain type of challenge for employees. Whereas vast amounts of textual data are a generic challenge related to the limited ability of humans to deal with large data volumes, dealing with event log might be challenging to its technical nature and lack of specific knowledge to be able to derive insights. In this dissertation, according to the MIT Total Data Quality Management (TDQM) framework [102], we contributed to the data quality research in line with its cycle of *Define*, *Measure*, *Analyze*, and *Improve*. We, first, defined a concept of BP complexity as the *fitness for use* data quality aspect from the user's point of view. Second, we developed a set of measures for two complexity types, followed by a thorough analysis of the measurement results. Finally, we suggested the BP improvement solutions obtainable based on our framework. Using our work, interested researchers can extend our measures by defining and analyzing new concepts.

Business Intelligence. Moreover, our framework serves as a basis for potential BI applications to develop automation approaches for decision-making support of BP actors. For example, we suggested a dashboard mock-up to visualize our BP complexity-related findings in [64]. Further, in line with our latest observations on dashboard research and existing gaps [62], one could work on dashboards as part of comprehensive BP complexity-based decision-making support solutions focusing on the end-user involvement in all stages of the dashboard design and implementation. For example, information regarding BP complexity and its counterparts can be cumulatively presented in a dashboard to be used by managers to assess the work of a department for a certain period and to make strategic Human Resources (HR) decisions. More specifically, depending on the case and available information, answers can be found to such questions as:

(i) What teams and/or employees with which skills are the most or least demanded in this quarter?

(ii) Are there any trends over 3–5 years in this context?
(iii) Which teams and/or employees are most or least overloaded?
(iv) What HR training and/or HR reallocation is needed?

Hence, dashboards capable of a user-centric presentation of such information and its dynamic adjustment to the needs of decision-makers can be developed in the future.

7.3 Limitations and Open Issues

Each of the chapters lists the limitations relevant to specific research work. In this section, we summarize and extend the discussed limitations and open issues in relation to the gradual framework development in this dissertation.

In Chapter 2, the literature review-based method to address complexity in organizations is presented. While conducting the SLR, our main focus was on complexity in organizations, in particular, complexity revolving around BPs. In addition to complexity, simplicity is another common phenomenon within organizations. Whereas complexity and simplicity are not necessarily antonyms, the interplay between these two notions has been extensively studied in organizational research [105]. For example, [105] assume that understanding organizational complexity requires taking into account the value of simplicity. [106] suggest a simplicity-based approach as a practical guide to reduce complexity in any organization. Thus, analyses on addressing or benefiting from simplicity in organizations might reveal relevant insights into dealing with complexity.

A novel BP complexity concept based on textual data is developed in Chapter 3. In this concept definition, DML taxonomy development is essential. Moreover, to develop a DML taxonomy, domain expert involvement is critical. Particularly, defining context-specific rules for determining DML levels and refining DML taxonomy are two core tasks for which expert knowledge is necessary. In other words, DML taxonomy development has the limitation of dependency on domain experts. Aside from that, as an open issue, the reusability of the developed DML taxonomies can only be determined and improved by applying them in more case studies of various organizations. A further limitation is related to the type of textual data. As defined in Section 1.6, we focus on textual data generated by BP actors in the BP execution and serving as an input to the BP. However, other textual data like legal documents, corporate standards, and interview transcriptions also represent a valuable source of knowledge that can support BP analysts at various phases of the BPM lifecycle. The analysis of the textual complexity of such data constitutes an open issue for future research. Another textual data-related limitation is that we considered only texts written in English, which makes the work on multilingual solutions a relevant open issue.

As explained in Chapter 5, a detailed analysis of the connection between textual data and event log-based complexities is performed. In this analysis, a further breakdown is applied to understand the connection between complexities for various subsets of BP instances using their attributes. For creating reasonable breakdowns and determining relevant attributes, domain expert involvement was necessary. To overcome this limitation, such attributes can be defined in a machine-readable form, and their relevance can be identified with machine learning.

Generally speaking, processes represent all of the work performed inside an organization [1]. As a result, process complexity and methods to measure it can imply a variety of aspects and factors, such as those emerging from the process context [107, 108], which are frequently challenging to collect. These also include human factors like culture, end-user profiles, background, and novice BP workers. Additionally, BP perceived complexity is influenced by how it is depicted in a BP model. In other words, BP perceived complexity depends on BP model quality and modeling notation. While event log complexity metrics are the main focus of the research presented in this dissertation, other BP-related complexity metrics and other approaches to complexity like those considering context represent a promising open issue.

Since textual data and event log are not combined in other works, no alternatives exist to compare the quality of our framework. Regardless of that, the framework can be further evaluated by applying it in several domains.

7.4 Future Work

Each of the chapters of this dissertation provides ideas for future research. In this section, we further enrich the ideas and sketch some of the interesting and promising directions for future research work.

Considering the aforementioned limitation of the SLR on complexity in organizations, simplicity can be put into the scope as a future research direction of the SLR. With this, one can analyze how organizations approach complexity while aiming simplicity in their setting. Such analysis can be used to extend the literature review-based method addressing complexity with a special focus on balancing complexity and simplicity. Thus, organizations can gain knowledge on what factors are important to determine a cost-efficient trade-off.

For future work regarding the novel BP complexity concept based on textual data, there are two prominent avenues: developing new linguistic features and real-time BP complexity identification. For the former, libraries such as [74] can be considered to develop linguistic features that reflect social, cognitive, and affective perspectives of textual data. In addition, conversations taking place in BP executions can be incorporated into textual data analysis. Regarding real-time BP complexity identification, the prototype of a recommender system for BP workers [55] that automatically identifies BP complexity and provides a recommendation can be taken as the basis. With real-time guidance for BP workers, the way they deal with complexity can be facilitated. Furthermore, multilingual solutions as well as other textual data types like legal documents, interview transcriptions, their semantics and writing style are worth considering to support BP improvement at all the stages of the BPM lifecycle.

Another interesting future research direction is an application of decision mining techniques [109–112] to obtain knowledge on decisions made by BP workers based on textual data. What features of textual data yield which decisions and how these decisions affect the execution of BPs can be deeply analyzed.

Moreover, Artificial Intelligence (AI) is a promising research field that can be used for automated decision-making. Decisions previously made in BP executions can be exploited for training AI technologies to imitate human intelligence in handling textual data. For example, patterns in historical decisions and corresponding textual data characteristics can be used to develop smart support agents, i.e., chatbots. With the help of the chatbots, BP workers can focus on complex decisions that require context knowledge in addition to textual data.

Additionally, the consideration of other complexity metrics like those of BP models and approaches to complexity, for example, related to BP context, can open multiple new research directions.

Finally, to draw attention to the framework, the design possibilities of a Graphical User Interface (GUI) can be studied in detail. Such a dashboard can facilitate the BP complexity-based decision-making support for BP workers.

Appendix A - Evaluation Questionnaire

Evaluation Questionnaire.

This questionnaire was used to introduce the CHM process workers to the research topic and get first feedback as a preparation for the follow-up interview in the form of online discussion.

Research Context Information (5 min. read):

Dear participant,

Within a research project conducted using the Company XXX Change Management (CHM) tickets dataset, the researchers aim to investigate the Decision-Making nature of the incoming Requests for Change (RfCs), i.e., to classify them into the *routine*, *semi-cognitive*, and *cognitive* requests (activities). Different methods from semantic and linguistic analyses are conceptually combined to extract the necessary knowledge from the initial RfC textual description, which would help classify the incoming request into routine, semi-cognitive, or cognitive one.

The *semantic analysis* of the RfC textual descriptions was performed using Text Mining algorithms to extract descriptive keywords from the incoming RfC textual data. Afterward, based on the state-of-the- art research approaches and literature, the researchers developed a set of indicators and, using these indicators, semantically classified the extracted keywords into three Decision-Making Logic (DML) levels: routine, semi-cognitive, or cognitive. The developed vocabulary of keywords[1] is used to classify the exemplary RfCs into the three DML levels.

The *linguistic analysis* of the RfC descriptions is a second step of the research. The knowledge about linguistic aspects will include the contextual structure (the RfC length and level of details), emotional coloring and writing style, vocabulary richness (repetitions, synonyms, new words not existing in the extracted vocabulary) in the RfC textual descriptions.

Hereby, the researchers' interpretation of the DML levels and related RfC classification is to be considered from the point of view of computer-/ technology-aided process support potential. Thus, the researchers use the following interpretations:

- *routine requests*: those expressible in rules so that they are easily programmable and can be performed by computers at economically feasible costs;
- *semi-cognitive requests*: those where no exact ruleset exists, and there is a clear need for information acquisition. Here, computer technology cannot substitute but has the potential to increase the productivity of employees by partial task processing;
- *cognitive requests*: those where not only an information acquisition but also evaluation and complex problem-solving are needed. Here, computer technology can offer minimal decision support leaving the task processing to the process worker.

Questionnaire (15-20 min. fill-in):

1. In the context of the present research applied on the case study of CHM at Company XXX, please briefly define using few (3-5) keywords, what does it mean for you as a CHM process worker to process:

(a) Routine request: Click or tap here to enter text.

(b) Semi-cognitive request: Click or tap here to enter text.

(c) Cognitive request: Click or tap here to enter text.

2. In the context of your daily RfC processing work, please define the complexity of the routine RfC and routine-cognitive RfC using the scale from 1 (minimal complexity) to 10 (maximal complexity). You can also suggest further subcategorization based on the complexity score level.

(a) Routine request: Click or tap here to enter text.

---

[1] The vocabulary can be found in appendix. The extracted descriptive keywords are organized into parts of speech (PoS) with the following semantic contextual meaning: ***R****esources* (nouns), ***T****echniques* (verbs and verbal nouns), ***C****apacities* (adjectives), ***C****hoices* (adverbs) (RTCC structure). Afterward, with the help of indicators, the keywords are classified into those having routine, semi-cognitive, and cognitive contextual meaning.

(b) Semi-cognitive request: Click or tap here to enter text.

(c) Cognitive request: Click or tap here to enter text.

Notes: Please include further subcategories if meaningful
Click or tap here to enter text.

3. In the context of your daily RfC processing work, how would you assess the dependency of the RfC complexity from the following criteria? Which criteria make the RfC more complicated to process? Please give the score from 1 (completely irrelevant criteria) to 10 (highly relevant criteria).

Textual length of incoming RfC (long texts): Click or tap here to enter text.

Level of details (overdetailed): Click or tap here to enter text.

Emotional coloring[2]: Click or tap here to enter text.

Writing style[3]: Click or tap here to enter text.

Usage of standard CHM vocabulary words: Click or tap here to enter text.

Usage of telegraphic style: Click or tap here to enter text.

4. <u>Give an example</u> from your practice of a routine, semi-cognitive, and cognitive RfC.

(a) Routine request: Click or tap here to enter text.

(b) Routine-cognitive request: Click or tap here to enter text.

(c) Cognitive request: Click or tap here to enter text.

Notes: Please include further examples of subcategories if meaningful
Click or tap here to enter text.

5. Here are the examples of how the researchers classified the RfCs using the developed vocabulary[4] of keywords based on the data set and the sentence-by-sentence approach. <u>Please agree or disagree, write a few comments, especially in case of disagreement.</u>

(a) Selected example of a routine request: X0012345678, Due to some errors in application and in database a restart has been recommended. Check if the application is up and running. Check if the DB is up and running.

Ticket ID X0012345678 (sentences)	**Values**	**Part of speech & its contextual semantic aspect**	**Keyword DML**	**Ticket DML**
Due to some errors in <u>application</u> and in <u>database</u> a re<u>start</u> has been recommended.	application	noun (resource)	routine	*<u>routine</u>* in **resources & techniques**
	database	noun (resource)	routine	
	start	verb (technique)	routine	
<u>Check</u> if the <u>application</u> is up and <u>run</u>ning.	check	verb (technique)	routine	
	application	noun (resource)	routine	
	run	verb (technique)	routine	
<u>Check</u> if the DB is up and <u>run</u>ning.	check	verb (technique)	routine	
	run	verb (technique)	routine	

[2] E.g. a lot of exclamation marks, polite adverbs "please" or urgency adverbs "immediately"

[3] Repetitions, synonyms

[4] Can be found in appendix

Click or tap here to enter text.

Please agree or disagree, write a few comments, especially in case of disagreement.

(b) Selected example of a mixed request: X00111222, Please stop-start BBB(EUTAD) databases mentioned below: BBB1, BBB2, BBB3, BBB4, BBB5. server: hfgssak. Please check mentioned databases if were stop-start successfully and if applications after start running properly.

Ticket ID X00111222 (sentences)	Values	Part of speech & its contextual semantic aspect	Word DML	Ticket DML
Please stop-start BBB(EUTAD) databases mentioned below: BBB1, BBB2, BBB3, BBB4, BBB5.	stop start database	verb (technique) verb (technique) noun (resource)	routine routine routine	*routine* in **resources & techniques** and *semi-cognitive* and *cognitive* in **choices**
server: hfgssak.	server	noun (resource)	routine	
Please check mentioned databases if were stop-start successfully and if applications after start running properly.	check database stop start successfully properly application run	verb (technique) noun (resource) verb (technique) verb (technique) adverb (choice) adverb (choice) noun (resource) verb (technique)	routine routine routine routine semi-cognitive cognitive routine routine	

Click or tap here to enter text.

Please agree or disagree, write a few comments, especially in case of disagreement.

(c) Selected example of a mixed request: X000888444, SAP R3 environment should be updated for TTT Light. The installation will affect DB part only. SAP R3_OM will deliver the software SAP R3_LL should execute the installation test should be done by SAP R3_LL and SAP R3_OM after installation.

Ticket ID X000888444 (sentences)	Values	Part of speech & its contextual semantic aspect	Word DML	Ticket DML
SAP R3 environment should be updated for TTT Light.	environment update	noun (resource) verb (technique)	semi-cognitive routine	*semi-cognitive* in **resources** & *routine* in **techniques**
The installation will affect DB part only.	install	verb (technique)	routine	
SAP R3_OM will deliver the software SAP R3_LL should do the installation test should be done by SAP R3_LL and SAP R3_OM after installation.	install test	verb (technique) verbal noun (technique)	routine routine	

Click or tap here to enter text.

Please agree or disagree, write a few comments, especially in case of disagreement.

6. Please provide a few comments, feedback, and critic to the researchers regarding the conducted research.

Click or tap here to enter text.

Bibliography

[1] Dumas, M., Rosa, M. L., Mendling, J. & Reijers, H. A., *Fundamentals of business process management*, 2nd ed., Springer, 2018 (cited on pages 3, 5, 19, 20, 182).

[2] Mulholland, B., *17 BPM statistics to help you increase efficiency in your business*, 2017, URL: https://www.process.st/bpm-statistics-increase-efficiency/ (visited on 09/01/2022) (cited on page 3).

[3] Stokdyk, J., *Rail franchising: excel error strikes again*, 2012, URL: https://www.accountingweb.co.uk/tech/excel/rail-franchising-excel-error-strikes-again (visited on 09/01/2022) (cited on page 3).

[4] DiDio, L., *Human error is top cause of downtime*, 2018, URL: https://techchannel.com/SMB/10/2018/Human-Error-Top-Cause-of-Downtime (visited on 09/01/2022) (cited on page 3).

[5] BBC, *British Airways says IT chaos was caused by human error*, 2017, URL: https://www.bbc.com/news/business-40159202 (visited on 09/01/2022) (cited on page 3).

[6] Laguna, M. & Marklund, J., *Business process modeling, simulation and design*, 3rd ed., Chapman and Hall/CRC, 2018 (cited on page 3).

[7] d.velop public sector GmbH, *Wir sind d.velop public sector*, 2021, URL: https://www.codia.de/software (visited on 09/01/2022) (cited on page 3).

[8] Axelos, *Information technology infrastructure library*, tech. rep., Axelos, 2019, URL: https://www.axelos.com/best-practice-solutions/itil (visited on 09/01/2022) (cited on pages 3, 19, 22, 23).

[9] Kobayashi, V., Mol, S. T., Berkers, H., Kismihók, G. & Hartog, D. N. den, "Text mining in organizational research", *Organizational Research Methods*, vol. 21, 2018, pp. 733–765 (cited on pages 3, 5).

[10] Augusto, A., Mendling, J., Vidgof, M. & Wurm, B., "The connection between process complexity of event sequences and models discovered by process mining", *Information Sciences*, vol. 598, 2022, pp. 196–215 (cited on pages 3, 5, 8, 9, 171, 173).

[11] Aa, H. van der, Carmona, J., Leopold, H., Mendling, J. & Padró, L., "Challenges and opportunities of applying natural language processing in business process management", in: *Proceedings of the 27th International Conference on Computational Linguistics (COLING2018), Santa Fe, New Mexico, USA, August 20-26, 2018*, Association for Computational Linguistics, 2018, pp. 2791–2801 (cited on pages 3–5, 8, 179).

[12] Rizun, N., Revina, A. & Meister, V. G., "Assessing business process complexity based on textual data: evidence from ITIL IT ticket processing", *Business*

Process Management Journal, vol. 27, no. 7, 2021, pp. 1966–1988, DOI: 10.1108/BPMJ-04-2021-0217 (cited on pages 4, 8, 9, 15, 16, 18, 70, 171, 173, 174, 179).

[13] Gigerenzer, G. & Gaissmaier, W., "Heuristic decision making", *Annual Review of Psychology*, vol. 62, 2011, pp. 451–482 (cited on page 4).

[14] Brocke, J. vom, "The power of text mining - how to leverage naturally occurring text data for effective enterprise information systems design and use", in: *Proceedings of the 18th International Conference on Enterprise Information Systems (ICEIS2016), Rome, Italy, April 25-28, 2016*, ScitePress, 2016, p. 23 (cited on page 4).

[15] Brocke, J. vom, *Class notes: power of text-mining in BPM*, 2017, URL: https://www.bptrends.com/class-notes-power-of-text-mining-in-bpm-2/ (visited on 09/01/2022) (cited on page 4).

[16] Gupta, V., Lehal, G. S., et al., "A survey of text mining techniques and applications", *Journal of Emerging Technologies in Web Intelligence*, vol. 1, no. 1, 2009, pp. 60–76 (cited on page 4).

[17] Brocke, J. vom & Mendling, J., "Frameworks for business process management: a taxonomy for business process management cases", in: *Business Process Management Cases, Digital Innovation and Business Transformation in Practice*, Springer, 2018 (cited on pages 4, 8).

[18] Baiyere, A., Salmela, H. & Tapanainen, T., "Digital transformation and the new logics of business process management", *European Journal of Information Systems*, vol. 29, no. 3, 2020, pp. 238–259 (cited on pages 4, 6, 8, 20).

[19] Schmidt, R., Möhring, M., Härting, R.-C., Zimmermann, A., Heitmann, J. & Blum, F., "Leveraging textual information for improving decision-making in the business process lifecycle", in: *Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015), Sorento, Italy, 17-19 June, 2015*, Springer, 2015, pp. 563–574 (cited on page 5).

[20] Midden, J. van, "Using process mining and event log analysis for better business strategy decision-making", B.Sc. thesis, University of Twente, 2021 (cited on page 5).

[21] Brocke, J. vom & Rosemann, M., *Handbook on business process management 2: strategic alignment, governance, people and culture*, Springer, 2014 (cited on pages 5, 19).

[22] Aalst, W. M. Van der, "Business process management: a comprehensive survey", *International Scholarly Research Notices*, vol. 2013, 2013, pp. 1–37 (cited on page 5).

[23] Dijkman, R., Dumas, M., Van Dongen, B., Käärik, R. & Mendling, J., "Similarity of business process models: metrics and evaluation", *Information Systems*, vol. 36, no. 2, 2011, pp. 498–516 (cited on page 5).

[24] Recker, J., "Suggestions for the next wave of BPM research: strengthening the theoretical core and exploring the protective belt", *Journal of Information Technology Theory and Application*, vol. 15, no. 2, 2014, pp. 5–20 (cited on page 6).

[25] Hevner, A. R., March, S. T., Park, J. & Ram, S., “Design science in information systems research”, *MIS Quarterly*, vol. 28, no. 1, 2004, pp. 75–105 (cited on pages 7–10).

[26] Hevner, A. & Chatterjee, S., *Design science research in information systems*, Springer, 2010 (cited on page 7).

[27] Denning, P. J., “A new social contract for research”, *Communications of the ACM*, vol. 40, no. 2, 1997, pp. 132–134 (cited on page 7).

[28] Markus, L., Majchrzak, A. & Gasser, L., “A design theory for systems that support emergent knowledge processes”, *MIS Quarterly*, vol. 26, no. 3, 2002, pp. 179–212 (cited on page 8).

[29] Revina, A. & Aksu, Ü., “An approach for analyzing business process execution complexity based on textual data and event log”, *Information Systems*, vol. 114, 2023, p. 102184, DOI: 10.1016/j.is.2023.102184 (cited on pages 8, 9, 15, 16, 19, 126).

[30] Revina, A., Buza, K. & Meister, V. G., “IT ticket classification: the simpler, the better”, *IEEE Access*, vol. 8, 2020, pp. 193380–193395, DOI: 10.1109/ACCESS.2020.3032840 (cited on pages 9, 15, 16, 106).

[31] Rizun, N., Revina, A. & Meister, V. G., “Method of decision-making logic discovery in the business process textual data”, in: *Proceedings of the 22nd International Conference on Business Information Systems (BIS2019), Seville, Spain, June 26-28, 2019*, Springer, 2019, pp. 70–84 (cited on pages 9, 15, 16, 18, 171).

[32] Rizun, N., Meister, V. G. & Revina, A., “Discovery of stylistic patterns in business process textual descriptions: IT ticket case”, in: *Proceedings of the 33rd International Business Information Management Association Conference (IBIMA2019), Education Excellence and Innovation Management through Vision 2020, Granada, Spain, April 10-11, 2019*, International Business Information Management Association (IBIMA), 2019, pp. 2103–2113 (cited on pages 9, 15, 16, 18).

[33] Rizun, N. & Revina, A., “Business sentiment analysis. Concept and method for perceived anticipated effort identification”, in: *Proceedings of the 28th International Conference on Information Systems Development: Information Systems Beyond 2020 (ISD2019), Toulon, France, August 28-30, 2019*, 2019 (cited on pages 9, 15, 16, 18).

[34] Zipf, G. K., *Selected studies of the principle of relative frequency in language*, Harvard University Press, 1932 (cited on page 9).

[35] Blei, D. M., “Probabilistic topic models”, *Communications of the ACM*, vol. 55, no. 4, 2012, pp. 77–84 (cited on page 9).

[36] Simon, H. A., *The sciences of the artificial*, MIT Press, 1996 (cited on pages 9, 10).

[37] Vessey, I. & Glass, R., “Strong vs. weak approaches to systems development”, *Communications of the ACM*, vol. 41, no. 4, 1998, pp. 99–102 (cited on page 10).

[38] Benbasat, I. & Zmud, R., “Empirical research in information systems: the practice of relevance”, *MIS Quarterly*, 1999, pp. 3–16 (cited on page 10).

[39] Cooper, H. M., “Scientific guidelines for conducting integrative research reviews”, *Review of Educational Research*, vol. 52, no. 2, 1982, pp. 291–302 (cited on page 12).

[40] Kitchenham, B., *Procedures for performing systematic reviews*, tech. rep. TR/SE-0401, Keele, UK, Keele University, 2004 (cited on page 12).

[41] Snyder, H., “Literature review as a research methodology: an overview and guidelines”, *Journal of Business Research*, vol. 104, 2019, pp. 333–339 (cited on page 12).

[42] Durbin, C. G., “How to read a scientific research paper”, *Respiratory Care*, vol. 54, no. 10, 2009, pp. 1366–1371 (cited on page 12).

[43] Wand, Y. & Weber, R., “Research commentary: information systems and conceptual modeling - a research agenda”, *Information Systems Research*, vol. 13, no. 4, 2002, pp. 363–376 (cited on page 12).

[44] Fettke, P., “How conceptual modeling is used”, *Communications of the Association for Information Systems*, vol. 25, no. 1, 2009, p. 43 (cited on page 12).

[45] Olivé, A., *Conceptual modeling of information systems*, Springer Science & Business Media, 2007 (cited on page 12).

[46] Recker, J., *Scientific research in information systems - a beginner’s guide*, Springer, 2013 (cited on pages 12, 13).

[47] Stake, R. E., *The art of case study research*, Sage, 1995 (cited on page 12).

[48] Yin, R. K., *Case study research: design and methods*, 4th ed., Sage, 2013 (cited on page 12).

[49] Ørngreen, R. & Levinsen, K., “Workshops as a research methodology”, *Electronic Journal of E-Learning*, vol. 15, no. 1, 2017, pp. 70–81 (cited on page 12).

[50] Biggs, S., *Resource-poor farmer participation in research: a synthesis of experiences from nine national agricultural research stations*, tech. rep. 3, OFCOR, Comparative Study, 1989 (cited on page 12).

[51] Cornwall, A. & Jewkes, R., “What is participatory research?”, *Social Science & Medicine*, vol. 41, no. 12, 1995, pp. 1667–1676 (cited on page 12).

[52] Rowley, J., “Designing and using research questionnaires”, *Management Research Review*, 2014 (cited on page 13).

[53] Meuser, M. & Nagel, U., “The expert interview and changes in knowledge production”, in: *Interviewing Experts*, Springer, 2009, pp. 17–42 (cited on page 13).

[54] Daelemans, W., “Explanation in computational stylometry”, in: *Proceedings of the 14th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing2013), Samos, Greece, March 24-30, 2013*, Springer, 2013, pp. 451–462 (cited on pages 15, 171).

[55] Revina, A. & Rizun, N., “Multi-criteria knowledge-based recommender system for decision support in complex business processes”, in: *Proceedings of the Workshop on Recommendation in Complex Scenarios co-located with 13th ACM Conference on Recommender Systems (RecSys2019), Copenhagen, Denmark, September 20, 2019*, vol. 2449, CEUR, 2019, pp. 16–22 (cited on pages 15, 16, 182).

[56] Revina, A., Rizun, N. & Aksu, Ü., "Towards a framework for context awareness based on textual process data", in: *Proceedings of the 26th International Enterprise Design, Operations and Computing Workshop (EDOCW2022), Bolzano, Italy, October 3-7, 2022*, Springer, in press (cited on pages 15, 17).

[57] Revina, A., Aksu, Ü. & Meister, V. G., "Method to address complexity in organizations based on a comprehensive overview", *Information*, vol. 12, no. 10, 2021, DOI: 10.3390/info12100423 (cited on pages 16, 18, 26, 171).

[58] Revina, A., "Business process management: integrated data perspective. A framework and research agenda", in: *Proceedings of the 29th International Conference on Information Systems Development: Crossing Boundaries between Development and Operations (DevOps) in Information Systems (ISD2021), Valencia, Spain, September 8-10, 2021*, 2021 (cited on pages 16, 19, 152, 171).

[59] Revina, A., "Assessing process suitability for AI-based automation. Research idea and design", in: *Proceedings of the 21st International Conference on Business Information Systems (BIS2018), Berlin, Germany, July 18-20, 2018*, Springer, 2018, pp. 697–706 (cited on page 16).

[60] Buza, K. & Revina, A., "Speeding up the SUCCESS approach for massive industrial datasets", in: *Proceedings of the 2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA2020), Novi Sad, Serbia, August 24-26, 2020*, IEEE, 2020, pp. 1–6 (cited on page 17).

[61] Revina, A., "Considering business process complexity through the lens of textual data", in: *Proceedings of the the 16th International Multi-Conference on Computing in the Global Information Technology (ICCGI2021), Nice, France, July 18-22, 2021*, ThinkMind, 2021, pp. 12–14 (cited on page 17).

[62] Revina, A., "Exploring dashboards as socio-technical artifacts: literature review-based insights", in: *Proceedings of the 18th International Conference on E-Business (ICE-B2021), Online, July 7-9, 2021*, ScitePress, 2021 (cited on pages 17, 180).

[63] Revina, A. & Aksu, Ü., "Towards a business process complexity analysis framework based on textual data and event logs", in: *Proceedings of the 17th International Conference on Wirtschaftsinformatik (WI2022), Nürnberg, Germany, February 21-23, 2022*, AIS eLibrary, 2022 (cited on page 17).

[64] Rizun, N., Revina, A. & Meister, V. G., "Analyzing content of tasks in business process management. Blending task execution and organization perspectives", *Computers in Industry*, vol. 130, 2021, p. 103463 (cited on pages 17, 180).

[65] Revina, A., Buza, K. & Meister, V. G., "Designing explainable text classification pipelines: insights from IT ticket complexity prediction case study", in: *Interpretable Artificial Intelligence: A Perspective of Granular Computing*, vol. 937, Springer, 2021, pp. 293–332 (cited on page 17).

[66] Meister, V. G., Hu, W., Revina, A., Cikus, M. & Müller, J., "Fachanwendung für digitale Modulkataloge. Eine Untersuchung zu Graph-basierter Daten-Modellierung und Navigation", in: *Proceedings of the 50. Jahrestagung Der Gesellschaft Für Informatik, Hochschule 2030, Karlsruhe, Germany, September 29 - October 1, 2020*, Gesellschaft für Informatik, 2020, pp. 467–480 (cited on page 17).

[67] Kalabina, E., Belyak, O., Meister, V. G. & Revina, A., “What kind of employees’ team is necessary for industrial digital transformation? Theoretical and practical analysis”, in: *Proceedings of the Digital Transformation in Industry: Trends, Management, Strategies (DTI2020), Ekaterinburg, Russia, November 27, 2020*, Springer, 2021, pp. 183–193 (cited on page 17).

[68] Belyak, O., Kalabina, E., Revina, A. & Meister, V. G., “Comparative analysis of the knowledge sharing process in cross-functional teams: management practices in Russia and Germany”, in: *Proceedings of the 7th International GSOM Emerging Markets Conference, Saint Petersburg, Russia, November 11-21, 2020*, 2020, pp. 167–171 (cited on page 17).

[69] Salminen, J., Mustak, M., Rizun, N., Revina, A., Nikiforova, A., Almerekhi, H., Jung, S. gyo & Jansen, B. J., “Integrating AI into customer service: improving the actionability of customer feedback analysis using machine learning”, in: *Proceedings of the AIRSI2022: Technologies 4.0 in Tourism, Services and Marketing, University of Zaragoza, Spain, July 11-13, 2022*, 2022, pp. 175–179 (cited on page 17).

[70] Salminen, J., Mustak, M., Rizun, N., Revina, A., Nikiforova, A., Almerekhi, H., Jung, S. gyo & Jansen, B. J., “Integrating AI into customer analysis deep learning experiments on customer feedback prioritization in social media”, *Technological Forecasting & Social Change*, 2022, submitted to the journal (cited on page 18).

[71] Revina, A., Kalabina, E. & Belyak, O., “Employee engagement and soft skills in the digitalization of the economy. Preliminary study results”, in: *Proceedings of the Digital Transformation in Industry: Trends, Management, Strategies (DTI2022), Ekaterinburg, Russia, October 28, 2022*, 2022 (cited on page 18).

[72] Revina, A., Kalabina, E. & Belyak, O., “Do employee soft skills influence their engagement in organizations? Insights from industry and academia”, in: *Proceedings of the 9th International GSOM Emerging Markets Conference, Saint Petersburg, Russia, October 5-8, 2022*, in press, 2022 (cited on page 18).

[73] Rizun, N., Revina, A. & Edelmann, N., “Application of text analytics in public service co-creation: literature review and research framework”, in: *Proceedings of the 24th Annual International Conference on Digital Government Research - Together in the unstable world: Digital government and solidarity, Gdańsk, Poland, July 11-14, 2023*, submitted to the conference, 2023 (cited on page 18).

[74] Pennebaker, J. W., Boyd, R. L., Jordan, K. & Blackburn, K., *The development and psychometric properties of LIWC2015*, tech. rep., University of Texas, Austin, 2015 (cited on pages 19, 182).

[75] Chambers, A. J., Stringfellow, A. M., Luo, B. B., Underwood, S. J., Allard, T. G., Johnston, I. A., Brockman, S., Shing, L., Wollaber, A. & VanDam, C., “Automated business process discovery from unstructured natural-language documents”, in: *Proceedings of the 18th International Conference on Business Process Management (BPM2020), Seville, Spain, September 13–18, 2020*, Springer, 2020, pp. 232–243 (cited on page 20).

[76] Bellan, P., Dragoni, M. & Ghidini, C., "Process extraction from text: state of the art and challenges for the future", *ArXiv preprint*, 2021, 2110.03754 (cited on page 20).
[77] López, H. A., "Challenges in legal process discovery", in: *1st Italian forum on Business Process Management (ITBPM@BPM2021), Rome, Italy, 10 September, 2021*, CEUR, 2021, pp. 68–73 (cited on page 20).
[78] Aalst, W. M. P. van der, *Process mining - data science in action*, 2nd ed., Springer, 2016 (cited on pages 20–22).
[79] Hayes, B., Curtiss, S., Szabolcsi, A., Stowell, T., Stabler, E., Sportiche, D., Koopman, H., Keating, P., Munro, P., Hyams, N., et al., *Linguistics: an introduction to linguistic theory*, John Wiley & Sons, 2013 (cited on page 21).
[80] Cabré, M. T., *Terminology: theory, methods, and applications*, vol. 1, John Benjamins Publishing, 1999 (cited on page 21).
[81] Cook, G., *Applied linguistics*, Oxford University Press, 2003 (cited on page 21).
[82] Kao, A. & Poteet, S., *Natural language processing and text mining*, Springer Science & Business Media, 2007 (cited on page 21).
[83] Cancho, R. F. i & Solé, R. V., "Least effort and the origins of scaling in human language", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, 2003, pp. 788–791 (cited on page 22).
[84] Trewatha, R. & Newport, G., *Management: functions and behavior*, Business Publications, 1976 (cited on page 22).
[85] Marques, G., Gourc, D. & Lauras, M., "Multi-criteria performance analysis for decision making in project management", *International Journal of Project Management*, vol. 29, no. 8, 2011, pp. 1057–1069 (cited on page 22).
[86] Galup, S. D., Dattero, R., Quan, J. J. & Conger, S., "An overview of IT service management", *Communications of the ACM*, vol. 52, no. 5, 2009, pp. 124–127 (cited on page 22).
[87] Perry, D. E., Porter, A. A. & Votta, L. G., "Empirical studies of software engineering: a roadmap", in: *Proceedings of the Conference on the Future of Software Engineering*, 2000, pp. 345–355 (cited on page 170).
[88] Easterbrook, S., Singer, J., Storey, M.-A. & Damian, D., "Selecting empirical methods for software engineering research", in: *Guide to Advanced Empirical Software Engineering*, Springer, 2008, pp. 285–311 (cited on page 170).
[89] Runeson, P. & Höst, M., "Guidelines for conducting and reporting case study research in software engineering", *Empirical Software Engineering*, vol. 14, no. 2, 2009, pp. 131–164 (cited on page 171).
[90] Global Knowledge, *2020 15 top-paying certifications*, 2021, URL: https://www.globalknowledge.com/us-en/resources/resource-library/articles/2020-15-top-paying-certifications/#gref (visited on 09/01/2022) (cited on pages 171, 173).
[91] Paramesh, S. & Shreedhara, K., "Automated IT service desk systems using machine learning techniques", in: *Data Analytics and Learning*, Springer, 2019, pp. 331–346 (cited on pages 171, 173).
[92] Saaty, T. L., "The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making", in: *Multiple Criteria Decision Analysis*, Springer, 2016, pp. 363–419 (cited on page 172).

[93] Vom Brocke, J. & Mendling, J., *Business process management cases. Digital innovation and business transformation in practice*, Springer, 2018 (cited on page 174).

[94] Byron, K. & Thatcher, S. M. B., "Editors' comments: "What I know now that I wish I knew then" - teaching theory and theory building", *Academy of Management Review*, vol. 41, no. 1, 2016, pp. 1–8 (cited on page 174).

[95] Fiske, S. T., "Mind the gap: in praise of informal sources of formal theory", *Personality and Social Psychology Review*, vol. 8, no. 2, 2004, pp. 132–137 (cited on page 174).

[96] Andrews, R., Wynn, M., Hofstede, A. H. ter, Xu, J., Horton, K., Taylor, P. & Plunkett-Cole, S., "Exposing impediments to insurance claims processing", in: *Business Process Management Cases, Digital Innovation and Business Transformation in Practice*, Springer, 2018, pp. 275–290 (cited on page 175).

[97] Van Looy, A. & Rotthier, S., "Kiss the documents! How the City of Ghent digitizes its service processes", in: *Business Process Management Cases, Digital Innovation and Business Transformation in Practice*, Springer, 2018, pp. 187–204 (cited on page 176).

[98] Ohno, T. & Bodek, N., *Toyota production system: beyond large-scale production*, Productivity Press, 2019 (cited on page 176).

[99] Suchy, J., Suchy, M., Rosik, M. & Valkova, A., "Automate does not always mean optimize: case study at a logistics company", in: *Business Process Management Cases, Digital Innovation and Business Transformation in Practice*, Springer, 2018, pp. 463–483 (cited on page 177).

[100] Mitra, S. & Jenamani, M., "Helpfulness of online consumer reviews: a multiperspective approach", *Information Processing & Management*, vol. 58, no. 3, 2021, p. 102538 (cited on page 179).

[101] McGilvray, D., *Executing data quality projects: ten steps to quality data and trusted information (TM)*, Academic Press, 2021 (cited on page 180).

[102] Madnick, S. E., Wang, R. Y., Lee, Y. W. & Zhu, H., "Overview and framework for data and information quality research", *Journal of Data and Information Quality*, vol. 1, no. 1, 2009, pp. 1–22 (cited on page 180).

[103] Acton, T., Halonen, R., Conboy, K. & Golden, W., "DeLone & McLean success model as a descriptive tool in evaluating the use of a virtual learning environment", in: *Proceedings of the International Conference on Organizational Learning, Knowledge and Capabilities, NUI Galway, Amsterdam, 2009*, 2009, pp. 1–16 (cited on page 180).

[104] Laumer, S., Maier, C. & Weitzel, T., "Information quality, user satisfaction, and the manifestation of workarounds: a qualitative and quantitative study of enterprise content management system users", *European Journal of Information Systems*, vol. 26, no. 4, 2017, pp. 333–360 (cited on page 180).

[105] Rego, A. et al., "Complexity, simplicity, simplexity", *European Management Journal*, vol. 28, no. 2, 2010, pp. 85–94 (cited on page 181).

[106] Collinson, S & Jay, M, *From complexity to simplicity: unleash your organisation's potential*, Palgrave Macmillan, 2012 (cited on page 181).

[107] Brocke, J. vom, Baier, M.-S., Schmiedel, T., Stelzl, K., Röglinger, M. & Wehking, C., "Context-aware business process management", *Business & Infor-*

mation Systems Engineering, vol. 63, no. 5, 2021, pp. 533–550 (cited on page 182).

[108] Rosemann, M., Recker, J. & Flender, C., "Contextualisation of business processes", *International Journal of Business Process Integration and Management*, vol. 3, no. 1, 2008, pp. 47–60 (cited on page 182).

[109] Teinemaa, I., Dumas, M., Maggi, F. M. & Francescomarino, C. D., "Predictive business process monitoring with structured and unstructured data", in: *Proceedings of the 14th International Conference on Business Process Management (BPM2016), Rio de Janeiro, Brazil, September 18-22, 2016*, Springer, 2016, pp. 401–417 (cited on page 182).

[110] Smedt, J. D., Hasic, F., Broucke, S. K. L. M. vanden & Vanthienen, J., "Towards a holistic discovery of decisions in process-aware information systems", in: *Proceedings of the 15th International Conference on Business Process Management (BPM2017), Barcelona, Spain, September 10-15, 2017*, Springer, 2017, pp. 183–199 (cited on page 182).

[111] Bazhenova, E., Zerbato, F., Oliboni, B. & Weske, M., "From BPMN process models to DMN decision models", *Information Systems*, vol. 83, 2019, pp. 69–88 (cited on page 182).

[112] Leewis, S., Smit, K. & Zoet, M., "Putting decision mining into context: a literature study", in: *Digital Business Transformation*, Springer, 2020, pp. 31–46 (cited on page 182).

Curriculum Vitae

Aleksandra Revina obtained a diploma in Linguistics and Technical Translation from the National Technical University of Sevastopol, Crimea, in 2007. From 2008 to 2012, she studied Business Administration and Linguistics at the University of Potsdam, Germany. In 2015, she successfully finished her master's study program with a focus on Business Process and Knowledge Management at the Brandenburg University of Applied Sciences (BUAS). Afterward, she worked as a research scientist and project manager in the research unit of a telecom company. At the end of 2017, she started her industrial Ph.D. at the Technical University of Berlin and BUAS in a cooperative procedure. She is currently working as an academic and research staff at BUAS.

Her research interests include but are not limited to diverse methods and tools for business process analysis and automation from such subject fields as Business Information Systems, Business Process Management, Text Analytics, and Linguistics. The ultimate goal is to develop efficient decision-making support for process workers. During her Ph.D. research, Aleksandra (co-)authored and presented more than 17 papers at renowned international conferences and published six articles, including one book chapter, in peer-reviewed international academic journals with high impact factor. She successfully completed her Ph.D. in January 2023 with a scientific defense resulting in an overall decision "with distinction" (summa cum laude).