Till Plumbaum, Funda Klein-Ellinghaus, Anna Reeske, Kristin Pelz, Frank Hopfgartner

Health Assistance for Immigrants

Chapter 3
Health Assistance for Immigrants

Till Plumbaum, Funda Klein-Ellinghaus, Anna Reeske, Kristin Pelz, and Frank Hopfgartner

Abstract Our personal health should be one of our main concerns but unfortunately, due to modern lifestyle, far too many people ignore their own well-being. Consequently, methodologies need to be developed that assist us in living a healthier life. In this chapter, we present a health assistance system for immigrants. The system consists of two parts: A health information system that allows users to search for health information using natural language queries composed of multiple languages and a prevention service that assists users in their cooking routine and motivates them to perform frequent physical exercises. The information system uses NLP techniques to understand the user query, matches it to a health ontology we developed, and offers the user a comprehensive answer. The prevention service is embedded in a smart home environment. We present the technical details of both systems and show a user study to demonstrate that the system works well in providing highly relevant health information.

Till Plumbaum
Technische Universität Berlin, e-mail: till.plumbaum@dai-labor.de

Funda Klein-Ellinghaus
Leibniz Institute for Prevention Research and Epidemiology, e-mail: kleinelf@bips.uni-bremen.de

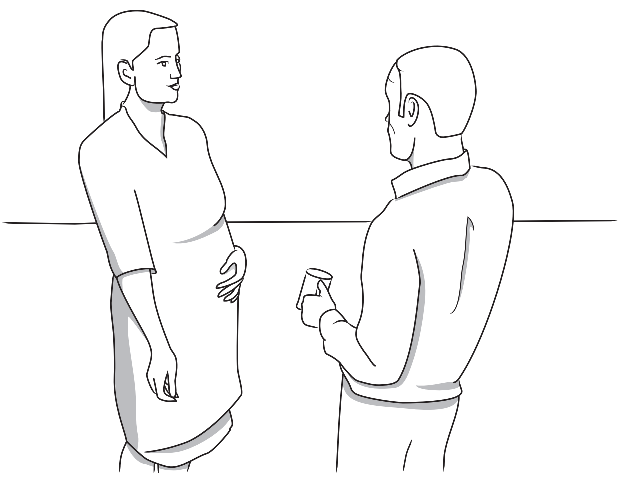
Anna Reeske
Leibniz Institute for Prevention Research and Epidemiology, e-mail: reeske@bips.uni-bremen.de

Kristin Pelz
AOK Bundesverband, e-mail: kristin.pelz@bv.aok.de

Frank Hopfgartner
Technische Universität Berlin, e-mail: frank.hopfgartner@tu-berlin.de

Towards a Healthy Life

There were two things that Steven liked a lot about his work: Free coffee and the view from the kitchen window. The office kitchen was in the thirteenth floor, the coffee machine right next to the window. Indeed, it happened quite a few times that he caught himself just gazing outside, observing the tiny looking people passing by in their tiny little cars. “We really live in a world of constant motion...”, he thought while he waited for this Latte Macchiato to come out of the coffee machine. “Good afternoon, Mr. Marks.” Steven turned away his eyes from the street outside to check who just welcomed him. It was Selma, the young new employee from Turkey. Steven turned around and straightened up. “İyi günler, Selma. Nasılsın?” – Good morning. How are you? These were pretty much the only few words he could say in Turkish. “Teşekkürler, iyiyim. Sen?”, Selma returned his greeting. “I am sorry, Selma, but I am way behind with my Turkish lessons. I am fine, if that’s what you have asked me?”, he responded, switching to English. Selma smirked. “Yes, I was asking how you are these days.” Steven “Okay, that’s an important sentence to know. I will try to remember it.” But you know, I am not the youngest any longer, picking up new languages is very hard.”



Of course Selma knew how hard it is to learn new languages. After all, it was only months ago that she and her husband decided to leave their beloved home in Ankara to start a new life in Germany. Although she enjoyed living in Ankara, the job offer that she received from here was just too good to reject. Working as a research engineer for a car manufacturer. Developing the cars of tomorrow. When they saw the job advertisement, both her husband and Selma were hooked. Cars had been their passion for a very long time. In fact, they first met each other at a motor show. And when they eventually got the job offer – both of them – there was no holding back anymore. In record time, they said good-bye to sunny Turkey and moved to Germany.

The first problems started shortly after though. Only a few weeks after settling in her new home, Selma started to feel sick in the morning and hence decided to see a medical doctor. His diagnosis was a little shock for them. Selma was pregnant! Her husband and she had tried for quite a while now to become pregnant but the idea of conceiving a child in a foreign country scared her. So many questions popped up in her mind: Would she have to change her lifestyle to ensure that the baby can develop properly? Would she have to give up her job now or would she be able to get financial aid from the state? After all, she didn't know anything about Germany and its health care system.

In fact, it was Steven Marks who managed to reduce her panic level significantly. Being the deputy manager of the research department, he was one of the first people at work that she informed about her pregnancy. "Don't worry Selma", he said back then. "You can get paid maternity leave and have a guaranteed right to return to work after that. There should also be courses and information services that your health insurance provider provides." Indeed, only a few days after she reported her pregnancy to her health insurance provider, they send her a brochure about their online health care information system. According to their brochure, the system was designed to provide information in different languages and should help her in answering most pregnancy-related questions that would arise in the next few months. She tried out the system on the same day. Being able to query in Turkish, German, or English, the system provided her a lot of information about pregnancy in general and information about services she could participate in. Given that all information is provided in her mother tongue Turkish, Selma grew confidence that they would manage their new life as young parents in this unfamiliar country.

Selma smiled. "Yes, I know very well, Mr. Marks", she responded and reached for a bottle of water from the fridge...

3.1 Introduction

Nowadays, more and more people suffer from so-called "Western diseases", i.e., health conditions that are caused by lack of exercise, poor eating habits and unhealthy lifestyles [31]. While these diseases of affluence are a serious threat to our personal well-being, they are even more critical for pregnant women who are not only responsible for their own health, but also for the health of their unborn children. Addressing the need for healthier living, large education campaigns and seminars have been funded to promote healthier lifestyles, advertising low-fat diets, weight management or physical exercises. With the increasing availability of the Internet, health care providers and governments extended the education campaign to the WWW, offering health information portals for multiple topics online. Two conclusions can be drawn from this provision of information. First of all, it is of high importance to provide accurate information, i.e., the content should be created by professionals and evaluated for accuracy [10, 9]. Secondly, such information sources should not be used for self-diagnosis; their only purpose should be to allow patients to better understand their physician's diagnosis. Although health care providers intend to provide health information services to all their clients, immigrants have been identified as vulnerable population [6] that benefit less from existing health care systems since language and cultural barriers prevent them from using existing prevention services. This is a problem especially for countries with significant numbers of immigrants such as Germany, where roughly 20% of the population has an immigrant background [38, 34] and consequently, a potentially large group of citizens may not able to understand the main language of the host nation.

Apart from providing users easy access to health-related content, an equally important approach to support people in living a healthier life is to actively educate them. In the context of food consumption, this is done via nutrition facts labels that are legally required on most packaged food. Although different regulations exists in individual countries on what has to be written on these labels, they serve as guidelines on different targets for nutrients such as energy, protein, or fat. A study on the effects of these nutrition fact claims on consumer product evaluation is presented by Keller et al. [19]. Although providing overviews of ingredients helps raising awareness about individual eating habits, the share of people who cook their own food is declining. At the same time, we can observe a significant growth of the ready-meals market. Van der Horst et al. [42] study the association between overweight, cooking skills and ready meal consumptions. Their study illustrates the importance to actively promote healthy food preparation.

The third important aspect of healthy living is frequent physical exercises. Various studies (e.g., [44]) report a direct connection between physical activities and personal well-being. Regular physical activity is a resource for body and soul [39]. On the one hand, a physical active lifestyle can contribute to reduce the risk of cardiovascular diseases, obesity and complaints of the muscular and skeletal system [1]. On the other hand, regular physical exercise can reinforce the mental wellbeing. The World Health Organization describes lack of exercise as the fourth important risk factor for mortality [46]. They recommend adults a medium intense activity of 2, 5 hours per week. As mentioned before though, lack of physical activity is one of the main consequences of sedentary lifestyle.

In order to address the issues mentioned above, we introduce a health assistant for immigrants [29] that consists of two subsystems, namely a multilingual health information service and a prevention service. Figure 3.1 illustrates the connection between these individual components. Combined, the services are a comprehensive approach to support people for healthier leaving by giving them information about health topics, supporting healthier eating and getting enough exercises.

The multilingual health information service [28] guarantees personalized access to professionally created health care content. The system, in the remainder of this chapter referred to as GID, enables immigrants to inform themselves about medical conditions and preventive health services. By providing them with available information in the language of their choice, GID helps those people who have language-

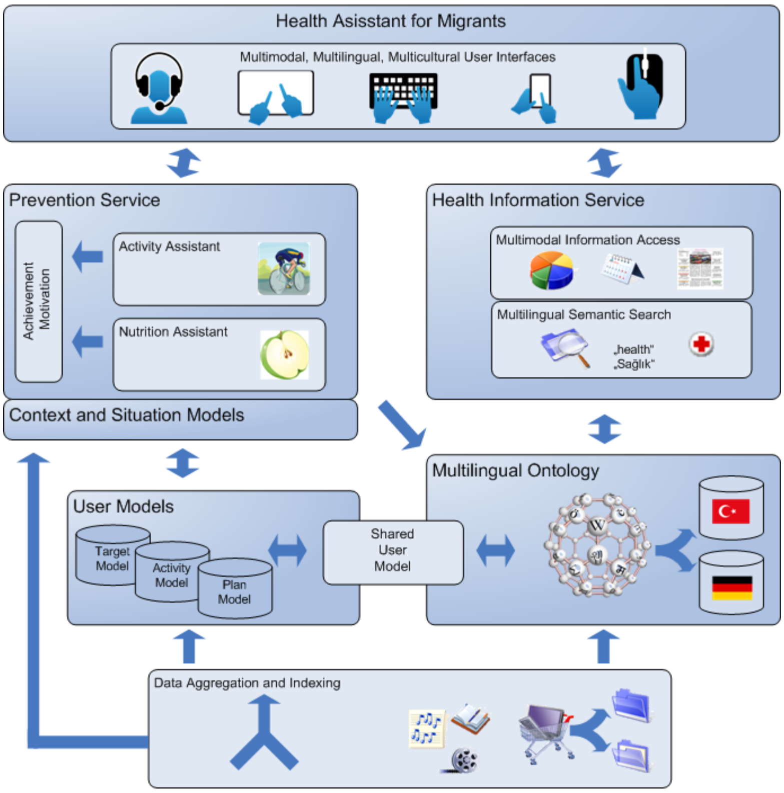


Fig. 3.1: System overview of the health assistant and its parts.

related difficulties in understanding their physicians. Information is adapted based on users' personal context such as pre-existing medical conditions and their location.

The prevention service, referred to as PS, consists of a cooking assistant and a virtual trainer. The cooking assistant provides the user a wide range of different recipes and nutrition information about different meals. There is also a possibility to personalize the recipe search with the aid of different criteria. The trainer supports the physical activity of the user with tree different activity sessions.

This chapter is structured as follows. In Section 3.2, we discuss the health services and survey existing work that is related to our system. In Section 3.3, we introduce the system infrastructure of the health information system and outline the scientific challenges that it addresses. We then present in Section 3.4 an initial evaluation of the system. Following that, we present in Section 3.5 the prevention service. A discussion and conclusion is given in Section 3.6.

3.2 Related Work

This work touches upon various research areas, including the use of online health information services, knowledge bases, the identification of user context, semantic information management, healthy nutrition and physical activity. In the remainder of this section, we discuss related work on these research aspects.

3.2.1 Online Health Information Services

With the growing importance of the Internet, one could also witness the growth of online health-oriented information platforms such as PubMed[1] and Scopus[2], which enable researchers and professionals to check up on latest research results on biomedical topics [7]. While these databases provide detailed access to state-of-the-art research results, they are less suitable for the general public who would like to check up on symptoms that they are experiencing. According to Morahan-Martin, up to 4.5% of all Internet searches are about health-related topics [24], indicating the significance of this topic in our life. An overview of different resources in the Web that can be used for this information gathering task is provided by Johnson et al. [18]. Generally, three types of services can be identified: (a) Professionally maintained health advice and information services where users can check their symptoms in a constantly updated database (e.g., the NHS Direct[3] service maintained by the English National Health Service) or check up on public service announcements (e.g., by the World Health Organization), (b) unsupervised sources such as Wikipedia and (c) discussions of similar cases in online forums, blog posts, or biased advertisements for specific products that can be retrieved by standard Web search engines. From a medical point of view, relying on such sources is not recommended and the consultation of a professional is highly advised. Therefore, the American Medical Library Association recommends to "trust your physician, not a chat room" [8]. Morahan-Martin suggests to approach this unconsidered information handling by asking physicians to point their patients to reliable health portals and to work on improving such sites, e.g., by improving search and retrieval techniques. Our works builds on her suggestion.

[1] `http://www.ncbi.nlm.nih.gov/pubmed/`
[2] `http://www.scopus.com/`
[3] `http://www.nhsdirect.nhs.uk/`

3.2.2 Computer-Supported Knowledge Bases

In the healthcare domain, an exact and unambiguous definition of diseases, symptoms, etc. is indispensable. Only when it is clear what the problem (disease) is, physicians can say how to treat it. Given the complex nature of this topic, physicians have to rely on an extensive knowledge base. Traditionally, this knowledge base has been maintained in books and journals but since the introduction of computer-based knowledge systems, methods have been developed that support physicians in the anamnesis process. A promising method that allows for a structured processing of data is the use of ontologies. Ontologies define "the concepts, relationships, and other distinctions that are relevant for modeling a domain" [11]. Ontologies, in the context of computer and information sciences, are also machine-readable and sharable. Thus, ontologies represent an ideal basis for an intelligent healthcare information system. In the following, we present some approaches to create a unified model in the health domain

One of the biggest computer-supported biomedical knowledge bases is the Unified Medical Language System (UMLS)[4], a platform that provides a unified vocabulary for biomedical and health terms. UMLS is maintained by the US National Library of Medicine and updated on a frequent basis. The knowledge base of UMLS consists of three components: (a) Biomedical terms from various controlled vocabularies (such as SNOMED-CT[5]) are defined in a Metathesaurus as semantic concepts. (b) The semantic relationship between these concepts are defined in a Semantic Network and the concepts are categorized into broad categories (semantic types). (c) Syntactical, orthographic and morphological information about the biometric terms are defined in a lexicon. UMLS has a strong focus on the US health market, and hence, multilingual aspects are not fully supported, making it difficult to apply it to the scenario that is outlined in this chapter.

Other health-related ontologies include MEDCIN and SNOMED-CT, two classification systems used to store patient health records. The International Health Terminology Standards Development Organisation (IHTSDO) promotes SMOMED-CT as a standard for health records. The aforementioned approaches define common vocabularies and their relation to each other for the medical domain, but they define the terms in a proprietary format that is not machine-readable and shareable. These constrains make re-using and sharing of data complicated though. OpenGALEN [32, 37] is a medical ontology developed in the European GALEN project. OpenGalen offers a comprehensive knowledge base of medical terms and relations. It provides three types of ontologies: A high level ontology, defining general structures, a common reference model defining the re-usable parts intended to be shared between ontologies, and extensions for sub-domains and specific use cases. The GALEN ontology is available under an open source license.

The presented approaches all contribute to the goal of a common, shareable and re-usable notion of the medical domain for building health related applications. Nev-

---

[4] `http://umlsinfo.nlm.nih.gov/`

[5] `http://www.ihtsdo.org/snomed-ct/`

ertheless, none of these approaches fulfill the requirements to serve as the basis for a multilingual health information system. In particular, they lack modeling of multilingual descriptions and a connection to related health information is not available.

3.2.3 Identification of user context

An important feature of online information systems is to adapt information based on users' individual requirements [17]. Requirements can depend on different factors such as the users' demands and context. In the health domain, context such as health condition of the user and their location plays an important role. For instance, recommendations for a healthy diet for a person with Type 2 diabetes and a pregnant person differ significantly and thus should be reflected in the information that is tailored to the users' needs.

Tailoring information based on these user contexts requires storage of user-centric information. In order to easily connect this information with ontology-based knowledge bases, i.e., for providing a personalized healthcare system, a promising approach is to model information using user-centric (semantic) ontologies. Semantic-based user ontologies are for instance:

- FOAF: The Friend-of-a-Friend model is a RDF-based user model mainly designed for the web. It defines demographic data such as name, age and friendship relations. Common applications of FOAF are social networks services [3].
- SWUM: The Semantic Web User Model defines a comprehensive user model designed for the needs of the modern social semantic web. It models information about the users' demographics, friendship, etc. [30].
- GUMO: The General User Modelling Ontology tries to provide a user model covering all aspects of life. GUMO models health related information such as blood pressure or temperature [15].

In this work, we introduce an approach to map users' demands (as expressed by the search query) with the computer-based knowledge basis under consideration of the context that is stored within a semantic-based user ontology, thus presenting a healthcare information system that can provide answers based on context.

3.2.4 Multilingual Semantic Information Management

As argued above, one main challenge in the described scenario is to deal with different languages that might be relevant to provide health-oriented information, e.g., the difference between the patients' mother tongue (*Language A*) and the language used by the physician (*Language B*). Multilingual information management has been studied extensively in literature. A straightforward approach to cross-lingual information management is either a direct translation approach [26] or using an interlingual mapping like EuroWordNet [43], which have been shown to work well at the CLIR task [5]. Imprecise translations are acceptable for the retrieval performance, as the document search itself is more important than disambiguation of individual translated terms [16]. Other approaches map individual documents into a higher dimension semantic feature space that is uniform for different languages. Thus, it is possible to map similar documents to nearby points in the feature space, even if they do not share a language. Sorg and Cimiano [33] have exploited the explicit links between related Wikipedia articles of different languages to map documents to a Wikipedia feature space in which documents are considered similar when they are semantically similar to the same set of articles. A similar approach is to map documents to multilingual ontology concepts, which can be represented as points in a feature vector space [12]. In this work, we employ such ontology mapping method to link concepts expressed in different languages.

3.2.5 Graph-Based Search

The advantage of ontology-based knowledge bases is that they can be exploited to identify related concepts. A common approach to identify these related concepts is to apply graph-based search algorithms. These algorithms infer links between ontology nodes that are not explicitly stored in the ontology (e.g., [36]). Most graph search algorithms (e.g., [40, 41]) are based on breadth-first search or depth-bounded depth-first search. These algorithms find related nodes by calculating a relatedness score to the input nodes. A high relatedness means that the nodes can be reached by several parallel short paths [21]. The scoring function should consider the specific properties of the ontology domain, taking into account edge weights and semantics while computing the path weights. The nodes with the highest relatedness score are considered to be most similar to the input nodes. Within this work, we apply graph-based search to retrieve relevant health information based on the users' search query and their personal context.

3.2.6 Healthy Nutrition

Unhealthy nutrition is one of the main reasons for diseases of affluence and consequently, there is a grand need to convince people to choose healthy food rather than convenience food. Unfortunately, people are constantly facing advertisement campaigns whose main purpose is not to sell healthy products [13]. Active measures against this situation are information campaigns. This includes approaches such as the traffic light rating system, i.e., the regulation that food producers and providers have to clearly state on their product how healthy their product is. Other approaches include marketing campaigns which are often financed by the government or health insurance companies. While these are considered to be rather passive information campaign, another approach is to actively assist the people to prepare their food. This can be done in form of cooking classes where exclusively healthy food options are taught [4], but also in the form of software systems (e.g., [14]) that take over the task of the chef instructor. In this work, we introduce a software-based nutrition assistant that assists users in preparing healthy food.

3.2.7 Physical Activities

Various studies suggest that personal achievement is one of the main driving forces behind sports activities. Nicholls [25] argues, for example, that one of the main reasons for the success of competitive sports such as running, tennis, or swimming is the possibility to directly compare one's physical abilities with others. Another motivation is to experience (and to extend) physical limitations. This can in particular be observed in extreme sports such as bungee jumping, base diving, or other dangerous sports. He refers in this context to task-oriented and ego-oriented sports. In both cases, individual achievement, either by outperforming others or by reaching new limits, is the main reason to perform sports. Hence, in order to motivate people to get more physically active, personal aims need to be identified and targeted. Members of the Quantified Self movement rely on the power of numbers to measure personal achievements. By recording their physical activities using step counters, accelerometers, or other wearable sensors, people can directly measure how far they are from reaching their personal aims (e.g., [35]). A promising method to help people in determining what to achieve is to rely on badges, points, or leaderboards, i.e., principles that have shown to be very successful in games (e.g., [45, 23]). The use of these principles in a non-gaming environment is commonly referred to as gamification.[6] Within this work, we aim to motivate people to get physically active by providing a virtual trainer. This trainer teaches the user how to perform various exercises. Whenever a user successfully repeats an activity, they receive points as an indication of personal achievement. Moreover, a personal activity list is created, thus allowing users to compare their achievements with other users of the system.

3.3 Semantic Health Information System

The next sections describe the main components of the multilingual health information system. The system builds on semantic technologies to perform the task of multilingual information supply. We introduce the underlying ontology that allows us to generate a semantic knowledge base of health information. Furthermore, we introduce the data acquisition and management tasks, the user modeling technique, supported querying modes and introduce the user interface.

[6] A detailed overview of gamification is provided in Chapter 9.

3.3.1 The Health Ontology

In order to provide a knowledge base that can easily be maintained by a computer system, we define a simple health ontology (HO) that defines basic health concepts and their relations. Figure 3.2 provides an overview of the ontology, detailed descriptions of the concepts and relations are listed in Table 3.1 and Table 3.2. The concepts in the Health Ontology are enriched with multilingual labels, e.g., the concept *pregnancy* is labeled with the Turkish term "gebelik" and the German expression "Schwangerschaft". Besides, relevant documents are attached to the concept nodes. In a health care scenario, information quality is of crucial importance. False information does not only lead to a loss of trust but also may lead to serious harms. Therefore, information in the HO should be maintained by a group of experts. In the GID project, physicians maintained information in the ontology.

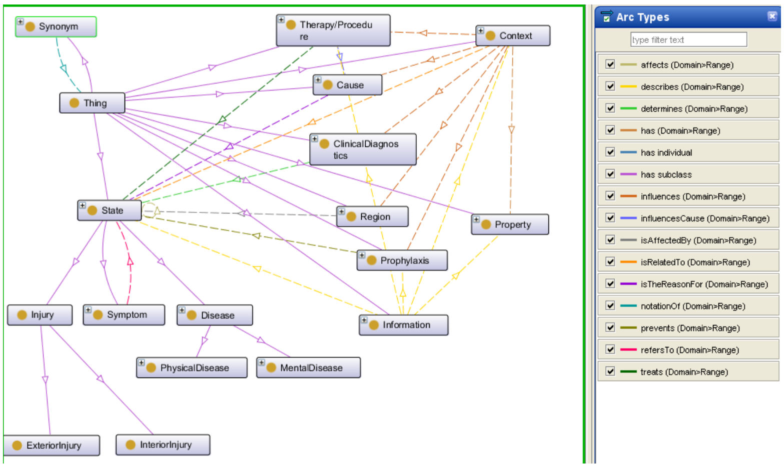


Fig. 3.2: The GID Health Ontology

3.3.2 User Model

As explained above, personalization services rely on the creation of user models to store personal information in a user profile. User profiles are then exploited to adapt information accordingly. In the health scenario, the most important factors are the users' demographic details (i.e., age, sex) and their personal context (e.g., their hometown, language knowledge and pre-existing medical conditions). In order to

Table 3.1: Concepts of the Health Ontology

Concept	Description
Context	This concept describes a composite information node composed of property nodes and state nodes. A context concept for instance can describe that certain therapies are only applicable for pregnant women.
Property	Is connected to a context node and specifies the context, for instance a certain pregnancy week.
Therapy/Procedure	This node describes possible therapies to cure a disease.
Cause	Cause nodes describe the reason for a medical condition, e.g., obesity is a cause for diabetes.
ClinicalDiagnostics	Describes an approach to make a diagnosis. For instance to take an x-ray of somebody.
Region	Part of the human body plus psych. Indicates where a disease or injury is situated.
Prophylaxis	This node describes prevention procedures, e.g., regular teeth-brushing helps preventing cavities.
Information	The information node is the node that is displayed to the user with information about the node it is connected too. This can be information about diabetes, sport, pregnancy or others. Information nodes can be a text or website, a video or picture.
State	State is the general node to describe a condition of the body. Sub-nodes are injury or disease.
Injury	Injury is a sub-node from State. It is splitted in inner and exterior injuries
InteriorInjury	Internal Bleeding is an example for an InteriorInjury node
ExteriorInjury	An abrasion is an example for exterior injuries.
Symptom	Symptom nodes define possible clinical signs that indicate a State.
Disease	Disease is a sub-node from State and is split into mental and physical disease.
MentalDisease	This node comprises all diseases connected to a human's psych. For example, depression is a MentalDisease.
PhysicalDisease	The PhysicalDisease node describes all body related diseases.
Synonym	The synonym node is important for all multilingual aspects of the described system. It is explained later in Section Multilingual Semantic Search.

receive personalized health information, users of the GID system are required to provide above information.

3.3.3 Multilingual Semantic Search

As explained above, concepts in the health ontology are enriched with multilingual labels. When a user enters a search query, we process the entered text by lowercasing it and pruning unusual characters. The search terms are then used as a query to retrieve matching concept labels contained in the Health Ontology. We do not perform language-specific stemming or remove stop-words of the search input since

Table 3.2: Relations of the Health Ontology

Relation	Description
isTheReasonFor	Connects Cause and State nodes. For example, obesity is caused by bad eating habits.
determines	Connects ClinicalDiagnostics and State. Application: Obesity is determined by measuring the BMI.
affects	The affects edge defines effects that one State has on anther. For example, Obesity affects Diabetes.
prevents	The prevents edge connects the Prophylaxis and State node. A healthy diet prevents obesity.
refersTo	refersTo connects Symptom nodes and State nodes, e.g., breathing problems indicate obesity.
treats	The influences edge connects the context node with the Region, Prophylaxis, Therapy/Procedure, Cause and ClinicalDiagnostics node.
influences	This node describes prevention procedures, e.g., regular teeth-brushing helps preventing cavities.
isAffectedBy	Connects Region node with State node, such as Psyche is affected by depression.
describes	The describes edge connects information with the Property, Context, Prophylaxis, Therapy, State node. Thereby, we can add information (pdfs, web sites etc.) to a node.
influencesCause	Connects Therapy/Procedure with Cause nodes. E.g., the treatment of respiratory distress is dependent on whether the cause is asthma or bronchitis.
has	has connects Context and Property nodes. This allows defining a special context. For example, the state pregnant can be constrained with the Property 9th month, to indicate that a treatment is only allowed in the 9th month.
isRelatedTo	Connects context with state. See example above.
Symptom	Symptom nodes define possible clinical signs that indicate a State.
notationOf	Connects Synonyms with all type of nodes to add multilingual information.

this enables us to easily extend the GID system with additional languages without requiring any changes to program code. Rather, we use a fuzzy search based on the open-source information retrieval system Lucene[7] to cover slight alterations in term surface forms. Taking the multi-lingual search query "*Hamilelikte hangi besinleri yememeliyim?*" (Translation: "*Which food must I not eat when I am pregnant?*") as an example, the language-independent concepts *Pregnancy* and *Malformation* are identified, since they have been labelled using the keyword "pregnant". Furthermore, the concepts *Nutrition* and *Alcohol* are identified by the label "besinleri".

Using the retrieved query-relevant ontology concept nodes, we employ a graph-search to find conceptually related information nodes. Different concepts from ontology classes like diseases, diagnostics, and treatments are semantically linked with weighted edges in our HO. Our algorithm performs a full graph search along these edges, bounded in depth. Information nodes found during this traversal are ranked

[7] http://lucene.apache.org/

based on the proportion to the edge-weights of the path of the originating concept and anti-proportional to that path's length. Information nodes found via multiple paths receive the sum of the relevance values of all of those paths. Using the example query above, we find amongst others an information node concerning *Alcohol Consumption during Pregnancy*. The information node is ranked highest because it is linked very closely to all of the three health concepts found for our keywords. As mentioned above, information nodes in the HO have documents attached to them, which are used as search results. These documents can be of any language, and have a language tag attached. Direct translations of documents are marked as copies. In order to gather search results for visualization, we collect the documents attached to the most relevant found information nodes during the graph search step and rank the list of documents depending on their relevance values combined with the user's language preferences. Depending on these values, it is possible for users to find documents in a different language near the top of the list, if the document is not available in the preferred language but is very highly relevant. Thus, users will always see a list of results balanced by relevance and their language preferences. The results on this list are independent of the language of the search query, since the results were found via a mapping to language-independent ontology concepts.

In above example, multiple documents of Turkish and German languages are attached to the information node *Alcohol Consumption during Pregnancy*. Assuming the user specified a Turkish language preference, the first search result presented to them would be the document "*Gebelikte alkol kullanımı*" (Translation: "*Alcohol consumption during pregnancy*").

3.3.4 Graphical User Interface

In addition to the search result list, the GID Web User Interface provides several supportive UI elements that help serve the users' information need and let them adjust the search. A screenshot is shown in Figure 3.3. In the remainder of this section, we introduce different features of this interface.

Entering the website, the user can log in to receive context-based search results. On top of the interface, the user can type in a search query. Since GID matches search terms with concepts in the health ontology, the search query can be formulated in different languages. In the screenshot, the user Selma has logged in and typed in the search query used in the example above: "*Hamilelikte hangi besinleri yememeliyim?*" (Translation: "*Which food must I not eat when I am pregnant?*") Under the search box, the interface lists the concepts that GID extracted from the search query. In the example screenshot, the following concepts have been detected from the search query: malformation (Biçimsizlik), pregnancy (Gebelik), non-alcoholic fatty liver disease (non-alkolik yağlı karaciğer hastalığı), and gout (Gut) in the context of nutrition (beslenme) and pregnancy (gebelik). Furthermore, the following context has been extracted from Selma's profile that will be considered when ranking search results: home town (Berlin) and pre-existing medical conditions (Diyabet

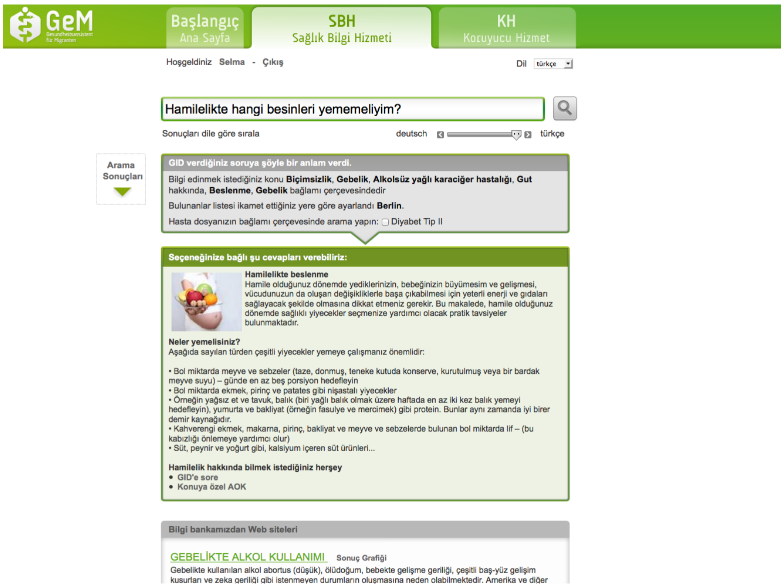


Fig. 3.3: Screenshot of the GID Search Interface

Tip II, Type II Diabetes). Under this visualization, the interface provides a topic box containing professionally edited information for the identified concepts. Under this topic box, the search results are displayed in descending order of relevance. The user can adjust their language setting by dragging a slider seamlessly between two languages. This sets a gradual preference of one language over the other and can exclude a language completely if dragged to the very end of the other language. Changing this setting immediately affects the search results and re-sorts the result list to reflect different language preferences. The user can seamlessly observe how search results are rearranged based on their language settings.

Each search result offers a button that when clicked shows a view of the paths of the Health Ontology that were traversed to find this search result. A screenshot is shown in Figure 3.4. The semantic graph allows the user to understand *why* this search result is relevant to their search query. It also makes transparent the underlying workings of the ontology. Additionally, the interface enables users to visually browse through the ontology concepts by expanding concept nodes and traversing semantic edges between concepts, and also finding documents attached to a concept's information nodes.

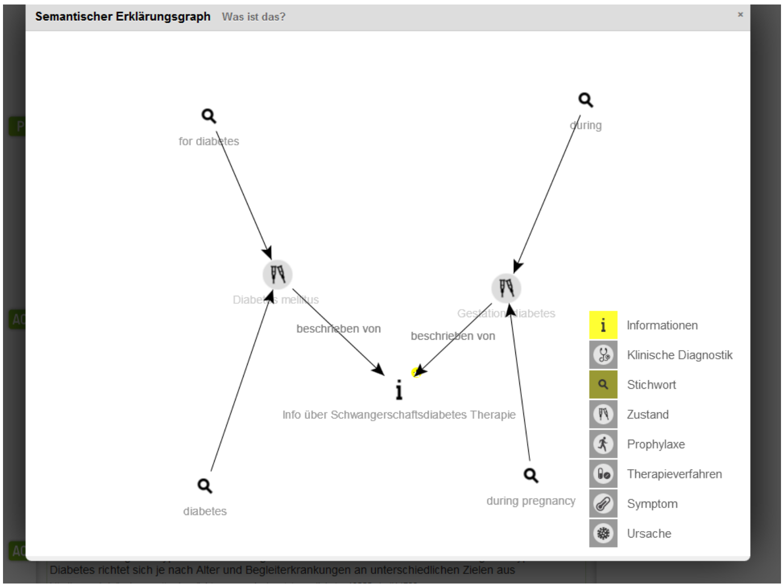


Fig. 3.4: Graph Search Explanation

3.4 Evaluation

In order to assess immigrants' expectations and needs for a personalized health information assistant, we conducted a demand analysis among Turkish migrants in a midsized town in Germany. In addition, we evaluated the development and implementation of the health assistant. Applying a qualitative research approach, we recruited a small study cohort from the local Turkish community as well as a German control group. In the second phase of the evaluation, they were asked to interact with the system in a supervised scenario and were further asked to assess the assistant with respect to its technical implementation, relevance of the presented content, usability in daily life and its potential to present health information in a structured manner. In particular, we focused on the following research questions: (a) How do the participants assess the adaptation of search results based on personal context? (b) How do they evaluate the system's handling of cultural differences? (c) How do they evaluate the usability of the graphical user interface (especially the ability to change language settings)? In the remainder of this section, we outline the set up of the user evaluation.

3.4.1 Participants and Recruitment

Building on experiences from former migrant studies [47], we applied a cultural-sensitive recruitment concept using key persons from existing social networks (family and youth centers) in order to establish the access to our study population. We recruited nine families with a Turkish migrant background and four German families as a control group. As suggested by Lamnek [20] and Pelz et al. [27], we invited a homogeneous focus group of six participants with similar socio-economic backgrounds from two different districts to participate in the user evaluation.

3.4.2 System Interaction

During the user-centric evaluation of the system, the study group was separated into a German-speaking group and a Turkish-speaking group due to limited knowledge of the German language of some participants. Both groups interactively tested the health assistant prototype in a supervised user-scenario. During the test, they were asked several questions concerning the design, usability and cultural sensitivity. The moderated focus group discussion was held in Turkish language. First, the participants had the opportunity to summarize their impressions on the health assistant. Second, the discussion was moderated according to our interview guideline and focused mainly on the usability and the potential for the implementation and use in daily life. We compiled the main messages in a mind map during the focus group discussion for discussion structuring and subsequent data analysis.

3.4.3 Evaluation Outcome

For data analysis, the bilingual interviews and the focus group discussion records were transcribed and translated. The data was analyzed using the method of content analysis by Mayring [22]. A summary of the group discussion is summarized in the remainder of this section.

How do the participants assess the adaptation of search results based on personal context?

Participants found the search results to be appropriate and clearly presented. It was suggested by some members of the group that the detailed answer box should contain specific advice on actions to be taken. Most participants found the adaptation of search results based on the context of a personal user profile useful and found that this feature helped in making an individualized search experience. However, some users were reluctant in providing certain pieces of profile information and wished to know specifically how they would be used.

How do they evaluate the system's handling of cultural differences?

Users wished for specific information about the Turkish governmental health system, especially compared to the German system.

How do they evaluate the usability of the graphical user interface (especially the ability to change language settings)?

Performing semantic search was found to be intuitive by a large majority of users in the study. Many users entered full health related questions, as intended, without needing any instructions. The feature to enter mixed-language queries was found to be very useful by all participants, and was considered to be of high importance for everyday use by the group. However, most users needed to be made aware of functionality, since it was not readily apparent from the interface itself.

Language preference settings were found to be especially helpful by people with little proficiency in the German language. At first however, the difference in functionality between the language preference slider and the language selection dropdown-box was not intuitive to some participants and needed to be explained.

Additional comments

Asking the participants for advice on how to improve the system further, they suggested that GID should also enrich their results with pictures and videos. They remarked that reading a lot of text makes interaction with GID unappealing. Additionally, we noted that the ontology graph helped younger users to understand how their search results were found, but did not help older users that much.

3.5 Prevention Service

The second service of the health assistant is the prevention service, referred to as PS. This service consists of two subsystems focusing on nutrition and activity support for people: a cooking assistant and an activity assistant. In this section, we first introduce the cooking assistant that uses structured knowledge about food, healthy eating habits and user information to build a personalized cooking and eating plan for a single user or a group of users. Additional services such as a food-shopping assistant complete a service that allows creating healthy eating behaviors. Then, we present the virtual trainer which uses 3D-camera techniques to track users doing exercises at home. To motivate users doing their sports, the trainer combines serious games to make it more fun with motivational parts such as the combination with the cooking assistant. If a user makes more sports, the cooking assistant gives positive feedback by allowing the user to choose 'unhealthier' menus.

3.5.1 Nutrition Assistant

The nutrition assistant, shown in Figure 3.5, helps users eat in a healthy way. Users can get a list of cooking recipes that are tailored to their physical activity, medical conditions and cultural background. For example, if the user has diabetes, recipes with high sugar content are avoided. Besides, the interface provides an overview on how healthy their current lifestyle is with respect to physical activities and nutrition habits. This helps users to learn about the food they eat.

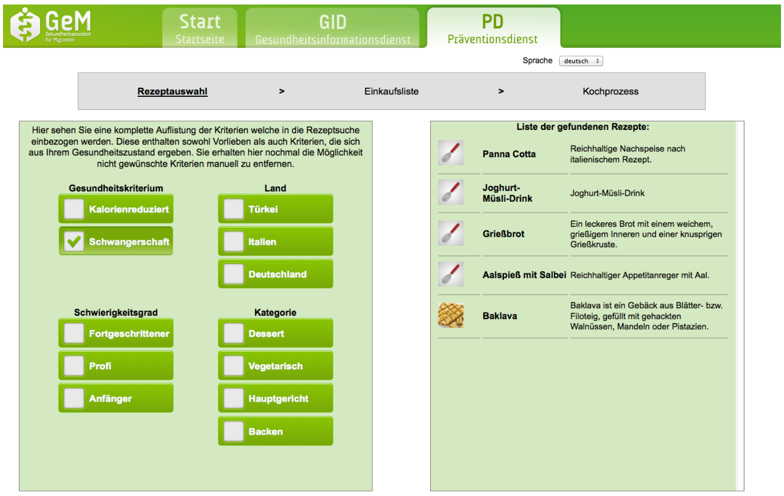


Fig. 3.5: Screenshot of the nutrition assistant.

The assistant provides the user with different cooking sessions. The user can select their meal from a range of different recipes. For the selection they can choose different health criteria, their country, the level of difficulty and the category of the dishes. In the category they can decide if they want a dessert, a main meal or a vegetarian meal. Furthermore they can select if they eat alone or with another family member. Therefore, they have the possibility to preset the likes and dislikes or diseases of the persons. When connected to a smart home, the system can also be used to control the domestic appliances such as the stove or the microwave. After the step-by-step cooking sessions the dinner will be ready. In some cases the cooking sessions are video-tailored. Furthermore the cooking assistant gives an overview of all ingredients needed for the meal. So the user has the possibility to create their own shopping list and can send it to one of the family members. The user also gets nutrition information about their meal so that they have an overview of the calories, protein, carbohydrates, sugar, fat and fibers.

3.5.2 Activity Assistant

The activity assistant intends to defeat one's weaker self. By using game mechanics and rewards, people are motivated to do physical exercises at home [2]. As shown in Figure 3.6, users see a digital trainer who demonstrates an exercise (e.g., Jumping Jacks) that should be imitated by them. The user has the possibility to select his trainer of three different figures – women, man or ogre. The trainer offers the user three different activity sessions. Users' motions are tracked using a XBOX Kinect 3D camera and compared to the instructed movements. During the training sessions the virtual trainer is giving a feedback, if the exercises are performed correct or not. For each activity, the users can earn activity points. The more exercises the users perform, the more activity points they earn. These points are directly fed into the nutrition assistant, i.e., the food suggestion depends on the users' individual energy expenditure.

3.6 Conclusion and Discussion

In this chapter, we presented a health assistant that addresses the specific needs of immigrants. The system consists of two parts: A health information system and a prevention service. The health information system provides a search facility on a semantic database to assist people in finding health-oriented information such as details about services provided by the health insurance provider answers for specific health-oriented questions. The system is mainly designed to assist immigrants with limited knowledge of the national language used in their host country in finding relevant information. Therefore, the system allows users to formulate search queries in their mother tongue, the host's language and in a mix of both languages. Addressing the specific immigration situation in Germany, we focused on the implementation of German and Turkish languages. Due to the structured data processing method that is introduced in this chapter, other languages can easily be incorporated. Search

Fig. 3.6: Activity Assistant: The user, shown in the upper right corner, is captured by the XBOX Kinect 3D camera and should follow the exercises of the trainer (left).

results can be adapted based on the users' preferred language and other personal contexts such as the users' location or pre-existing medical conditions. The underlying technology of this system is a health ontology that has been introduced in this chapter. Relevant information is retrieved by exploiting semantic relations between different concepts in this ontology. In order to evaluate the usability of this system, we followed a qualitative analysis scheme. Discussions in a focus group following this evaluation indicate that the system can be employed to assist immigrants to find information in their own and in the host nations' language.

The second part of the system, the prevention service, consists of two parts. The first part is a cooking assistant that, based on users' profiles, assists the users in selecting healthy food as well as in preparing the meal. Apart from providing detailed information about the ingredients of various dishes, the system recommends meals on the users' individual requirements. The system provides a step-by-step guide on how to prepare the meal and allows the user to control the home appliances needed for cooking. Moreover, the prevention service offers an activity assistant that motivates users to perform exercises in a fun setting. Choosing from different virtual trainers, users have to repeat physical exercises. Users can earn points by accurately performing these exercises.

Acknowledgements This work was funded by the Federal Ministry of Education and Research (BMBF) under funding reference number 01IS10055A-C.

References

1. Steven N. Blair, Harold W. Kohl, Ralph S. Paffenbarger Jr, Debra G. Clark, Kenneth H. Cooper, and Larry W. Gibbons. Physical fitness and all-cause mortality. a prospective study of health men and women. *JAMA*, 262:2395–2401, 1989.
2. Maged N. Kamel Boulos. Xbox 360 kinect exergames for health. *Games for Health Journal*, 1:326–330, 2012.
3. D. Brickley and L. Miller. Foaf vocabulary specification 0.91. namespace document, foaf project, 11 2007.
4. B.J. Brown and J.R. Hermann. Cooking classes increase fruit and vegetable intake and food safety behaviors in youths and adults. *J Nutr Educ Behav.*, 37, 2005.
5. P. Clough and M. Stevenson. Cross-language information retrieval using eurowordnet and word sense disambiguation. *Adv. in Information Retrieval*, pages 327–337, 2004.
6. KP. Derose, JJ. Escarce, and N. Lurie. Immigrants and health care: Sources of vulnerability. *Health Affairs*, pages 1258–1268, 09 2005.
7. Matthew E. Falagas, Eleni I. Pitsouni, George A. Malietzis, and Georgios Pappas. Comparison of pubmed, scopus, web of science, and google scholar: strengths and weaknesses. *The FASEB Journal*, 22(2):338–342, 2008.
8. S. Fox and L. Raine. Vital decisions: How internet users decide what information to trust when they or their loved ones are sick. plus a guide from the medical library association about smart health-search strategies and good websites., 2002.
9. O Kuss G Eysenbach, J Powell and ER Sa. Empirical studies assessing the quality of health information for customers on the world wide web: a systematic review. *The Journal of the American Medical Association*, pages 2691–2700, 05 2002.
10. C. Glenton, EJ. Paulsen, and AD. Oxmanm. Portals to wonderland: Health portals lead to confusing information about the effects of health care. *BMC Medical Informatics and Decision Making*, 03 2005.
11. Thomas R. Gruber. Toward principles for the design of ontologies used for knowledge sharing. *Int. J. Hum.-Comput. Stud.*, 43(5-6):907–928, 1995.
12. J. Guyot, S. Radhouani, and G. Falquet. Ontology-based multilingual information retrieval. *CLEF Workhop, Working Notes Multilingual Track*, pages 21–23, 2005.
13. Jason C.G Halford, Jane Gillespie, Victoria Brown, Eleanor E Pontin, and Terence M Dovey. Effect of television advertisements for foods on food consumption in children. *Appetite*, 42(2):221–225, 2004.
14. Reiko Hamada, Jun Okabe, Ichiro Ide, Shin'ichi Satoh, Shuichi Sakai, and Hidehiko Tanaka. Cooking navi: Assistant for daily cooking in kitchen. In *Proceedings of the 13th Annual ACM International Conference on Multimedia*, MULTIMEDIA '05, pages 371–374, New York, NY, USA, 2005. ACM.
15. Dominik Heckmann, Eric Schwarzkopf, Junichiro Mori, Dietmar Dengler, and Alexander Kröner. The user model and context ontology gumo revisited for future web 2.0 extensions. In Paolo Bouquet, Jrme Euzenat, Chiara Ghidini, Deborah L. McGuinness, Luciano Serafini, Pavel Shvaiko, and Holger Wache, editors, *Proceedings of the International Workshop on Contexts and Ontologies: Representation and Reasoning (CO:RR)*, volume 298 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2007.
16. D. Hiemstra and F. De Jong. Disambiguation strategies for cross-language information retrieval. *Research and Advanced Techn. for DL*, pages 852–852, 1999.
17. Frank Hopfgartner and Joemon M. Jose. Semantic user profiling techniques for personalised multimedia recommendation. *Multimedia Syst.*, 16(4-5):255–274, 2010.

18. Pamela T. Johnson, Jennifer K. Chen, John Eng, Martin A. Makary, and Elliot K. Fishman. A comparison of world wide web resources for identifying medical information. *Academic Radiology*, 15(9):1165–1172, 2008.
19. Scott Keller, Mike Landry, Jeanne Olson, Anne Velliquette, Scot Burton, and J. Craig Andrews. The effects of nutrition package claims, nutrition facts panels, and motivation to process nutrition information on consumer product evaluation. *Journal of Public Policy and Marketing*, 16:256–269, 1997.
20. S. Lamnek. *Qualitative Sozialforschung*. Beltz Verlag, 2005.
21. Andreas Lommatzsch, Till Plumbaum, and Sahin Albayrak. A linked dataverse knows better: Boosting recommendation quality using semantic knowledge. In *Proc. Advances in Semantic Processing*, pages 97–103, Wilmington, DE, USA, 2011. IARIA.
22. P Mayring. *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Beltz Verlag, 2007.
23. Jane McGonigal. *Reality is broken: Why games make us better and how they can change the World*. Penguin Press, 2011.
24. Janet M Morahan-Martin. How internet users find, evaluate, and use online health information: a cross-cultural review. *Cyberpsyc. and Beh.*, 7(5):497–510, 2004.
25. J. G. Nicholls. *The competitive ethos and democratic education*. Harvard University Press, 1989.
26. D. Oard. A comparative study of query and document translation for cross-language information retrieval. *Machine Translation and the Information Soup*, pages 472–483, 1998.
27. C. Pelz, A. Schmitt, and M. Meis. Knowledge Mapping als Methode zur Auswertung und Ergebnispräsentation von Fokusgruppen in der Markt- und Evaluationsforschung, 2004.
28. Till Plumbaum, Sascha Narr, Elif Eryilmaz, Frank Hopfgartner, Funda Klein-Ellinghaus, Anna Reese, and Sahin Albayrak. Providing multilingual access to health-oriented content. In *Proceedings of the 25th European Medical Informatics Conference*, MIE, pages 393–397. IOS Press, 2014.
29. Till Plumbaum, Sascha Narr, Veit Schwartze, Frank Hopfgartner, and Sahin Albayrak. An intelligent health assistant for migrants. In *Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare*, PervasiveHealth, pages 307–308. IEEE, 2013.
30. Till Plumbaum, Songxuan Wu, Ernesto William De Luca, and Sahin Albayrak. User modeling for the social semantic webs. In *Proc. Workshop on Semantic Personalized Information Management: Retrieval and Recommendation*, pages 78–89, 2011.
31. Tessa M. Pollard. *Western Diseases: An Evolutionary Perspective*. CUP, 2009.
32. J Rogers and A Rector. Galen's model of parts and wholes: experience and comparisons. *Proc AMIA Symp*, pages 714–718, 2000.
33. P. Sorg and P. Cimiano. Cross-lingual information retrieval with explicit semantic analysis. In *Working Notes of the Annual CLEF Meeting*, 2008.
34. Statistisches Bundesamt. Bevölkerung und Erwerbstätigkeit. Technical report, 09 2012.
35. Melanie Swan. Emerging patient-driven health care models: An examination of health social networks, consumer personalized medicine and quantified self-tracking. *International Journal of Environmental Research and Public Health*, 6(2):492–525, 2009.
36. Benjamin Taskar, Ming Fai Wong, Pieter Abbeel, and Daphne Koller. Link prediction in relational data. In *NIPS*, 2003.
37. H. ten Napel and J. Rogers. Assessment of the galen methodology on holistic classifications for professions allied to medicine. *Medinfo*, 10(Pt 2):1369–1373, 2001.
38. United Nations, Department of Economic and Social Affairs. World population policies 2005. Technical report, 03 2006.
39. U.S. Department of Health and Human Services. Physical activity and health: A report of the surgeon general. Technical report, CDC, 2006.
40. David Vallet, Frank Hopfgartner, and Joemon M. Jose. Use of implicit graph for recommending relevant videos: A simulated evaluation. In *ECIR*, pages 199–210, 2008.
41. David Vallet, Frank Hopfgartner, Joemon M. Jose, and Pablo Castells. Effects of usage-based feedback on video retrieval: A simulation-based study. *ACM Trans. Inf. Syst.*, 29(2), 2011.