PLATE

Product Lifetimes And The Environment

3rd PLATE Conference

September 18–20, 2019

Berlin, Germany

Nils F. Nissen

Melanie Jaeger-Erben (eds.)

Poppe, Erik: **Time in market: using data mining technologies to measure product lifecycles**. In: Nissen, Nils F.; Jaeger-Erben, Melanie (Eds.): PLATE – Product Lifetimes And The Environment : Proceedings, 3rd PLATE CONFERENCE, BERLIN, GERMANY, 18–20 September 2019. Berlin: Universitätsverlag der TU Berlin, 2021. pp. 661–667. ISBN 978-3-7983-3125-9 (online). https://doi.org/10.14279/depositonce-9253.

Universitätsverlag der TU Berlin

3rd PLATE 2019 Conference
Berlin, Germany, 18-20 September 2019

PLATE
Product Lifetimes And The Environment

Time in Market: Using Data Mining Technologies to Measure Product Life Cycles

Poppe, Erik
Technische Universität Berlin, Chair of Transdisciplinary Sustainability Research in Electronics, Berlin, Germany

Keywords: Data Mining; Web Crawling; Time in Market; Obsolescence; Smartphones.

Abstract: Time in Market (TIM) is a metric to describe the time period of a product from its market entry to its decline and disappearance from the market. The concept is often used implicit to describe the acceleration of product life cycles, innovation cycles and is an essential part of the product life cycle concept. It can be assumed that time in markets is an important indicator for manufacturers and marketers to plan and evaluate their market success. Moreover, time in markets are necessary to measure the speed of product life cycles and their implication for the general development of product lifetime. This article's major contributions are to presenting (1) time in markets as a highly relevant concept for the assessment of product life cycles, although the indicator has received little attention so far, (2) explaining an automated internet-based data mining approach to gather semi-structured product data from 5 German internet shops for electronic consumer goods and (3) presenting initial insights for a period of a half to one year on market data for smartphones. It will turn out that longer periods of time are needed to obtain significant data on time in markets, nevertheless initial results show a high product rollover rate of 40-45% within one year and present a time in market below 100 days for at least 16% of the captured products. Due to the current state of work, this article is addressed to researchers already engaged in data mining or interested in the application of it.

What is TIM good for?

Time in Market (TIM) as a product metric is considered to measure the time period of products entry into a market and its eventual withdrawal.

Did you know that there are at least over 500 different types of toasters, more than 800 water boilers and over 1.600 smartphones being traded in Germany within one year (data by the author)? In today's era of mass consumption and customization there are hundreds, if not thousands, of new electronic consumer products introduced every week worldwide (Cox/Alm 1998). It's astonishing how shop operators, consumers and market authorities can deal with all this complexity (Schwartz 2009, Schneider et al 2007). Especially modern online distribution systems enable a world of abundance with unlimited shelf spaces and an enormous choice of products. (Anderson 2008: 143ff). "Yet abundance is the driving force in all economic growth and change" (Gilder 2006: 6).

Little empirical evidence of shortening of product life cycles

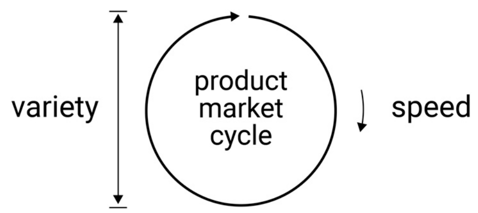

Figure 1. Product variety and life cycle speed in the market.

There have been many discussions on premature obsolescence and accelerated product life cycles, but not much research on the actual speed of the market cycle and variety of products in the market (Figure 1).
Bayus was one of the first to critically pointed out that while there is a lively discussion about the shorting of product life cycles, nobody cares about the empirical evidence (Bayus, B., 1998). Even reference books on product life cycle management, which suggest holistic approaches, ignore TIM as a metric and objective of the product life cycle (Stark, J., 2015).

The more products, the more obsolescence

This is striking given the fact that shorter TIM and an ever-increasing variety of products can be seen as a main driver for obsolescence. According to the linear assumption and thesis: the higher the variety of products in the market (e.g. by product proliferation, mass customization) and the shorter their timespan in the market, the higher the absolute number of products that will become obsolete at some point in time (Trentmann 2017: 672 ff., Slade 2006:29-55, Packard 1960:29-40). In short, the more products, the more obsolescence. The main advantages of the TIM indicator i.a.:

- *Business relevance:* TIM is an important factor and objective for the success of companies in an increasingly competitive consumer electronic market, especially when general product life cycles are shortening (Cooper 2005:19-20, Brown/Lattin 1994).

- *Optimized product strategies:* Better insights on TIM can mitigate the risk of misplanning, overproduction that cause negative environmental effects and business damage.

- *Accessibility:* In many cases semi-structured product data can be openly accessed and automatically processed afterwards.

Operationalization of TIM

According to a general market-oriented concept of product life cycles, there are different stages a product goes through from its initial launch until completion of the life cycle at the point the product becomes obsolete (Kotler/Armstrong 2012:273). Figure 2 shows an adapted and simplified product life cycle to give a better understanding on the difficulties to measure TIM.

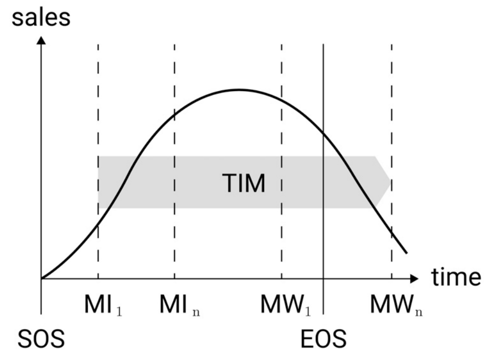

Figure 2. TIM within the product life cycle and cumulated sales rate.

The key parameters shown in figure 2 are:

- *Start of sales (SOS):* The manufacturer or market representative start sales, directly to customers, to trade intermediaries or to trading companies.

- *Market introduction (MI):* The products become available in the market through different vendors at different point of times (MI_{1-n}). The cumulated sales rate typically goes up.

- *End of sale (EOS):* At some point in time the manufacturer discontinues the production and distribution of the product.

- *Market withdrawal (MW):* Manifold reasons induce vendors to withdraw the products from their marketplace at different point of times (MW_{1-n}). The cumulated sales rate typically declines.

Differentiation at multiple scales

The model outlined above is to be understood as an idealized superposition of several microcycles on different levels. On the one side, there is a complex differentiation of vendors on the market stretching from retailers to second hand re-commerce marketplaces. On the other side, there is a complex differentiation at the product level. A smartphone of the type iPhone 8 Plus from the manufacturer Apple is offered in at least 14 corresponding configurations due to different colors and size of the internal memory. The same problem of distinction occurs with different colors of toasters, notebook keyboards in different languages or televisions display size.

Regarding the bulk of product variety and single product market cycles it seems rational to itemize the observable data along the taxonomy presented in figure 3.

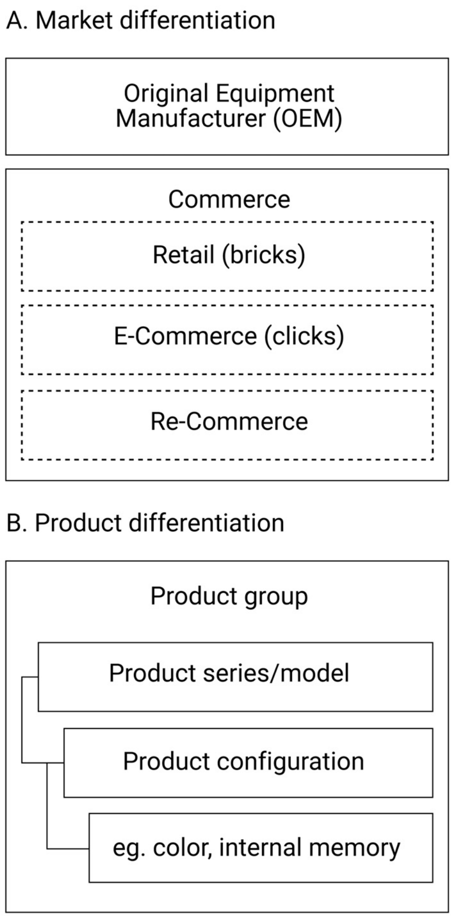

Figure 3. Market and product specific taxonomy.

The taxonomy proposed in figure 3 offers a basic ontology to describe market and product specific TIM. The e-commerce sector at the market level (A) is assumed to achieve a longer timespan in the markets as they don't have the same scarcity of physical shelf space such as retailers in physical stores. Moreover, re-commerce providers probably realize expectably longer timespans in markets due their business model. Inevitably, for further analytics it becomes necessary to distinguish between different product levels (B). As already mentioned, products are often manufactured in series or models and have different configurations. A smartphone model of the type of an Apple iPhone 8 with 64 gigabyte internal memory can have a longer TIM than its smaller equivalent with only 32 gigabyte memory. Where possible, it is necessary to analyze TIM at different product levels to enable detailed analysis of the results.

Data mining approach

Inevitably, an article of this scope cannot cover every technical detail. For this reason, the following section will focus on the concept of data mining and major challenges of web crawling to give even non-experts an insight in this complex field. For a broad overview the author recommends Fan and Bifet (Fan/Bifet 2014) and for a detailed study of the general concept and application Sumathi and Sivandam (Sumathi/Sivanandam 2006).

Definitions and key concepts

Data mining refers here to the process of acquiring and analyzing large amounts of data in order to discover patterns and other information (Sumathi/Sivanandam 2006).

Web crawler are software applications commonly known as bot, web spider, web scraping or data harvesting. They can browse hyperlinks, collect pages from online resources and extract the relevant information (Patil/Patil, 2016, Castrillo-Fernández 2015).

Technical requirements

The basic task was to setup a system that automatically extracts, preprocess and stores online product data from different product categories of selected German e-commerce and retail shops on a weekly basis. In addition, the system must be able to match products from one shop to another. The implementation process and a simplified recursive workflow are shown in figure 4 and explained below.

Step 1: Selection

The online shops in Table 1 were selected because of the provider's market share in the consumer electronics sector and the possibility to automatically retrieve data. The latter point is important because some online shops limit access to websites through legal and technical restrictions. In order to avoid double counting of single shops, aggregators such as amazon.de, google.de, or idealo.de where excluded from the selection.

Shopname	Market place	Period since
saturn.de	bricks & clicks	2018-06
mediamarkt.de	bricks & clicks	2018-06
medimax.de	bricks & clicks	2018-09
euronics.de	bricks & clicks	2018-09
conrad.de	bricks & clicks	2019-02

Table 1. Selected online shops.

The five selected shops accounted in the year 2017 at least for 50-60% of the net sales of the leading companies in the German consumer electronics retailing sector (Estimation based on LZ 2018). Thus, the presented data can be considered as relatively representative for the predefined marketplace (see figure 3).

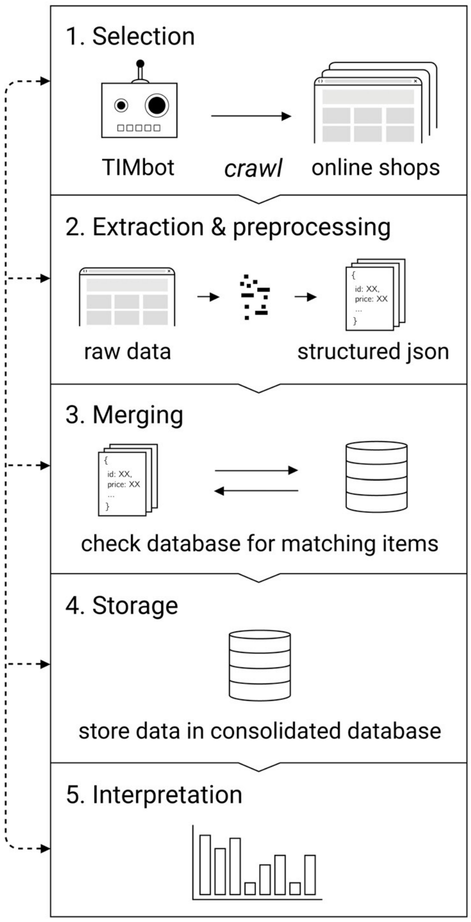

Figure 4. Simplified workflow of the data mining and webcrawling process.

The general process of webcrawling is realized through various libraries, frameworks and software written with the programming language Python (Python 2019). Due to the different technical setups and structure of the shop websites various crawling approaches are used. All together they build a software system that will be named in the following as *TIMbot*.

Step 2: Data extraction and preprocessing
TIMbot is crawling predefined product categories by the shops internet address (URL) and extracts the raw semi-structured data from the HTML structure or via occupying the database interface (API). The data will be filtered through predefined patterns that select target information such as the name, brand, product link and additional product information. The most important information is the European Article Number (EAN) which identifies each product and its different configurations by a unique number. The preprocessed data is then stored as structured data string in a json-File.

Step 3: Merging
All product data will be stored in a consolidated database. In order to generate time series from one week to another the product data from the json-File will run through a specific merger script (*Merger*), to check first if the EAN is already present in the database. If not, TIMbot will entry a completely new product in the database. If the product already exists, only some information such as price and page rank will be stored as a separate new data point (*event*).

Step 4: Storage
The data will be stored in a MySQL-Database system which provides a highly reliable and almost unlimited dataspace. MySQL is widespread, recommended for structured data and well documented. So far, the database contains over 40.000 unique product items and captures over 2 million corresponding datapoints (*events*) as time series.

Step 5: Interpretation
The Database can be accessed by a variety of different analysis and statistic programs such as MS Excel, Tableau or Stata.

Initial insights into the smartphone market

The following statistics focuses on the product group of smartphones and a limited time period from July 2018 to June 2019. Due to the current status of the research project the results must be seen as a first step for a proof of concept and do not claim to be fully representative.

Apart from that, the data gives a hint about the possibilities of the introduced data mining approach above. Table 2 gives a basic market overview, while figure 5 shows a negative correlation of price changes and number of days for the total sample of products tracked within in the period.

Market Overview

Classification	Ratio
First data	July 2018
Last data	June 2019
Month (total)	12
Product models (total)	580
Unique products (total)	1.603
Brands (total)	55
Price range unique products	50 – 1.179 €
Median price unique products	249 €

Table 2. Selected online shops.

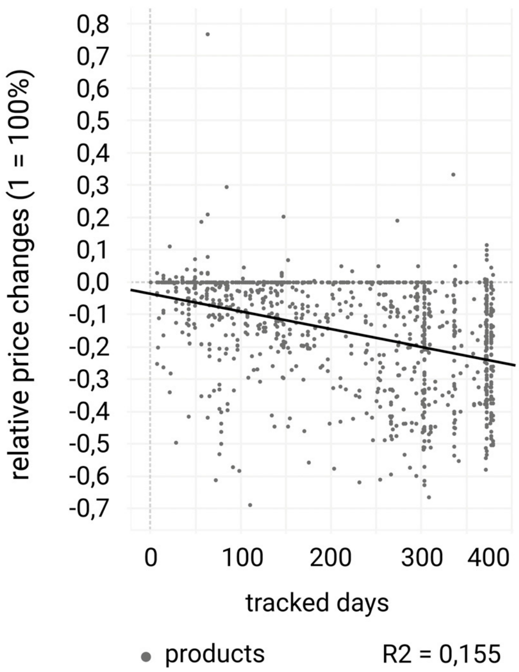

Figure 5. Price change within the tracked period.

Total market entries and withdrawals

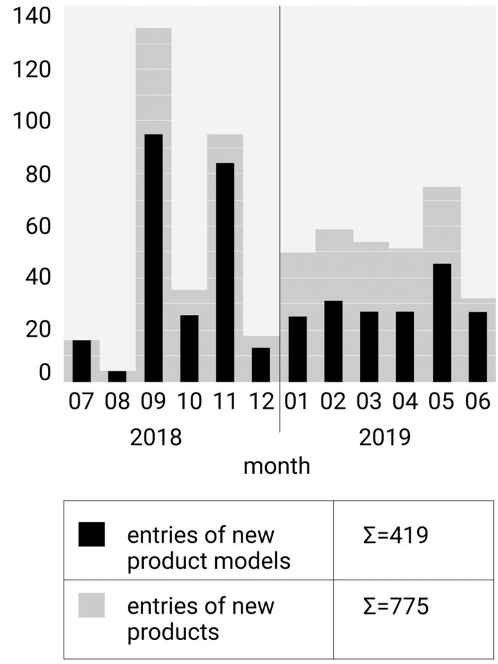

Figure 6. Total market entries of smartphones from July-2018 to June-2019.

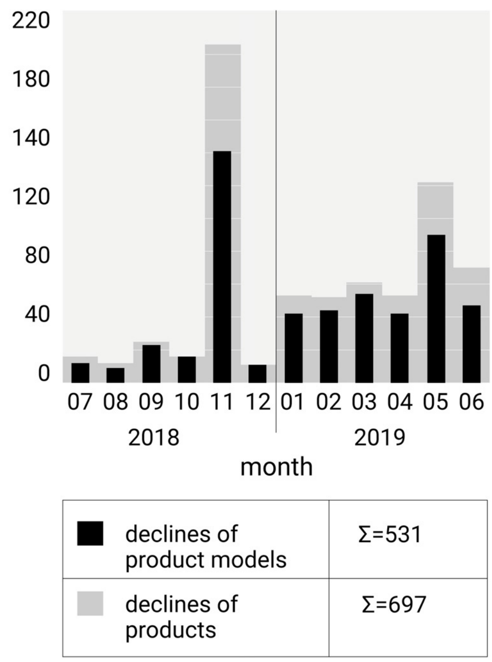

Figure 7. Total market declines of smartphones from July-2018 to June-2019.

Figure 6 and 7 are showing the total market entries and declines of smart phones on the level of product models such as an iPhone Xs or Huawei P10 (*product models*) and its various configurations such as color and memory size (*products*). The data indicate a cyclical trend with a high in the pre-winter season and decline in the holiday season in winter and summer. According to worldwide smartphone sales figures, sales always peak in the winter season (Business Wire 2019), so it can be assumed that the increased product rollovers from September to October will be realized in advance in order to prepare for the high season.

Within the tracked time period of 1 year and a total of 1.603 gathered unique products a total of 775 new products entered the market. Within this group almost 53% stay in the market for longer than the tracked time period of 12 month and the rest of 46% disappear (*fast mover*).

By contrast 697 products disappeared from the market within one year. Within this group 51% belong to the group of new product entries in this period and 49% to the stock that was already listed before the data tracking started.

Taken together the total market entries and declines suggest a high product rollover rate in the overall market. Around 40-45% of the total stock of products in the tracked shops are replaced within one year.

TIM for fast moving smartphones
Within the product group that entered the market and disappeared from it within one year (fast mover), it can be observed that almost 70% of this group have a shorter TIM than 100 days (16% of total products).

Figure 8 shows especially short TIM for products from one to 10 days (25%) in the fast mover group. A first qualitive analyses suggest that these products belong to the lower price segment (median 206 €, mean 296 €). Further qualitative research could examine whether short market cycles are caused through products specific criteria, market circumstances or the shop owners market strategy. However, such short cycles are certainly not induced by acute innovation leaps, thus provoking the fundamental question if they are economically rational. The observed data shows a significant proportion of short traded goods in the market (16%).

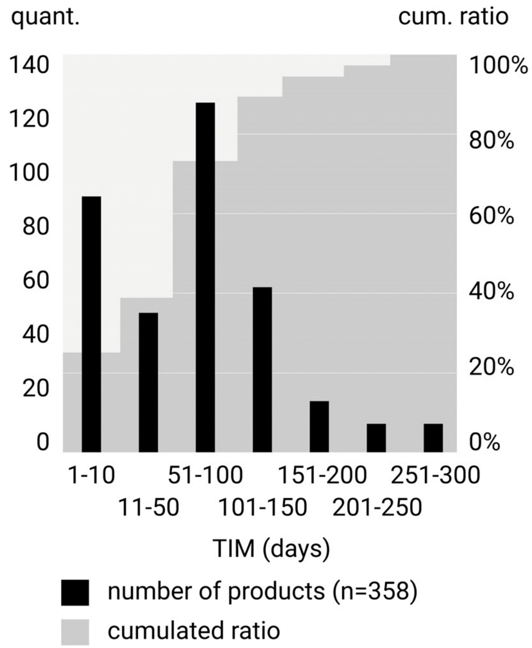

Figure 8. TIM (days) of fast moving smartphones.

Conclusion
The aim of this paper was to introduce the concept of time in market (TIM) as a new indicator for the assessment of product lifecycles. The concept has been shown to address important issues in assessing the shortening of product lifecycles and obsolescence issues.

In order to distinguish between different markets and products a model was presented to operationalize TIM. It becomes obvious, that TIM highly depends upon the selected marketplace and product group. By analogy with stock markets, the indicator must be considered as an index for specific markets and their product portfolios. For instance, re-commerce vendors are expected to realize longer TIM than retailers with a focus on new products.

Furthermore, an automated internet-based data mining approach was introduced that can gather highly available semi-structured product data from 5 German internet shops for electronic consumer goods. Product data were captured for a period of one year, including 40.000 unique product items of different product groups and corresponding timeseries with over 2 million datapoints (*events*).

Due to the limited time period an exemplary analyses of product data for smartphones provided a first proof of concept. The data mining approach works, and above all, it is possible to identify unique products from one store to another using their unique European Article Number (EAN). Furthermore, a first brief analysis shows that of the 1.603 products captured, around 40-45% are replaced within one year. Around 16% of smartphones in the online shops tracked have a remarkably short TIM of less than 100 days.

Fast market cycles can be seen as a main driver for obsolescence and need to be addressed to mitigate the risk of misplanning, overproduction and business damage. The applied data mining concept for TIM can help to identify priorities for virtually unlimited product groups and can provide long term insights into shortening product life cycles.

Acknowledgments

The author is a fellow of the interdisciplinary researcher group "Obsolescence as challenge for sustainability", which is funded by the German Federal Ministry of Education from 2016-2021. Special thanks to Alexander Schlegel for his technical advice on Python programming issues and my colleague Eduard Wagner (TU Berlin) who inspired me to web crawling.

References

Anderson, C. (2008). The long tail, Why the future of business is selling less of more (2nd ed.). New York: Hachete Books.

Business Wire (2019). Absatz von Smartphones weltweit nach Hersteller vom 4. Quartal 2009 bis zum 1. Quartal 2019 (in Millionen Stück) [Graph]. In Statista. Access on 31. Juli 2019, from https://de.statista.com/statistik/daten/studie/173048/umfrage/weltweiter-absatz-der-smartphone-hersteller-nach-quartalen/.

Castrillo-Fernández, O. (2015). Web Scraping: Applications and Tools. European Public Sector Information Platform, Topic Report No. 2015/16.

Cooper, R.G. (2017). Winning at new products, Creating value through innovation (5rd ed.). New York: Basic Books, pp. 19 – 20

Cox, W.M., & Alm, R. (1998). The right stuff: America's move to mass customization, Annual Report, Federal Reserve Bank of Dallas, pp. 3 – 26

Fan, W & Bifet, Albert. (2014). Mining big data: Current status, and forecast to the future. ACM SIGKDD Explorations Newsletter. 16. 1-5

Gilder, G. (2000): Telecosm, How Infinite Bandwidth Will Revolutionize Our World. New York: The Free Press, 200, p.6-7.

Kotler, P. & Amstrong G. (2012). Principles of Marketing (14th ed.), Prentice Hall, 2006, p. 273

LZ. (2018). Bruttoumsatz der führenden Unternehmen im Elektrofachhandel in Deutschland im Jahr 2017 (in Milliarden Euro) [Chart]. In Statista. Access on 22. Juli 2019, from https://de.statista.com/statistik/daten/studie/315124/umfrage/umsatz-der-fuehrenden-unternehmen-im-elektrofachhandel-in-deutschland/

Packard, V. (1960). The Waste Makers, New York/USA: David McKay, p.29 – 40

Patil, Y., Patil, S. (2016). Review of Web Crawlers with Specification and Working, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, Issue 1, 2016, pp. 220 – 223

Python (2019): Python 3.7.4 documentation, Access on 30.07.2019, from: https://docs.python.org/3/

Schneider, L., Rickert, B., Lücker, M., Rücker, A., Schilde, A., Schütte, G., Zoephel, S., Lindloff, K. (2007). Nachhaltigkeit im Alltag: Gut leben mit weniger Neuprodukten. In: Rabelt, V., Simon, K.-H., Weller, I., Heimerl, A. (1998). Nachhaltiger_nutzen. Möglichkeiten und Grenzen neuer Nutzungsstrategien, München: oekom Verlag, pp. 150 – 165

Schwarz, B. (2016): The Paradox of Choice: Why More Is Less, How the culture of abundance robs us of satisfaction. Revised Edition Harper Collins, 2009.

Slade, G. (2006): Made to break: technology and obsolescence in America, Cambridge: Harvard University Press, pp. 29 – 55

Sumathi, S., Sivanandam, S.N. (2006). Introduction to Data Mining Principles, Studies in Computional Intelligence (SCI) Vol. 29, Berlin/Heidelberg: Springer-Verlag, pp. 1 – 20

Trentman, F. (2017). Empire of things, How we became a world of consumers, from the 15th century to the 21st, UK: Penguin Books, pp. 672 ff.