Novel data structure and visualization tool for studying technology evolution based on patent information: The *DTFootprint* and the *TechSpectrogram*

December 4, 2020

Abstract. We introduce a new, bespoke data structure to analyze and visualize the evolution of a technology. The technology under analysis is defined by a set of patents corresponding to a technical field, owned by a company or invented by a team of research. Our data structure, the Dynamic Technology Footprint –*DTFootprint*–, facilitates the analysis and visualization of trends and dynamics of a given technology, and therefore the evolution of a technical field, a company, or a team of people. A graphical tool based on our data structure is defined, it is named *Technology Spectrogram –TechSpectrogram–*, because it is inspired by the acoustic frequency spectrograms: as the acoustic frequency spectrograms visualise the dynamics of an acoustic wave showing the evolution of its frequency components our tool shows the dynamics of a technology showing the evolution of its technological components, which are represented by the whole set of IPC-codes. Our graphical tool, the *TechSpectrogram* is shown for some study cases, and its application to the history of technology and technology management are disclosed.

Keywords: Technology evolution analysis; patent classification; patent citations

1. Introduction

Science and technology (S&T) has been a topic of study and academic research for decades now [1]. Scientific discoveries, new technologies and innovations are disclosed publicly in documents such as the company's bulletins, technical journal papers, theses or patents, and recently also in research centres or companies' web pages, internet technical forums and internet information repositories ([2], [3] and [4]).

In their quest to understand S&T, researchers have created a myriad of indicators and tools with the aim of speeding up the grasp of some aspects of S&T such as history of technology, policy making or technology monitoring ([5], [6] and [7]). S&T indicators usually collect information from a large collection of documents and transform it into a tiny set of values in order to drastically reduce the amount of information to be taken into account. On the other hand, some tools provide the very same information contained in the set of documents but transformed or presented in a way that makes clearer to observe the information, and therefore –hopefully– to note interesting elements, trends or bias within the set [8]. The data structure and the graphical tool that we have developed in our work belongs to this last kind of tools.

The recognized channel of S&T knowledge dissemination is published documents. Non-patent literature (NPL), such as books, papers, theses and conference proceedings, convey information about S&T innovations mainly produced by academic institutions, whereas patents disclose exclusively technological innovations, which are mainly developed by industrial corporations [2]. We focus on patent literature because we are interested in applied technology innovation. Additionally, patents are documents rich in reliable structured information [9] which is easy to reach through public databases such as among others *Espacenet*, *Patstat* or *Patft* [1]. This feature of patent literature improve the easiness of aggregation and processing of patent publications and their citations. Moreover, the fact that patents are produced largely by industrial corporations makes the body of patent literature particularly interesting to study technology.

The purpose of this work is to disclose a new data structure, the *Dynamic Technology Footprint* (*DTFootprint*) based on the *Technology Footprint* introduced by Perez-Molina[10]. Our dynamic data structure will be embodied in a graphical tool, the *Technology Spectrogram* (*TechSpectrogram*) visualizing the evolution in time of a specific technology. The *DTFootprint* is founded on the following structured patent information: the prior art patent citations and the allocated classification codes. These two structured informations have very exclusive features in relation with citations and classification in *NPL*, firstly in patents prior art citations which are produced by the patent examiner (a person necessarily different to the author), and secondly a universal classification scheme – the *IPC* (International Patent Classification)– is used for allocating classification codes to any published patent document [10]. Our aim is to analyse the evolution of a given technology by first producing a dataset – the *Technology Footprint* – corresponding to a specific year, and then generating a global dataset – the *DTFootprint* – for a lapse of years produced by combining the sequence of yearly datasets. The information contained in each yearly dataset is the technological components of the technology under analysis. These components form a description of the given technology in terms of the whole breadth of existent technology. The consequence of this transformation of a given technology in its components is to get direct information about the technical fields present in that technology and their influence. The technological components, their weights (level of influence) and its change over time facilitates the analysis of trends and the identification of the emergence or declination of some technological activities.

Our data structure is complete and ordered. It is complete because every IPC-bin (thus, any technology) and every quantum of time is present, and it is ordered according to the IPC (at IPC-Section level A, B, C, ..., H) and sequentially in time (1984, 1985, 1986, ...2015). The formal similarity of our tool with the frequency spectrogram analysis of

[1] ***Espacenet*** is a public – free of charge – worldwide patent database provided by the EPO with bibliographic and administrative data at `https://worldwide.espacenet.com`. ***Patstat*** is a worldwide patent database for statistical analysis provided by the EPO at `https://www.epo.org/searching-for-patents/business` and ***Patft*** is a public – free of charge – fulltext patent search database for American patent publications provided by the USPTO at `http://patft.uspto.gov/`

acoustic waves [2] has deep implications because it brings the potentiality of straightforward use of the whole range of processing available for frequency spectrograms to our tool, and therefore to the analysis of patent information.

Just as an indication of the potentiality of the direct use (or translation) of a processing from frequency spectrograms lets see for example the characterization by frequency spectrogram analysis of sound recordings applying landscape ecology techniques disclosed by Villanueva et al. [11]. They characterised a sound recording by computing in time three parameters of the spectrogram, namely Band Diversity, Band Evenness and Band Dominance. This characterisation can straightforwardly be translated to our tool in order to characterise technology generating the three equivalent parameters, namely Classification Diversity, Classification Evenness and Classification Dominance. Of course, these three parameters can always be applied to the classification data for statistical characterisation as was partially done by Leydesdorff [12] but our tool facilitates this characterisation because once the frequency spectrogram analysis is available, it requires a mere translation to be used by our tool.

We also highlight the versatility of our tool at graphical level. The straightforward use in our data structure of frequency spectrogram's image processing analysis represent a big potential, which is not the case for more complex visualisations such as patent landscapes.

The rest of the paper is organized as follows: Second section discloses related works. Third section discloses the *DTFootprint* data structure. Fourth section presents its graphical tool, the *TechSpectrogram* and some study cases, additionally a brief overview of the potential applications of our tool is given. Finally, the last section points out future works and summarizes our work.

2. Related works

The concept of S&T indicators has its roots back to the 1930s [13]. The outcome of scientific activity is – in principle – conveyed to the community by publications [13]. Thus, the analysis of S&T publications gives us indications of different aspects of S&T production [14].

The value of patent citations in the analysis of technology, in particular to highlight topic connections was already identified by E. Garfield in 1957 [15], N. Reingold [16] pointed out in 1960 the interest of patent collections as source material for historians of technology, and numerous authors have noted the value of patent analysis as tools for studying industrial corporation generated technology ([17] and [18]). Carpenter et al. in 1983 [19] discloses the

[2] Our tool orders the complete range of classification codes sequentially in time in a similar way that the frequency spectrogram order the complete range of frequencies sequentially in time, and the classification bins play the role of the frequency bins in the acoustic frequency spectrograms

use of patents to build S&T indicators . Since then, patent information has been analyzed from several perspectives, first statistically [20], then structurally by citations' network analysis [21] and semantically by text mining [9]. Recently, computing techniques such a Big Data or Artifical Intelligence (AI) has also been used to extract information from patent collections [22].

Technology trends and evolution is a crucial subject for technology researchers and managers. In 1978 P. Ellis et al.[23] disclosed a novel technique for displaying the history of technological subjects and their key turning points using patent citation networks. Martinelli has processed patent citations to understand the nature of technological changes [24]. Verspagen illustrates technological trajectories using networks of patent citations [25], and Yoon and Kim disclose a system for identifying technology trends from patents [22].

Historians of technology have examined the aggregation of classification codes allocated to a set of patents representing a specific technology and published within certain intervals of time for analyzing a particular technology within the span of years. For instance Spear [26] in 2002 and Perez Molina [27] use the distribution of classification codes for analyzing the evolution of Virtual Reality, and to identify the roots of Computer Graphics respectively.

Classification codes allocated to S&T publications provide useful structured information about technical content, and therefore these codes have been extensively used by bibliometrics researchers [28]. For example, Leydesdorff et al. have used them to generate maps of science [29], Liu and Liao exploited the codes for analysing a research topic such as *fuzzy decision research* [30] , and Arora et al. for identifying new developments in emerging technologies [31]. Perez-Molina has identified some unique features of the body of patent literature and he has defined a new data structure to analyze particular technologies based on the classification codes allocated to a set of patents [10]. The data structure in our work builds upon Perez Molina's datastructure, namely the Technology Footprint which is defined by the whole set of IPC-codes at a specific classification level (IPC-Section, IPC-Class, IPC-SubClass, ...) ordered according to the IPC (for the IPC-Section that is first A, then B, and so on until H). For a particular technology (company or group of inventors) formed by a set of patent publications, the prior-art citations are collected and the allocated IPC-codes are aggregated and stacked into bins corresponding to the whole set of classification codes. Our Footprint is formed by the IPC-codes allocated to the citations and to the set of selected patents, whereas Perez-Molina's is formed exclusively by the codes allocated to the citations. But more importantly we have limited the Footprints to a quantum of time of one year and we have taken the per-year footprints on a yearly interval. In consequence, we have defined a new data structure as a matrix, wherein each column corresponds to the Technology Footprint of a specific year, and the columns are sequentially ordered according to the yearly time interval (see Figure 1).

Graphic tools have been popular in order to visualize networks of information. Bibliometrics tools have produced different sorts of graphs to facilitate the identification of important items or the underlying structure of a collection of publications. The diversity of existing visualizations goes from relatively simple graphs to visually rich and complex rendering such as Yoon and Magee for exploring technology opportunities [32] and of Boyack et al. [33] to visualize landscapes, respectively. Yoon and Park disclose an analytical tool for high-technology trend which visualise patent networks based on text mining [34]. Liu and Zhu collect patent information to visualize patent citations using web mining [35]. In the technical field of social robotics, Mejia and Kajikawa identify research trends using colourful visualization of clusters produced from citations networks analysis [36]. Moreover, *maps of science* or *technology* have been also produced by some authors such as Yan and Luo for measuring patent distances [37], and the global maps of science of Leydesdorff and Rafols [29] .

Verspagen [25] generates a graph representing the network formed by a collection of patents showing information about the different year's intervals. Song et al. analyze the dynamics of a particular technical field – bioinformatics – over time using clusterization techniques. The collection of data is grouped and processed for a lapse of three years, and a different graph is produced for visualizing the relation between topics within the lapse [38].

Multiple authors have been generating graphs with timeline references such as Wu et al. to trace the evolution of electrochemical energy storage developments displaying clusters of citations networks [39], or Chen et al. to visualize the evolution of smart grid technology [40]. Several tools have been developed to visualize the evolution over time of technology explicitly displaying a timeline reference such as *Crossmaps* [41], *DIVA* [42] or patent maps produced by growing cell structure neural networks [43]. Sandal et al disclose two-dimensional *trend patent maps* providing the time scale map of the IPC [44]. The *trend patent maps* of Sandal et al. visualise only the assigned IPC codes, and therefore the whole range of technologies is not manipulate and represent , the whole set of existing IPC-codes. This fact implies that the visualisation doesn't represent areas or technologies of non activity which is an important information for history of technology or patent analytics.

The visualization of the evolution in time of a particular technology, company or group of people was identified as a topic worthy of future scrutiny [36]. To the best of our knowledge, there are no works identifying, for a specific technology, the technological components, its weights along an interval of time, and the corresponding visualisation in a simple and easy to read way.

In this work, given a specific technical field (a company or a group of inventors) defined by a set of patents, its technological components are the whole set of technologies represented by the aggregation of the assigned classification codes to the patents and its prior art citations, the classification codes combined in different proportions form the given technical field. And therefore, we have created a new data structure and visualization not only to provide a straightforward visual inspection of the evolution of a specific technology but to facilitate the identification of important factors or key moments of its evolution by the analytical processing of the related patent information by analytical tools such as data clustering, filtering or pattern recognition.

3. A dynamic technology footprint

Our objective is to generate a new data structure containing information from a collection of selected patents defining a technology, a company or institution, or a group of people for facilitating the analysis and visualization of its technological dynamics, such as its diversity, rate of change and evolution along an interval of years. In order to produce this new data structure we have mimicked the physicists who study the speech signal. As physicists, studying the acoustic wave produced by a sound, are interested in visualizing the dynamics of the signal showing the evolution of its frequency components, technology researchers are interested in visualizing the dynamics of a technology showing the evolution of its technological components. Thus in the same way that the latter used the frequency spectrogram the former could use a sort of spectrogram based on technological components, and therefore we have built our data structure accordingly and named our visualization tool: *Technology Spectrogram* (*TechSpectrogram*), these *technological components* play in the *TechSpectrograms* the role of the frequency bins in the frequency spectrograms.

We use the concept of technological components of a technology as disclosed by Fleming [45], who argues that the technology subclasses of a patent reflect the technological components of the patented technology. The allocated classification codes acts as proxies of the technologies combined in a particular patent. We have extended Fleming's concept to any classification code at a certain level of classification resolution as suggested by Perez-Molina [10]. Given a specific technical field (a company or a group of inventors) defined by a set of patents, its technological components are the set of technologies represented by the aggregation of the assigned classification codes to the patents and its prior art citations, which combined in different proportions form the given technical field.

In order to analyze the evolution of a specific technology, we have characterized it first for a quantum of time, i.e. one year time, and then we have examined the change of the characterization along all the years within the interval of time. In order to implement the yearly characterization of the specific technology our work draws on the data structure disclosed by Perez-Molina [10]; namely, the *Technology Footprint*, which will provide the *technological components* of a technology in each year. Then, we have defined a new data structure, the *Dynamic Technology Footprint* (*DTFootprint*), as a sequence of *Technology Footprint*s along the years within the time interval (see Figure 1).

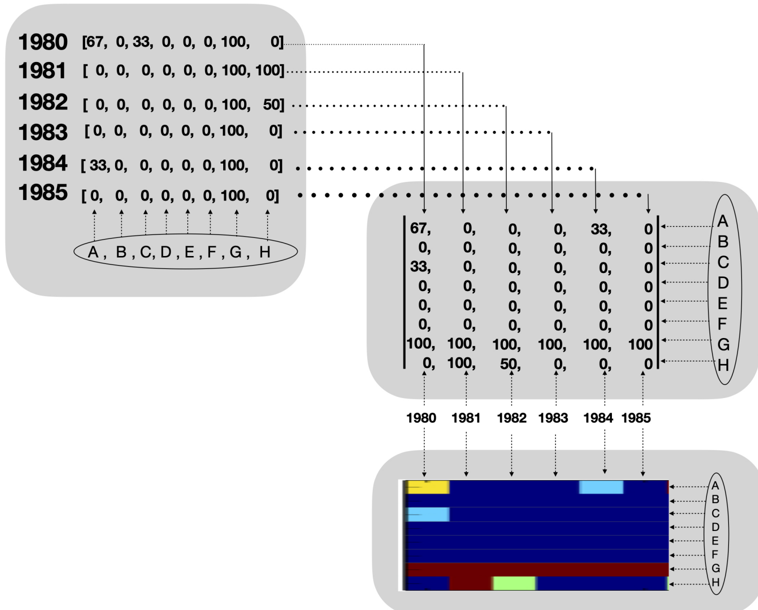

FIGURE 1. Example of the DTFootprint data structure from 1980 to 1985 (the matrix within the grey-area on the middle) and the corresponding *TechSpectrogram* visualisation (the grey-area on the bottom) generated from the yearly *Technology Footprint* data structure (the grey-area on the up-left hand side)

The data extraction to form the collections to process was done using the EPO's[3] database Patstat[4]. This is a relational database containing bibliographic and legal status information of more than 100 million patent documents from industrialized and developing countries ([46], [47])[5].

The procedure to build the *DTFootprint* performs the following six stages (see figure 2):

---

[3]EPO: European Patent Office

[4]At: www.epo.org/searching-for-patents/business/patstat.html

[5]For exact information about the PatStat coverage see at: https://public.tableau.com/profile/patstat.support#!/vizhome/CoverageofPATSTAT2018AutumnEdition/CoveragePATSTATGlobal

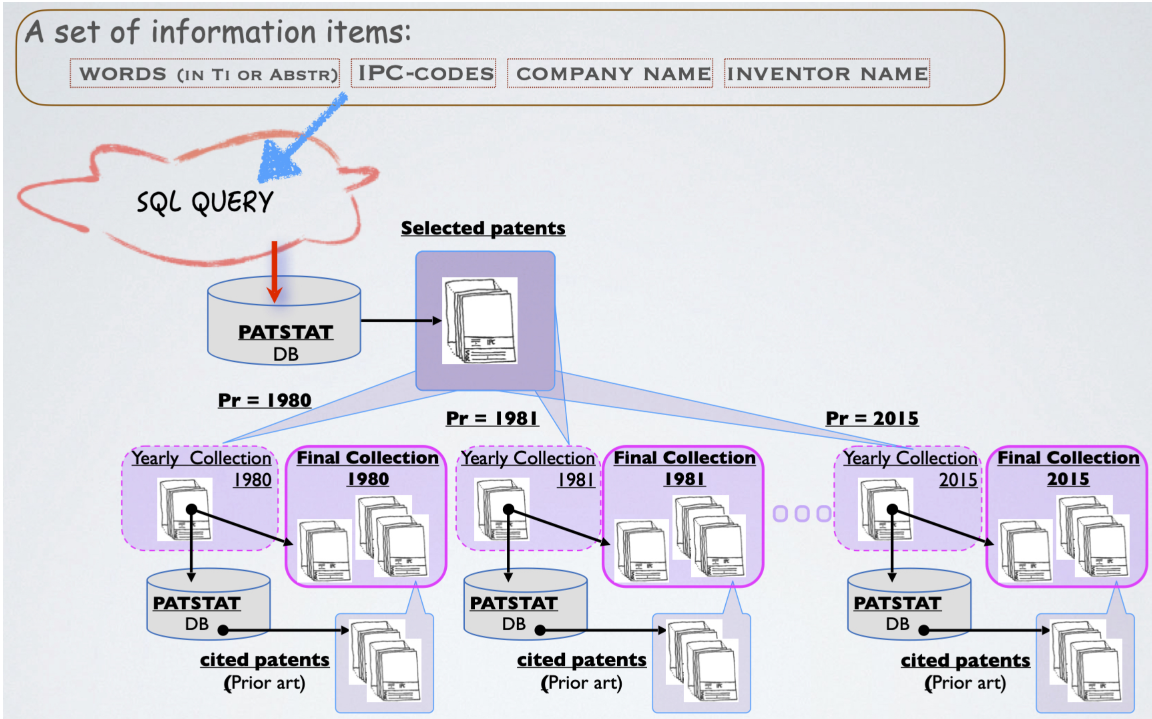

FIGURE 2. Overview of the data collection path

Stage One: Query Formulation. Firstly, patent documents disclosing the aimed technology are selected with some *SQL* queries in Patsat (see the "*selected patents" collection* in Figure 2). In the case of analysing a technical field, the technology is outlined by queries combining IPC classes, terms (in title or abstract) or a combination of both. In case of analyzing the technology produced by a company – research or academic institution – the technology is outlined by queries containing the organization as a patent applicant. Finally, in case of analysing the technology produced by a team of researchers, the queries must select the patents in which any of the researchers appears as an inventor. Of course, these queries outlining the targeted technology can be composed by any combination of those three basic selections, in this way a technology can represent, for example a company constrained to a technical field or a team of researchers limited to a particular technical expression. On top of these queries, we have limited the selection to patents having its priority date within a specific year, and this is done for every year along a time span. As an example, if we would like to study the technology of *Computer Graphics* then we select all the patent publications allocated to the IPC groups G06T11, G06T13, G06T15, G06T17 or G06T19[6]. In case that we aimed at the medical technology developed by Toshiba Corp.,

[6]G06T11: 2D image generation, G06T13: Animation, G06T15: 3D image rendering, G06T17: 3D modelling for computer graphics and G06T19: Manipulating 3D models or images for computer graphics

we select all the patents having as applicant Toshiba Corp. and having at the same time IPC codes in the A61 class [7].

Stage Two: Create Initial Collections. Secondly, we form a set of *yearly collections* according to the priority year of the patent documents of the "*selected patents*" collection (see Figure 2). At this point we get as many collections as years in the year span of analysis. Thus, if our time interval goes from (and including) 1980 to 2015 then we will create 36 *yearly collections*, namely a collection for 1980, another for 1981 and so on until the last collection for 2015. For instance, the 1980's collection will contain all the patents from the *selected patents* collection with priority date in 1980.

Stage Three: Create Citation Collections. Thirdly, the prior art cited against the patents contained in each *yearly collection* is collected and the set of corresponding yearly *cited patents* collections is created. This is done executing in PatStat an SQL set of queries which selects the citations of the patents gathered within the different *yearly collections* (see the bottom of Figure 2). The Citations are collected to permit a more detailed analysis of the technological influences; the yearly *cited patents* collections multiply the number of documents in relation to the *yearly collections* and therefore it gives us more granular detail. Citations do not change fundamentally the main technological components but enrich our data with a range of small weighted components.

Stage Four: Creating the Final Collections, Fourth, the set of yearly *cited patents* collections is added to the respective *yearly collections* forming the yearly *final collections*. Note that a collection for a specific year, first contains only patents first–filed on this precise year. However, the prior art cited against this set of publications is necessarily older, and possibly, from some years before. Thus, after adding its prior art to the collection of each year, the yearly collections contain patents from a plurality of years. The rationale behind this choice, as opposed to adding the prior art to the yearly collection corresponding to its year of priority, is that the technology of a year is characterized by the patents with priority date of this very year and the prior art cited by the examiner, although the cited technology comes from the past years.

Stage Five: Grouping by IPC. Fifth, the IPC codes allocated to every patent publication in each yearly collection are gathered and binned according to the IPC code at different levels of classification resolution, such as IPCs' *section*, *class*, *subclass*, *groups* and *subgroups* (see figure 3). Each of these bins will be interpreted as a *technological component* and the aggregated value of each bin would be in consequence the strength or importance of the corresponding *technological component* of the technology under study for the year of the collection.

[7] A61: medical or veterinary science, hygiene

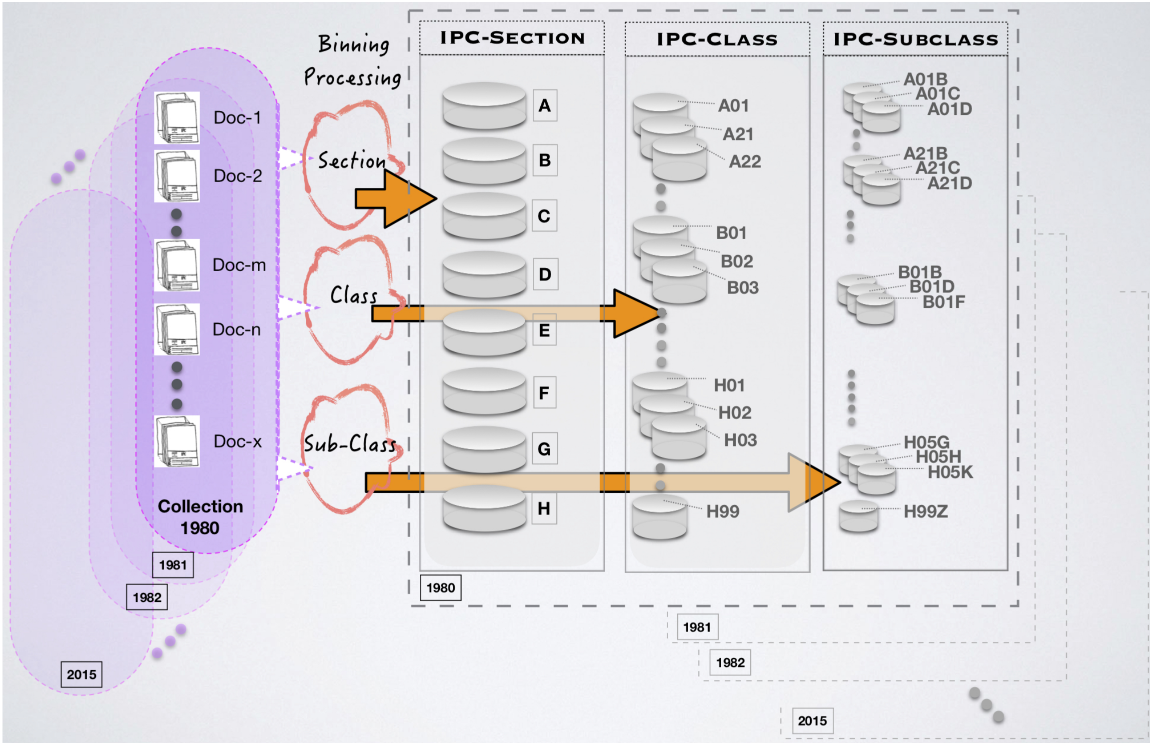

Figure 3. Overview of the *Binning* processing

Stage Six: Finalize DTFootprint. Finally, our new data structure, the *DTFootprint* is formed by bringing together in yearly order the resulting *technological components* at the different levels of the IPC. The *DTFootprint* will be organized as a two-dimensional array per IPC level (*Section*, *Class*, *sub-Class*, *Group* and *sub-Group*), and each of these two-dimensional arrays will be conformed by the year number as column and the IPC code bins as rows. Accordingly, the analysis of the data structure makes straight forward the observation of both the variety of technologies and their change over time, in terms of – in reference to – the panoply of existing technologies, as well as the identification of the emergence of some components or the decline of some-others.

4. A New Visualization Tool: The *Technology Spectrogram*

The idea is to convey the variety, importance (weight) and evolution of the *technological components* of a given technology in a rather simple graph, hence we have implemented a visualization instance of the *DTFootprint* data-structure as a coloured rectangle. The colours of the dots inside the rectangle are assigned in function of the technological component weight and presented with respect to the *technological components* and years as vertical and horizontal axes respectively (see Figure 4).

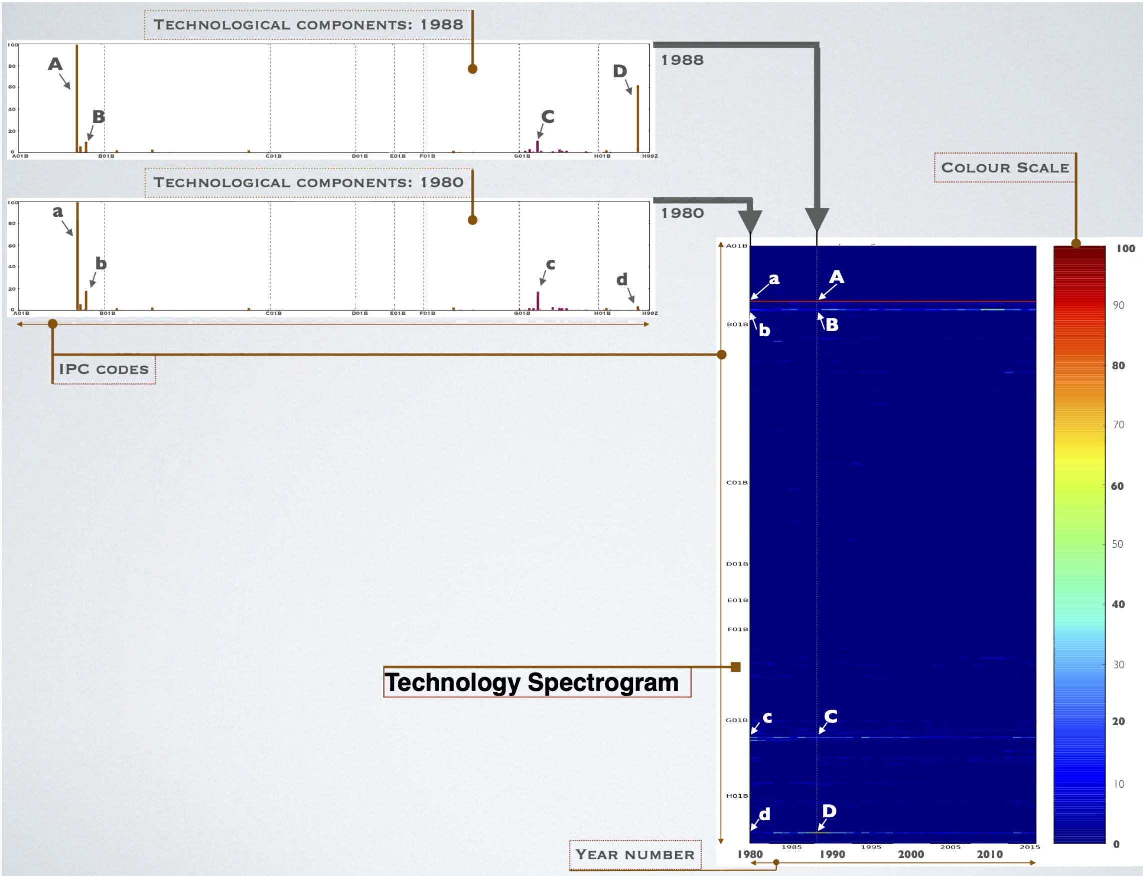

Figure 4. Relation between a ***TechSpectrogram*** and the ***Technology components*** of 1980 and 1988.

The *TechSpectrogram* dataset, in accord to the *DTFootprint* datastructure, is implemented as a 2D array having a column per year in increasing yearly order (for a time interval of 1980-2015, the 1st column would be 1980, the 2nd 1981, ..., the last 2015) and a row per IPC code[8] in increasing alphanumerical order [9]. Each cell of the *TechSpectrograms* contains the value of the *technological component* corresponding to its row code and its column year [10].

[8]`http://www.wipo.int/export/sites/www/classifications/ipc/en/guide/guide`

[9]For *Section* level, the 1st row would be **A**, the 2nd **B**, ..., the last **H**. For *Class* level, the 1st row would be **A01**, the 2nd **A20**, the 3rd **A21**, ..., the last **Z99**

[10]Although the examples presented here are limited to the years 1980-2015. Our data-set and tool can take into account a much more exhaustive coverage of years

The final visualisation of the dataset was executed, firstly normalizing the values for each column – so, per year – to the interval [0:100][11] in order to equalize it with respect to the variations of the yearly number of patent publications in the collection. Secondly, to facilitate the reading of the visualisation, the normalized values were displayed with a colour going from dark-blue to red corresponding to a value of the technological component equal to 0, and 100 respectively. Figure 4 displays an example of a *TechSpectrogram* and its *technology components* of two years (1980 and 1988), note that the most important components of these two years (See A and a in Figure 4) appear in the *TechSpectrogram* in red and a high-mid weighted component of 1988 (See D in Figure 4) appears in yellowish.

4.1. **A worked example.**

As an example of the use of *TechSpectrograms* let us take the example of computerized tomography (CT) technology. The lapse of visualization time was set from the year 1980 to 2015 (both included) in order to have a large interval of analysis. We have chosen publications from a single authority – the USPTO – in order to avoid duplications. The selection of patent documents was done by executing the following pseudo queries[12]:

TI: (*tomograph+* **AND** *imag+* **AND (***generat+* **OR** *reconstruct* **)) OR**
AB: (*tomograph+* **AND** *imag+* **AND (***generat+* **OR** *reconstruct* **)) NOT**
(TI: (*distribut+* **OR** *transmi+* **OR** *communicat+* **OR** *compress+* **OR** *uncompress+* **OR** *transfer+***) OR**
AB: (*distribut+* **OR** *transmi+* **OR** *communicat+* **OR** *compress+* **OR** *uncompress+* **OR** *transfer+***)) AND**
PrYear: (FROM*1980* **UNTIL** *2015*) **AND**
AUTHORITY: *US*

The idea with these queries is to select patent publications disclosing the generation or reconstruction of pictures using computerized tomography from 1980 to 2015. In order to be sure that the documents focus on the image generation computation, we exclude from the first set all the documents related in a certain extent to the transmission or storage of the –already – generated pictures. A total of 3906 documents were selected. These documents were then put in the yearly collections according to their priority dates. The next step, was to collect the prior art citations for each patent, a total of 27559 documents were cited, and they were then added to the corresponding yearly collections. Specifically, the 1980s *final collection* contained 57 patents (8 documents and 49 cited documents), the 1981s *final collection* contained 46 patents (11 documents and 35 cited documents), and so on until the 2015s *final collection*, which contained 1234 patents (240 documents and 994 cited documents)[13]. Thereafter, for each yearly collection, the classification codes

[11]The interval can be adjusted by the investigator to reflect a smaller or larger range and thus produce less or more detail

[12]For the exact SQL queries on Patstat see Annex 1 - SQL queries in PATSTAT database

[13]All the data were collected from ***Patstat*** in September 2019

for every patent document (both cited and citing documents) were identified and binned according to the IPC. Finally, a bunch of five datasets was built, one per IPC resolution level – *Section*, *Class*, *sub-Class*, *Group* and *sub-Group* –, from which the corresponding *TechSpectrograms* were generated.

Figure 5 shows the TSpectrogram of CT technology at IPC-Section level, the coarser level of classification resolution, and the colour scale. The vertical axe of the graph contains the IPC sections, and the horizontal axe contains the year number within the time interval. Looking at the graph it appears clearly that IPC-sections A and G are the most important components along all those years, followed by IPC-section H. At first view, these components seems in agreement with a CT device because section A — human necessities — (contains medicine) and section G — physics — (contains computers). It is may be interesting to note that section H —electricity —had a certain importance in the 1980s which decreased in time.

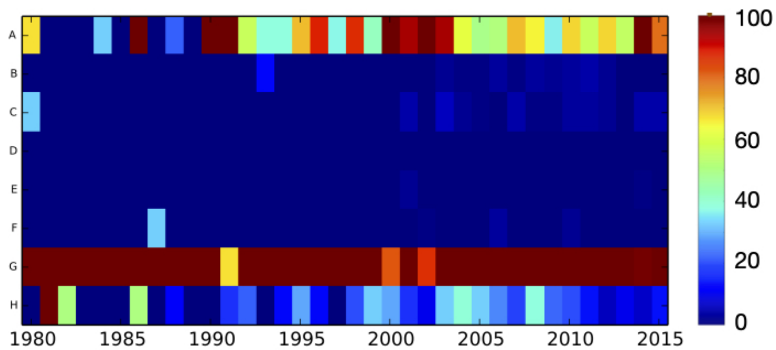

Figure 5. *TechSpectrogram* for Computer Tomography at IPC-section level, and the colour scale (left-hand side graph)

Figure 6 shows the *TechSpectrograms* obtained for CT technology at the three first levels of classification resolution, namely *Section* – left graph–, *Class* – middle graph– and *subClass* – right graph –.

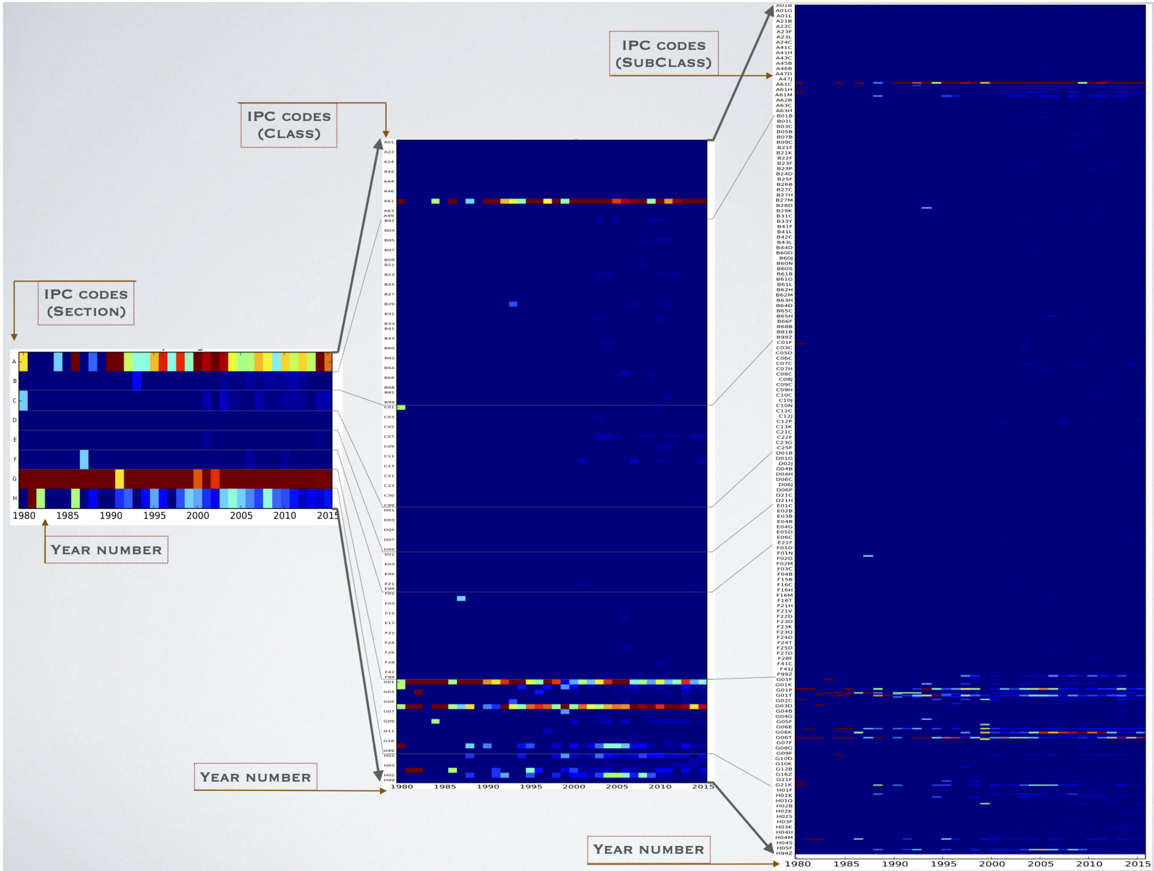

FIGURE 6. *TechSpectrograms* for Computer Tomography at IPC levels: *Section* – left –, *Class* – middle – and *subClass* – right – .

With the change of visualization level the conceptual – classification – resolution is zoomed in by a factor of about 16x between the *Section* and the *Class* levels, to include 8 and 131 bins – IPC codes – respectively (left-hand and middle graph in Figure 6). Changing from *Class* to *Sub-Class* (middle and right-hand graph in Figure 6) results in a conceptual resolution zoomed in factor of about 5x , including 131 bins and 645 bins respectively

Looking at the CT's *TechSpectrogram* at *Section* level (see left-hand graph in Figure 6), it appears clear that Sections A and G are the most important components along all those years, with H mainly "present" at the very beginning of the interval. Going to the *Class* level (middle graph in Figure 6) to have more technical details, it shows three main components together with a small set of minor ones (the bright bluish and greenish area at the bottom of the graph).

The main components corresponds to medicine – **A61** –, measuring instruments – **G01** –, and computing – **G06** –. This is consistent with the broad idea we can have for a technology such as CT, which is a computing machine generating medical pictures from some measured X-ray signals. What is more interesting is the fact that the *TechSpectrogram* reveals the dynamics of each of these components, and reveals multiple changes (colours) in the dynamics of the field. Indeed, medicine and computing are increasing its weight along the second half of the interval whereas measuring instruments is decreasing it. If now we go to the next step in resolution, the *SubClass* level (right-hand graph in Figure 6), the main component appears to be, systems for diagnosis, surgery and identification – **A61B** –, and the computing techniques gathered by two components related to image generation; namely, image data processing – **G06T** –, and presentation and recognition of data – **G06K** –. An area with relatively "active" components appears in measurement techniques at analysis of materials – **G01N** – and measurement of nuclear or X radiation – **G01T** –.

Observing the changes of each *technolgy component* along these years some patterns appear in the graphs, such as between others, peaks or, increasing and decreasing ramps. See for example at *SubClass* level (right-hand graph in Figure 6) the increasing ramp in the last years of the time interval of the image data processing – **G06T** – component, or the peak of presence at the end of the 2010s of the recognition of data – **G06K** – component. The identification of these patterns could highlight some specific situation, such as a front of development or the stagnation of a technical area in a particular moment in time.

Our visualization tool displays the *technological components* of a technology which can also be defined as a company or a group of people, thus it could be illustrative of its development activities, and thereby to figure out its development interests – or the lack of it –. See, for example in Figure 7, the *TechSpectrograms* at *Class* level of three corporations, namely Philips[14], Olympus [15] and Medtronic [16]– left, middle and right graph respectively–.

[14]https://www.philips.com , accessed on 14/08/2020
[15]https://www.olympus-global.com , accessed on 14/08/2020
[16]https://www.medtronic.com , accessed on 14/08/2020

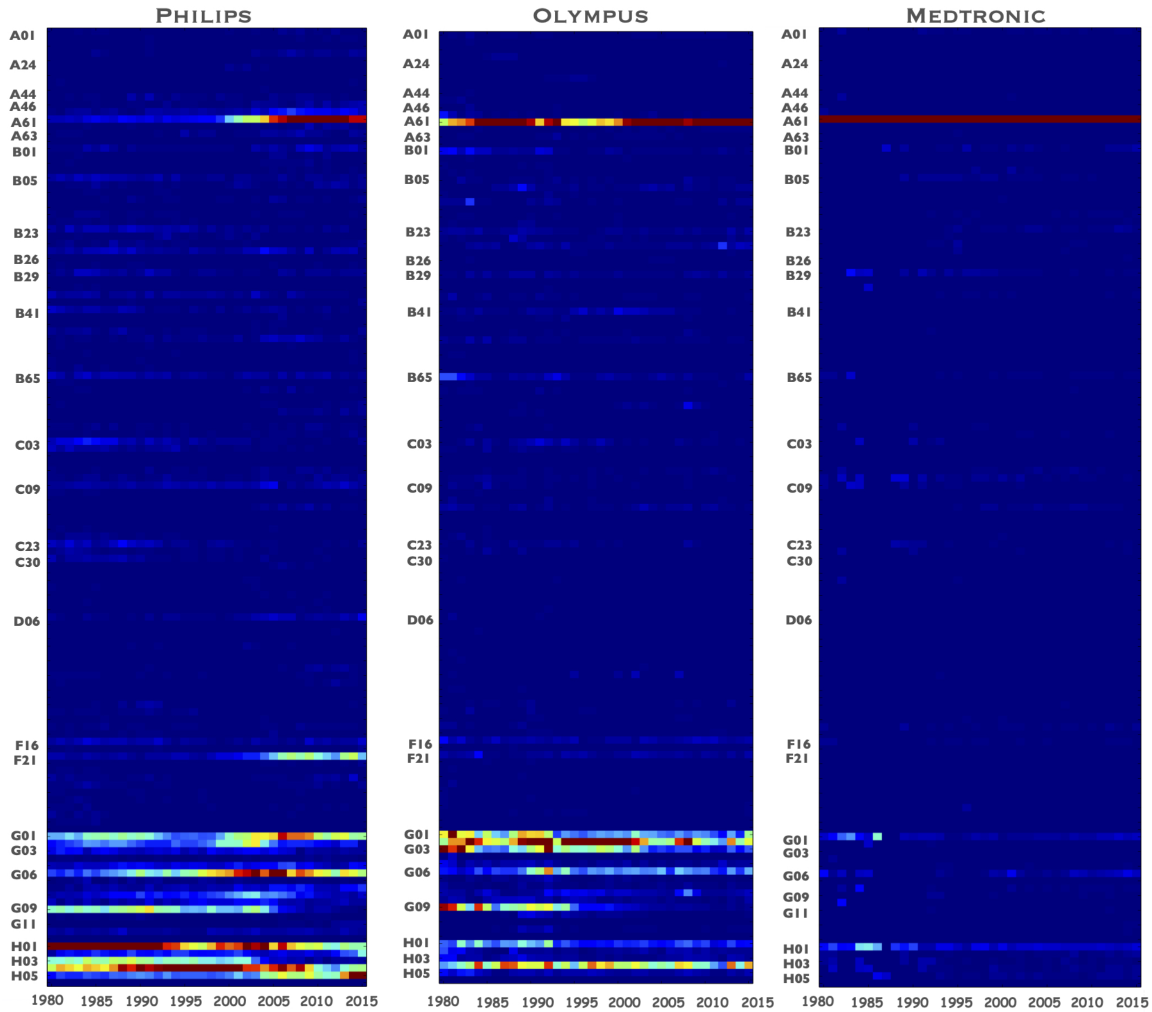

FIGURE 7. *TechSpectrograms* at IPC *Class* level of: Philips – left graph –, Olympus – middle graph – and Medtronic – right graph –.

Lets now compute the *TechSpectrograms* of three companies, such as Philips, Olympus and Medtronics. Looking at their *TechSpectrograms* it appears clear the differences in the range of technologies that they develop. The *TechSpectrogram* of Medtronic shows few *techonogical components*, and they focuss basically on medical technology – **A61** – with some secondary *components* in measuring techniques – **G01** – and electric elements – **H01** –, pointing out a company very specialized in medical technology. Whereas Philips and Olympus show *TechSpectrograms* with developments in numerous different technologies all along the monitored years.

Note that looking at these three *TechSpectrograms*, the three companies have important activities in medicine but with considerably different dynamics for the three of them. The medical technology *component* is for Medtronics constant along the time interval. For Olympus is not constant but all along the years it stays at a high level (never blueish, and few green, meaning at least 50%). On the contrary, for Philips it appears a change of dynamic around 2000, where the developments in medical technology increased until reaching the maximum within the company from 2007.

This dynamic of Philips highlighted in its *TechSpectrogram* probably represents the reinforcement of its medical and health-care division in the second half of the 2000s when Philips bought Lifeline Systems in 2006, Ximis and Respironics in 2007 and VISICU in 2008 [48].

The Philips' *TechSpectrogram* at *Class* level (see in Figure 7 the left-hand side graph) also reveals that after a strong "presence" of the electric devices developments – **H01** – *component* these technologies clearly decay from 2005-2006. To get more conceptual resolution, we can study the *Sub-Class* level (see the left-hand side graph in Figure 8). We observe that there are in fact two main technologies within electric devices, namely discharge tubes – **H01J** – and semiconductor devices – **H01L** –, both have the same dynamic pattern lightly shifted (see the *technology components* **A–A'** and **B–B'** on the right-hand graph in Figure 8). Concerning the semiconductors, the dynamic of this component is consistent with – and probably determined by – the fact that in 2005 Philips has sold its semiconductor division, which became an independent company named *NXP Semiconductors* [48], and therefore the semiconductor-related activities fade drastically out.

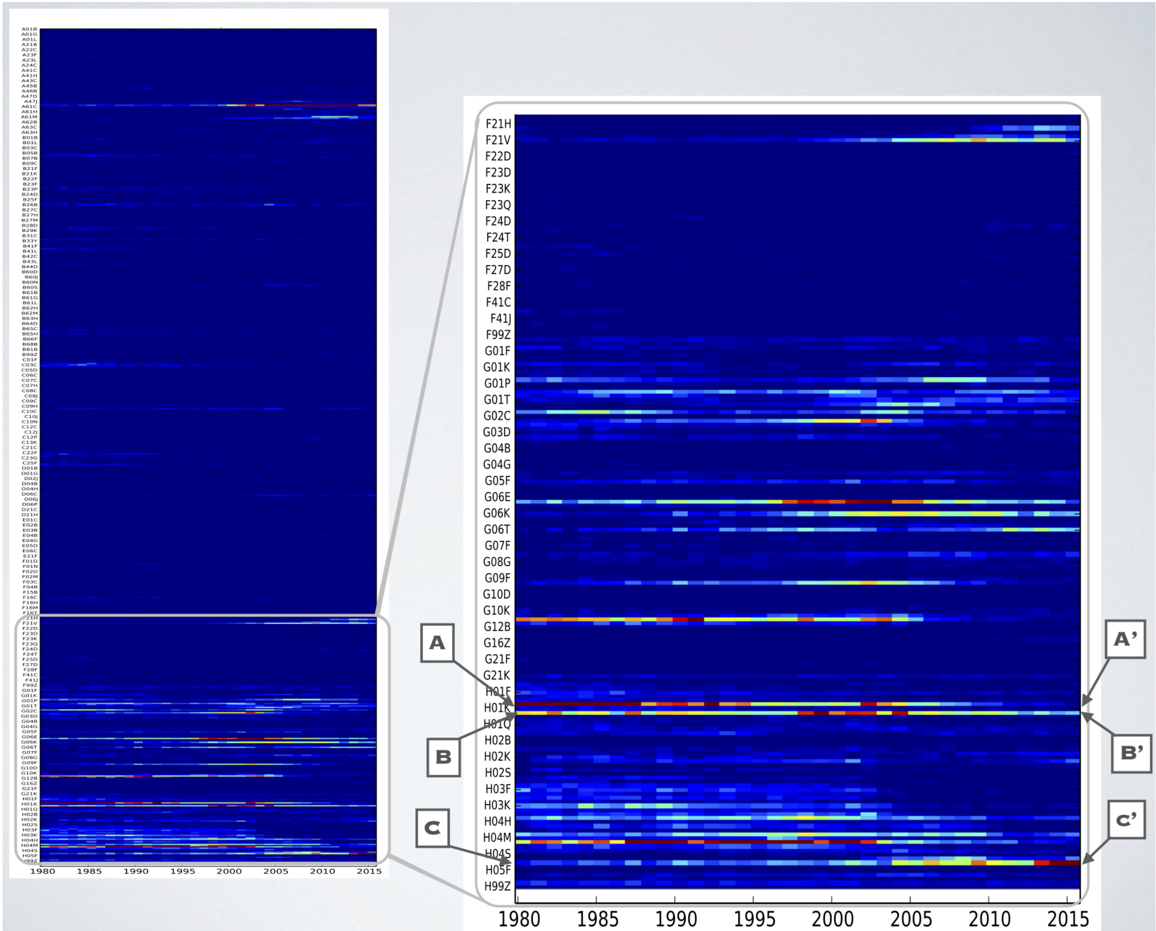

FIGURE 8. *TechSpectrograms* at IPC *SubClass* level of: Philips – left graph –, and a partial zoom-in view of it – right graph –.

4.2. **Applications: history of technology and technology management.**

In the field of history of technology, the *TechSpectrogram* of a technology – a company or a team of researches – covering a certain interval of time could highlight some milestones and transition events in the time span of study that could help the historian of technology to focus directly on them, and digging for evidences and correlations. We highlight again to the reader how the changes in the *technological components* of medicine and electric devices around the 2000s in the *TechSpectrogram* of the Philips company point to key moments in the history of this company.

In the field of technology management, the generation of the *TechSpectrogram* of a technology – a company or a team of researches – covering a certain interval of time could help to the discovery of innovation opportunities, and to the technological analysis of corporations in view of its merge and acquisition.

Traditionally, researchers looking for new innovation opportunities are interested in identifying trends in technology. The identification of growing techniques could help to focus on potential technology fronts. Our visualization tool highlights the evolution of activities by its colouring, and therefore facilitates the identification of areas of interest. For example, locating areas changing rapidly from blueish or greenish to reddish in the last years of the time span will point to candidate technology fronts. Let's analyse again the Philips *TechSpectrogram* at *SubClass* level (see the right-hand side graphs of Figure 8), it presents this sort of change in its lower area, corresponding specifically to *electric heating and lighting*[17] – **H05B** – (see the *technology component* **C–C'** in the right-hand side of figure 8. An analysis of the next level of conceptual resolution – the IPC *Group* – shows that the patents are mainly in devices for controlling light.

Innovation opportunities are also present in cases where a technology has been abandoned or its evolution has stagnated at a certain moment in time, and the reasons for this stagnation are some technical problems. Sometimes, years later the evolution of another technology make possible to bypass the ancient blocking technical problems, and therefore an innovation opportunity appears for the people linking both situations. To be aware of these innovation opportunities, it becomes capital to easily identify these areas of technological stagnation. Our visualization tool can be used to highlight these areas of stagnation identifying reddish areas that become, and stay blueish. So inactive technological areas after being highly active, and in consequence areas of interest for potential innovation.

Merge and acquisition of technological companies is a complex procedure that could be improved by the analysis of complementarity and duplicity of the technological activities of both companies [49], [50]. The generation, comparison and analysis of the *TechSpectrograms* of the companies to be merged could point to common techniques – duplicity –, or to draw attention to complementary areas between them.

5. Summary and Future Work

In this work we create a bespoke data structure defined to analyze the evolution of a technology, a company or a group of researchers based on the distribution of the IPC codes assigned to a collection of selected patents (and to its cited prior art) along an interval of years. The data structure takes the form of a 2D array. A visualization tool derived from this data structure is built, namely the *TechSpectrogram*, our tool shows a coloured rectangular graph showing different tonalities according to the activity of the respectif *technological components*. In a rather simple way, the variety and evolution in time of the

[17]Electric heating; electric lighting not otherwise provided for

technologies related to a given technique, company or team of people is highlighted. Study cases of techniques and companies are presented, and the application of our tool to the history of technology and to technology management has being briefly disclosed.

We are now identifying patterns in the *TechSpectrograms* graphs such as, among others, increasing and decreasing ramps, peaks, *bell* shapes, *plateau* or *step -up and -down* shapes. The identification of these patterns could highlight some specific situation, such as a front of development or a loss of interest for a technical area in a particular moment in time, the appropriation through some acquisition of a specific know-how, and so on. We will identify all these shapes to catalogue them, and to correlate the different patterns with technological and business events. We are also translating and testing image processing developed for frequency spectrograms analysis to our tool in order for automatic identification of common patterns between companies. Lastly, we are also investigating the effects of translating our images and outputs in a grey-scale compatible way, for increased accessibility.

At present we are also trying to produce animated maps of technologies and corporations using its *TechSpectrogram* as a temporal-multidimensional basis. We are producing yearly maps using the bins of IPC codes of the respective technologies and corporations as coordinates of a multidimensional space, and we are taking the yearly datasets for sequencing the generated yearly maps. We are trying to generate the animated maps by reducing the number of dimensions of the system, which are 8, 131 and 645 dimensions for the IPC level *Section*, *Class* and *sub-Class* respectively, to 2D using a *MultiDimensional Scaling – MDS* – algorithm [51]. We are testing different concepts of distance to implement the *MDS* algorithm. We will also investigate the automatic clusterization of the 2D array dataset, and then modelling its dynamics in order to forecast the evolution of the *technological components*.

6. Annex 1 - SQL queries in PATSTAT database

PATSTAT SQL queries of US Pat Publications with Priority Year = 1980, and title or abstract containing: tomograph+ AND imag+ AND (reconstruct+ OR generat+) AND-NOT (transmi+ OR communicat+ OR compress+ OR uncompress+ OR distribut+)

```
Select tls211_pat_publn.publn_auth , tls211_pat_publn.publn_nr , tls209_appln_ipc.ipc_class_symbol
FROM tls209_appln_ipc
JOIN tls211_pat_publn ON tls209_appln_ipc.appln_id = tls211_pat_publn.appln_id
WHERE tls209_appln_ipc.appln_id IN
  (SELECT tls201_appln.appln_id
   FROM tls201_appln
   JOIN tls202_appln_title ON tls201_appln.appln_id = tls202_appln_title.appln_id
   JOIN tls211_pat_publn ON tls201_appln.appln_id = tls211_pat_publn.appln_id
   JOIN tls203_appln_abstr ON tls201_appln.appln_id = tls203_appln_abstr.appln_id
   WHERE tls201_appln.appln_auth LIKE 'US' – Limited to USPTO applications
   AND tls201_appln.earliest_filing_year = '1981' – Year definition
   AND tls201_appln.ipr_type = 'PI'
   AND (tls202_appln_title.appln_title LIKE '%generat%'
     OR tls203_appln_abstr.appln_abstract LIKE '%generat%'
```

```
    OR tls202_appln_title.appln_title LIKE '%reconstruct%'
    OR tls203_appln_abstr.appln_abstract LIKE '%reconstruct%')
  AND (tls202_appln_title.appln_title LIKE '%imag%'
    OR tls203_appln_abstr.appln_abstract LIKE '%imag%')
  AND ( tls202_appln_title.appln_title LIKE '%tomograph%'
    OR tls203_appln_abstr.appln_abstract LIKE '%tomograph%')
  AND tls201_appln.appln_id NOT IN
   (SELECT tls201_appln.appln_id
   FROM tls201_appln
   JOIN tls202_appln_title ON tls201_appln.appln_id = tls202_appln_title.appln_id
   JOIN tls203_appln_abstr ON tls201_appln.appln_id = tls203_appln_abstr.appln_id
   WHERE (tls202_appln_title.appln_title LIKE '%distribut%'
    OR tls202_appln_title.appln_title LIKE '%transmi%'
    OR tls202_appln_title.appln_title LIKE '%communicat%'
    OR tls202_appln_title.appln_title LIKE '%compress%'
    OR tls202_appln_title.appln_title LIKE '%uncompress%'
    OR tls203_appln_abstr.appln_abstract LIKE '%distribut%'
    OR tls203_appln_abstr.appln_abstract LIKE '%transmi%'
    OR tls203_appln_abstr.appln_abstract LIKE '%communicat%'
    OR tls203_appln_abstr.appln_abstract LIKE '%compress%'
    OR tls203_appln_abstr.appln_abstract LIKE '%uncompress%')))
GROUP BY tls211_pat_publn.publn_auth , tls211_pat_publn.publn_nr , tls209_appln_ipc.ipc_class_symbol
ORDER BY tls211_pat_publn.publn_auth , tls211_pat_publn.publn_nr , tls209_appln_ipc.ipc_class_symbol
```

PATSTAT SQL queries for the citations of US Pat Publications with Priority Year = 1980, and title or abstract containing: tomograph+ AND imag+ AND (reconstruct+ OR generat+) ANDNOT (transmi+ OR communicat+ OR compress+ OR uncompress+ OR distribut+)

```
SELECT p1.publn_auth , p1.publn_nr , tls209_appln_ipc.ipc_class_symbol , tls211_pat_publn.appln_id
FROM tls211_pat_publn
JOIN tls212_citation ON tls211_pat_publn.pat_publn_id = tls212_citation.pat_publn_id
JOIN tls211_pat_publn AS p1 ON tls212_citation.cited_pat_publn_id = p1.pat_publn_id
JOIN tls209_appln_ipc ON p1.appln_id = tls209_appln_ipc.appln_id
WHERE tls211_pat_publn.appln_id in
  (SELECT tls201_appln.appln_id
  FROM tls201_appln
  JOIN tls202_appln_title ON tls201_appln.appln_id = tls202_appln_title.appln_id
  JOIN tls203_appln_abstr ON tls201_appln.appln_id = tls203_appln_abstr.appln_id
  JOIN tls211_pat_publn ON tls201_appln.appln_id = tls211_pat_publn.appln_id
  WHERE tls201_appln.appln_auth LIKE 'US'
  AND tls201_appln.earliest_filing_year = '1980'
  AND tls201_appln.ipr_type = 'PI'
  AND (tls202_appln_title.appln_title LIKE '%generat%'
    OR tls202_appln_title.appln_title LIKE '%reconstruct%'
    OR tls203_appln_abstr.appln_abstract LIKE '%generat%'
    OR tls203_appln_abstr.appln_abstract LIKE '%reconstruct%')
  AND (tls202_appln_title.appln_title LIKE '%imag%'
    OR tls203_appln_abstr.appln_abstract LIKE '%imag%')
  AND (tls202_appln_title.appln_title LIKE '%tomograph%'
    OR tls203_appln_abstr.appln_abstract LIKE '%tomograph%')
  AND tls201_appln.appln_id NOT IN
   (SELECT tls201_appln.appln_id
   FROM tls201_appln
   JOIN tls202_appln_title ON tls201_appln.appln_id = tls202_appln_title.appln_id
   JOIN tls203_appln_abstr ON tls201_appln.appln_id = tls203_appln_abstr.appln_id
   WHERE (tls202_appln_title.appln_title LIKE '%distribut%'
        OR tls202_appln_title.appln_title LIKE '%transmi%'
        OR tls202_appln_title.appln_title LIKE '%communicat%'
        OR tls202_appln_title.appln_title LIKE '%compress%'
        OR tls202_appln_title.appln_title LIKE '%uncompress%'
        OR tls203_appln_abstr.appln_abstract LIKE '%distribut%'
        OR tls203_appln_abstr.appln_abstract LIKE '%transmi%'
```

```
        OR tls203_appln_abstr.appln_abstract LIKE '%communicat%'
        OR tls203_appln_abstr.appln_abstract LIKE '%compress%'
        OR tls203_appln_abstr.appln_abstract LIKE '%uncompress%')))
GROUP BY p1.publn_auth , p1.publn_nr , tls209_appln_ipc.ipc_class_symbol , tls211_pat_publn.appln_id
ORDER BY p1.publn_auth , p1.publn_nr , tls209_appln_ipc.ipc_class_symbol , tls211_pat_publn.appln_id
```

We have collected the patent documents of the following years (until 2015) by changing the year string in the corresponding SQL query, namely: **AND** tls201_appln.earliest_filing_year = '1980'

References

[1] S. H. Cutcliffe, "The historical emergence of sts as an academic field in the united states," *Argumentos de Razon Tecnica*, vol. 4, no. 1, pp. 281–292, 2001.

[2] C. Henry and J. E. Stiglitz, "Intellectual property, dissemination of innovation and sustainable development," *Global Policy*, vol. 1, no. 3, pp. 237–251, 2010. doi:https://doi.org/10.1111/j.1758-5899.2010.00048.x.

[3] H. K. Schachman, "From "publish or perish" to "patent and prosper"," *Journal of Biological Chemistry*, vol. 281, no. 11, pp. 6889–6903, 2006. doi:https://doi.org/10.1074/jbc.X600002200.

[4] H. P. Van Dalen and K. Henkens, "Intended and unintended consequences of a publish-or-perish culture: A worldwide survey," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 7, pp. 1282–1293, 2012. doi:https://doi.org/10.1002/asi.22636.

[5] C. Freeman and L. Soete, "Developing science, technology and innovation indicators: What we can learn from the past," *Research policy*, vol. 38, no. 4, pp. 583–589, 2009. doi:https://doi.org/10.1016/j.respol.2009.01.018.

[6] E. Geisler, *The Metrics of Science and Technology.* Quorum Books, Greenwood Publishing Group Inc, 1962.

[7] K. Smith, "Measuring innovation," *in The Oxford Handbook of Innovation*, 2006. doi:https://doi.org/10.1093/oxfordhb/9780199286805.003.0006.

[8] H. F. Moed, W. Glänzel, and U. Schmoch, *Handbook of quantitative science and technology research. The Use of Publication and Patent Statistics in Studies of S&T Systems.* Springer, 2004. doi:https://doi.org/10.1007/1-4020-2755-9.

[9] W. Seo, N. Kim, and S. Choi, "Big data framework for analyzing patents to support strategic r&d planning," *2nd Intl Conf on Big Data Intelligence and Computing*, pp. 746–753, 2016. doi:https://doi.org/10.1109/DASC-PICom-DataCom-CyberSciTec.2016.131.

[10] E. Perez-Molina, "The role of patent citations as a footprint of technology," *Journal of the Association for Information Science and Technology*, vol. 69, no. 3, pp. 610–618, 2018. doi:https://doi.org/10.1002/asi.23979.

[11] L. Villanueva-Rivera, B. Pijanowski, J. Doucette, and B. Pekin, "A primer of acoustic analysis for landscape ecologists.," *Landscape Ecology*, vol. 26, no. 9, pp. 1572–9761, 2011. doi:https://doi.org/10.1007/s10980-011-9636-9.

[12] L. Leydesdorff, "Diversity and interdisciplinarity: how can one distinguish and recombine disparity, variety, and balance?," *Scientometrics*, vol. 116, pp. 2113–2121, 2018. doi:https://doi.org/10.1007/s11192-018-2810-y(01234.

[13] B. Godin, "The obsession for competitiveness and its impact on statistics:the construction of high-technology indicators," *Journal of the Patent Office Society*, vol. 33, no. 8, pp. 1217–1229, 2004. doi:https://doi.org/10.1016/j.respol.2004.07.005.

[14] Y. Okubo, "Bibliometric indicators and analysis of research systems: Methods and examples," *OECD Science, Technology and Industry Working Papers*, vol. 1997/01, 1997. doi:https://doi.org/10.1787/18151965.

[15] E. Garfield, “Breaking the subject index barrier - a citation index for chemical patents,” *Journal of the Patent Office Society*, vol. 39, no. 8, pp. 583–595, 1957.

[16] N. Reingold, “U.s. patent office records as sources for the history of inventions and technological property,” *Technology and Culture*, vol. 1, pp. 156–167, 1960.

[17] H. Ernst, “Patent information for strategic technology management,” *World Patent Information*, vol. 25, pp. 233–242, 2003. doi:https://doi.org/10.1016/S0172-2190(03)00077-2.

[18] M. Mogee and R. Kolar, “International patent analysis as a tool for corporate technology analysis and planning,” *Technol. Anal. Strateg. Manag*, vol. 6, no. 4, pp. 485–504, 1994. doi:https://doi.org/10.1080/09537329408524188.

[19] M. Carpenter and F. Narin, “Validation study. patent citations as indicators of science and foreign dependence,” *World Patent Information*, vol. 5, no. 3, pp. 180–185, 1983. doi:https://doi.org/10.1016/0172-2190(83)90139-4.

[20] U. Schmoch, F. Laville, M. Pianta, and G. Sirilli, “The measurement of scientific and technological activities: Using patent data as science and technology indicators,” *Patent Manual*, 1994.

[21] A. Martinelli and O. Nomalerr, “Measuring knowledge persistence: a genetic approachto patent citation networks,” *J Evol Econ*, vol. 24, pp. 623–652, 2014. doi:https://doi.org/10.1007/s00191-014-0349-5.

[22] J. Yoon and K. Kim, “Trendperceptor: A property function based technology intelligence system for identifying technology trends from patents.,” *Expert Systems with Applications*, vol. 39, pp. 2927–2938, 2012. doi:https://doi.org/10.1016/j.eswa.2011.08.154.

[23] P. Ellis, G. Hepburn, and C. Oppenhein, “Studies on patent citation networks,” *Journal of Documentation*, vol. 34, no. 1, pp. 12–20, 2000. doi:https://doi.org/10.1108/eb026650.

[24] A. Martinelli, “An emerging paradigm or just another trajectory? understanding the nature of technological change using engineering heuristic in the telecommunications switching industry,” *Applied Sciences*, vol. 41, no. 2, pp. 414–429, 2012. doi:https://doi.org/10.1016/j.respol.2011.10.012.

[25] B. Verspagen, “Mapping technological trajectories as patent citation networks: A study on the history of fuel cell research.,” *Advances in Complex Systems*, vol. 10, no. 1, pp. 93–115, 2007. doi:https://doi.org/10.1142/S0219525907000945.

[26] B. Spear, “Virtual reality: patent review,” *World Patent Information*, vol. 24, pp. 103–109, 2002. doi:https://doi.org/10.1016/S0172-2190(02)00002-9.

[27] E. Perez-Molina, “The technological roots of computer graphics.,” *Annals of the History of Computing*, vol. 36, no. 3, pp. 30–41, 2014. doi:https://doi.org/10.1109/MAHC.2014.47.

[28] K. Borner and A. Scharnhorst, “Visual conceptualizations and models of science,” *Journal of Informetrics*, vol. 3, pp. 161–172, 2009. doi:https://doi.org/10.1016/j.joi.2009.03.008.

[29] L. Leydesdorff, S. Carley, and I. Rafols, “Global maps of science based on the new web-of-science categories,” *Scientometrics*, vol. 94, pp. 589–593, 2013. doi:https://doi.org/10.1007/s11192-012-0784-8.

[30] W. Liu and H. Liao, “A bibliometric analysis of fuzzy decision research during 1970 – 2015,” *International Journal of Fuzzy Systems*, vol. 19, no. 1, pp. 1–14, 2017. doi:https://doi.org/10.1007/s40815-016-0272-z.

[31] S. Arora, A. Porter, J. Youtie, and P. Shapira, “Capturing new developments in an emerging technology: an updated search strategy for identifying nanotechnology research outputs,” *Scientometrics*, vol. 95, pp. 351–370, 2013. doi:https://doi.org/10.1007/s11192-012-0903-6.

[32] B. Yoon and C. Magee, “Exploring technology opportunities by visualizing patent information based on generative topographic mapping and link prediction,” *Technological Forecasting and Social Change*, vol. 132, pp. 105–117, 2018. doi:https://doi.org/10.1016/j.techfore.2018.01.019.

[33] K. Boyack, B. Wylie, and G. Davidson, “Domain visualization using vxinsight® for science and technology management,” *Journal of the American Society for Information Science and Technology (JASIST)*, vol. 53, no. 9, pp. 764–774, 2002. doi:https://doi.org/10.1002/asi.10066.

[34] B. Yoon and Y. Park, "A text-mining-based patent network: Analytical tool for high-technology trend," *Journal of High Technology Management Research*, vol. 15, pp. 37–50, 2004. doi:https://doi.org/10.1016/j.hitech.2003.09.003.

[35] Z. Liu and D. Zhu, "Web mining based patent analysis and citation visualization," *Proceedings*, pp. 19–23, 2009. doi:10.1109/WMWA.2009.33.

[36] C. Mejia and Y. Kajikawa, "Bibliometric analysis of social robotics research: Identifying research trends and knowledgebase," *Research Policy*, vol. 7, pp. 1–17, 2017. doi:https://doi.org/10.3390/app7121316.

[37] B. Yan and J. Luo, "Measuring technological distance for patent mapping," *JASIST*, vol. 68, no. 2, pp. 423–437, 2017. doi:https://doi.org/10.1002/asi.23664.

[38] M. Song, G. Heo, and S. Kim, "Analyzing topic evolution in bioinformatics: investigation of dynamics of the field with conference data in dblp," *Scientometrics*, vol. 101, pp. 397–428, 2014. doi:https://doi.org/10.1007/s11192-014-1246-2.

[39] F. Wu, R. Li, L. Huang, H. Miao, and X. Li, "Theme evolution analysis of electrochemical energy storage research based on citnetexplorer," *Scientometrics*, vol. 110, pp. 113–139, 2017. doi:https://doi.org/10.1007/s11192-016-2164-2.

[40] S. Chen, M. Huang, and D. Chen, "Visualization of the technology evolution in smart grid," *Proceedings of PICMET 2011*, vol. 5, pp. 1–7, 2011.

[41] S. Morris and G. Yen, "Crossmaps: Visualization of overlapping relationships in collections of journal papers," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, pp. 5291–5296, 2004. doi:https://doi.org/10.1073/pnas.0307604100.

[42] S. Morris, C. DeYongb, Z. Wua, S. Salmanb, and D. Yemenu, "Diva: a visualization system for exploring document databases for technology forecasting," *Computers and Industrial Engineering*, vol. 1, no. 43, pp. 841–862, 2002. doi:https://doi.org/10.1016/S0360-8352(02)00143-2.

[43] H. Sung, H. Yeh, J. Lin, and S. Chen, "A visualization tool of patent topic evolution using a growing cell structure neural network," *Scientometrics*, vol. 11, no. 3, pp. 1267–1285, 2017. doi:https://doi.org/10.1007/s11192-017-2361-7.

[44] N. Sandal and A. Kumar, "Searching and analysing patent document to solve r&d problems," *DESIDOC Journal of Library & Information Technology*, vol. 36, pp. 65–72, 2016. doi:https://doi.org/10.14429/djlit.36.2.9504.

[45] L. Fleming and O. Sorenson, "Technology as a complex adaptive system: evidence from patent data," *Research Policy*, vol. 30, pp. 1019–1039, 2001. doi:https://doi.org/10.1016/S0048-7333(00)00135-9.

[46] G. de Rassenfosse, H. Dernis, and G. Boedt, "An introduction to the patstat database with example queries," *ArXiv*, vol. abs/1404.7447, 2014. doi:https://doi.org/10.1111/1467-8462.12073.

[47] EPO, "Patstat," *At: www.epo.org/searching-for-patents/business/patstat.html*, 2019.

[48] C. Philips, "History," *At: https://www.companieshistory.com*, 2019.

[49] A. Breitzman and P. Thomas, "Using patent citation analysis to target/value m&a candidates," *Research–Technology Management*, vol. 45, no. 5, pp. 28–36, 2016. doi:https://doi.org/10.1080/08956308.2002.11671519.

[50] B. Clayton, "Understanding the unpredictable: Beyond traditional research on mergers and acquisitions," *Emergence: Complexity and Organization*, vol. 12, no. 3, pp. 1–19, 1975.

[51] I. Borg and P. Groenen, *Modern multidimensional scaling: Theory and applications*. Springer, 2005.