論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Interactions between patterns of social attention, funding, and emerging technologies: case study of robotics research
著者(和文)	MejiaCaballero Cristian Andres
Author(English)	Cristian Andres Mejia Caballero
出典(和文)	学位:博士(学術), 学位授与機関:東京工業大学, 報告番号:甲第10927号, 授与年月日:2018年6月30日, 学位の種別:課程博士, 審査員:梶川 裕矢,日髙 一義,橋本 正洋,中丸 麻由子,仙石 愼太郎
Citation(English)	Degree:Doctor (Academic), Conferring organization: Tokyo Institute of Technology, Report number:甲第10927号, Conferred date:2018/6/30, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

INTERACTIONS BETWEEN PATTERNS OF SOCIAL ATTENTION, FUNDING, AND EMERGING TECHNOLOGIES

- CASE STUDY OF ROBOTICS RESEARCH -

Cristian Andres Mejia Caballero

Dissertation submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Department of Innovation
Graduate School of Innovation Management
Tokyo Institute of Technology

Supervisor
Professor Yuya Kajikawa

March 2018

ABSTRACT

This thesis explores the underlying dynamics between social attention, funding, and emerging technologies in order to discover patterns that may improve funding allocation strategies when it comes to select emerging research trends.

It is argued that social expectations of technologies represented by levels of attention extracted from printed media may play a role in the configuration of emergent research topics. To discover if such relation exist data from academic publishing and news articles are used, and analyzed by applying methods grounded in network theory, artificial intelligence, and text mining. Objectives and research background are discussed in the introductory Chapter 1.

The target of analysis is robotics research. Robotics is known for its social engagement and broad discussions in the media, and also for the greater interest of governments in funding robotics related technologies. Making it an exemplary case study. In chapter 2 it is identified the taxonomy of robotics, and research trends. This is done through a comprehensive bibliometric study of academic literature.

Even though bibliometrics help us to reveal the landscape of robotics, little can be inferred on how specific research clusters are connected to society as a whole. Chapter 3 takes steps in that direction. It is explored the interconnection of publishing patterns of academic articles in the previously identified clusters to actual levels of social expectations. News data is incorporated as representation of social opinion, from where levels of attention and sentiment polarity are observed. Peaks of inflated expectations in relation to robotics were revealed. Moreover, it was found how positive discussions in the media are reflected as an increase of publications for specific robotics topics in a posterior time.

In the interface of solving social issues and promoting academic research, lies funding. Public and private organizations reconfigure their funding allocation strategies in response to what they believe are the most pertinent topics to target. Chapter 4, develops a method to assess whether funding organizations target innovative research.

Finally, Chapter 5 brings together emergence, expectations, and funding. Parting from observations collected in the previous chapters it is attempted to stablish a theory of funding of emergent technologies based on social attention. It is expected that the findings discussed in this research serve to identify technologies that timely bring solution to the social concerns at hand, and provide a reference for better funding allocation strategies.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Professor Yuya Kajikawa for his guidance, sharp criticism, and encouragement. But above all, for granting me trust and independence. That freedom became the greatest asset during my 4 years living in Japan.

Also, I would like to express my gratitude to Professor Kazuyoshi Hidaka whose support allowed me to attend the PICMET conferences in 2016 and 2017. An invaluable contribution for expanding my knowledge and network. And Professor Kumiko Miyazaki, who despite our little interaction her works on emergence of technology and robotics kept constantly appearing in my explorations of previous literature. Her pioneering endeavors became part of the building blocks of the present dissertation, and stimulated me to go beyond.

I am thankful to Ichiro Okabe and Yasutomo Takano, members of Kajikawa Laboratory. Always active, creating opportunities to improve our technical skills. I learnt a lot from them.

Being most of the time the only foreigner in the student room, it was good to count with the colleagues of Miyazaki Laboratory in Ookayama campus. In conversations during my visits there, we could share the adventures of living in a country that far, both in distance and culturally. They summed up to the enjoyment of being here. Thank you all.

This dissertation could not have been possible without the financial support of the Ministry of Education, Culture, Sports, Science and Technology of Japan. My greatest appreciation to the government of Japan for keeping the scholarship program going for so many years.

Finally, I thank to my family and friends in my country, Honduras. The computer screen that separated us all this time was no barrier to our strong affection. They are my source of motivation that keeps me moving forward to an even brighter future.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Funding organizations are key actors driving science and technology. It is expected for funding organizations to support new developments that leads to prominent impact in society. While these stakeholders put their best effort on achieve that, now and then examples appears of their missoperations. For instance, groundbreaking discoveries that have led to Nobel prizes have also been subject of funding rejection (Berg 2008). Because at the very early stage the uncertainty on the future contribution of some topics makes them, on the eyes of those evaluating grant applications, a risky endeavor.

The task of selecting emerging research topics that may bring greater socioeconomic impact is not an easy one. Every day more and more papers, books and reports are published, patents filed, and news reported. The pool of knowledge from where novel trends should be detected is abundant and overwhelming. Better tools are needed to make sense of that knowledge, find meaningful patterns and make better informed decisions. The need of those tools is particularly relevant when it comes the time to distribute scarce resources.

Fortunately, the conceptualization and operationalization of emergent technologies is an active field of research (Rotolo et al. 2015). Several methodologies are on the table to assess features among technologies in large portfolios, enabling to separate potential breakthroughs from matured ones. Most of those methods start on the premise that emergence can be observed by exploring characteristics of scientific outputs (e.g. academic articles), dismissing other features and actors that also participate in the configuration of new technologies.

Among them, it is acknowledged that the wider society takes an important role in shaping the directions of new technologies and industries (Borup et al. 2006; Geels & Verhees 2011; Deuten et al. 1997). However, the mechanisms for understanding that relationship are complex in nature. And most of the studies of the social shaping of technologies are based in qualitative, ex-post manner (Borup et al. 2006). Notwithstanding, some features of the wider society and their influence in the creation of new knowledge might be at reach of measure. In particular the notion of social expectations.

Social expectations have been measured through several frameworks (Alkemade & Suurs 2012). However, the latent connection of those lies in the measurement of social attention. An indicator of the intensity on how topics are discussed in society, usually using as proxy printed media (Melton et al. 2016; Dedehayir & Steinert 2016).

However, there is no understanding on how social attention, funding, and the emergent technologies observed in scientific knowledge production interact among them. Can we say that funding encourages the creation of new topics, and then those are discussed in society? Or is society who takes the first step? In this thesis, we extract starting evidence to point how these three actors (society, researchers, and funders) interact.

If social attention can be claimed as an early indicator of new developments in science, then funding organization may take advantage of that knowledge to leverage their funding strategies. As it will be possible to anticipate some upcoming targets and fund them accordingly. Researchers may also get benefit in understanding these dynamics, by knowing the stance or interest of funding organizations in regards the topics of science making possible to look for those with greater chances to receive grants.

In the following sections we further explain objectives, research background, and the data and methods of this thesis.

1.1 Objective

The objective of this thesis is to elucidate the dynamics that connect social attention, funding, and emerging technologies. To address this main objective, we independently assess how social attention and funding are related to emergent research by answering the following questions:

- Do social attention of technologies plays a role in defining new emergent research? Here, it is argued that, at least partially, academic research is determined by social expectations as represented in the imaginings of technologies in the media.
- Do funding organizations target innovative research? This thesis argues that while funding organizations attempt to allocate their resources to groundbreaking research, mechanisms in their selection strategies may constraint them to target properly.

To answer those, we apply methods from network theory and text mining to analyze patterns hidden in the large corpus of data that represent the science space and the society space. We demonstrate the feasibility of applying those methods to identify levels of expectations, and emotional response of society towards new developments.

This thesis finalizes with a discussion about the relationship between social attention, funding, and emergent research. Aiming to help policy makers, and researchers in making informed decisions for better use of limited funds, by targeting research and technologies that potentially lead to practical innovations with a greater impact in a faster pace.

1.2 Research background

This section explores the theoretical background of the main topics in this thesis. First, it is discussed the corpus of literature that address the problem of detecting emerging technologies under the science and technology policy domains. Next, the connections between social attention and funding are discussed.

1.2.1 Emergent technologies

Until recently, there was no consensus on the definition of what is an emerging technology. However, that has not been an impediment for the study of attributes of emerging technologies across the history of the innovation studies. Seminal research can be found as early as 1875 (Lewes 1874). But the conceptualization and quantification of emergence comes from more recent years, in particular from the work of Martin (Martin 1995), Porter (Porter et al. 2002), and Cozzens (Cozzens et al. 2010), whose work emphasize the use of bibliometrics, or the quantification of features from publications like academic articles or patents.

Emergence can be measured from several attributes. For instance, Avila-robbinson and Miyazaki (Avila-Robinson & Miyazaki 2011) made a comprehensive exploitation of bibliometrics attributes from publications on *micro/nano electromechanical systems*. They separated attributes based on dynamism (speed of growing), variety (of categories), diffusion (citation counts received), complexity (actors involved), chronological (novelty) to identify emergent trends. However, those are not the only attributes that can be observed from bibliometric features and is not uncommon that different authors create their own indicators.

Because of that, and with the purpose of unifying the concept of emergence Rotolo, Hicks and Martin (Rotolo et al. 2015) reviewed how emergence have been operationalized in the science and technology policy domain literature until 2015, coming to the following definition:

"relatively fast growing, radically novel technology characterized by a certain degree of coherence persisting over time and with the potential to exert a considerable impact on the socio-economic domains which is observed in the composition of actors, institutions and patterns of interaction among those[…] it's most prominent impact, however, lies in the future and so in the emergence phase it is somewhat uncertain and ambiguous" (Rotolo et al. 2015).

Where 5 attributes are present. Radical novelty, relatively fast growth, coherence, prominent impact, and uncertainty and ambiguity. The former three can be observed through quantitative methods.

First, novelty or newness. Emergent technologies are expected to be new. This newness can be observed from publication years of a cohesive group of literature. Some researchers may first classify the literature by using methods like citation network, and identify groups within that network, those groups that are newer can be considered emergent (Garfield et al. 1964). In this line, Takano et al. (Takano et al. 2017) built a classification schema based on novelty. by exploiting the average year of the groups, and the novelty of the leading paper in those groups, papers can be classified into four categories: breakthrough, change maker, incremental, and matured.

Fast growth is other well expected attribute of emergence, and the one most operationalized. Referring to the number of articles on the same topic that are aggregated in subsequent periods of time. The technology-life cycle, or s-shape trend literature is part of this feature (Ho et al. 2014; Abercrombie et al. 2012).

Then, coherence. Which is the extent on how topics become part of a unified narrative. This can be observed from an analysis of the text content (Furukawa et al. 2015), or by studying the network properties when a network is involved in the analysis (Fujita et al. 2014).

The present thesis built upon the above-mentioned framework to measure indicators of emergence for novelty, growth, and coherence. The other two attributes of emergence are impact, and uncertainty. The assessment of prominent impact, which is usually expected but rarely measured (Rotolo et al. 2015), is part of a more qualitative corpus of research (Borup et al. 2006). Few literature exist on the latter one from a bibliometric perspective (Lucio-Arias & Leydesdorff 2009).

1.2.2 A transdisciplinary approach for funding science

The involvement of stakeholders outside the academia in the pursue of shared goals refers to transdisciplinary research (Tress et al. 2005). The stakeholders, also referred as practitioners, serve as counterpart to the scientific community all along the process of knowledge production. Because both parts takes a relevant role in the goal setting and process it is said that the outputs of transdisciplinary research are socially robust, meaning that aim to solve real-world problems and are expected to be socially legitimated upfront. Besides of also being aware of the local and historical context (Nowotny 1999; Klein 2015).

Tress et al. schematize transdisciplinary research like in Figure 1-1 where all participants work together to define and accomplish a common goal. In its core transdisciplinarity refers to different academic disciplines working jointly with practitioners aiming to solve real-world problems (Klein 2015). And is considered a research approach, not a method or theory, that enables mutual learning between science and society (Jahn et al. 2012).

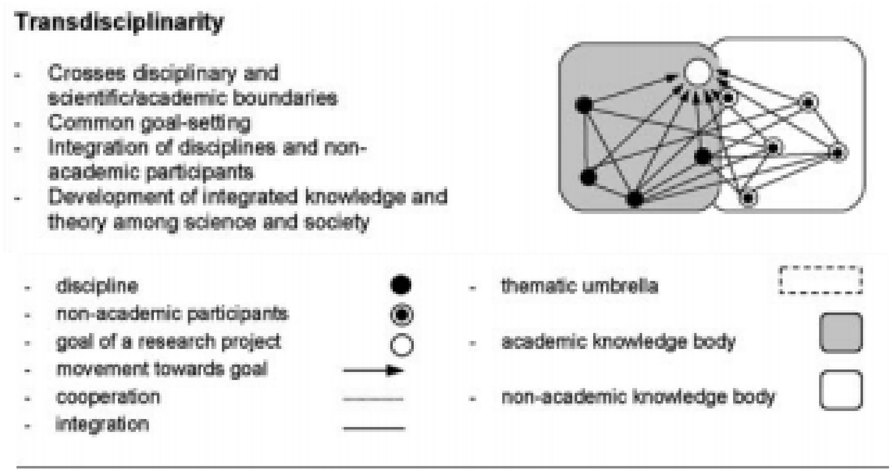

Figure 1-1 Characteristics of Transdisciplinary research as in Tress et al. (2005)

Transdisciplinarity is a response to the increasing specialization and granular partition of scientific fields, as it is natural result to the ever blurrier line that separates science and society. According to Gibbons we moved to a "Mode-2" of science. In a previous model, science acted as a provider of reliable knowledge to society, and thus, being unidirectional. The driver of this model is the individual creativity of the researcher (Gibbons 1999). In the new model, however the outputs of science require a dialogue with society, and thus the creativity comes up as a group phenomenon being the outputs more sensible to social implications (Gibbons 1999).

What is the role of funding and funding organizations in this narrative of transdisciplinarity research? While funding organizations may be taken as stakeholders within society, in the present study we define them as a separate group of actors, being active between the researchers and the rest of stakeholders and practitioners. And in particular, being the actors who listen to society and channel resources to the academia in order to accomplish transdisciplinary goals.

The relevance in putting funding organization as brokers of needs and sources is to take advantage of the better decision making capabilities that the transdisciplinary approach provides. As stated by Walter et al. (2007) "transdisciplinarity aims to increase decision making capacity of stakeholder in an effective, transparent, and reliable way through mutual learning between researchers and stakeholders and by providing robust future orientations for practice".

Moreover, transdisciplinarity is linked to uncertainty, controversy, and complexity; the fifth attribute of emergent technologies. By aiming social recognition of scientific outputs, philosophy and direction can be lost (Aeberhard & Rist 2009). This lead to a recognized gap regarding transdisciplinary frameworks or methodologies that work in the interface of science and society for the benefit of knowledge production (Lawrence 2015; Klein 2015), in particular little is known on structural factors in the social shaping of technological development in the transdisciplinary ecosystem (Klein & Kleinman 2002), a gap that we attempt to contribute in partially fill.

1.2.3 Society, social attention and funding

Society as a concept can be defined in multiple ways (Dean 2010; Nakano 2012). However, in this thesis society is seen as the group stakeholders beyond academia, funding organizations, and governments. Those stakeholders or users are aware of the introduction of technologies in their social context and thus have an opinion, well informed or not. Following the tradition of social studies of science and technology, what can be assessed from society are their expectations. In this regard, expectations are defined as "*real-time representations of future technological* *capabilities*" (Borup et al. 2006). Those representations can be extracted by several means, like public opinion polls, surveys, and news. In the mode-2 production of knowledge the space for dialogue between academics and society might take place anywhere. But the representation of society in printed media has been found as a good resource of information of public opinion (Gibbons 1999; Nowotny et al. 2001). In particular, it is aimed to capture the level of attention of the general publics as represented in the news media in relation to emerging technologies.

One of the attributes of emerging technologies is their prominent impact. Emerging technologies are expected to generate impact by reconfiguring the patterns of knowledge production, or changing the interactions among actors or institutions. This changes are bidirectional, in consent to the literature of the coevolution of technology and society (Nelson & Winter 1982). This is, new technologies emerge partially due to the expectations and visions of actors and users, however, in subsequent stages those actors and user change in their interactions because of the impact exerted by the technology envisioned, and then new expectations appears.

The cycles of expectation and technology impact are assessed from mixed frameworks of quantitative and qualitative nature. Being news articles, reports, reviews among others the sources from where statements of expectations are extracted for study. For instance, Geels and Verhees (Geels & Verhees 2011) explained the rise and fall of nuclear energy in the Netherlands by revising its social representations in printed media, reveling the power of expectation in legitimating the innovation journeys of technology. Similar approaches have been used by Alkemade and Suurs (Alkemade & Suurs 2012) to explain the emergence of sustainable technologies.

One of the tools used to operationalize expectations is the hype cycle (Fenn & Raskino 2008). Hype cycle proponents argue that technologies follow a fixed pattern of social expectations, including an initial period of over expectations, followed by a fall of disappointment, from which just few stakeholders will survive. However, the existence of the hype cycle is disputed and instead a more generalized form of hype dynamics have been proposed (Van Lente et al. 2013). Those hyped dynamics are a result of measurement of the level of attention. Which is the amount statements referencing a topic. Or simply put, the amount of news (etc.) articles related to a given technology (Melton et al. 2016).

Those expectations reconfigure the path of technology development, and thus funding organizations are expected to respond accordingly. So far, there is no literature that address the relationship of social attention and funding. However, they might share an indirect linkage as they both belong to the coevolving system. The present thesis attempts to reveal that connection.

1.3 Data sources and methods

One of the outputs of technology research comes in the form of articles in academic journals, from where solutions to social problems can be extracted. The constant evolution and diversification of technology research makes the comprehensive understanding and discovery of plausible solutions a difficult task. Even experts may overlook solutions that appear in distant disciplines. Data mining enables researchers and decision makers to overcome such concerns by applying techniques that reveal the intricate connections among several fields of knowledge otherwise hidden in the large databases of information.

Certainly, academic papers are not the sole output of technology research, but they have characteristics that facilitate large-scale analysis to reach conclusions faster. One of those characteristics are the cited references they contain, and the knowledge network that can be created of. These citation networks have been broadly studied (Radicchi et al. 2012; Borner & Scharnhorst 2009) and the use of them dates back to 1965 with the work of de Sola Price (de Solla Price 1965). They have been used to map the landscape of science (Kajikawa et al. 2014), create taxonomy of knowledge (Klavans & Boyack 2017), find research and technology fronts (Shibata et al. 2008) among other uses (Persson 2010; Leydesdorff & Rafols 2011). Innovation research has also been subject of citation network analysis by Kajikawa et al (Kajikawa et al. 2012). In their work, research on "innovation" was studied through a network of papers, that was divided into tightly connected clusters that represented specific knowledge of the discipline. This analysis helped to understand the development of innovation studies and drawn recommendations for future directions. Citation networks have proven to be a sound methodology in the science and technology studies, and innovation.

On the other hand, newspapers are a good source of information about society. They cover topics that deserve debate, action and legal regulation (Janssen et al. 2008), influence the formation of public opinion (DiMaggio et al. 2013) and they are always up to date. Newspaper data has been used in innovation research to measure public expectations towards technology and to understand legitimacy in technology adoption (Geels & Verhees 2011). In the work of Melton et al (Melton et al. 2016) international newspaper articles were used to measure the hype and disappointment cycles in relation to seven alternative fuel vehicle technologies to give recommendations towards the displacement of fossil fuels. Nevertheless, the use of this type of data has been scarce in the literature, but its value and crescent interest of application in technology studies has been identified (Rotolo et al. 2015).

The challenge when analyzing newspaper articles is their unstructured nature, they lack bibliographic references, and vary in size and vocabulary usage. To deal with that, we can apply topic models, a methodology that has been used to analyze this type of data in the social science and humanities (Mohr & Bogdanov 2013). Topic models are algorithms that extract the themes, or topics, that compose a collection of documents. Their automated nature facilitates the exploration of large corpus of text, an advantage over previous research that has been relied on the actual reading and manual coding for classification of news articles to reach conclusions.

News articles also embed other type of information. Like the emotional response of the public, or sentiment. News articles or part of them, can be classified, for instance as positive or negative, depending on the views of the writer or the target topic. To evaluate sentiment in news researchers have had to realize content analysis, creating tagging rules and reading each article to label accordingly, and this is subject of time and cost limitations. However, recent advancement in artificial intelligence and machine learning have helped to conduct sentiment analysis faster and accurate.

In summary, this thesis analyze emergent research from academic articles, and social aspect from news. The former are studied by applying citation networks, and features of the latter are studied with topic model and sentiment analysis.

1.4 Outline

This thesis is organized as shown in Figure 1-2. The present introductory chapter discussed the research background and objectives pursued. In Chapter 2, we take a look to robotics research. We state the reason of selection of robotics as object of study, and also bring an exploratory study based on bibliometric techniques to reveal the main research trends. Chapter 3 and Chapter 4, are the main topics of this dissertation. Chapter 3, reveals how social attention is connected to research topics in the robotics domain. We elaborate a methodology to stablish such connection, and demonstrate the role of news in the configuration of new topics in the robotics. Chapter 4 is about funding organizations. Topics in the robotics research landscape are assessed in terms of novelty, an attribute of emerging technologies, that help us to identify matured, incremental, breakthrough and change making technologies. This classification helps us address the question whether funding agencies are promoting innovative research. We identified the leading organizations, and their topics of interest. Chapter 5, we discuss the findings of previous chapters, and their implications in light of emerging technologies, here we argue that a pattern is observable in regards on how funding organizations react to emergent topics in science. After revealing the starting evidence of that pattern we conclude the thesis summarizing the contributions, spotting improvement opportunities, and gearing to future directions.

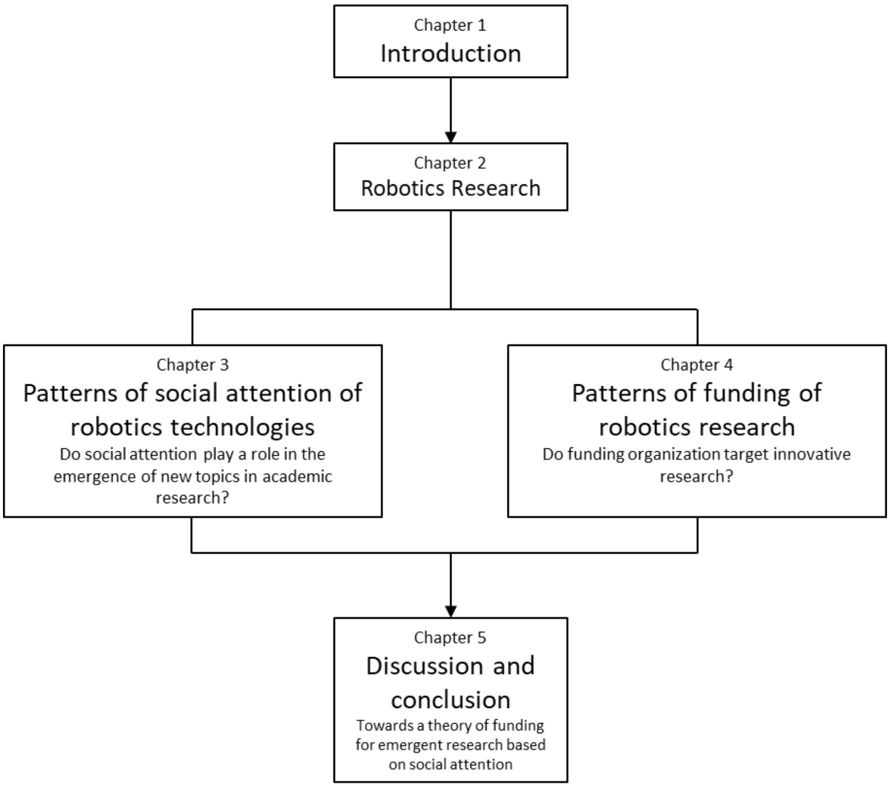

Figure 1-2 Outline of this thesis

2 ROBOTICS RESEARCH

2.1 Introduction

This chapter explores the target of study, the robotics research. Through a bibliometric study of academic articles main trends are observed. The objective of this chapter is to present the structure of robotics research, represented by the knowledge network of academic articles. The methodology applied, and the academic landscape presented in this chapter are constant for the remaining of the thesis.

Three main reasons make the robotics an exemplary case study. First is historical, robotics as a topic of social interest has been discussed in literature and media long before its appearance in industry and academia. The term *robot* was spread internationally due to the theater play Rossum's Universal Robots by Karel Capek in 1920 (Capek 1920), derived from the Slavic language word *robota* meaning forced labor or servitude. The keyword "robot" became catchy at that time to the extent of enduring several decades before being used in patents and academic articles. The term have been used in news articles since then, as a metaphor of an advance technology able to take jobs and exceed human capabilities (Woodburry 1927). This thesis studies the interactions between academic outputs and society indicators through methodologies based on text analysis of papers and news articles. Hence, the presence of social discussion of robotics since the earliest stage of robotics research makes possible a comprehensible analysis.

Secondly, robotics is a growing field. Both research and development represented in papers and patents show steady growths. Market penetration of robotics is also optimistic. The International Federation of Robotics estimates that around 5.4 million service robots for personal or domestic use where sold in 2015, being an increase of 16% in relation to the previous year. With projections pointing to 42 million units sold by 2019 (International Federation of Robotics 2016b), and even higher increase rates for industrial robotics to 20% annually, reaching an estimation of $35 billion of global market value for 2015 (International Federation of Robotics 2016a). With such positive odds and tech leaders "dreaming on robots in every home" (Gates 2007), the studies of the implication of the adoption of robotics technologies in the society has never been more necessary.

Finally, Robotics has drawn worldwide attention and is integrated into the technology development strategies of several countries (Ministry of Economy Trade and Industry of Japan 2015; National Science Foundation 2016; SPARC 2016), suggesting that it receives governmental support and warrants public agency involvement. In this thesis we correlate funding, and funding agency involvement, to indicators of emergence of technology and prominent impact. The field of robotics collects those attributes.

In order to understand what is happening in the robotics research we explore the knowledge structure, or academic landscape that can be built by applying bibliometrics methods. This is the construction of knowledge structure based on the characteristic or properties of published materials; academic articles for this case.

Bibliometrics has become an important tool to grasp the entire perspective of a research domain, especially when the amount of publication is large and its delimitation is complex. We can expose the taxonomy of a field by creating maps of science or academic landscapes. Technique that has been used to map large corpus of literature, like maps of all scientific knowledge (Boyack et al. 2005), or mapping specific ones like nanobiotechnology (Takeda et al. 2009).

Bibliometric techniques have been applied before to the field of robotics. An overview of the field have been conducted to discover the strengths and weakness of Japanese robotics (Naito et al. 2013), Also focused in Japan, this time to analyse the system of innovation Kumaresan and Miyazaki (Kumaresan & Miyazaki 1999) analyzed patents, papers, and market data revealing trends towards the servitization of robotics. Goeldner et al. analyzed patents and publications on care-robotics (Goeldner et al. 2015), and Ittipanuvat et al. (Ittipanuvat et al. 2014) used academic landscapes to establish connections between robotics and gerontology. This chapter, however, offer a more comprehensive view of the field.

2.2 Data and methods

A summary of the data collection and treatment is shown in Figure 2-1. Data were collected from the Web of Science, a widely used database of bibliographic information. The Web of Science incorporates the Science Citation Index Expanded, the Social Sciences Citation Index, and the Arts & Humanities Citation Index, covering more than 18,000 journals, and conference proceedings. We searched for articles of social robotics by querying "robot* in the web of science search engine. The asterisk serves as a wild card, allowing to extract articles having words starting with "robot" (e.g. robots, robotics) in either the title, abstract or keywords. 200,139 articles were found. However, such approach is prone to retrieve articles containing the query that may not be directly related to the field. To filter out those unrelated articles we used the properties of the citations networks. We focus in those articles which include at least one reference to another article in the retrieved dataset, and neglected the rest. 142,587 (71.2%) articles accomplished this condition. Data were retrieved on July 1st. 2017.

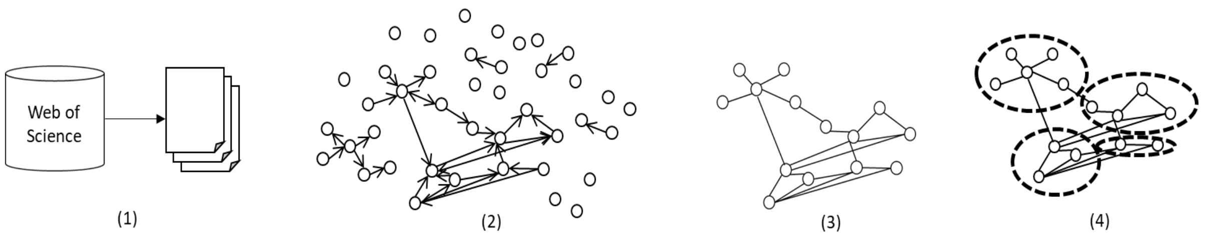

Figure 2-1 Overview of the methodology in chapter 2.
(1) Data retrieval. (2) A citation network is created based on the references of the articles. (3) The largest connected component is extracted. (4) Clusters are obtained from the network.

As mentioned above, the list of references was used to create a citation network of articles. A direct linkage is established between two articles if one of them mentions the other in its references. This is known as intercitation or direct citation. Compared to other linkage methodologies like co-citation (Small 1973) and bibliographic coupling (Kessler 1963), direct citation networks have been found to bring most accurate representation of knowledge taxonomies (Klavans & Boyack 2017), and be better at identifying research fronts (Shibata et al. 2008). We extracted the largest connected component, which is the one having the tightly connected structures of knowledge related to our research target. Within the largest component we identify clusters of papers densely connected by applying a topological algorithm based on modularity maximization. Modularity is a measure of the strength of connections within partitions or clusters. A high modularity value implies that intra-cluster connections are dense whereas inter-cluster connections are sparse. It is defined as follows (Clauset et al. 2008):

$$\boldsymbol{Q} = \sum_{s=1}^{N}\left[\frac{\boldsymbol{l_s}}{\boldsymbol{l}} - \left(\frac{\boldsymbol{d_s}}{\boldsymbol{2l}}\right)^{2}\right] \qquad \text{Eq. 2-1}$$

Where N is the number of clusters, l_s and d_s are the number of links, and the sum of the degrees of nodes within cluster s, respectively. The algorithm iterates until the maximum modularity is reached, thus obtaining the number of cluster automatically. The network is then visualized by applying a spring layout suitable for large networks, which places similar cluster close each other (Adai et al. 2004). To facilitate the interpretation of the network we visualized only the internal edges of each cluster, and assigned them different colors.

The above-mentioned steps complete the first level of clustering. In large networks, where partitions are obtained from modularity optimization, the issue of resolution limit is present. Resolution limit is the phenomena in which several small clusters are aggregated as single ones. To avoid such condition the clustering algorithm is applied recursively. For each level, top clusters that collect up to 90% of articles were subdivided if they fit the condition of the resolution limit (Fortunato & Barthélemy 2007):

$$\boldsymbol{l_s} - \sqrt{\boldsymbol{2L}} > \boldsymbol{0} \qquad \text{Eq. 2-2}$$

Where l_s is the number of edges within the cluster, and L the edges in the network or superior level sub-cluster. By following this condition, the network was sub-clustered until the 4th level, thus obtaining fine grained resolution.

Once the cluster are obtained, we revised the articles within them and labeled each cluster based on the contents of their most cited academic articles.

2.3 Results

2.3.1 Academic landscape of robotics research

Robotics is a growing field. Publication trends are shown in Figure 2-2, The earliest article in our dataset is from 1969 (Halpern 1969), scattered publications exist all across until the mid-80’s when the field started growing. Reaching 14,638 papers published in 2016, with an annual increase rate of 12.7% on average for the last five years. Hence, the need to better grasp knowledge structure by examining its academic landscape.

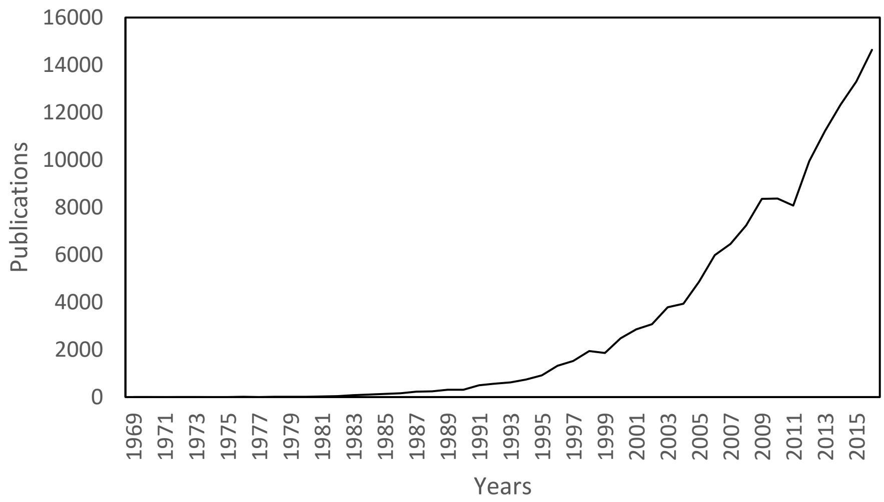

Figure 2-2 Trend in publications of robotics research

By exploiting the information on the references of the articles a citation network was created. We extracted the maximum component that contained 142,587 articles (71.2% of the total) and divided the network into clusters based on its topological characteristics. Figure 2-3 shows the citation network. We found 4 clusters collecting 92% of the robotics network. We labeled those as *Automation and control systems*, *Robot locomotion*, *Autonomous robots,* and *Robotic surgery*.

These 4 cluster can be considered as the main streams of research. To reach fine grained cluster resolution, the clusters were recursively subdivided by applying the same algorithm. We repeated this division up to the fourth level of depth, where the sub-clusters were small enough to continue the splitting. Table 2-1, provides a summary of the number of clusters in each level and their sizes.

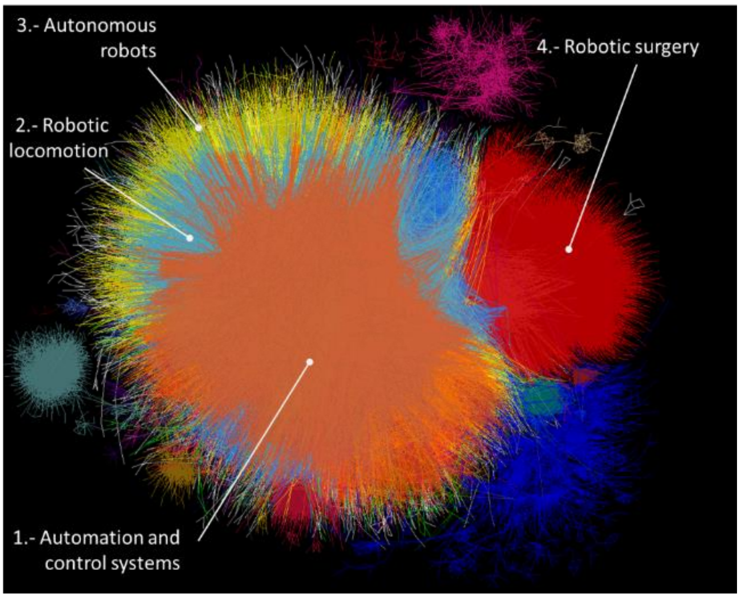

Figure 2-3 Citation networks of robotics research. (First level of clustering)

Table 2-1 Granularity of recursive clustering

Clustering depth	Number of clusters	Size (articles)		
		Min	Average	Max
Level 1	4	16,792	31,890	40,641
Level 2	30	371	3,796	14,885
Level 3	122	25	792	4,935
Level 4	450	9	180	1,329

To bring an overview of the robotics research we will focus on the four clusters in the first level, and their three largest sub-clusters in the second level. Table 2-2 summarizes their contents.

Table 2-2 First and second level clusters

	Cluster	Sub-cluster	Articles	%	Ave. Year
1	Automation and control systems	1.1 Path planning / Obstacle avoidance	14885	10.4%	2009.1
		1.2 Visual servoing	10884	7.6%	2006.8
		1.3 Robot manipulators	9020	6.3%	2006.3
2	Robotic locomotion	2.1 Walking robots	11431	8.0%	2010.0
		2.2 Robot learning	10020	7.0%	2009.6
		2.3 Rehabilitation Robotics	8588	6.0%	2011.6
3	Autonomous robots	3.1 Simultaneous localization and mapping	8220	5.8%	2009.2
		3.2 Human-Robot Interaction	6481	4.5%	2011.3
		3.3 Tactile sensing and robotic grasping	6368	4.5%	2008.5
4	Robotic surgery	4.1 Robotic assisted surgery	5061	3.5%	2009.9
		4.2 Robot assisted radical prostatectomy	3859	2.7%	2012.5
		4.3 Transoral robotic surgery	3585	2.5%	2011.7

At the first level, robotics research seems to resemble a gradient from basic to applied research. Automation and control system is the collection of articles about actuators, sensing, programming devices and frameworks that enable robotics systems in general. Robotics locomotion, tackle on mobile robotics. Mobile robotics as research field can be considered the originator of what is known today as service robotics. Therefore, a more specialized field. Third, autonomous robots. A step further on the mobile robotics research, containing articles of technologies for robotics in unstructured environments. And finally, robotics surgery. Articles focused in research and development of robotics systems that helps doctors or patients during medical interventions.

At the second level, the average year of clusters range from 2006.3, robot manipulators, to 2012.5, Robot assisted radical prostatectomy. In general, sub-clusters of automation and control systems are larger and older than the rest, hinting their maturity in the field. Following, we identify key players for this sub-clusters as listed in Table 2-3.

Table 2-3 Key actors in robotics research

Sub-cluster	Country	Institution	Journals / Conferences	Authors
1. Automation and control systems				
1-1 Path planning / Obstacle avoidance	USA	Carnegie Mellon Univ	IEEE Int Conf Robot	Kumar, V
	Peoples R China	Univ Penn	Lect Notes Comput Sc	Yang, SX
	Canada	Georgia Inst Technol	IEEE Int C Int Robot	Sukhatme, GS
	Japan	MIT	Robot Auton Syst	Savkin, AV
	France	Univ Illinois	Int J Robot Res	LaValle, SM

Sub-cluster	Country	Institution	Journals / Conferences	Authors
1-2 Visual servoing	Peoples R China	Nanyang Technol Univ	IEEE Int Conf Robot	Ohnishi, K
	USA	Keio Univ	P Amer Contr Conf	Cheah, CC
	Japan	Harbin Inst Technol	IEEE Decis Contr P	Santibanez, V
	Canada	Chinese Acad Sci	IEEE T Robotic Autom	Chaumette, F
	Taiwan	Shanghai Jiao Tong Univ	Robotica	Dixon, WE
1-3 Robot manipulators	Peoples R China	Harbin Inst Technol	IEEE Int Conf Robot	Zhang, YN
	USA	Tsinghua Univ	Mech Mach Theory	Gautier, M
	Canada	Shanghai Jiao Tong Univ	Robotica	Korayem, MH
	Japan	Sun Yat Sen Univ	J Robotic Syst	Liang, B
	France	Chinese Acad Sci	IEEE T Robotic Autom	Xu, WF
2. Robotic locomotion				
2-1 Walking robots	USA	Carnegie Mellon Univ	IEEE Int Conf Robot	Takanishi, A
	Japan	Univ Tokyo	IEEE Int C Int Robot	Kajita, S
	Peoples R China	MIT	IEEE-Ras Int C Human	Caldwell, DG
	Germany	Natl Inst Adv Ind Sci & Technol	Robotica	Yokoi, K
	South Korea	Korea Adv Inst Sci & Technol	Int J Robot Res	Kim, JH
2-2 Robot learning	USA	Carnegie Mellon Univ	IEEE Int Conf Robot	Dorigo, M
	Japan	Ecole Polytech Fed Lausanne	Lect Notes Comput Sc	Peters, J
	Germany	MIT	Lect Notes Artif Int	Rus, D
	England	Univ Tokyo	IEEE Int C Int Robot	Tani, J
	Peoples R China	Univ So Calif	Robot Auton Syst	Nolfi, S
2-3 Rehabilitation Robotics	USA	Northwestern Univ	IEEE Int Conf Robot	Riener, R
	Japan	MIT	Int C Rehab Robot	Krebs, HI
	Italy	Scuola Super Sant Anna	IEEE Eng Med Bio	Reinkensmeyer, DJ
	Peoples R China	Rehabil Inst Chicago	IEEE Int C Int Robot	Albu-Schaffer, A
	Germany	ETH	P IEEE Ras-Embs Int	Hogan, N
3. Autonomous robot				
3-1 Simultaneous localization and mapping	USA	Carnegie Mellon Univ	IEEE Int Conf Robot	Burgard, W
	Peoples R China	Univ Freiburg	Robot Auton Syst	Siegwart, R
	Germany	MIT	Lect Notes Comput Sc	Thrun, S
	Spain	Chinese Acad Sci	IEEE Int C Int Robot	Dissanayake, G
	South Korea	Korea Adv Inst Sci & Technol	Lect Notes Artif Int	Lee, S
3-2 Human-Robot Interaction	USA	Osaka Univ	Lect Notes Artif Int	Ishiguro, H
	Japan	Carnegie Mellon Univ	ACMIEEE Int Conf Hum	Kanda, T
	Germany	Univ Hertfordshire	Int J Soc Robot	Hagita, N
	England	Tokyo Metropolitan Univ	Lect Notes Comput Sc	Dautenhahn, K
	Italy	MIT	IEEE Int Conf Robot	Kubota, N
3-3 Tactile sensing and robotic grasping	USA	Shanghai Jiao Tong Univ	IEEE Int Conf Robot	Kragic, D
	Peoples R China	Univ Tokyo	IEEE Int C Int Robot	Chen, SB
	Japan	Chinese Acad Sci	Ind Robot	Arimoto, S
	Germany	Ritsumeikan Univ	Int J Adv Manuf Tech	Xi, N
	Italy	Harbin Inst Technol	Robot Cim-Int Manuf	Dillmann, R
4. Robotic surgery				
4-1 Robotic assisted surgery	USA	Yonsei Univ	Surg Endosc	Oleynikov, D
	Germany	Cleveland Clin	Int J Med Robot Comp	Bonatti, J
	Italy	Stanford Univ	Ann Thorac Surg	Chitwood, WR
	South Korea	Univ Illinois	J Thorac Cardiov Sur	Bonaros, N
	Japan	Johns Hopkins Univ	J Laparoendosc Adv S	Schachner, T
4-2 Robot assisted radical prostatectomy	USA	Mayo Clin	J Endourol	Menon, M
	Italy	Univ Calif Irvine	Bju Int	Patel, VR
	Germany	Yonsei Univ	Urology	Zorn, KC
	France	Mem Sloan Kettering Canc Ctr	Eur Urol	Ahlering, TE
	South Korea	Harvard Univ	J Urology	Tewari, AK
4-3 Transoral robotic surgery	USA	Johns Hopkins Univ	IEEE Int Conf Robot	Yang, GZ
	Germany	London Imperial Coll	Laryngoscope	Taylor, RH
	South Korea	Yonsei Univ	Lect Notes Comput Sc	Okamura, AM
	England	Univ Pittsburgh	Int J Med Robot Comp	Chung, WY
	Peoples R China	Univ Penn	Head Neck-J Sci Spec	Duvvuri, U
	South Korea	Harvard Univ	J Urology	Tewari, AK

For autonomous and control systems, China and U.S. take the leading roles, followed by Canada and Japan. In the remaining clusters US heads, the research landscape. Being constants China, Japan, and Germany. This last one having an important role in robotics surgery. For the institutions there is more variety. Universities seems to excel within specific topics. However, the MIT and Carnegie Mellon University are the two institutions whose research is frequently appearing in a broad of topics.

Journals and conference as specialized venues of publishing can be expected of being focused on particular topics. And in the case of robotics is mostly so. Particular is the case of the IEEE International Conference of Robotics and Automation, which papers tops 10 of the sub-clusters. Followed by EEE/RSJ International Conference on Intelligent Robots And Systems, Lectures Notes on Computer Science. Finally, as expected authors have a narrow focus of research, they do not go beyond their respective clusters.

2.3.2 Main research streams

Next, sub-clusters are explored in detail. Cluster contents are discussed in light of the most cited papers within them.

Cluster 1: Automation and control systems

Sub-cluster 1-1 Path planning / Obstacle avoidance

Algorithms for decision making in road mapping are discussed in this cluster. They are applied to less complex architecture of robotics, like holonomic robots where the estate of components is known in every moment. It also contains, path planning algorithm for robot teams (Khatib 1986; Kavraki et al. 1996; Balch & Arkin 1998).

Sub-cluster 1-2 Visual servoing

Visual servoing refers to the study of road mapping by means of tracking features in a sequence of images captured on a camera. Commonly the task of visual servoing is to estimate 3D shapes of what is captured, and then interpreta a path based on those estimations (Hutchinson et al. 1996; Slotine & Weiping Li 1987; Espiau et al. 1992).

Sub-cluster 1-3 Robot manipulators

Here, a set of articles tackling on the computation of states of robotic manipulators. In particular, metrics to assess the accuracy of those computations in reference to the Jacobian matrix and condition number (Yoshikawa 1985; Bobrow et al. 1985; Cannon & Schmitz 1984).

Cluster 2: Robot locomotion

Sub-cluster 2-1 Walking robots

This cluster discuss the many technologies for legged-movement. In particular bipeds and hexapods. Biped walking is inextricably related to humanoids, thus highly cited paper includes those on algorithms and systems for robots like Honda humanoid robot. Also, those towards the achievement of natural walking by applying passive-dynamics. And the pursue of energy efficient, with less components, and more reliable and robust designs (Hirai et al. 1998; Collins 2005; Saranli et al. 2001).

Sub-cluster 2-2 Robot learning

The robot learning sub-cluster contains research on robotics learning from demonstration (Argall et al. 2009), self-reconfigurable robot systems (Yim et al. 2007), and artificial intelligence methods for movement planning and world representation (Brooks 1991). It covers research on machine learning techniques that can be applied in robotics systems to decide actions and trajectories (Schaal 1999).

Sub-cluster 2-3 Rehabilitation Robotics

In this cluster, researchers report on the benefits of applying robotics for rehabilitation of patients in several types of scenarios. Most cited articles focus on the rehabilitation of upper-limb movement after stroke (Kwakkel et al. 2008), randomized clinical trials comparing robotics, usually exoskeletons, to conventional techniques (Colombo et al. 2000; Lum et al. 2002).

Cluster 3: Autonomous robots

Sub-cluster 3-1 Simultaneous localization and mapping

This cluster gathers papers tackling on the problem of simultaneous localization and map building (SLAM). Which a challenging task in autonomous robots. Robots are located in an unknown starting location of an unknown environment, then they incrementally build a map of this environment while simultaneously computing their location. Methods cover the identification and representation of shapes, to statistical methods for movement decision (Besl & McKay 1992; Gamini Dissanayake et al. 2001; Thrun 2001).

Sub-cluster 3-2 Human-Robot Interaction

Is a diverse cluster collecting research on social and service robotics, here we found research on human factors for the design of robotics, care robotics, educational and recreational robotics, and emotion detection. The central topic of study is *robots as partners* and *robots for the elderly* (Fong et al. 2003; Kanda et al. 2004; Breazeal 2003a). Also covers methods on gesture recognition, gesture commands, and robot understanding of pointing behavior. And also military robotics and ethical issues of robots (Sparrow 2009).

Sub-cluster 3-3 Tactile sensing and robotic grasping

Hand-like design, grasping and manipulation robotics devices are the subject of study of this cluster. It covers the development of tactile sensors to capture properties that can be measured through contact, including the shape of an object, texture, temperature, hardness, moisture content, etc. (Dahiya et al. 2010; Lee & Nicholls 1999; Shimoga 1996).

Cluster 4: Robotics surgery

Sub-cluster 4-1 Robotic assisted surgery

This cluster is in itself an overview of robotics assisted surgery, review articles on the benefits, clinic trials, and incorporation of new technologies were discussed. In particular, laparoscopic surgery is well documented (Giulianotti et al. 2003; Ballantyne 2002; Marescaux et al. 2001).

Sub-cluster 4-2 Robot assisted radical prostatectomy, and

Sub-cluster 4-3 Transoral robotic surgery

Both cluster targets robotic technologies for the cases prostate removal (Ficarra et al. 2009; Ahlering et al. 2003; Tewari et al. 2003) and treatment of oral (mouth, throat, tongue, tonsil) cancer respectively. Researcher articles in these cluster describe progress in robotics systems that help surgeon identify damaged tissues, and enhance their precision of movement in narrow spaces (O'Malley et al. 2006; Gregory S Weinstein et al. 2007; G S Weinstein et al. 2007).

As explained, the results above describe the contents of the network of robotics research at first and second level. However, to test the indicators of emergence further clustering was conducted to achieve topics as specific as possible, as long as the topological features of the network allowed subdivisions. Summary of the contents of the 122 and 468 clusters of the third and fourth level can be as an interactive dashboard here and here.

2.4 Summary

Robotics research is an exemplary case study for the task of elucidating the interactions between social attention, funding, and emerging technologies. This is given by its historical presence and social discussion, a growing market, and public funding involvement. A descriptive analysis of major trends was made by exploring the citation network of academic articles in the domain. It is dived into 4 larger clusters: automation and control systems, robot locomotion, autonomous robot, and robot surgery. In the following chapters we explore how such cluster relates to social attention, and how funding agencies participated within them.

3 PATTERNS OF SOCIAL ATTENTION OF ROBOTIC TECHNOLOGIES

3.1 Introduction

The increment of use in robotics technologies have been accompanied with greater social attention, that can be observed from an increase in the media coverage. This chapter takes advantage of that coverage in printed media to study how society responds to, or motivate the creation of, new technology knowledge. Therefore, it is explored the role of social attention in the creation of new research in the academia.

First off, is needed to study the role of news in the context of the dialogue between science and technology. As we are situated in a "mode-2" of knowledge production where ideas from practitioners in society are situated at the same level of those of scientist to reach unified goals (Gibbons et al. 2010), we need a place for discussion to happen between the 2 actors, researchers and social practitioners. This place has been called "the agora", which along with the transdisciplinarity approach is one of the pillars supporting the new social contract of science with society (Gibbons 1999; Nowotny et al. 2001).

The agora has a broad meaning referring to any setting, physical or not, for the dialogue to happen. More important than the place or means of communication, it is what is expected to occur in the agora. Scientist are expected to participate actively in order to have their outcomes constantly legitimized, while society is allow to react or participate in the knowledge production process. Society stakeholders, help in the process by sharing either concealed or contested values towards the direction of science (Meppem 2000).

From the point of view of the scientific community, society may take any of two roles one being "consulting" where social actors responds and react to the scientific outputs. Or "participatory" transdisciplinary research where society takes an active participation in the process of knowledge production (Mobjörk 2010). Either case, in a transdisciplinary research approach is often said that through the agora society speaks back to science its agreements or disagreements (Nowotny et al. 2001). However, that doesn't necessarily mean that scientist listen.

Empirical evidence show that despite a broad acknowledgement of the importance of conducting transdisciplinary research not always researchers engage in it. In particular "choice of research strategy is associated with its perceived benefits for idea generation, publication opportunities, intellectual effort required, the costs of team coordination, and satisfaction with organizational resources" determine whether scientist, younger scientist in particular, will listen to society, and thus engage in transdisciplinary research (Lauto & Sengoku 2015).

Therefore, we have moved to a "mode-2" production of knowledge where scientist are aware of the importance of social engagement and socially robust outcomes. However, even though society speaks back to science, sometimes there are no incentives or easy way for scientist to actually listen to it. The good news is that new methodologies may help scientist to get involved in the agora and listen to society, at least in a "consulting" manner.

This is possible given that the dialogue between science and society has moved to new places, in particular it has moved beyond research lab and governmental institutions to a broader set of locations. Media and the new communication technologies has taken a distinguished role (Gibbons 1999). News media might not entirely represent the public opinion in relation to technological developments, but helps as a window to check some impacts of research like motivation and encouragement, production and distribution of ideas, and discussion and opinion formation (Walter et al. 2007). In relation of the discussion of scientific outputs in society, media may tell a story of the intensity or relevance an academic topic has, or whether society has concealed or contested values towards it (Melton et al. 2016). Particularly social expectations take an important role as it is known that “expectations may contribute to the emergence and shaping of socio-technical configurations by attracting interest and resources of producers, financial actors and also users.” (Truffer et al. 2008). Thus, the study and monitoring of those expectations may allow us to understand the position of a technology and its future.

One of the tools used to evaluate the maturity of a technology in the market is the hype cycle (Fenn & Raskino 2008), a curve that describe social expectations along with the technology life cycle. It is argued that some technologies generate a clear pattern of expectations after their introduction. At first, new technologies are discussed vividly in society, oftentimes exaggerating their benefits, producing a peak of inflated expectations. Then, when expectations and reality do not match, disappointment happens, this produce a fall from where just a few stakeholders get the technology to rise again to fit the now adjusted societal expectations.

Hype cycle studies have shown different results across technologies, usually pointing on how the real shape of the curve moves away from that theoretically depicted. Is being reported that the hype cycle tends to appear when the technology of analysis is specific (a product), and start to disappear at higher levels of aggregations (Borup et al. 2006), patterns in the shape also differs when mapping the expectations of different actors (Jun 2012). These and other empirically found incongruences have led to suggest moving the study of expectations to a more general “hyped dynamics”, instead of the fixed hype cycle (Dedehayir & Steinert 2016).

Studies in hype cycles usually rely in newspaper information, as they are rich of social information and have been also used to predict market behavior and volatility (Engle & Ng 1993), and evaluate legitimacy (Geels & Verhees 2011). For the hype cycle the number of articles in a period of time becomes an indicator of social attention (Ruef & Markard 2010). So that, some researchers have use *hype* as synonym of *attention* (Melton et al. 2016; Van Lente et al. 2013), and this paper follows those same steps. The second characteristic of the hypes, beside level of attention, is the sentiment. To evaluate sentiment in news researchers, have to realize content analysis, creating tagging rules and reading each article to label accordingly, and this is subject of time and cost limitations. Instead of content analysis in this paper we performed automated sentiment analysis.

Sentiment analysis is the task of automatically coding text entries as positive or negative, and sometimes neutral (Medhat et al. 2014). This can be done at sentence or document levels. Several techniques have been proposed to realize the tagging, including lexicon based methods where the text is compared to curated vectors of positive or negative words in order to get a sentiment score (Taboada et al. 2011). Other techniques include the use of neural networks and deep learning, which reportedly perform better for some types of text (Moraes et al. 2013). Research and application of sentiment analysis is on the raise and is being used to monitor customer behavior (Liu 2010), find relation between social networks and finance (Engle & Ng 1993), among others.

In this chapter the motivation is to scale up the amount of data we can analyze for the understanding of social expectations. We apply sentiment analysis in two directions. First, we aim to observe the relation of sentiment and the inflated levels of attention. We want to compare if what is observed in media coverage of robotics technologies, fits the theory described in the hype cycle literature. Concretely, we analyze if the pattern of hypes for technology concepts at higher levels of aggregations, correspond to a positive rise, and negative falling of expectations. Second, we discuss the dynamics observed between topics in academic articles and similar discussions found in the media. New discoveries may lead to greater discussion in the news. However, the opposite direction may also hold true, an incipient technology that is positively discussed in the mass media, may lure researchers to join that field. Then media will serve as a trigger for the generation of academic outputs. Concretely, the objective of this paper is to investigate the relationship and dynamics between sentiment in technology news and academic publications.

Previous studies have also analyzed newspaper articles to evaluate hyped dynamics and sentiment. For instance, sustainable technologies in the Netherlands (Alkemade & Suurs 2012), and alternative fuel vehicles in US (Melton et al. 2016). In those articles, news about biofuels, hydrogen, and natural gas related technologies were analyzed with the purpose of better understanding social expectation patterns. Both studies identified expectations curves that resemble the hype cycles at different level of specificity, including the characteristic positive rise, and negative falling of expectations.

Other studies focused on healthcare robotics in German print media (Laryionava & Gross 2012) and gene technology in Australian newspapers (Petersen 2001) by doing content analysis, studying the news sentiment without taking into consideration the hype cycle itself. All aforementioned research have in common the study of expectations in the form of quantity and sentiment of news. However, they tackle on specific technologies in a well-defined geographical area. Also, the task of tagging sentiment in news was done manually by creating labeling rules, and reading the entire corpus of news. For that reason, the number of articles analyzed in those researches remained small.

This chapter separates from the line of those previous studies as we analyze the hype cycle at a higher level of aggregation accumulating all type of robotics technologies that appeared in our source of news. Sentiment analysis is also introduced for the understanding of the relation between media attention of technology and academic outputs in the form of journal articles.

The rest of the chapter is composed as follows. In the next section we explain the hype cycle, sentiment analysis and topic models applied to news, the construction of citation networks to understand knowledge structure of academic research, and how semantic linkage can be obtained by text mining. Then the process of data acquisition is shown, to move to the results section where the patterns of sentiment, hype, and academic production are discussed. We conclude summarizing our findings.

3.2 Data and methods

Figure 3-1 depicts the overview of the present research. We analyzed two types of data: newspaper articles, and academic articles, each of them following a specific treatment. From the news, we measured the levels of social attention and sentiment, and then, the corpus was sub-grouped by their semantic relation using topic models, a well-known methodology to categorize unstructured text. On the other hand, we extracted the knowledge structure of academic articles using citation networks. Finally, we found semantic similarities between the clusters of papers and the topics in the news. These similarities also served to identify patterns in time, and the role of the news sentiment in regards of academic production. Detailed explanations are provided in the next sections.

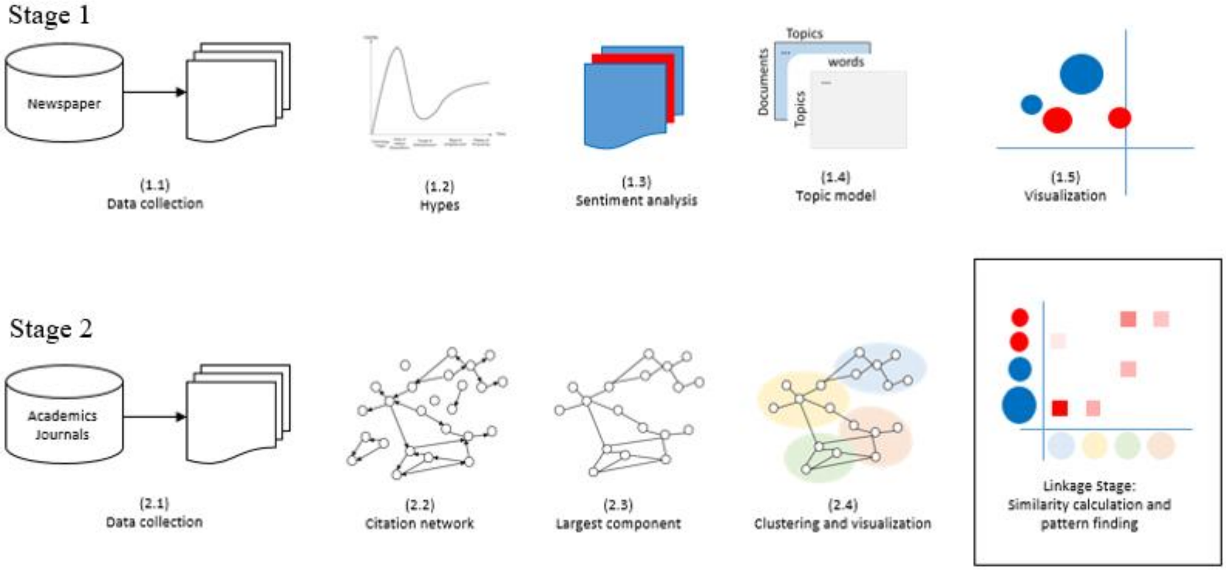

Figure 3-1 Overview of the methodology in chapter 3

Positive and negative topics in news are compared to knowledge clusters of papers about robotics

3.2.1 Topic model and sentiment analysis of news articles

We started collecting the data of our target technology from an international newspaper (1.1). For each year, the amount of news was normalized against the total number of news articles published in that newspaper in the same year, this serves as measure of the media coverage of robotics in the overall landscape of news. Then such coverage is plotted in a timeline, showing us the ups and downs of public attention, hence, the hypes are revealed (1.2).

Next, we performed sentiment analysis (1.3) at the document level by using a commercially available tool. Sentiment is extracted based on the overall textual content of the news, rather than persons, products, or events. We used the Alchemy API which is integrated in IBM's Watson Developer Cloud. IBM reports an accuracy of 86.1% (IBM 2015). Third party independent research also tested several web based services in different scenarios concluding that Alchemy API is one of the best multipurpose sentiment analyzers (Serrano-Guerrero et al. 2015).

IBM's Watson labeled each news article in our dataset as positive, negative, or neutral based in the news headline, lead paragraph, and abstract when available. Because we are interested in the polarized view on robotics technology, neutral news were neglected from analysis. Having each news labeled with its sentiment, we looked for patterns in timeline and hypes. Then, the complete corpus of news was classified into topics by using topic models (1.4).

A topic models is an unsupervised machine learning technique that classify documents according their underlying themes (Blei et al. 2010). The themes, or topics, can be seen as sets of words that tend to co-occur forming multinomial probabilistic distributions over all the vocabulary found in the corpus. As a simple explanation, documents are composed of topics, and topics are composed of word probabilities. However what is observable in reality are the news and their words, therefore an algorithm is used to extract the hidden topics.

Before applying the topic model, trivial preparation tasks were completed. First, we removed stopwords and infrequent terms that appeared in 5 news or less. Punctuation and numbers were also removed, and the text was transformed to bag-of-words, and then, represented as a document-word matrix filled with the word frequencies. Those frequencies were translated to probability distributions which is needed for the algorithm that solves Eq. 3-1, the topic model.

$$p(w|d) = \sum_{t=1}^{T} p(w|t)p(t|d). \qquad \text{Eq. 3-1}$$

t stands for each of all possible T topics, d and w are the document (news) and words probabilities respectively.

In this research, we use the Gibbs Sampling algorithm to solve it. Gibbs sampling is a Markov chain Monte Carlo algorithm that iterates over all words in the corpus to classify them into the desired T number of topics (Griffiths & Steyvers 2004). The Gibbs sampling algorithm is broadly used in topic modeling and has been applied on patents, journal articles, and newspaper corpus (Hu et al. 2014; Rosen-Zvi et al. 2004; DiMaggio et al. 2013).

The main decision before running the model is selecting a proper number of topics, while is not uncommon to arbitrarily select it according the volume of data, we preferred to take advantage of the natural statistical properties of our data by selecting the number of topics that gives the maximum log-likelihood, a statistical approach for the selection of the best models (Wallach et al. 2009). Once the number of topics is selected, we can run the model and classify the news articles. This type of classification is fuzzy, meaning that a news get classified into several topics according to probability scores. News articles having little probability in each topic might be regarded as noise. To overcome this, we classified each news to its most probable topic that is 0.5 or higher. Since there is no possibility for a document to have 2 or more topics scoring higher than that, news article gets classified to only one single topic.

Finally, we label each topic as positive or negative by finding the dominant sentiment from the collection of news it aggregates. This is done by counting the number of positive and negative news within it, and taking the largest share.

3.2.2 Citation network of academic articles

We compare news to knowledge production as represented by academic articles in robotics (2.1). Differently from the news and their purely text driven nature, papers have other characteristics that can be studied, particularly citation references are of value when we need to understand the structure of science. We collected bibliographic data of journal articles about robotics and created a citation network based on their references. Each paper was linked to the papers in its references if they were present in the dataset (2.2), this can be understood as arrows pointing back in time generating a so-called direct citation network (de Solla Price 1965), where the papers are nodes and the arrows are the edges. From this, some group of papers that are tightly knit can be observed, and those groups, or clusters, represent units of knowledge that embed specific ideas, concepts, technologies, or applications (Borner & Scharnhorst 2009).

Besides direct citation networks, there are other methods to build networks using references, as bibliographic coupling (Kessler 1963) and co-citations (Small 1973). However direct citation networks have proven to give the best partition of clusters when the data is large and spans through several years (Fujita et al. 2014; Shibata et al. 2009). Direct citation networks are being used for science mapping and technology forecasting (Shibata et al. 2008).

3.2.3 Semantic similarity between topics and clusters

Once the network is obtained, we neglect the direction of the links, and collected only the largest component (2.3). Finally, an algorithm is applied to identify clusters (2.4). In this research we applied the Louvain algorithm (Blondel et al. 2008) which has shown high performance when splitting citation networks (Šubelj et al. 2015). The algorithm first look for small, densely connected papers by locally and greedily optimizing the modularity, a network property that indicates whether a collection of nodes have an organized structure in comparison to randomness. When the network has been split in several small clusters, those clusters are treated as nodes and the process start again, aggregating communities in a multilevel way. The algorithm finishes when the modularity gets maximized. By using this method, the network is partitioned in a number of clusters that comes up naturally from the network, and each paper belongs to a single cluster only.

3.2.4 Data sources

For newspaper articles we selected The New York Times as a source of information because is worldwide known, regarded as a quality provider of news, and several academic research has been done on its corpus (Newman et al. 2006; Melton et al. 2016). The New York Times is acknowledged to have liberal views and being on the left side of political spectrum (Jutel 2016). However, it was not found evidence that such views affected technology news. The data was extracted by using their application programming interface, that allows flexible and comprehensible search on their archive of news dating back to September 1851 (The New York Times 2016). This newspaper also classifies each news based on a controlled vocabulary called Times Topics referring to specific people, events, places, and subjects. This is a helpful feature that allows to control the scope of the query. We searched for news articles having the word “robot*” in the headlines, or contents (the asterisk stands for a wildcard so that we can get news about “robots”, “robotic”, and “robotics”) that belong to the subjects "Robots*", "Technology" or "Artificial Intelligence". In doing so, we could filter out news related to robot movies and T.V. series. Major newspaper’s sections included: Technology, Science, Opinion, Business, U.S., Health, and others. It should be noted that Times Topics subjects are not the same as the news sections. The subjects are metadata, at higher level of specificity, and human tagged. We extracted all available news in the archive, which dated as far as 1927. However, we only analyzed the subset of 1,218 news articles about robotics from 1976 to 2015.

Journal articles on robotics were extracted from the Web of Science Core Collection. The Web of Science is known as a reliable source of academic articles, indexing more than 12,000 journals in basically all disciplines, it is also regarded as one data provider having a very strict and high standard indexing. We queried articles having the word “robot*” in the title, abstract, or keywords, retrieving 160,361 articles from 1976 to 2015.

1976 was selected as the lower boundary because is the year in which news articles about robotics started to rise (Figure 3-2). 2015 is selected as upper boundary for research reproducibility matters; while news articles can be retrieved to exact dates, less control is available with academic articles which are searchable in rounded years only, not having complete data for 2016 by the date of retrieval (February 10th, 2016).

3.3 Results

In this section, the data is studied from two perspectives. First, hype and sentiment of robotics is analyzed in the light of the hype cycle literature. Secondly, it is studied the configuration of research topics in relation to similar topics found in the media.

3.3.1 Social attention and sentiment

Figure 3-2 shows the normalized frequencies of news and academic articles about robots. The oldest registry in the news is in 1927 followed by a small interest in the usage of the “robot” word. On the other hand, the oldest academic article in our dataset appeared until 1969, taking more than 40 years for the term to get into academic publishing, and other 10 years to gain momentum, raising as a field in the mid-80s. News about robotics started to generate higher level of attention from 1976, thus, we focused the subsequent results from that date.

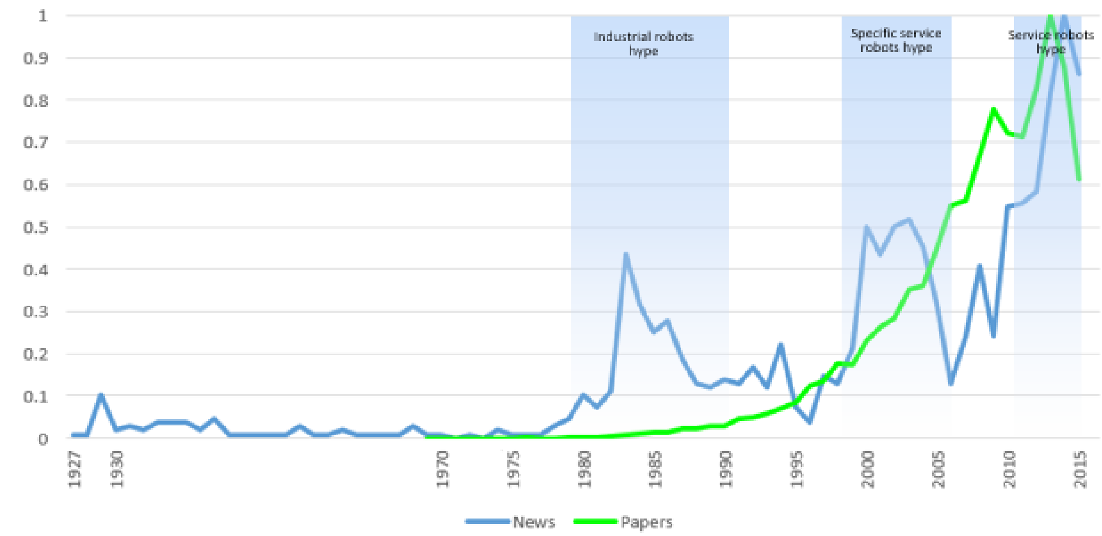

Figure 3-2 Normalized frequency of news and academic articles on robotics.
The first news and academic articles using the word “robot” appeared in 1927 and 1969 respectively. An increasing level of attention on news about robots is observed from 1976.

From the news we can observe three peaks of inflated social attention. The first peak corresponds to the hype generated from the industrial robots, it spans all the 80s to the early 90s, having the greater peak in 1983. The second hype collects news about heterogeneous robotics technologies that showcased the news from 2000 to 2005, this period was dominated by Honda’s humanoid ASIMO, Sony’s pet robot AIBO, Nasa’s mars rover missions, and the introduction of the vacuum cleaner robot Roomba. The final peak corresponds to the service robotics, a wave of discussions on the implication of driverless vehicles, drones, and the general spread of robotic technologies. In appendix 7.1 we show examples of the headlines that hit the three hypes.

We labeled each news article as positive or negative by using sentiment analysis to find linkages between sentiment and level of attention. Hype cycle related literature points that the raising portion of the curve from the innovation trigger to the peak of inflated expectations, is characterized by an overall positive sentiment, whereas after the peak to through of disillusionment is marked by negative sentiment caused by unaccomplished expectations (Fenn & Raskino 2008). Such affirmations are based on case to case analysis, and we wanted to test them in a higher level of aggregation for all the corpus of robotic technologies. Figure 3-3 and Figure 3-4 summarizes the findings.

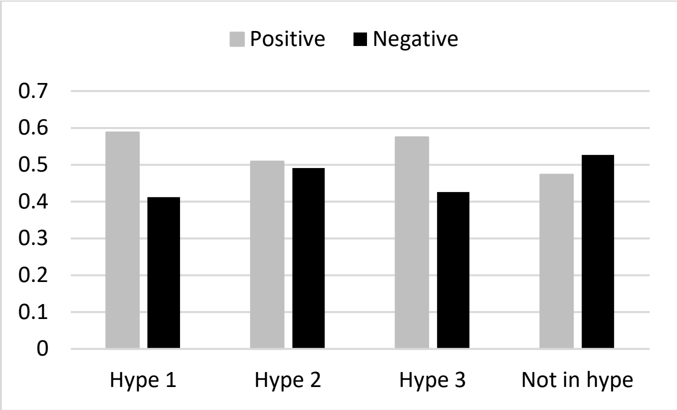

Figure 3-3 Proportion of positive and negative news in each hype.
Hype 1 (1978 – 1990); Hype 2 (1999 – 2006); Hype 3 (2010 – 2015); Not in hype (rest of years)

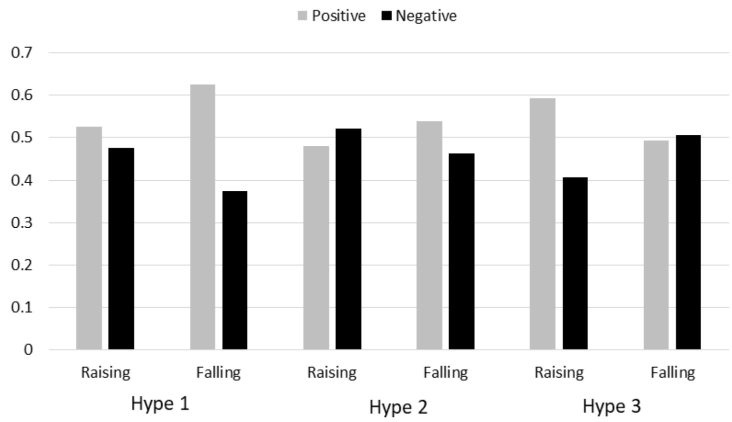

Figure 3-4 Proportion of positive and negative news in rising or falling years of each hype.

We computed the proportion of positive and negative news in each hype, and for the un-hyped years as shown in Figure 3-2.In general, the industrial, and the service hypes holds a larger share of positive news. While the second hype is balanced. If we split the hypes to observe the sentiment in the rising and falling portions we end up having the proportions in Figure 3-4. Contrary to that expected from the hype cycle literature robot news do not show a clear pattern of positive raising expectations versus a falling negative disappointment. Moreover, in the case of the industrial robotics hype, news articles in the decreasing years shown a larger share of positive news. The second hype, even though is balanced gets to show an opposite trend having more negative news on the rise, and more positive news in the falling. The third hype, however, is not conclusive. The falling portion of the hype is covering years 2014 and 2015 only, which is a shorter period than the others. Also, its trend is still under development based on recent inventions, and discoveries. It will be needed more time to obtain a complete view of the "through of disillusionment". We conclude that robotic technologies in the news are discussed in a balanced form when it comes to sentiment.

3.3.2 Topics and sentiment

Newspaper articles were classified into specific topics by using topic models. The maximum log-likelihood estimation gave us an optimum partition when the topics are 22. Topics were assigned a sentiment according the largest share of positive or negative news within them. Also, the average year of each topic was calculated. On the other hand, the citation network of academic articles ended up having 105,061 connected articles, classified in 46 clusters of knowledge.

Topics and clusters were compared in terms of their text similarity. Topics and clusters can share textual similarity without having a real semantic connection. For instance, in our corpus the words “robot” and “technology” appear frequently in several topics and clusters alike, and our similarity metric will signal such connection. Therefore, to overcome that issue we only take into consideration those connections that score an above average similarity, giving as result 298 connections.

The highest similar pair is topic 13 “robot competition” and cluster 18 “rescue robots”, news articles include coverage of robotic competitions in college. The semantic connection is stablished by the vocabulary related to obstacle avoidance, and the coverage of *Robotics Challenge* which also include a rescue robot challenge competition. The second connection is about space robotics; however news articles covers particularly NASA robot projects. The third connection is about surgery.

Finally, we looked for patterns by using the set of highly similar pairs. We are interested to know if an idea, or technology, is first discussed in the news, or if it waits for certain academic maturity before being discussed further in the society. Also, we want to clarify if positive or negatives topics takes a role in that relationship.

Once having the connections from the similarity matrix, we compared the average year of the topics and the clusters to understand which of them was formed first. We also separated the connections according to their sentiment, resulting in the contingency Table 3-1. By inspecting the raw frequencies, the pattern gets visible, in the case of the robotics technologies, positive news precede papers about the same technology. And, negative news tend to occur more when the paper cluster are formed first.

Table 3-1 Evaluation of similarities between topics and clusters

	Similar connections			**Non-similar connections**		
News Topics	***Positive***	***Negative***	***Total***	***Positive***	***Negative***	***Total***
Before papers	135	83	218	330	263	593
After papers	30	50	80	98	93	191
Total	165	133	298	428	356	784
Chi Square	14.1304			1.0979		
P-value	0.000171			0.29473		
Significance	**Significant** at p < 0.10			Not significant at p < 0.10		

To understand whether that distribution is statistically significant we tested our results using the Chi square test of independence, suitable given the categorical nature of the data. The test shows a p-value of 0.000171 which is statistically significant, meaning that there is dependency between the averages year of topics and sentiment when the similarity is above average. We also, wanted to know whether such significance also holds for the bellow average connections, however in that scenario the p-value is of 0.29473, being not significant. Then, we can conclude that there is a connection between positive news discussed first and similar academic output coming in a later period.

This section explored how similar topics and cluster have a relationship based on the sentiment of the news. The present results, are observed from the exclusive case of robotics, it might deserve further analysis in other scientific topics also discussed in society (e.g. GMO). Regional and sub-topical variations are also expected. For instance, military robotics is a topic mostly discussed negatively in printed media; having a share of 51% of negative news.

3.4 Summary

This chapter aimed to reveal the relationship between social attention and the development of new research topics in academia. It was found that positive news precedes science cluster when the contents are similar. This is observed by the amount of articles published on defined topics. While our study is retrospective, it hints the possibility of using such pattern in a predictive and actionable way. By monitoring positive technology news, it may be possible to forecast nascent fields of technology. Therefore, we are also showing the application of sentiment analysis for science and technology studies. Previous research had to rely in manual sentiment tagging, which is costly and time consuming. By using sentiment analysis, it is possible to analyze large volume of data rapidly, and the outputs are accurately reproducible. It allows us to give an objective face to the subjectivity of sentiment.

It was also studied sentiment analysis in robot news in relation to their hyped dynamics, and to specific topics. In relation to the hype cycle studies we shared two findings. First, hypes also occur at higher levels of aggregations, and not only to specific technologies. Second, the rising and falling portion of the hypes not necessary stand for positive and negative narratives respectively. In the case of robotics technologies, the coverage is balanced. This is contrary to what is discussed in other studies. However, our intention is not to disprove them, rather to highlight the heterogeneous nature of social attention towards technology.

However, there are some limitations to overcome before generalizing the results. First, we observed balanced sentiment patterns in each of the hypes of robotics, without a clear distinction that positive sentiment dominates significantly the rising of the hype, or that negative news dominates the falling. This may be related to the level of aggregation used in this study. By looking through all robot news at once in the same window of time, patterns may be lost. Inspecting more specific robotic technologies like drones, or self-driving vehicles separately may provide another perspective. Secondly, we took advantage of the metadata and availability of data provided publicly by The New York Times, however utilizing other sources of information may be of interest to generate more robust comparison between news and papers. Such increase of data sources must be accompanied with good practices to remove noise (as robot’s movie news), and decisions on how to handle repeated news across providers. Third, we took as example robotics technologies, which are known to generate expectation and news with sentiment cues, but other case studies are necessary before generalizing that positive news precede science clusters. Fourth. Trends were analyzed by the number of articles. However, it was not investigated if the contents of news may have impact in the actual contents of the articles, or how the media shapes the vies of the scientist involved in such discoveries.

In a general context we studied how society speaks back to science. However, there is implied that science speaks to society as well. In aiming a better social engagement science may fall in the game of showcasing its results as amenable as possible, sometimes in an exaggerated manner diluting the line between fact and fiction. Nowotny calls this type of communication as “low-cost realities”, and while it is intentionally good to bring information to society and engage it in academic discussion, it may also cause confusion (Nowotny 2005). This type of communication may raise hype over some emerging technologies, or delegitimize the academic outputs. Therefore, we may expect that science communication also plays an important role in defining the imagining of science in society and how this imaginings are represented in the media. We let those tasks opened to be tackled on our own as future research.

4 PATTERNS OF FUNDING OF ROBOTICS RESEARCH

4.1 Introduction

In this chapter we explore the dynamics between funding and academic outputs. It is aimed to answer the question of whether funding organizations target innovative research. Funding organizations identified in the acknowledgement of articles are characterized based on the level of participation in regards of breakthroughs, change maker, incremental or matured stages of research. Properties of the citation network of academic articles are used to infer features related to the emergence of technology and then classify the articles accordingly. Additionally, strategies of selected funding agencies are explored with respect of specific scientific fields.

The call for funding agencies to effectively target innovative research has been executed in the past, encouraging organizations to embrace risk (Berg 2008; Muller 1980). The challenge comes in the form of giving researchers the freedom to pursue curiosity-driven research on their own and the demand for bringing tangible outputs to the society to ensure that the return of investment can be measurable. Thus, funding strategies followed by organizations may determine the directions of innovation in the academic domain (Braun 1998; Lok 2010). They exercise authority over the production of knowledge (Whitley et al. 2010), and shape and control the direction of science by influencing public policy, generating incentives towards specific research topics, or created mission-oriented programs (Gläser & Velarde 2018).

However, The need for better tools to measure the impact of funding for encouraging the emergence of new knowledge in the frontiers of science is a recognized gap (Lane 2009).

So far, bibliometricians have played a role in developing quantitative indicators that assess research quality, mainly in terms of the number of publications and citations received. However, as methodological tools gets refined and data coverage gets broader, new venues of exploration have opened up. In this research, we take a step forward from the evaluation of academic impact of research funds to the study of fund allocation for innovative research. To explore the role of funding organizations in driving research towards more innovative and path-breaking directions, we used quantitative methods such as citation networks and the research classification schema (Takano et al. 2017), along with the funding information found in the acknowledgement section of academic articles. This paper particularly attempts to identify methods that help understand whether funding organizations are targeting new research fronts and identify the state of funding for specific knowledge fields.

4.1.1 Funding and Acknowledgements

To analyse the participation on funding agencies in the production of new knowledge, we take a look to the acknowledgement section of academic articles. Studies on the acknowledgment section have been traced back to 1972 to the work of K. H. Mackintosh (Cronin & Overfelt 1994). In early stages of acknowledgment studies, the extraction and classification was done entirely manually. As bibliographic databases improve, the linkage between publications and grants can be found directly. The key piece of information lies in the acknowledgement section of the articles. The academic courtesy of mentioning the persons and organizations who collaborated during the research process is also being subject of study (Cronin & Overfelt 1994), and interest in this field is gaining attention to the extent of being identified as *influmetrics* (Costas & Leeuwen 2012). The patterns and usage of acknowledgements vary across fields. Cronin and colleagues in their study of acknowledgements in Psychology and Philosophy literature identified seven types of acknowledgments: Conceptual, editorial, financial, technical, moral, reader, and unknown (Cronin et al. 2003). In general, the financial type of acknowledgement has received more attention. One of the advantages of financial acknowledgements is that there is a direct linkage between the paper and its sponsors.

Linking funding inputs with academic outputs has not been an easy task. Studies on funding and academic impact have used separated data sources to establish this connection (Hosotsubo & Nishii 2016). These studies particularly tried to connect the grantee data from funding organizations reports to metadata (i.e. funding codes) in other bibliographic databases by author name or paper-title matching, in a process that presented several caveats. For instance, the principal investigator is not always the one writing the papers, and the time that is taken to generate the output varies across fields, making the linkage process difficult. (Boyack & Börner 2003; Boyack & Jordan 2011).

However, more recently, the refinement of machine learning techniques for text mining have facilitated the task of entity recognition, and algorithms have been developed for this concrete task (Giles & Councill 2004). Data providers of bibliographic information such as Clarivate Analytics through the Web of Science have also started including such information in a cleaner and curated format (Web of Science 2008). This has provided richer data that spans across disciplines and countries allowing comprehensive assessments. Among the earliest applications of such specific acknowledgment data for analyzing funding patterns is the study of Lewison and Markusova (2010), a bibliometric analysis of cancer research in Russia, and the study of Wang and Shapira (2011), which includes an overview of funding in nanotechnology. The incorporation of funding data from the acknowledgements section of academic articles into the bibliographic databases has opened up possibilities for new venues of research.

4.1.2 Funding and Academic Impact

Previous literature has explored the value of funding in academic research from different perspectives. As a stream of research, funding has been studied as an explanatory variable of high-quality outputs given the nature of grant awards being accompanied by a peer review process (Gillet 1991; Lewison & Dawson 1998). The dependent variable in these studies generally takes the form of citations received by funded papers versus non-funded papers. In the work of Gök, Rigby, & Shapira (2016), acknowledgement information was used to identify funding organizations and group them in different categories such as governmental, non-governmental, and international funding for six European countries. These categories, among other features, were used as independent variables in a regression model explaining the number of citations received by the articles. A high correlation was found between the citations received and funding, particularly from national sources, a correlation that also holds when academic articles are analyzed by topic instead of country level. Shen, Hu, Lin, Tsai, and Ke (2016) observed that highly cited articles in computer science conferences are also those that have funding acknowledgements. Academic articles that have received funding are also linked with being accepted by journals with relatively high impact factors (Wang & Shapira 2015), related to a greater number of co-authors (Yegros-Yegros & Costas 2013), and motivating interdisciplinary knowledge (Lyall et al. 2013). In summary, this stream of research study includes funding as an explanatory factor of any of the dimensions of academic impact.

The other stream of research targets the organizations themselves, addressing the concern about whether they achieve good performance in their allocation strategies. Again, performance is evaluated on the basis of the accumulation of citations received by the sponsored papers. However, other features may also be evaluated using visualization tools (Boyack & Börner 2003; Zoss & Börner 2012), by exploring alternative metrics (Thelwall et al. 2016), or incorporating qualitative assessments (Donovan et al. 2014). Beyond this, few bibliometric research have considered the evaluation of funding organizations with respect to their performance in targeting innovative research and shifting research trend.

4.1.3 Funding and emerging technologies

Research at the interface of detecting emerging technologies and using acknowledgement data is scarce. Wolcott et al. (2016) used the sponsor data found in the acknowledgements as an explanatory variable to predict the emergence of innovative research. However, they found only a low correlation with their definition of breakthroughs. Further, to address the issue of detecting emerging technologies during the peer review of grant applications, Hörlesberger et al. (2013) applied a logistic regression model to selected features of each proposal, showing that interdisciplinarity and similarity to previously identified emerging research are defining factors. Both these articles contribute significantly by linking funding to the concept of emergence; however, a research gap between the overall landscape of a specific research field and its funding sources still exists.

As discussed above, bibliometric research on funding and acknowledgements has focused on the evaluation of academic impact based on the number of publications and citation counts. However, while article production and citation count may suggest high-quality research, this is not necessarily true as an indicator of innovation. Therefore, the notion of whether funding agencies target cutting-edge topics remains under-researched.

Parallel to studies of funding, an increasing number of bibliometrics studies address the issue of detecting emerging topics. For example, Rotolo, Hicks, & Martin (2015) identified five attributes of emerging technologies: radical novelty, relatively fast growth, coherence, prominent impact, and uncertainty and ambiguity. The former three can be observed by adopting quantitative methods, and several bibliometric methodologies may be used to identify one or more of those attributes.

For instance, radical novelty and coherence can be captured by citation network analysis (Shibata et al. 2009) by observing how new clusters appear in the network over time as well as by observing the close connections of the nodes. Purely text-based (Yan 2014) and hybrid methods (Glanzel & Thijs 2012) have been used in a similar manner. On the contrary, fast growth can be detected by comparing the document counts of a topic with a function, usually based on an S-shaped curve (Ho et al. 2014).

This chapter uses the attributes of coherence and novelty that can be observed in the clusters of citation network to classify research topics into four categories: breakthrough, change maker, incremental, and matured. In particular, this framework distinguishes between truly radical emerging technologies based on relatively new knowledge (the change maker category) and emergence from well-understood research fields (the breakthrough category) (Takano et al. 2017). We build upon those methodologies to characterize the funding organizations identified in the Acknowledgements section of articles based on their level of participation within the four abovementioned technology categories. We also use the properties of the citation networks of academic the novelty of emerging topics and then classify the articles accordingly. Therefore, our research incorporates a new layer of analysis of publication records and citation counts.

The intentions of funding organizations to target research focuses can be revealed by looking at the objectives posted in their guidelines, websites, and call-for-grants. However, the challenge is knowing whether the outputs achieved from receiving the funding have remained aligned with those original objectives. Studying this gap between outputs and objectives as well as the lack of tools for such an assessment is the main motivation of the present research. Also, the methodology discussed in this chapter aims to bring a comprehensive view of the organizations participating in a single academic domain. Their topical interest and intensity of funding can only be seen from the guidelines of those organizations that have an open disclosure of their strategies. By looking funding from the acknowledgement information of academic articles, we can observe those agencies participating even if they do not have an open guideline, as long they allow their sponsored researchers to reveal their sources. Our contribution is twofold. Firstly, we link acknowledgements to research categories. Secondly, funding organizations obtain an assessment tool with which to compare their research targeting as reported by sponsored authors.

4.2 Data and methods

Figure 4-1 provides an overview of the methodological approach in this chapter. First, data was extracted from the academic articles database and funding organization names were extracted from their acknowledgements. By exploiting the cited references, we created a direct citation network from where clusters of knowledge were identified. These clusters were classified as change maker, breakthrough, incremental, and mature. Finally, the funding organizations were characterized by their participation in each of these four categories and the results were mapped. Additionally, we inspected the areas of interest of selected organizations by measuring the subject categories they support the most. In the following section, we explain the method in detail.

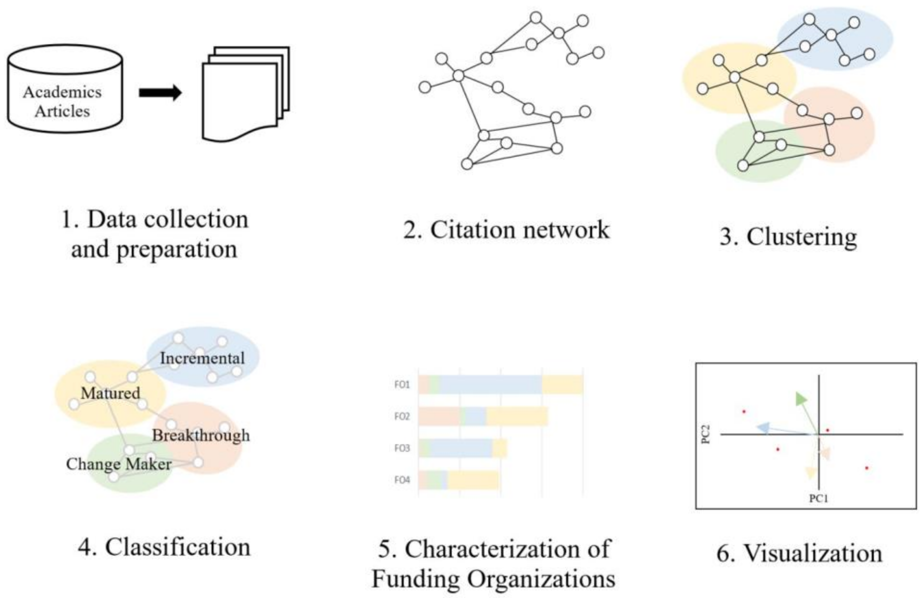

Figure 4-1 Overview of the methodology in chapter 6

For evaluating funding organizations, we define a funding organization as any entity mentioned in the funding acknowledgement of academic articles. These may be ministries, agencies, NPOs, companies, projects, or programs. In Table 1, we list the organizations that appear frequently in this study and the acronym we choose for them.

The Web of Science takes funding information directly from the acknowledgement section of papers. It provides two fields, one having the raw text as written in the paper and the other, a curated field showing only the names of organizations and when available, the funding code. In this research, we used the latter field. This information is only available for papers published from 2008, however that first years does not have complete acknowledgement data (Paul-Hus et al. 2016). Therefore, For purpose of this study, we only evaluate funding in relation to the share of articles from 2009 onwards within each cluster.

While the inclusion of funding acknowledgement by the data provider greatly facilitated the identification of funding agencies, a good deal of effort was still required in data cleaning and name disambiguation. For instance, funding entities with long names, such as "Ministry of Education, Culture, Sports, Science and Technology of Japan," also appear as "Ministry of Education, Sports, Culture, Science and Technology of Japan" or any other variation, omission, or abbreviation. We used programming to help identify similar instances; however, ultimately all names were manually checked and standardized. Hierarchical relationships in the funding organizations were also observed as in the case of MEXT as a parent of JSPS and JST. However, some papers may acknowledge both, son and parent, as funding organizations. Given that parent organizations may simultaneously support the work of some researchers (Lepori 2011), we could not always regard this as a redundant acknowledgement. Therefore, in terms of institutional structure, we respected the entries as reported in the dataset with no further merging.

In Table 4-1 we list the organizations that appear frequently in this study and the acronym we choose for them. We focus our study on two groups of funding organizations, the international top funding agencies by the number of sponsored articles, and the five most frequent Japanese organizations found in the dataset. The latter is used to investigate how different agencies and programs having different characteristics will affect research direction and performance. Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT), a ministry of Japan Government, has two funding agencies, JSPS and JST. JSPS is responsible for pure and basic science, while JST is applied research. GCOE is a program in MEXT which aims to develop research, education, and international partnerships. NEDO is a funding agency under another ministry, Ministry of Economy, Trade and Industry (METI), and supports applied and industrial research.

Table 4-1 Selected funding organizations in the field of robotics

Funding Organization	Country	Acronym
863 State High-Tech Development Plan	China	863-CH
973 National Basic Research Program	China	973-CH
Defense Advanced Research Projects Agency	USA	DARPA
European Commission	EU	EC
European Union	EU	EU
The Fundamental Research Funds for the Central Universities	China	FRFCU-CH
Global Center of Excellence	Japan	GCOE
Japan Society for the Promotion of Science	Japan	JSPS
Japan Science and Technology Agency	Japan	JST
Ministry of Education, Culture, Sports, Science and Technology of Japan	Japan	MEXT
New Energy and Industrial Technology Development Organization	Japan	NEDO
National Institute of Health	USA	NIH

National Science Council	Taiwan	NSC-T
Natural Science and Engineering Research Council	Canada	NSERC
National Natural Science Foundation	China	NSF-CH
Swiss National Science Foundation	Switzerland	NSF-SZ
National Science Foundation	USA	NSF-US

To collect the data, we searched the Web of Science Core Collection for academic articles on the topic “robot*” published between 2009 and 2016 inclusively. In this chapter, the network was created again in response to the fact that the database comprehensively collect information of funding organizations since that year (Paul-Hus et al. 2016). Funding agency names are taken from the curated field provided by the Web of Science. However, intense cleaning and name disambiguation was performed to avoid misspellings or name variations introduced by authors when acknowledging their sponsors.

We the dataset cleaned we proceeded to create the citation network, including all articles regardless if they mentioned financial support or not. It is worth noting that we could have created a citation network using only the set of papers that reported funding. However, such an approach is biased against the detection of innovative technologies, because it is known that some highly cited papers, which are the constitutive blocks of clusters, may not have funding acknowledgements (Zhao 2010). Our approach classifies the technologies into categories using all the available data that is connected. Thus, by classifying the clusters, we can see how different funding organizations participate within them, not necessarily sponsoring the core paper, but at least demonstrating their ability of allocating funds in promising new and emerging fields.

Clusters are the unit of measurement for the following step. We used a research classification schema (Takano et al. 2017) to label each cluster as change maker, breakthrough, incremental, or matured. This classification responds to two main features of the clusters: the average age of the cluster, which is the average of the publication years of all the papers within the cluster, and the age of the core paper, which is the paper having the maximum connections in the cluster. The classification computed as follows: the novelty of a cluster is evaluated with Eq. 4-1, the difference between the average year of its articles, $\bar{x}_{cl}$, and the average year of all articles in the network, $\bar{x}_{net}$. Positive values of X refer to clusters which are novel on average. Then, it is computed the novelty of the leading article in each cluster. This leading article, or hub paper is the article having the highest degree within the cluster. Is novelty is obtained by applying Eq. 4-2. Where y is the publication year of the most connected article, and $\bar{y}_{cl}$ is the average year of the cluster. A positive value of Y means that the most relevant paper in the cluster, in terms of citations, is novel in average.

$$X = \bar{x}_{cl} - \bar{x}_{net} \quad \text{Eq. 4-1}$$

$$Y = \mathrm{y} - \bar{y}_{cl} \quad \text{Eq. 4-2}$$

An interpretation of the four categories is provided in Table 2. The research classification schema has correspondence to the technology life cycle and has been presented as a good representation of the state of research for the Internet of things and related technologies (Takano et al. 2016).

Table 4-2 Research classification schema (Takano et al. 2017)

Type	Condition	Interpretation
Change maker	X and Y > 0	Research target is active, and its core literature is relatively new
Incremental	X > 0, Y < 0	Research target is active, but its core literature is relatively old
Breakthrough	X < 0, Y> 0	Research target is inactive, but its core literature is relatively new
Matured	X and Y < 0	Research target is inactive, and its core literature is relatively old

Finally, for each funding organization mentioned in the dataset, we computed the amount of papers they sponsored in each of the four categories. In order to investigate funding patterns of among different funding agencies, the distributions obtained were standardized and used as input for principal component analysis and plotted as a biplot after multidimensional scaling, generating a map of funding organizations and their trends. This analysis was conducted for both, the amount of papers and the academic impact measured by the citations received in those papers.

Additionally, we studied the funding strategies of the organizations by inspecting the subject categories they focus on. Each academic article was assigned to one or more specific scientific subject area depending on its topic. Thus, the interests of a given organization may also be revealed by looking at the distribution of categories its papers fall in. In a similar process, in the previous step, we mapped the funding organizations using subject categories as variables for the principal component analysis.

4.3 Results and discussion

4.3.1 Overall landscape of robotics research and funding

From the 79,209 academic articles having the keyword "robot" in any of their text fields, 59,088 (75%) were connected because the citation relationships. The rest were papers using the word "robot" that did not necessarily belong to the field. 31.92% of the articles in the network reported at least one funding organization in the acknowledgement, a proportion close to that reported for the entire Web of Science Core Collection database (Paul-Hus et al. 2016).

460 clusters were detected and classified as breakthrough, change maker, incremental, or matured. However, 157 were small clusters, in which none of their articles reported funding acknowledgement. Those small clusters accounted for only 1.22% of all nodes in the network and were neglected from the analysis. Figure 4-2 shows the network of robotic research along with the largest change maker clusters in terms of academic articles.

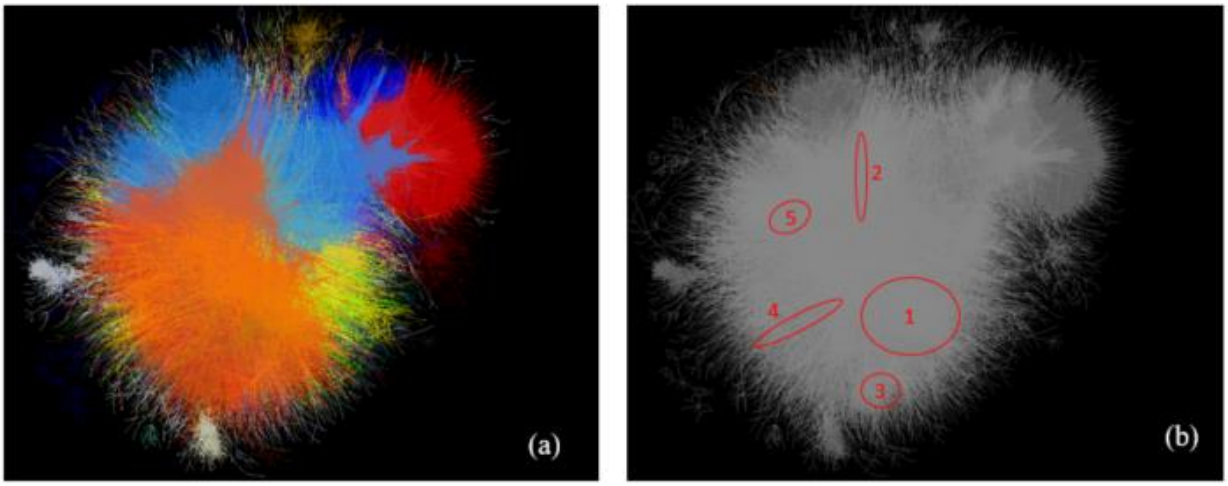

Figure 4-2 Citation network of robotics research (a) and change maker clusters (b).
Change maker cluster tend to be small and hidden behind the larger incremental-type clusters. Five out of 38 change maker clusters' positions are shown as an example. #1 Research on navigation for service robotics in agriculture; #2 Multi-legged robot walking; #3 Wireless technologies for sensing and charging; #4 Machine-learning algorithms for navigation; #5 Motion control for hanging robots

Figure 4-3 shows the aggregated number of articles falling into the four categories in RCS. Academic articles on robotics were largely classified into incremental clusters. The breakthrough and change maker categories, without surprise, collected a small portion of publications, owing to the difficult nature of producing new cutting-edge knowledge with relatively large academic impact. We found that the breakthrough and change maker clusters were small, having an average size of 21 and 32 articles, respectively. On the contrary, clusters representing incremental and matured technologies comprised an average of 548 and 134 articles, respectively. Following this, the papers were aggregated by categories to compute the participation of the funding organizations within each of these four categories. This difference in size was also present in the average size of clusters. As more researchers joined the topics, the incremental and matured clusters got larger.

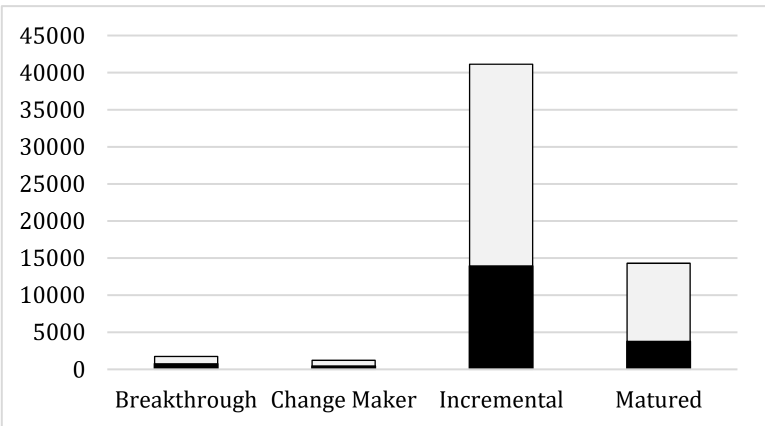

Figure 4-3 Classification of articles of robotic research by type of technology.
The dark shade represents the proportion of articles with financial acknowledgement.

The largest number of clusters in the network were classified as incremental. These are clusters having a new knowledge base, but the core academic article is relatively old. Robot-assisted therapy was one of many examples found. This includes the usage of robots for the rehabilitation of motor functions after strokes. (Lo et al. 2010). Further, in the medical field, we found research on robotic thyroid surgery and transoral surgery (Kang et al. 2009). Within the pure engineering spectrum, we found literature addressing issues such as motion systems for spherical robots (Joshi et al. 2010), among several others.

Next, we found that robotics research was focused on matured technologies, corresponding to clusters in the older end of the dataset. Among them, we found micro-robots (Nelson et al. 2010) and tele-operated robots (Sanders et al. 2011).

Breakthrough clusters are those having a core paper that has been receiving attention in recent years, but the knowledge base they build upon is relatively old. Relevant literature exists in this case, which attempts to solve the problem of robot selection when there is a broad availability of robotic systems for a defined task (Parameshwaran et al. 2015). This cluster draws ideas from operations research and management concepts such as multiple criteria decision-making, which is as old as 1974, but its application in the context of robot system selection has received attention lately. Similarly, we found the concept of tensegrity, a structural principle broadly applied in architecture, but now used for the improvement of dynamics or locomotion by integrating the principle in a robotic structure (Caluwaerts et al. 2014). We found 83 breakthrough clusters having at least one funding organization engaged.

Finally, we found 38 change maker clusters, those having a relatively new knowledge base and core papers. The technologies discussed in this category are primarily oriented to the service robotics, being the robots for agriculture the largest cluster. This includes robotic devices, or algorithms dealing with the automation of tasks in open and unstructured environments, and the handling of live produce which requires gentle and accurate treatment (Bechar & Vigneault 2016). Other research include progress in the speed and autonomy of legged robots (Tedeschi & Carbone 2015), and machine learning algorithms for autonomous navigation (Fathinezhad et al. 2016), and others.

The blue portion of the bars in Figure 4-3 corresponds to the proportion of papers with funding acknowledgement; from this perspective, the breakthrough and change maker clusters performed better than the other two, which is a positive overall indicator for the funding organizations, given that they may be expected to contribute more actively in these categories. However, it should be noted that under our methodology, breakthrough refers to new technologies that were born from old, well-understood research topics. While breakthroughs are desirable outcomes, from the funding perspective, they may be considered as playing safe because the funding tends to be allocated to a well-understood knowledge base. More interesting is the case of change makers. Because both the core literature and knowledge base are newer, the risk of fund allocation may be perceived higher.

4.3.2 Characterization of funding organizations

More than 11,000 organizations were acknowledged in the network; out of these, 8,432 were acknowledged only once. To avoid such sparsity, we decided to focus on those organizations mentioned 10 times or more, obtaining 445 funding organizations. For each organization, we computed the number of papers they funded in each of the four categories such that the distribution obtained served as the characterization itself. The top 10 international funding organizations are shown in Figure 4-4 along with their allocation across categories. This characterization offers an snapshot of the papers published during the range of study (2009 – 2016). If funding agencies decided to change their funding policies during that period, then the present result is should be considered as the final accumulated output of their policies. Notably, incremental technology was found to be the main target of organizations. Funding organizations with a larger share of papers in the change maker clusters were NSF-CH, NSF-US, EC, EU, and 863-CH. Such distributions can help reveal the underlying patterns across funding organizations, taking into account even small variances in the distribution. Further, we conducted principal component analysis with the same input data. After standardization and via multidimensional scaling of the principal components, we mapped the state of funding for the field of robotics as depicted in Figure 4-5.

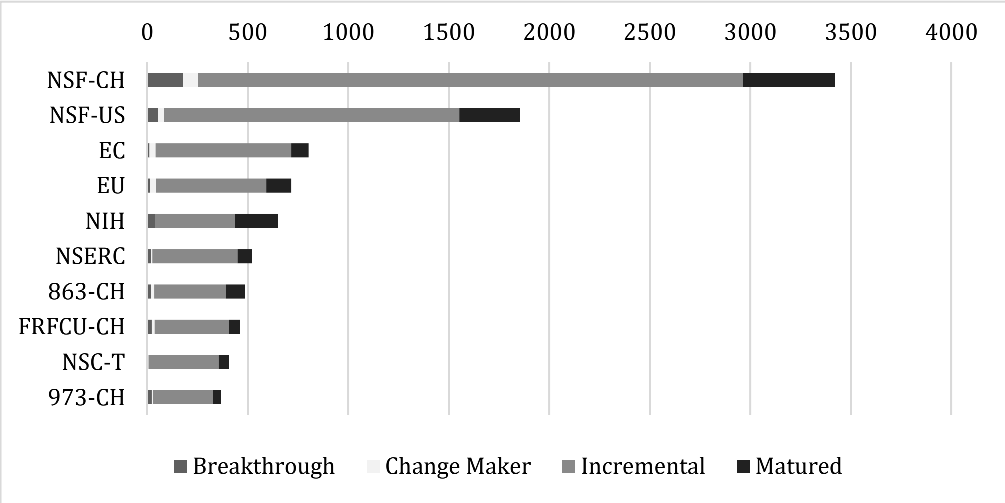

Figure 4-4 Characterization of funding organizations by number of articles and type of technology they sponsor.

It is worth noting that the top funding organizations detected are government agencies. The private sector has little participation in funding research compared to public funding. The company having the maximum mentions in the acknowledgement is American company Intuitive Robotics Inc. that is responsible for the da Vinci Surgical System, a robotic system for minimally invasive complex surgeries. It ranked at position 26, being mentioned in 153 articles (22% incremental, 78% matured). The misrepresentation of the private sector in the acknowledgement section of academic articles may be explained by several reasons such as greater interest in patenting than that in academic research or a strategic preference for anonymity (Rigby 2011).

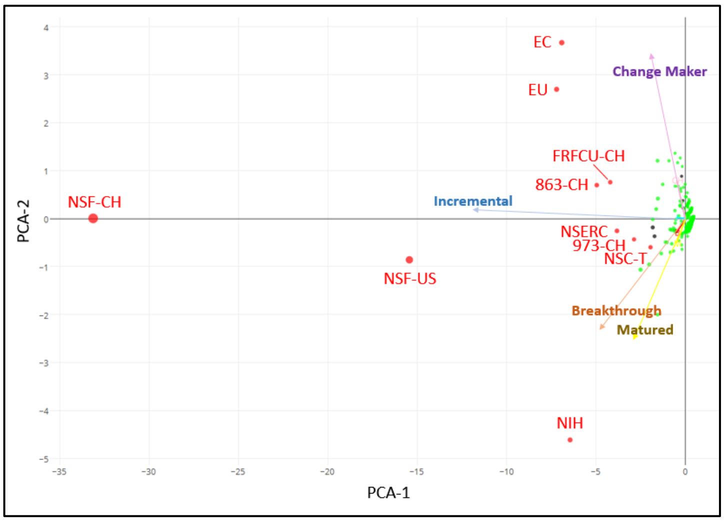

Figure 4-5 Funding organizations of robotic research by number of papers sponsored in each category. The top 10 funding organizations are shown in red.

The axes in Figure 4-5 represent the principal components, and the dots refer to the relative distance among the funding organizations. The four variables are shown as arrows pointing from the origin. The direction of the arrows is interpreted as the intensity of that attribute. Finally, when two or more funding organizations appear close, it is an indication that they share a similar pattern in their funding. The top 10 organizations that are marked in red clearly show their outlier nature as responsible for most of the funding in the network of robotic research. Particularly, the NSF-CH plays a prolific role in terms of the number of publications, allocating funds mostly in the incremental category. Following the NSF-CH, the NSF-US also focuses on the incremental category, but its funding in the incremental category is lower than the former, given the larger participation of matured and breakthrough technologies in its distribution. Down below is the NIH, which mainly targets incremental and matured categories and has almost null participation in change maker technologies, resulting in its existing position. Its conservative nature for fund allocation is well-known in the academic community (Berg 2008; Rangel et al. 2002); therefore the position is expected. The EC and EU are close each other; as discussed above, this may be due to redundant acknowledgement practices or merely the interchangeable use of the name. Further, as can be seen, their closeness serves as confirmation that they allocate funds following the same strategy. Although they fund less change maker articles in comparison to NSF-CH, in terms of their own distribution, they are better at targeting change maker technologies. The reasons for European funds to lead to change making technologies, may partially be explained by the guidelines imposed by their "Framework Programmes". Since the starting of the frameworks in 1984, they have passed through 8 iterations. Concretely, Framework Programmes 6 and 7 running from 2002 to 2013, and contributing with more than 30 billion euros in Research and Development projects (European Comission 2015), ask for grant applicants to join efforts among at least 3 institutions of different countries, encouraging multidisplinary teams. Multidisciplinarity is recognized as an attribute innovation (Olson et al. 2001; Campbell et al. 2017). In the 8th iteration, also called Horizon2020, there were also defined three research targets, basic research, industrial applied research, and "societal challenges" (Anon n.d.). This last one is a "challenge-based approach" aiming to integrate different disciplines (Anon n.d.). As it is born from a real-world problem, driven by policy makers, society, and researchers its transdisciplinary nature is revealed. The remaining organizations follow different patterns as revealed in the chart. It is worth noting that among the two Chinese programs, 863-CH and 973-CH, the objective of the former is funding cutting-edge technology, while that of the latter is funding basic research. Their positions in the map represent the abovementioned objectives.

We also visualized the pattern of funding output in terms of the academic impact measured by the number of citations received by their sponsored papers in . While the characteristics of the two maps are similar, this time, the matured attribute is present over the horizontal axis close to the incremental arrow, implying a high correlation between the two attributes. Therefore, the projections over the vertical axis are determined by the breakthrough and change maker attributes, which happen to be orthogonal.

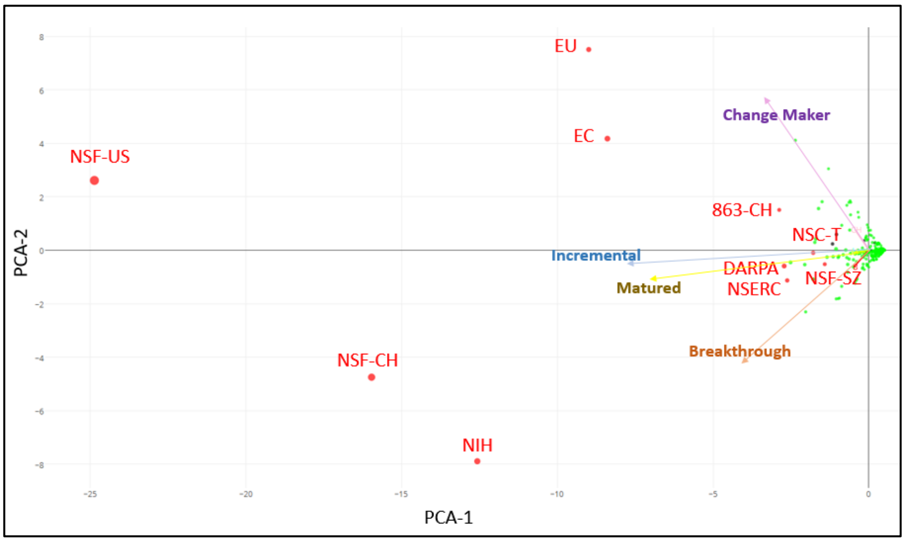

Figure 4-6 Funding organizations for robotic research by citations received in each category. The top 10 funding organizations are shown in red.

Again, NSF-CH and NSF-US dominate the map. However, NSF-US has a greater academic impact. Although this impact comes mainly from incremental papers, the organization's position is pulled up, showing that change maker technologies play a better role in capturing citations. The NIH, EC, EU, and 863-CH remain in similar positions to those shown based on the number of papers in Figure 5. Citations received by papers funded by DARPA and NSF-SZ were found to be higher than those funded by FRFCU-CH and the 973-CH, displacing them in the chart. The performance of Swiss funding in relation to other European countries has also been acknowledged before (Gök et al. 2016).

4.3.3 Japanese funding agencies

We also developed a detailed view of Japanese funding organizations in robotics for the purpose of contrasting our methodology with the expected outputs from their actual guidelines and allocation strategies. Further, to examine their role within their own institutional structure, we investigated the five most mentioned organizations from our data. Figure 4-7 shows a close-up of the funding maps by the number of papers and citations received.

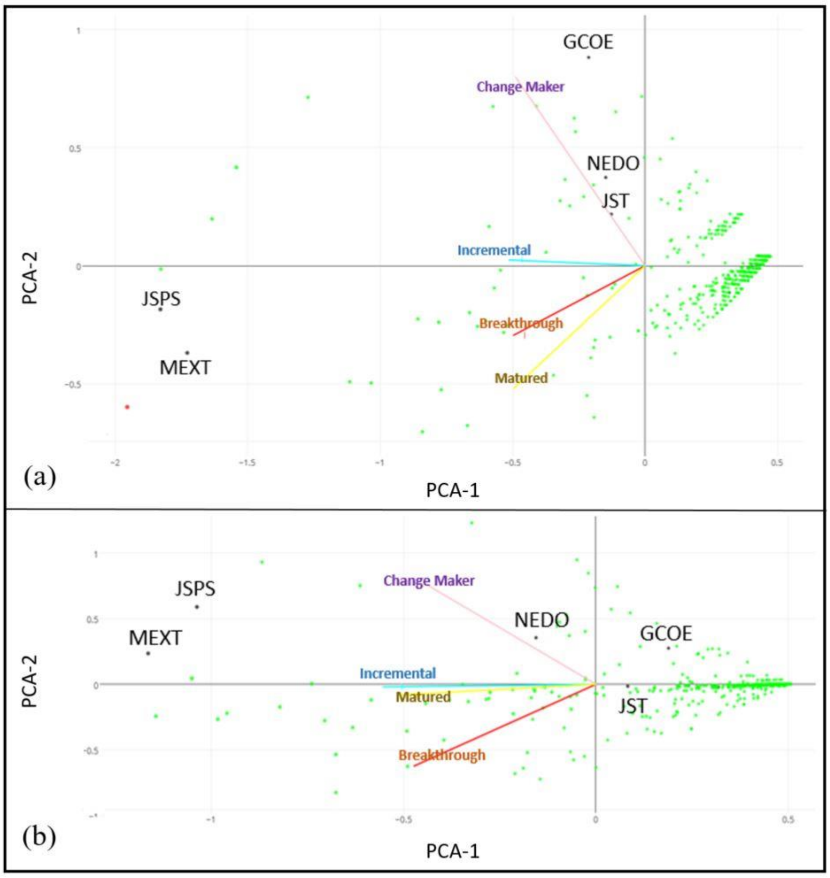

Figure 4-7 Japanese funding organizations of robotic research in relation to the four types of technology. (a) by papers sponsored and (b) citations received, in each category

Funded papers from the JSPS mostly belong to incremental-type technologies. JST, on the other hand, has better positioning in targeting change maker technologies. This is curious result because JSPS is mentioned to support curiosity-driven basic research, while JST is for mission-oriented applied research. This might be due to the application and examination process of the grants. Applications to Grants-in-Aid for Scientific Research, which is the major funding program in JSPS, are submitted to very specific academic category and examined by peer-review in that category. And thus inter-disciplinary usually faces the difficulty in application and examination. And examination is based on past achievements in addition to the proposal. It will lead to accept incremental research based on applicant's past achievements and well-developed capabilities, and suppress emerging and change-maker research. On the other hand, JST allocates its funds through a project manager and committee from various backgrounds, implying more freedom when it comes to funding change-maker research as can be seen in Figure 4-7(a), while it does not lead to high citations as in Figure 4-7(b).

The GCOE is a special program within the JSPS oriented to "highly creative and vanguard research" wherein the proposal must embed a future concept. Along with the spirit of multidisciplinary and international collaboration, which are known indicators of better academic performance, it is not surprising that GCOE gets a higher relative position towards change maker technologies. NEDO, similar to the case of JST, supports applied research but is responsible to development of industrial technologies and companies join the projects. It is examined by a similar system of JST but is also along with policy makers to decide fund allocation. NEDO produces higher citation impact, which might because applied characteristics of robotics technologies and superior capability of Japanese companies in this field. Finally, MEXT is also acknowledged in the articles, although core fund allocation for research are decided by JSPS and JST, as mentioned earlier. The appearance of MEXT may come from other types of funding, such as scholarships, that may be administered directly by the ministry, or perhaps, by the natural practice of a researcher to mention both the funding agency and the parent ministry. Given the proximity to JSPS, this may be true, especially for researchers funded by that agency.

Interestingly, both JSPS and JST seems to switch roles from the perspective of citations (Figure 7b). Even though peer-review system is conservative, the curiosity-driven research by academic will lead to high impact research. The position of JSPS suggests that the citations received from its change maker papers have better impact that that in the case of JST. NEDO and GCOE remain in a position where change maker technologies play an important role. Finally, MEXT continues to be close to JSPS, as the second evidence of double acknowledgement.

4.3.4 Funding strategies by targeted subject areas

Within the scope of robotic research, funding organizations also have more specific fields of interest. We examined those specific targets by mapping the distribution of subject fields assigned to the articles. Given that the data under analysis was about robotics, we neglected four subject areas that were the direct representation of this field, namely "Robotics," "Engineering," "Computer Science," and "Automation and Control" and focused on the subsequent most frequent categories. Figure 8 shows the distribution for selected agencies, the top international and Japanese organizations. We found that those categories vary from purely basic to medical research, "instrument & Instrumentation" and "Surgery" being the most common. Top organizations seem to have a similar pattern of interest, especially for the most frequent categories. However, smaller categories may also play a significant role. Therefore, we repeated the process of multidimensional reduction of principal components to bring out the patterns at the surface.

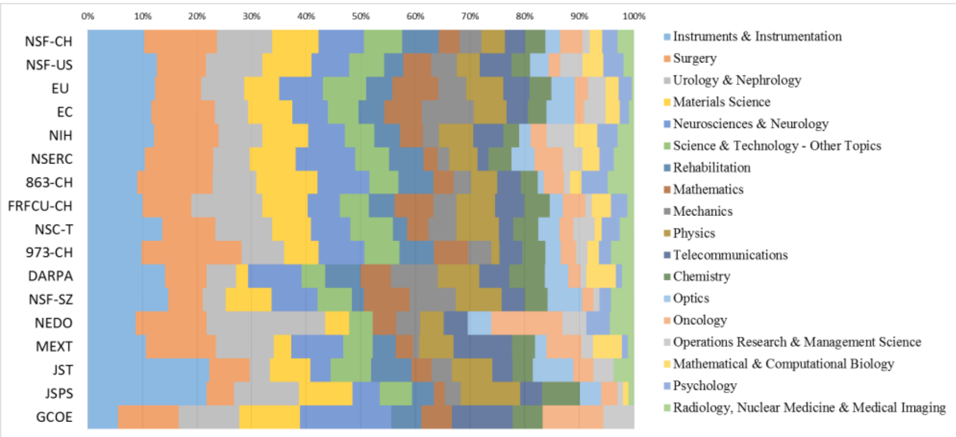

Figure 4-8 Distribution of the most sponsored subject fields within robotics research for selected funding organizations

Figure 4-9 shows the funding agencies with respect to specific research field targets. Similar to previous figures, the axes represent relative distance. However, unlike the biplots based on the four technology-type categories, this time, the attributes are spanned all over the chart. This is owing to the different nature and skewness of the distributions, being not as sharp as those produced by the large participation of the incremental category in the previous charts. Through the inspection of attributes, it can be observed that medical research dominates the right-hand side panel, while basic and engineering fields represent the left one. Funding organizations of the same country seem to be grouped by region in the chart. This is not necessarily expected, because funding organizations may have different objectives. We can infer, to some extent, from the chart that robotic research follows funding polices set at the country (or regional for EU) level. This may not hold for all the agencies; for instance, DARPA seems to have a far different strategy than NSF-US and NIH. In addition, the FRFCU-CH is closer to targeting the interest of Japanese organizations rather than the Chinese organizations.

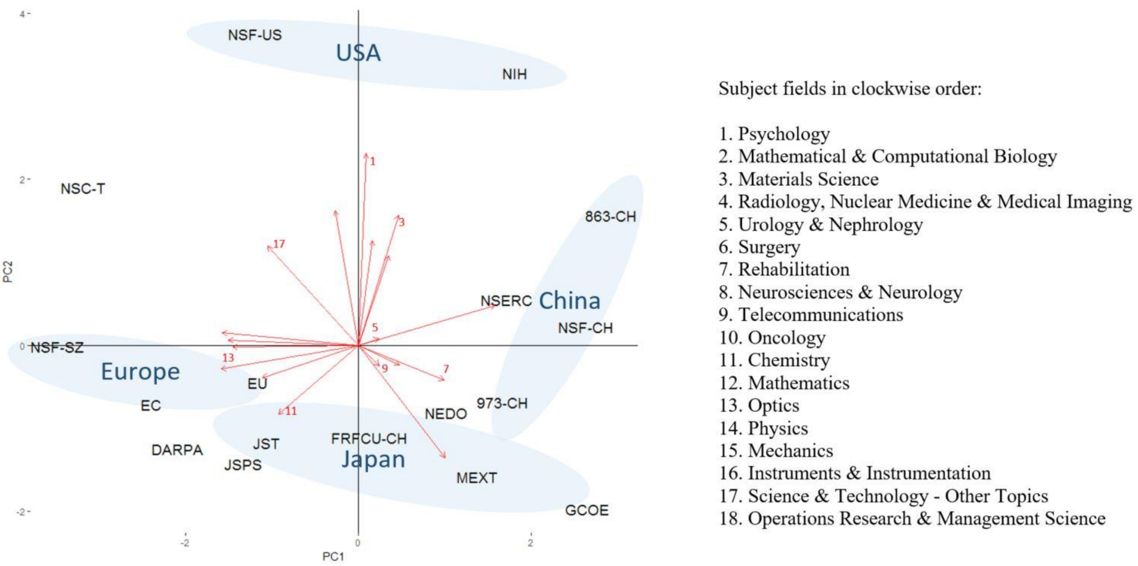

Figure 4-9 Target subject of funding organizations

Despite those trends, little else can be inferred of the role of governments. Acknowledgements seem to contain the name of the organizations only, and details of projects and programs is barely present. Table 4-3 shows a summary of the presence of projects, programs, frameworks, or ministries participating in the funding of robotics as reported in the acknowledgements.

Table 4-3 Summary of different types of sponsoring entities reported in the acknowledgements.

	Count	Papers	Coverage	Cites
Ministries	54	2,005	10.63%	10,010
Programs	43	2,220	11.77%	10,565
Projects	10	201	1.07%	1,129

For instance, 2,005 articles out of the 18,860 (10.63%) reporting funding acknowledge to have received sponsorship from any of the 54 ministries identified in the robotics dataset. The most acknowledged ministries are the Ministry of Knowledge and Economy of Korea, the MEXT of Japan, and the Ministry of Education, Science and Technology of Korea with 302, 241, and 127 academic articles respectively.

Some authors may straightforwardly acknowledge their government. "Spanish Government" was the one most reported having 144 articles. However, there were also found "Catalonian Government" and "Valencia Government" as sponsoring entities. Both Catalonia and Valencia are provinces of Spain, thus revealing that researchers may recognize local sponsorship rather than national.

4.4 Summary

In this chapter it was found that funding agencies do target innovative research, but mostly in the incremental stage. Also, the participation of funding agencies in matured topics is higher than in breakthroughs, or change making clusters. Through, a closer inspection of Japanese funding agencies it was also found, that their practices at the time of allocating resources may be the determinant of their position in the funding landscape. This is, those agencies more flexible, with participative selection committees like the JST may target emergent topics better than other organizations were the applications include fixed categories, or weight up previous achievement of the applicants, thus favoring incremental research, like JSPS.

To reach those findings the information provided in the acknowledgement section of papers was used to identify funding agencies. Then, by means of citation networks, we extracted the features that allowed the classification of research articles into breakthrough, change maker, incremental, and matured-type of technologies.

The National Natural Science Foundation of China and the National Science Foundation of US play a leading role in promoting robotics research. Although also largely incremental, European funds have been proven to be among the best in spotting change maker technologies. Other funding agencies such as the National Institute of Health of US follow known strategies focused on both incremental and matured categories, over other most innovative categories.

Japanese funding agencies were also investigated to understand how their different characteristics affect research direction. It was found that the Japanese Society for the Promotion of Science, even though having a conservative peer review-based funding allocation system, when it comes to academic impact they do target change marker research. However, in terms of amount of papers they focus on incremental technologies. On the other hand, articles funded by the Japan Science and Technology Agency may be regarded as innovative, but receiving fewer citations. Other organizations as the New Energy and Industrial Development Organization were positioned as expected from its known practices, towards change maker research.

Within the scope of robotics, funding agencies also target other specific fields of knowledge, where medical research is particularly relevant, “Surgery” being one of the most frequent targets. By mapping the funding agencies corresponding to their fields of interest, we found that they tend to concentrate at the country level, a pattern easily explainable given that most of them are government-dependent agencies, but also suggesting homogeneous guidelines within each country.

We started with the premise that research fronts can be observed from citation networks of academic articles. Among the methodologies for creating citation networks it has been found that direct citation networks work better. We also know, that novelty and coherence are two attributes of emergence, both applied to our framework, through clustering and a classification based on their age. We expect the academic landscape provided represent the state-of-research for the time range in study. We then measured the participation of funding organizations as reported in the acknowledgement section of the articles.

Researchers acknowledge financial sponsors voluntarily or because of grants requirements. In both cases this serves as a signal of the openness of funding organizations in being linked to the outputs brought by the sponsored researchers. Therefore, Acknowledgements can be used as an indication of that disclosure. Under the proposed method, only those organization openly engaged in academic production can be assessed. Moreover, given the shortcoming on how researchers acknowledge their sponsors, it can only be measured their participation, and not the amount of funding received. In fact the mere presence of Acknowledgments does not necessarily correlates with funding (Cronin & Shaw 1999).

Some limitations were found. The proposed method is supposed to assess funding for academic research only. Funding agencies may pursue different objectives and have different funding modalities whose impact is beyond academic research. The assessment of the overall impact of public or private funding is beyond the scope of our study, and several other studies and frameworks are already available (Guinea et al. 2015; Kostoff 1993; Bloch et al. 2014). Also, for analyzing the contribution of governments it will be needed clear definitions of boundaries in identifying local or national funding, and having the names of the agencies or organizations that may be considered as governmental. In a more general trend, dealing with entity names that change in time, as in the case of ministries changing as new governments are elected, is still a challenge. However, our method may serve as an indicator that can easily be integrated with those frameworks. With respect to the data and methods, the classification model adopted in this research depends on relative bibliometric features. The refinement of bibliometric tools for classifying technologies into different categories of innovation is a field rich of opportunities. Notwithstanding, given the results and their correspondence to known funding patterns, our case study also sums up the suitability of such categories. It may be pertinent for future research to translate the methodology for mapping the innovative participation of institutions or authors, or even companies by using citation networks of patents.

The mapping of funding agencies through the proposed method may help policy makers and funding organizations assess whether actual outputs are in line with their proposed objectives, set new directions, and establish benchmarks. Analyzing different windows of time, will also reveal if changes in policies correspond to new outputs generated. Researchers may also benefit from the study of funding patterns to establish international collaborative strategies or to spot the agencies that more likely sponsor their type of research.

5 DISCUSSION AND CONCLUSION

This thesis applied methodologies from network theory and text mining to elucidate the interactions between social attention and funding in the configuration of emerging technologies. Our findings are circumscribed in the robotics domain, were both active participation of society, and funding from public and private sources is expected. Here, it is brought a discussion in relation to the finding in previous chapters, their connection to emerging technologies, challenges, and future directions.

5.1 Findings

Chapter 3 addressed the issue of elucidating whether levels of social attention play any role in the configuration of new research topics. The amount of news was used as proxy of social attention, and social sentiment, as positive or negative was attached to the news. The most relevant finding was that topics positively discussed in media may lead to further developments of similar topics in academia. Therefore, the methodology applied suggest, that it might be possible to forecast advancement in science by checking the landscapes of social attention of technologies.

However, news did not discuss all topics in science. By inspecting the connections, we observe news concentrated in practical technologies. This is applied research over basic research. Which may be expected to some extent, as basic research topics can be still uncertain in their role in society.

Also, the discussion was around on how the representation of social attention in news articles reflects the patterns observed in the hype cycle literature. We observed hypes for robotics, however, they reflect a balanced sentiment contrary to the positive rise and negative delusion argued by those frameworks. Therefore, we cannot say that robotics follows a hype cycle, but that hype dynamics are present, and follows different features to those originally proposed by Gartner (Fenn & Raskino 2008).

In chapter 4, we aim to understand whether funding organizations target innovative research. We parted from the premise that funding organizations indeed attempt to target innovative research but other forces may constraint them to target them properly. We take advantage, of the network properties to identify novel clusters in robotics research, and classify them into matured, incremental, breakthroughs or change makers.

By analyzing the participation of funding organizations in the clusters we found that, funding organizations do target innovative research, but mostly incremental. And, that participation of funding agencies in matured technologies is even larger than in the change maker and breakthroughs. The largest contributors were the National Science Foundation of China and the U.S. However, among the international organization those of Europe were found to fit a pattern towards change making research.

By addressing the cases of the Japan Society for the Promotion of Science and the Japan Science and Technology Agency, we discussed that their allocation strategies may be determining their locations in the funding maps. As for the JSPS which the application process is more controlled tend to be incremental, while JST being more participative tends to change making research.

From a methodological perspective, we bring forth the possibility of using acknowledgement data for the creation of funding landscapes in science, which may serve as benchmarking tool, or assessment of funding policies.

The studies in chapter 3 and 4 are independent. However, the underlying connection relates to emergent technologies. In the following sections we discuss about how to leverage the detection of emergent research, how that is related to social attention and funding, and finally, what are the dynamics that connects them.

5.2 Towards a more comprehensive indicator of emergence

The study of social attention stablished relations between topics in news and cluster of academic articles. However, differently from the approach on chapter 4, we did not make a typology of innovation, as the objective was to identify if discussion in news appeared before to those similar topics in science. Which it was found to be true.

It is worth noting that while the approach undertaken in chapter 4 takes into consideration the coherence of the network (when forming clusters) and novelty (from where the typology is obtained), those are not the only attributes of emergence. As explained before, fast growth, impact, and uncertainty are attributes also in play. Fast growth is measurable, while the other two are regarded to be qualitative. Can we derive new emergence indicator that takes into consideration the quantifiable attributes of emergence? And how such indicator may be related to social attention and funding?

To answer those questions indicators of emergence were computed for the fourth level of sub-clusters (see Chapter 2), which offer partitions of the 4 streams of research to higher level of resolution. For each of those 450 sub-clusters we computed three types of indicators of emergence that can be obtained by their bibliographic features as explained in Table 5-1, while Figure 5-1 Shows correlations between those indicators.

Table 5-1 Quantifiable attributes of emergence

Indicator	***Description***
Novelty	
Cluster newness	Average publication year of the papers in the cluster.
Hub newness	Difference between the publication year of the hub-paper of the cluster and the cluster newness. A hub-paper is defined as the paper with the highest degree or connections within the cluster.
Growth	
Speed	Average of the slopes of the cumulative publication trend of papers in the cluster between the years 2014 to 2017.
Acceleration	Average of the slopes of the cumulative publication trend of papers in the cluster between the years 2014 to 2017. Normalized in relation to cumulative trends of all robotics publication for those years.
Coherence	
Network coherence	Density of edges within the cluster. Computed as the proportion of the actual number of edges over the total possible edges in the cluster.
Text coherence	Average of the cosine similarity among the papers in the clusters. Term vectors were computed based on the TFIDF of the words found on titles and abstracts.

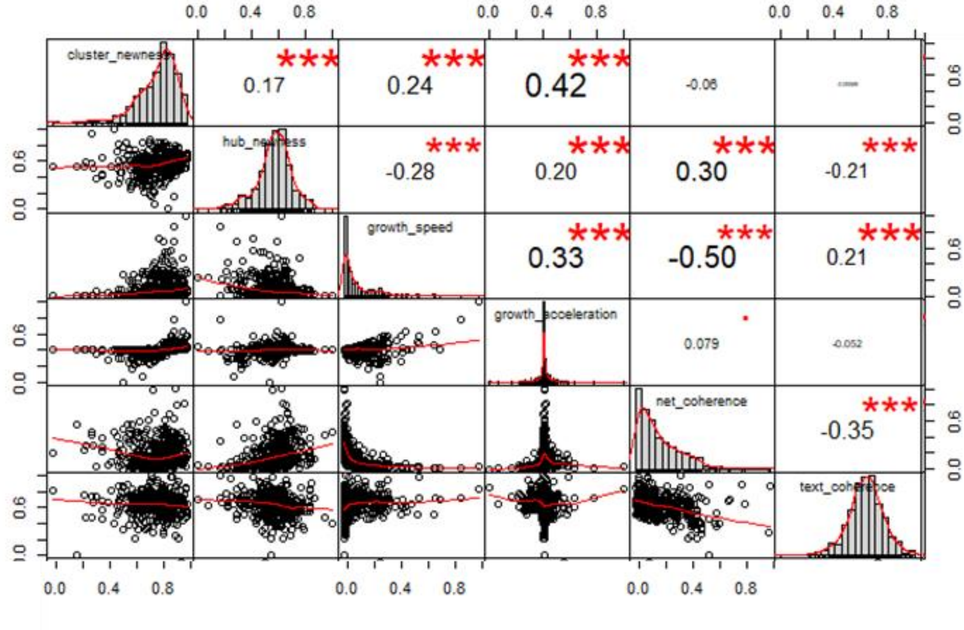

Figure 5-1 Correlation between indicators of emergence.
Size and average citations per paper are also incorporated.
Stars refer to the p-values "***" = 0, "**" = 0.001, "*" = 0.05, "." = 0.1
All indicators are normalized, from 0 to 1, being 1 the representation of larger or better.

The diagonal shows the distribution of each indicator. Growth speed, network coherence, meaning that most of the clusters perform poorly on those. Growth acceleration is concentrated in the middle as reflection that clusters grow proportionally to the overall publishing of robotics papers. Hub newness and text coherence resemble a bell-shaped distribution. On average the publication year of hub papers is close to the average of the publication years of the rest of articles in the cluster. For text coherence, means than on average cluster have text similarity around 0.7.

Bellow the diagonal, are bivariate scatter plots with a fitted trend line. Corresponding to the correlation score for the same intersection above the diagonal (The stars signal confidence intervals). The correlation among the emergence indicator is poor even between pairs of the same categories. This is due to the fact that we are dealing with all the clusters that came up from our network, including well known matured clusters that by definition must score low.

It is known that emergent clusters will perform well in one or more indicators (Avila-Robinson & Miyazaki 2011). For that reason we computed an *emergence score*, which is the simple average of the 6 indicators. Figure 5-2 shows the score of each cluster sorted from the highest. A set of clusters clearly differs from the rest. Therefore, the emergent robotics cluster were identified. High scoring clusters are described in the appendix 7.2.

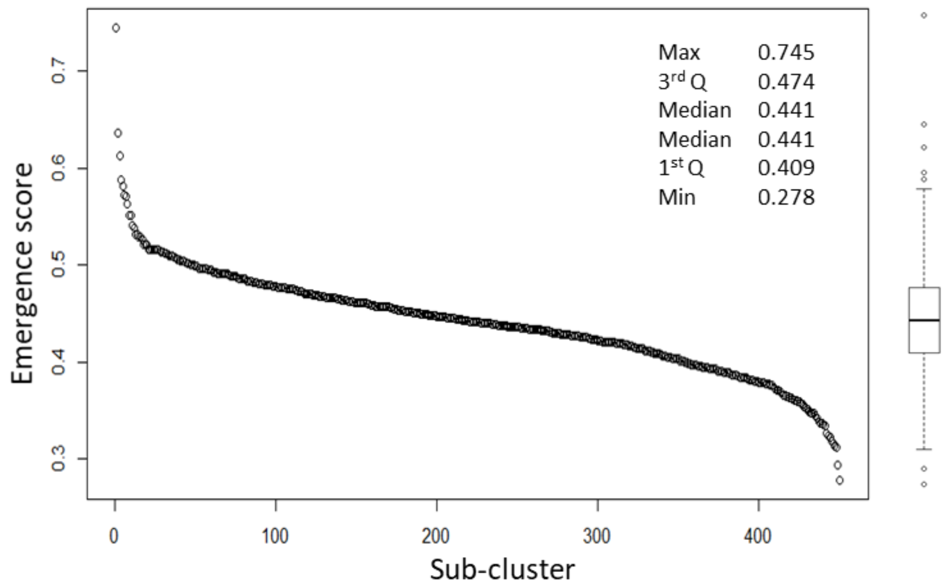

Figure 5-2 Emergence scores for the 450 sub-clusters of robotics research

To understand the relationship of the emergent clusters to social attention and funding, we derive indicators based on the learnings of chapter 3 and 4. To finally correlate all of them. Figure 5-3 shows the set of indicators.

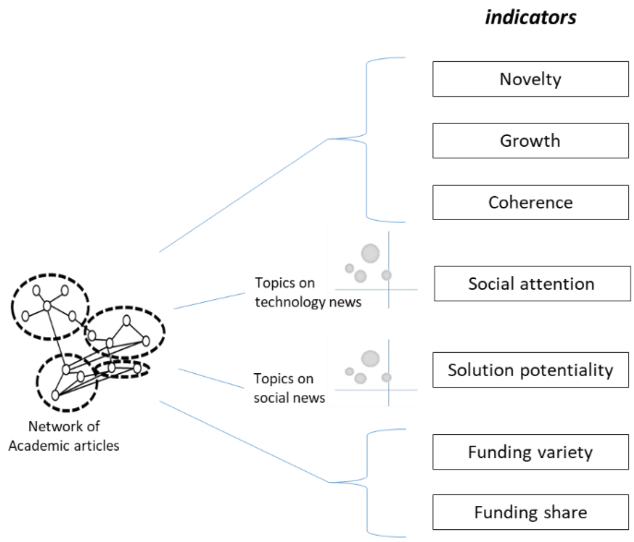

Figure 5-3 Indicators of emergence, social attention, and funding.

Two indicators are derived from the methodology in Chapter 3, by using both news and academic articles. One of the limitations found there, was the use of a single newspaper. Therefore, to overcome that issue in this computation we used a more comprehensive source of news articles. The Ebscohost database. It is like the Web of Science in purpose and usability, however, it also indexes international newspapers and magazines.

For each set, we created topic models in the same manner as described in chapter 3. Then, it was computed the similarity between clusters in the network of academic articles and the topic models of news. The similarity matrices are the input to compute two social aspect indicators.

The first is *social attention score*, which is the weighted sum of similarities of the cluster to the topics in the robotics news (11,972 articles). And the second, is the *potential solution score*, which is the sum of the similarities to topics in the country-news articles, the one containing social issues. This last indicator is a derivation of the methodology explained in chapter 3, where the key difference lays in the change of the dataset of news (4,068 articles). In this case, covering social issues news of a given country (i.e. Japan), thus bringing a landscape of topics discussed about that country, where social issues can be found (The mechanics of that approach are in the appendix 7.3).

For each cluster we computed two indicators of funding. The *variety of organizations*, which is the count of different organizations sponsoring research in one cluster. And the *share of funding*, which the proportion of articles (from 2009 to 2017) that mention financial acknowledgement.

Topic models of technology news and society news are shown in Figure 5-4. Despite the increase of data in relation of previous chapter, the topic model outputted solutions with fewer topics. 30 and 25 respectively. The number of topics is defined by the maximum likelihood of the models and is independent of the amount of data. Here, topics represent only a mean for the computation of the scores.

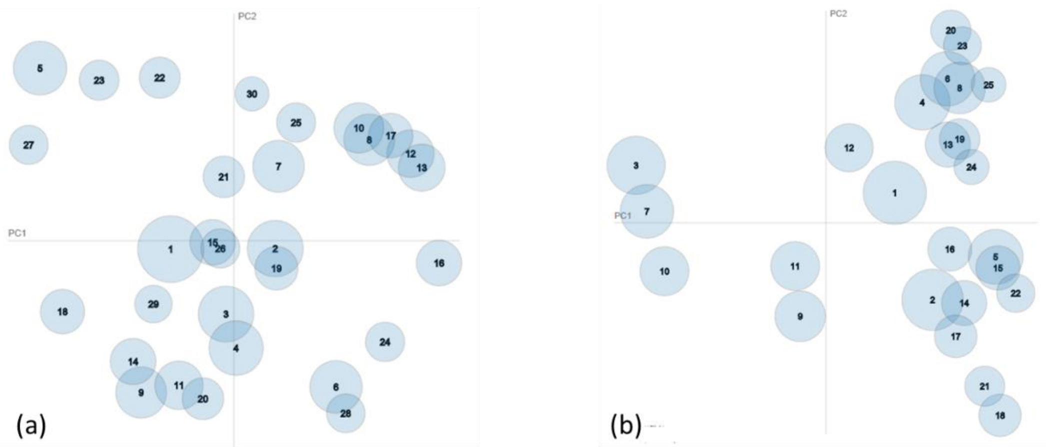

Figure 5-4 Topic models for (a) technology news and country specific news (b)

Cosine similarity was calculated between clusters and topics. This similarities were used to measure the attention score, and solution potentiality score, which distributions are shown in Figure 5-5. On average both indicator conform weak signals as is observed by the right-skewed distributions. Notwithstanding, the score serves to detect cluster whose topics are similar to the discussion found in the media.

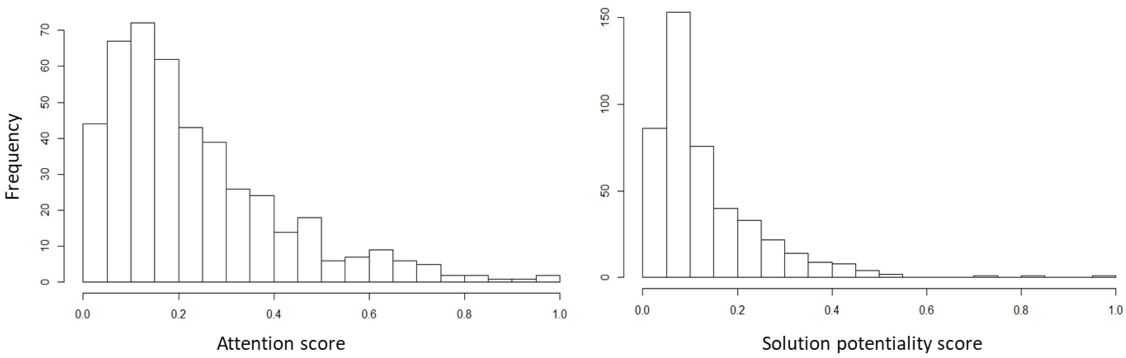

Figure 5-5 Distributions of the social indicators

Correlation among them are shown in Table 5-2. We want to know the correlation scores between indicators affecting most emergent clusters. Therefore, to compute the correlation matrix, we considered only those clusters having an emergence score in the third quartile. Being the higher clusters. The highest correlation is between the variety of funding organizations participating in the cluster and growth of speed. Besides that, corelations are low for the rest of pairs.

Table 5-2 Correlation of indicators of emergence, social attention indicators and funding

		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
1	Cluster newness	1.00											
2	Hub newness	-0.12	1.00										
3	Growth speed	0.05	-0.50	1.00									
4	Acceleration	0.35	-0.35	0.75	1.00								
5	Net coherence	-0.13	0.21	-0.63	-0.42	1.00							
6	Text coherence	-0.25	-0.19	0.13	0.04	-0.23	1.00						
7	Cites per paper	0.14	-0.25	0.37	0.30	-0.29	-0.05	1.00					
8	*Emergence score*	0.19	-0.07	0.42	0.62	0.19	0.30	0.07	1.00				
9	Attention score	0.09	-0.08	-0.02	-0.02	0.00	0.01	-0.18	-0.03	1.00			
10	Potential solution	-0.04	-0.05	-0.02	-0.04	0.05	0.16	-0.17	0.07	0.69	1.00		
11	Organizations	-0.07	-0.47	0.90	0.53	-0.57	0.10	0.48	0.27	-0.07	-0.03	1.00	
12	Prop funded	-0.14	-0.12	0.23	0.11	-0.18	0.05	0.29	0.00	-0.34	-0.21	0.41	1.00

Therefore, it can be said that a relationship exist between the number of funding organizations participating in fast growing clusters.

5.3 Towards a theory of funding for emergent scientific research

From the previous chapters, and the above analysis of correlations this thesis elucidates some underlying dynamics between emergent technologies, social attention, and funding. Previously we argued that positive news about a technology topic do have a reflection in clusters of academic articles having a younger age and having similar textual contents. This suggest that researchers are not immune to the seduction of hot or trending topics, and when those topics are positively discussed they do try to pursue research in those directions. However, we cannot say that those hot topics were born in newspaper articles alone. There should be a trigger, that motivates the social discussion.

We also found a high correlation between the amount of funding organizations involved in a cluster and its growth, being the only high correlation in relation to the other attributes of emergence.

By putting all pieces together, this thesis leans towards a *theory of funding for emergent scientific research*, which mechanism of working is attempted to be depicted in Figure 5-6.

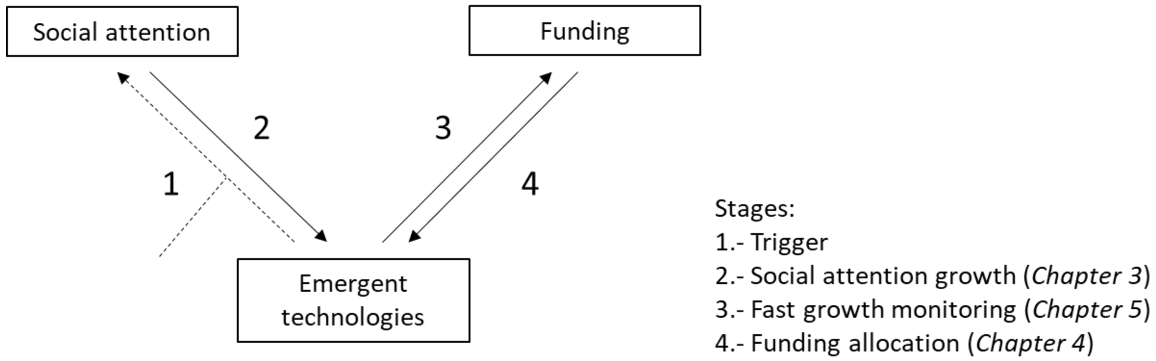

Figure 5-6 Dynamics of emergent technologies, social attention, and funding

In the academic landscape of a research domain, at any given moment, there is a set of emerging research topics. Some of them get to receive a high social attention becoming trending topics, which discussion happens to be positive. However, what triggers this social discussion may also come from outside the academic landscape, for instance in the form of a product or patent (Denoted with the dashed line). As researchers realize such trends they jump in, and drive their efforts towards those directions. However, being the path of science slower of that of news, the increase of academic publishing in the related topic only gets noticeable at a later time. When the field start growing.

This is the point when most funding organizations jump in. As they may perceive fewer risk when the capabilities of researchers become evident by the amount of publishing. However, at this point the path of research have been set, thus the efforts are rather of incremental nature.

The generalization of that mechanism would lead to one straightforward implication: speeding up the pace of innovation in scientific research. Instead for funding organizations to wait until identifying fast growing research topics to invest in, they may look at patterns of positive social attention obtained by measuring printed media. Figure 5-7 is the suggested model.

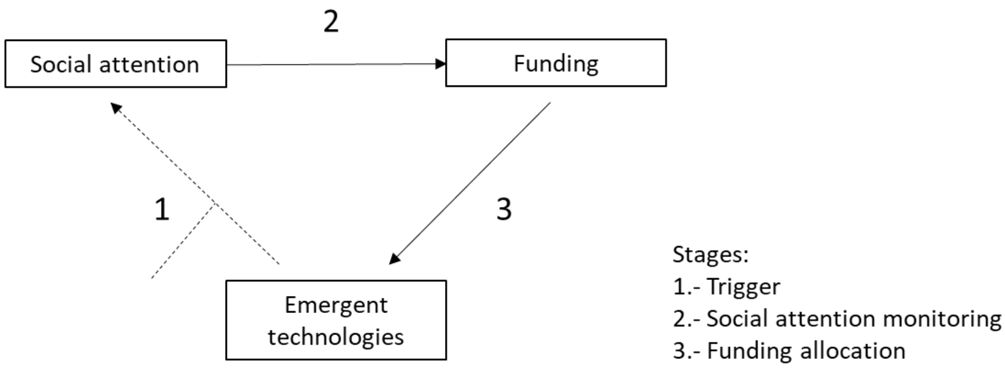

Figure 5-7 Proposed model for funding

The triggering phase is the same. Society gets the source of discussion either from emerging topics in science, or outside it. However, funding organization may react faster by monitoring the levels of social attention. Moreover, the integration of social attention monitoring and the emergence scores previous described may leverage their funding allocation strategies targeting emergent research timely. Notwithstanding, for such a theory come to be, it is needed to address the following challenges.

First, we joined the pieces for robotics research, where social involvement and funding is granted. Nowadays, other topics as deep learning, internet of things, or blockchain may be enjoying of similar conditions. But this cannot be said of all emergent topics of science.

Shifting from one paradigm to another, is not easy. To look at the levels of social attention in the news instead of the fast-growing trends of science carries a greater sense of risk. In this research we represented social attention as coherent narratives found in the news by applying the unsupervised machine learning technique of topic models. However, not all topics, even though well-defined found higher scores of similarities with the academic landscape of robotics. Moreover, topics model may also produce cryptic topics when the parameter is set without rigor.

This research studied the interaction between social attention and academic research. But understanding the origin (i.e. the trigger) was beyond scope. We delimited our scope to the measurement of two features, attention by the amount of news, and the social mood by sentiment analysis. It is still needed more research on how the technology hypes appear, and the underlying forces or actors that keep them lasting enough to generate impact in the academic domains.

It was also found that social attention indicators did not correlate with the classic attributes of emergence. Which seem to partially contradict the findings in Chapter 3. However, it is not the case. A closer look to the highly similar topics found in chapter 3 and the expanded study of social issues in the appendix 7.3, it is understood that topics discussed in the news are closely related to applied research. Basic research is still in a too early stage to generate momentum in society. In the third quartile of cluster having higher emergence score, it was not drawn any distinction between applied or basic research. Given that both of them were part of the correlation study is easy to see that this may have caused the lack of matching. However, is needed more evidence in this regard.

Also, the pattern and sequence of events revealed through the present methodology was observed at the aggregated levels of clusters or topics. In other words, groups of documents sharing an idea or having a conceptual linkage. Intuitively, there should also exist specific and explicit linkages between a single paper, researcher, or technology and the news it generated. However such direct connection was beyond the scope of the study, and limited to the general representation of technology in the imaginings of the people.

Chapter 4 also revealed that while most funding organization involved in this study, tend to be allocated in incremental clusters, some of them are better in spotting change making research, which according to the used framework correspond to the most novel clusters in the network. Understanding the allocation strategies of such organizations may leverage the way funding is conducted as a whole. From a transdisciplinary view, with the participation of parties outside the academia, and having starting points in real-world problems (Walter et al. 2007), it was observed that funding organizations that pay attention to social concerns and integrate those in their funding guidelines, as was the case for research sponsored by Europeans funds, get better at targeting change making research. Although is still needed more studies to compare other organizations in which the social component may also be present.

The reason for researchers to choose some topics over others can vary. The present thesis, however, shed light in a partial understanding of those reasons, as some topics in science and technologies can be observable first in society. On the other hand, basic research is argued to be borne by self-driven curiosity (Salter & Martin 2001). While some research in social sciences and humanities may respond to hypes of other nature, for instance high level attention to social issues topics whose are perceived as negative. Instead of positive rise of techno-hypes, those negatives hypes might be the drivers of social research. Opportunities of translating the methodologies, and theory developed in this thesis abound.

We took the narrative of the “mode-2” knowledge production and its transdisciplinary nature, to situate funding organizations as mediators between science and society. Arguing that a “consulting” transdisciplinary approach can help set the direction of knowledge by extracting the concealed or contested reactions of society from the polarity and intensity of discussion that academic topics rise in news media.

The findings of this dissertation also support the need of clarifying the position of the several actors in determining the direction of science. Each actor being researchers, funding organisms, governments or practitioner share an unbalanced stake of power and control, they may also follow hidden agendas that determine the direction of research (Rosendahl et al. 2015). Oftentimes, science and society are seen as a dichotomy where researchers are meant to serve as bridges in carrying the knowledge and come up with ideas for a successful implementation in society. However, other stakeholders also have that bridging power (Nowotny 2015; Lawrence 2015). As discussed, we point to the quota of power behold by funding organizations who may observe the reactions of society and act accordingly, proposing incentives or changing the approaches in which the sponsored scientist engage to the community. As the decision of whether pursue transdisciplinary research is also determined by the incentives set at institutional levels (Lauto & Sengoku 2015).

Funding organizations acting as watchers of society incorporate other benefits into the knowledge production process. When researchers have a broad spectrum of choices to grasp funding they may get free from some agendas impose by governments or other institutions. This means that researcher may willingly subscribe to their preferred approach to science (Musselin 2014). Although this is possible for fields were funding opportunities abound, in practice funding is getting scarce in relative terms (Gläser & Velarde 2018). And thus a better understanding of other constraints is still needed.

Also funding organization in general are seen as complementary counterpart to other social stakeholders. They complement funding from contracts and consulting, contributing to increasing the collaboration of universities with industry, making the knowledge transfer processes smoother (Muscio et al. 2013). This is relevant in those cases were other modalities of funding, like some coming from industry, may act as an impediment in the knowledge production process. Because, under some settings and for some research domains industry funding appear to slowdown or delay the diffusion of knowledge due to secrecy or withholding information on the research process (Campbell et al. 2000; Sismondo 2009). Notwithstanding, research outputs coming from industry funding are of more applied nature (Lam 2010), and thus being of more expedite benefit to society.

In the words of Costanza (2003) “having a shared vision of a sustainable and desirable society is one of the most challenging tasks facing humanity”, this can only be achieved if society legitimate this vision by participating in its creation, and the process has been transparent (Gibbons 1999). The theory developed in this thesis contributes to bring a clearer image of the participation of the actors involved in dialogue between science and society. By no means, it is argued that theory holds entirely. What this thesis is offering, is the starting evidence that the dynamics occur in that way. Concretely for *applied emerging topics* in scientific publishing. In this manner, the results presented in this thesis suggest that they can be replicated in technology domains expected to have a direct application into society. In particular, for technology topics where society is aware of its introduction, thus triggering a social discussion. While the main contribution of this thesis was to bring quantitative evidence of the dynamics of funding for emerging research based on social attention, this thesis did also provide methodological contributions.

The origin of the methodologies described in this research is rooted in the computer sciences and library sciences. However, what is highlighted is the potentiality of integrating those methods in a manner that may derive practical implications in the development of better science and technology policies. Also, we explored the possibilities of using alternative sources of information as news articles to leverage those results supported only by academic publishing. The both reflect two perspectives of reality. News articles signal the mood of the society in real time while being limited by the simplicity of vocabulary and depth of understanding. On the contrary, academic articles represent fronts of research, analyzed in depth, but constrained by the relative slow pace of publishing. It can be regarded that the first represents the problem space, and the second the solution space. Thus their integration might imply a powerful tool in the policy making.

The results obtained through this case study are expected to be used by stakeholders in the government, funding agencies, and the academia. In the first place, the findings and methods proposed in this thesis target policy makers, and decision makers in the government or funding organizations. Expecting for them to take better informed decisions conducing to sponsoring change making technologies, or to benchmark their performance.

In second place, researchers. From a transdisciplinary standpoint researchers might find their topics to be contested or concealed in society and by realizing that, try to proceed accordingly. This is, bringing members of society to the process of goal setting when negative attention is present. Also researchers may benefit by knowing the funding landscape of their topic. Identify funding organizations prone to take risks when the results are at early stage.

5.4 Future directions

Given the findings and challenge we came across, the opportunities to reach to a robust theory of funding can be as follows: First, further evidence is a must. The replication of the proposed methods to other fields of research is necessary to generalize the findings. Some topics in fact, can be arguing to be in the half of those purpose. For instance, nanotechnology is a field with studies in emergent technologies and funding, thus a complimentary study on social attention may help to reveal the complete view. On the other hand, topics like alternative fuel vehicles have been subject of hype cycles studies, and what is missing is the funding analysis.

Following, the refinement of mapping techniques for unstructured text may help to bring accuracy in the metrics obtained. Also, on the methodological side, frameworks for the identification of change making research can be improved by unifying the indicators of emergent technologies based on its agreed attributes. This constitute an interesting path of research because such frameworks may be applied at different levels (countries, institutions, authors). Also, the causalities of both, social attention and emergent research needs more study. We gave a glance on the role of social attention in the defining of some emerging research topics. But the understanding of development of basic research, or emerging topics in beyond the hard science is still unknown.

Finally, it was found little participation of the private sector in the configuration of research topics represented in academic articles. A comprehensive view of their role triggering emerging topics, or their overall dynamics in the framework of funding need more studies, which can also be covered by the inclusion of the analysis of patents.

6 References

Abercrombie, R.K., Udoeyop, A.W. & Schlicher, B.G., 2012. A study of scientometric methods to identify emerging technologies via modeling of milestones. *Scientometrics*, 91(2), pp.327–342.

Adai, A.T. et al., 2004. LGL: Creating a map of protein function with an algorithm for visualizing very large biological networks. *Journal of Molecular Biology*, 340, pp.179–190.

Aeberhard, A. & Rist, S., 2009. Transdisciplinary co-production of knowledge in the development of organic agriculture in Switzerland. *Ecological Economics*, 68(4), pp.1171–1181.

Ahlering, T. et al., 2003. Successful transfer of open surgical skills to a laparoscopic environment using a robotic interface: initial experience with laparoscopic radical prostatectomy. *The Journal of urology*.

Alkemade, F. & Suurs, R.A.A., 2012. Patterns of expectations for emerging sustainable technologies. *Technological Forecasting and Social Change*, 79(3), pp.448–456.

AlSumait, L. et al., 2009. Topic significance ranking of LDA generative models. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. pp. 67–82.

Andrade, A.O. et al., 2014. Bridging the gap between robotic technology and health care. *Biomedical Signal Processing and Control*, 10, pp.65–78. Available at: http://dx.doi.org/10.1016/j.bspc.2013.12.009.

Andrew Bagnell, J., 2014. Reinforcement Learning in Robotics: A Survey. *Springer Tracts in Advanced Robotics*.

Anon, Horizon 2020 sections - European Commission. Available at: https://ec.europa.eu/programmes/horizon2020/h2020-sections [Accessed November 28, 2017a].

Anon, Societal Challenges - European Commission. Available at: https://ec.europa.eu/programmes/horizon2020/en/h2020-section/societal-challenges [Accessed November 28, 2017b].

Argall, B.D. et al., 2009. A survey of robot learning from demonstration. *ROBOTICS AND AUTONOMOUS SYSTEMS*, 57(5), pp.469–483.

Avila-Robinson, A. & Miyazaki, K., 2011. Conceptualization and operationalization of emerging technologies: A complementing approach. In *2011 Proceedings of PICMET'11: Technology Management in the Energy Smart World*. pp. 1–12.

Bainbridge, W.A. et al., 2011. The benefits of interactions with physically present robots over video-displayed agents. *International Journal of Social Robotics*, 3(1), pp.41–52.

Balch, T. & Arkin, R.C., 1998. Behavior-based formation control for multirobot teams. *IEEE Transactions on Robotics and Automation*.

Ballantyne, G.H., 2002. Robotic surgery, telerobotic surgery, telepresence, and telementoring: Review of early clinical results. *Surgical Endoscopy and Other Interventional Techniques*.

Bechar, A. & Vigneault, C., 2016. Agricultural robots for field operations: Concepts and components. *Biosystems Engineering*, 149, pp.94–111. Available at: http://www.sciencedirect.com/science/article/pii/S1537511015301914 [Accessed April 15, 2017].

Beer, R.D., Chiel, H.J. & Drushel, R.F., 1999. Using autonomous robotics to teach science and

engineering. *Communications of the ACM*.
Berg, J.M., 2008. A Nobel lesson: the grant behind the prize. *Science (New York, N.Y.)*, 319(5865), pp.900–901.
Besl, P.J. & McKay, N.D., 1992. A Method for Registration of 3-D Shapes. *IEEE Trans. Pattern Anal. Mach. Intell. ()*.
Bilotta, E. et al., 2009. Edutainment robotics as learning tool. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*.
Blei, D., Carin, L. & Dunson, D., 2010. Probabilistic topic models. *IEEE Signal Processing Magazine*, 27, pp.55–65.
Bloch, C. et al., 2014. Developing a methodology to assess the impact of research grant funding: a mixed methods approach. *Evaluation and program planning*, 43, pp.105–117.
Blondel, V.D. et al., 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), p.P10008. Available at: http://arxiv.org/abs/0803.0476 [Accessed July 9, 2014].
Bobrow, J.E., Dubowsky, S. & Gibson, J.S., 1985. Time-Optimal Control of Robotic Manipulators Along Specified Paths. *The International Journal of Robotics Research*.
Borner, K. & Scharnhorst, A., 2009. Visual conceptualizations and models of science. *Journal of Informetrics*, 3(3), pp.161–172.
Borup, M. et al., 2006. The Sociology of Expectations in Science and Technology. *Technology Analysis & Strategic Management*, 18(September), pp.285–298.
Boyack, K.W. & Börner, K., 2003. Indicator-assisted evaluation and funding of research: Visualizing the influence of grants on the number and citation counts of research papers. *Journal of the American Society for Information Science and Technology*, 54(5), pp.447–461.
Boyack, K.W. & Jordan, P., 2011. Metrics associated with NIH funding: a high-level view. *Journal of the American Medical Informatics Association : JAMIA*, 18(4), pp.423–431.
Boyack, K.W., Klavans, R. & Borner, K., 2005. Mapping the backbone of science. *Scientometrics*, 64(3), pp.351–374.
Braga, R.A.M. et al., 2011. IntellWheels: Modular development platform for intelligent wheelchairs. *The Journal of Rehabilitation Research and Development*.
Braun, D., 1998. The role of funding agencies in the cognitive development of science. *Research Policy*, 27(8), pp.807–821.
Breazeal, C., 2003a. Emotion and sociable humanoid robots. *International Journal of Human Computer Studies*, 59, pp.119–155.
Breazeal, C., 2003b. Toward sociable robots. *Robotics and Autonomous Systems*, 42, pp.167–175.
Broadbent, E., Stafford, R. & MacDonald, B., 2009. Acceptance of healthcare robots for the older population: Review and future directions. *International Journal of Social Robotics*, 1, pp.319–330.
Brooks, R.A., 1991. Intelligence without representation. *Artificial Intelligence*, 47(1–3), pp.139–159. Available at: http://linkinghub.elsevier.com/retrieve/pii/000437029190053M [Accessed June 10, 2016].
Burke, J.L. et al., 2004. Moonlight in Miami: Field Study of Human-Robot Interaction in the Context of an Urban Search and Rescue Disaster Response Training Exercise. *Human-Computer Interaction*.
Calinon, S., Guenter, F. & Billard, A., 2007. On learning, representing, and generalizing a task in a humanoid robot. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*.
Caluwaerts, K. et al., 2014. Design and control of compliant tensegrity robots through simulation and hardware validation. *Journal of The Royal Society Interface*, 11(8), p.20140520.
Campbell, D. et al., 2017. Impact of multidisciplinary research on innovation. In *16TH*

INTERNATIONAL CONFERENCE ON SCIENTOMETRICS & INFORMETRICS.
Campbell, E.G. et al., 2000. Data withholding in academic medicine: characteristics of faculty denied access to research results and biomaterials. *Research Policy*, 29(2), pp.303–312. Available at: https://www.sciencedirect.com/science/article/pii/S0048733399000682 [Accessed February 24, 2018].
Candelas, F.A. et al., 2003. A Virtual Laboratory for Teaching Robotics. *International Journal of Engineering Education*.
Cannon, R.H. & Schmitz, E., 1984. Initial Experiments on the End-Point Control of a Flexible One-Link Robot. *The International Journal of Robotics Research*.
Capek, K., 1920. Rossum's Universal Robots. *Prague, CZ*.
Chang, W.L., Šabanović, S. & Huber, L., 2013. Use of seal-like robot PARO in sensory group therapy for older adults with dementia. In *ACM/IEEE International Conference on Human-Robot Interaction*. pp. 101–102.
Clauset, A., Moore, C. & Newman, M.E.J., 2008. Hierarchical structure and the prediction of missing links in networks. *Nature*, 453(May), pp.98–101.
Collins, S., 2005. Efficient Bipedal Robots Based on Passive-Dynamic Walkers. *Science*.
Colombo, G. et al., 2000. Treadmill training of paraplegic patients using a robotic orthosis. *Journal of rehabilitation research and development*.
Costanza, R., 2003. A vision of the future of science: Reintegrating the study of humans and the rest of nature. *Futures*, 35(6), pp.651–671.
Costas, R. & Leeuwen, T.N., 2012. Approaching the reward triangle: General analysis of the presence of funding acknowledgments and peer interactive communication in scientific publications. *Journal of the American Society for Information Science and Technology*, 63(8), pp.1647–1661.
Cozzens, S. et al., 2010. Emerging technologies: Quantitative identification and measurement. *Technology Analysis and Strategic Management*.
Cronin, B. & Overfelt, K., 1994. The Scholars Courtesy - a Survey of Acknowledgment Behavior. *Journal of Documentation*, 50(3), pp.165–196.
Cronin, B. & Shaw, D., 1999. Citation, funding acknowledgement and author nationality relationships in four information science journals. *Journal of Documentation*, 55(4), pp.402–408.
Cronin, B., Shaw, D. & La Barre, K., 2003. A cast of thousands: Coauthorship and subauthorship collaboration in the 20th century as manifested in the scholarly journal literature of psychology and philosophy. *Journal of the American Society for Information Science and Technology*, 54(9), pp.855–871.
Dahiya, R.S. et al., 2010. Tactile sensing-from humans to humanoids. *IEEE Transactions on Robotics*, 26(1), pp.1–20.
Dautenhahn, K., 2007. Socially intelligent robots: dimensions of human-robot interaction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 362(April), pp.679–704.
Dean, M., 2010. What is society? Social thought and the arts of government. *The British Journal of Sociology*, 61(4), pp.677–695. Available at: http://doi.wiley.com/10.1111/j.1468-4446.2010.01336.x [Accessed April 7, 2018].
Dedehayir, O. & Steinert, M., 2016. The hype cycle model: A review and future directions. *Technological Forecasting and Social Change*, 108, pp.28–41.
Deuten, J.J., Rip, A. & Jelsma, J., 1997. Societal embedding and product creation management. *Technology Analysis & Strategic Management*, 9(2), pp.131–148. Available at: http://dx.doi.org/10.1080/09537329708524275.
DiMaggio, P., Nag, M. & Blei, D., 2013. Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government

arts funding. *Poetics*, 41(6), pp.570–606. Available at: http://dx.doi.org/10.1016/j.poetic.2013.08.004.
Donovan, C. et al., 2014. Evaluation of the impact of National Breast Cancer Foundation-funded research. *The Medical journal of Australia*, 200(4), pp.214–218.
Duffy, B.R., 2003. Anthropomorphism and the social robot. *Robotics and Autonomous Systems,* 42, pp.177–190.
Engle, R.F. & Ng, V.K., 1993. Measuring and Testing the Impact of News on Volatility. *Journal of Finance*, 48(5), pp.1749–1778. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1993.tb05127.x/full.
Espiau, B., Chaumette, F. & Rives, P., 1992. A New Approach to Visual Servoing in Robotics. *IEEE Transactions on Robotics and Automation*.
European Comission, 2015. *Horizon: 30 years EU research framework programmes*,
Fathinezhad, F., Derhami, V. & Rezaeian, M., 2016. Supervised fuzzy reinforcement learning for robot navigation. *Applied Soft Computing*, 40, pp.33–41. Available at: http://www.sciencedirect.com/science/article/pii/S1568494615007486 [Accessed April 15, 2017].
Fenn, J. & Raskino, M., 2008. *Mastering the hype cycle: How to choose the right innovation in the right time* First. G. Inc., ed., Harvard Business Press.
Ficarra, V. et al., 2009. Retropubic, Laparoscopic, and Robot-Assisted Radical Prostatectomy: A Systematic Review and Cumulative Analysis of Comparative Studies. *European Urology*.
Fong, T., Nourbakhsh, I. & Dautenhahn, K., 2003. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42, pp.143–166.
Fortunato, S. & Barthélemy, M., 2007. Resolution limit in community detection. *Proceedings of the National Academy of Sciences of the United States of America*, 104(1), pp.36–41.
Fujita, K. et al., 2014. Detecting research fronts using different types of weighted citation networks. *Journal of Engineering and Technology Management - JET-M*, 32, pp.129–146. Available at: http://dx.doi.org/10.1016/j.jengtecman.2013.07.002.
Furukawa, T. et al., 2015. Identifying the evolutionary process of emerging technologies: A chronological network analysis of World Wide Web conference sessions. *Technological Forecasting and Social Change*, 91, pp.280–294. Available at: http://www.sciencedirect.com/science/article/pii/S004016251400105X.
Gabriele, L. et al., 2017. An educational robotics lab to investigate cognitive strategies and to foster learning in an arts and humanities course degree. *International Journal of Online Engineering*.
Gamini Dissanayake, M.W.M. et al., 2001. A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Transactions on Robotics and Automation*.
Garfield, E., Sher, I.H. & Torpie, R.J., 1964. *The use of citation data in writing the history of science,* Available at: http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=AD0466578.
Gates, B., 2007. A robot in every home. *Scientific American*, 296(1), pp.58–65.
Geels, F.W. & Verhees, B., 2011. Cultural legitimacy and framing struggles in innovation journeys: A cultural-performative perspective and a case study of Dutch nuclear energy (1945-1986). *Technological Forecasting and Social Change*, 78(6), pp.910–930. Available at: http://dx.doi.org/10.1016/j.techfore.2010.12.004.
Gibbons, M., 1999. Science's new social contract with society. *Nature,* 402(6761 Suppl), pp.C81-4.
Gibbons, M. et al., 2010. *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies,* 1 Oliver's Yard,~ 55 City Road,~ London~ ~ EC1Y 1SP~ United Kingdom: SAGE Publications Ltd.
Giles, C.L. & Councill, I.G., 2004. Who gets acknowledged: Measuring scientific contributions

through automatic acknowledgment indexing. *Proceedings of the National Academy of Sciences of the United States of America*, 101(51), pp.17599–17604. Available at: http://www.pnas.org/content/101/51/17599.full.

Gillet, R., 1991. Pitfalls in Assessing Research Performance by Grant Income. *Scientometrics*, 22(2), pp.253–263.

Giulianotti, P.C. et al., 2003. Robotics in general surgery: personal experience in a large community hospital. *Archives of surgery (Chicago, Ill. : 1960)*.

Glanzel, W. & Thijs, B., 2012. Using "core documents" for detecting and labelling new emerging topics. *Scientometrics*, 91(2), pp.399–416.

Gläser, J. & Velarde, K.S., 2018. Changing Funding Arrangements and the Production of Scientific Knowledge: Introduction to the Special Issue. *Minerva*, 56(1), pp.1–10.

Goeldner, M., Herstatt, C. & Tietze, F., 2015. The emergence of care robotics — A patent and publication analysis. *Technological Forecasting and Social Change*, 92, pp.115–131. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0040162514002753.

Gök, A., Rigby, J. & Shapira, P., 2016. The impact of research funding on scientific outputs: Evidence from six smaller European countries. *Journal of the Association for Information Science and Technology*, 67(3), pp.715–730.

Griffiths, T. & Steyvers, M., 2004. Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America*, 101 Suppl, pp.5228–5235.

Guinea, J. et al., 2015. Impact oriented monitoring: A new methodology for monitoring and evaluation of international public health research projects. *Research Evaluation*, 24(2), pp.131–145.

Haidegger, T. et al., 2013. Applied ontologies and standards for service robots. *Robotics and Autonomous Systems*, 61(11), pp.1215–1223. Available at: http://dx.doi.org/10.1016/j.robot.2013.05.008.

Halpern, B., 1969. The Robot and the Rabbit--A Pursuit Problem. *The American Mathematical Monthly*, 76(2), pp.140–145. Available at: http://www.jstor.org/stable/2317260.

Hegel, F. et al., 2009. Understanding Social Robots. *The Second International Conferences on Advances in Computer-Human Interactions (ACHI)*, (Section II), pp.169--174.

Hirai, K. et al., 1998. The development of Honda humanoid robot. *Robotics and Automation, 1998. Proceedings. 1998 IEEE International Conference on*.

Ho, J.C. et al., 2014. Technological barriers and research trends in fuel cell technologies: A citation network analysis. *Technological Forecasting and Social Change*, 82(1), pp.66–79.

Hörlesberger, M. et al., 2013. A concept for inferring `frontier research' in grant proposals. *Scientometrics*, 97(2), pp.129–148.

Hosotsubo, M. & Nishii, R., 2016. Relation between awarding of Grants-in-aid for scientific research and characteristics of applicants in Japanese universities. *Scientometrics*, 109(2), pp.1097–1116.

Hu, Z., Fang, S. & Liang, T., 2014. Empirical study of constructing a knowledge organization system of patent documents using topic modeling. *Scientometrics*, 100(19), pp.787–799.

Hutchinson, S. a et al., 1996. A tutorial on visual servo control. *IEEE Transactions on Robotics and Automation*.

IBM, 2015. Sentiment analysis with AlchemyAPI: A hybrid approach. *IBM Watson*. Available at: http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=LBW03035USEN.

Ijspeert, A.J. et al., 2013. Dynamical Movement Primitives: Learning Attractor Models for Motor Behaviors. *Neural Computation*.

Ilievski, F. et al., 2011. Soft robotics for chemists. *Angewandte Chemie - International Edition*.

International Federation of Robotics, 2016a. *Executive Summary. World robotics 2016 Industrial robots*, Available at: https://ifr.org/img/uploads/Executive_Summary_WR_Industrial_Robots_20161.pdf.

International Federation of Robotics, 2016b. *Executive Summary. World robotics 2016 Service Robotics*,
Ittipanuvat, V. et al., 2014. Finding linkage between technology and social issue: A Literature Based Discovery approach. *Journal of Engineering and Technology Management - JET-M*, 32, pp.160–184. Available at: http://dx.doi.org/10.1016/j.jengtecman.2013.05.006.
Jahn, T., Bergmann, M. & Keil, F., 2012. Transdisciplinarity: Between mainstreaming and marginalization. *Ecological Economics*, 79, pp.1–10.
Janssen, S., Kuipers, G. & Verboord, M., 2008. Cultural Globalization and Arts Journalism: The International Orientation of Arts and Culture Coverage in Dutch, French, German, and U.S. Newspapers, 1955 to 2005. *American Sociological Review*, 73(5), pp.719–740. Available at: http://asr.sagepub.com/content/73/5/719%5Cnhttp://asr.sagepub.com/content/73/5/719.abstract?patientinform-links=yes&legid=spasr;73/5/719.
Jara, C.A. et al., 2011. Hands-on experiences of undergraduate students in Automatics and Robotics using a virtual and remote laboratory. *Computers and Education*.
Joshi, V.A., Banavar, R.N. & Hippalgaonkar, R., 2010. Design and analysis of a spherical mobile robot. *Mechanism and Machine Theory*, 45(2), pp.130–136.
Jun, S.P., 2012. A comparative study of hype cycles among actors within the socio-technical system: With a focus on the case study of hybrid cars. *Technological Forecasting and Social Change*, 79(8), pp.1413–1430. Available at: http://dx.doi.org/10.1016/j.techfore.2012.04.019.
Jutel, O., 2016. The liberal field of journalism and the political - The New York Times, Fox News and the Tea Party. *Journalism: Theory, Practice & Criticism*, 17(8), pp.1129–1145. Available at: http://journals.sagepub.com/doi/10.1177/1464884915599949 [Accessed April 7, 2018].
Kajikawa, Y. et al., 2012. Academic Landscape of Innovation Research and National Innovation System Policy Reformation in Japan and the United States. *International Journal of Innovation & Technology Management*, 9(6), p.1. Available at: 10.1142/S0219877012500447%5Cnhttp://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=86463487&site=ehost-live.
Kajikawa, Y., Tacoa, F. & Yamaguchi, K., 2014. Sustainability science: the changing landscape of sustainability research. *Sustainability Science*, pp.431–438.
Kanda, T. et al., 2004. Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. *Human-Computer Interaction*, 19, pp.61–84.
Kang, S.W. et al., 2009. Robotic thyroid surgery using a gasless, transaxillary approach and the da Vinci S system: The operative outcomes of 338 consecutive patients. *Surgery*, 146(6), pp.1048–1055.
Kavraki, L.E. et al., 1996. Probabilistic Roadmaps for Path Planning in High Dimensional Configuration Spaces. *IEEE Transactions on Robotics and Automation*.
Kessler, M., 1963. An experimental study of bibliographic coupling between technical papers (Corresp.). *IEEE Transactions on Information Theory*, 9(1), pp.49–51. Available at: http://apps.webofknowledge.com/full_record.do?product=WOS&search_mode=GeneralSearch&qid=1&SID=P1q7ppFQ7Uv6dZHWaYz&page=1&doc=1 [Accessed December 17, 2015].
Khatib, O., 1986. Real-Time Obstacle Avoidance for Manipulators and Mobile Robots. *The International Journal of Robotics Research*.
Kim, S., Laschi, C. & Trimmer, B., 2013. Soft robotics: A bioinspired evolution in robotics. *Trends in Biotechnology*.
Klavans, R. & Boyack, K.W., 2017. Which Type of Citation Analysis Generates the Most Accurate Taxonomy of Scientific and Technical Knowledge? *Journal of the Association for Information Science and Technology*, 68(4), pp.984–998. Available at: http://arxiv.org/abs/1511.05078.

Klein, H.K. & Kleinman, D.L., 2002. The Social Construction of Technology: Structural Considerations. *Science, Technology, & Human Values*, 27(1), pp.28–52. Available at: http://journals.sagepub.com/doi/10.1177/016224390202700102 [Accessed November 30, 2017].

Klein, J.T., 2015. Reprint of "Discourses of transdisciplinarity: Looking back to the future." *Futures*.

Kobayashi, Y. et al., 2015. Formations for Facilitating Communication Among Robotic Wheelchair Users and Companions. In A. Tapus et al., eds. *Social Robotics: 7th International Conference, ICSR 2015, Paris, France, October 26-30, 2015, Proceedings*. Cham: Springer International Publishing, pp. 370–379. Available at: https://doi.org/10.1007/978-3-319-25554-5_37.

Kostoff, R.N., 1993. Semiquantitative methods for research impact assessment. *Technological Forecasting and Social Change*, 44(3), pp.231–244.

Kozima, H., Michalowski, M.P. & Nakagawa, C., 2009. Keepon: A playful robot for research, therapy, and entertainment. *International Journal of Social Robotics*, 1, pp.3–18.

Kumaresan, N. & Miyazaki, K., 1999. An integrated network approach to systems of innovation—the case of robotics in Japan. *Research Policy*, 28(6), pp.563–585. Available at: http://www.sciencedirect.com/science/article/pii/S0048733398001280%5Cnhttp://linkinghub.elsevier.com/retrieve/pii/S0048733398001280.

Kwakkel, G., Kollen, B. & Krebs, H., 2008. Effects of robot-assisted therapy on upper limb recovery after stroke: a systematic review. *Neurorehabilitation and Neural Repair*.

Lam, A., 2010. From "Ivory Tower Traditionalists" to "Entrepreneurial Scientists"? *Social Studies of Science*, 40(2), pp.307–340. Available at: http://journals.sagepub.com/doi/10.1177/0306312709349963 [Accessed February 24, 2018].

Lane, J., 2009. Science innovation. Assessing the impact of science funding. *Science (New York, N.Y.)*, 324(5932), pp.1273–1275.

Laryionava, K. & Gross, D., 2012. Deus ex machina or e-slave? Public perception of healthcare robotics in the German print media. *International journal of technology assessment in health care*, 28(3), pp.265–70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22980703.

Lauto, G. & Sengoku, S., 2015. Perceived incentives to transdisciplinarity in a Japanese university research center. *Futures*.

Lawrence, R.J., 2015. Advances in transdisciplinarity: Epistemologies, methodologies and processes. *Futures*, 65, pp.1–9.

Lee, M.M.. & Nicholls, H.H.., 1999. Tactile sensing for mechatronics—a state of the art survey. *Mechatronics*, 9(1), pp.1–31. Available at: http://www.sciencedirect.com/science/article/pii/S0957415898000452 [Accessed September 20, 2017].

Leite, I., Martinho, C. & Paiva, A., 2013. Social Robots for Long-Term Interaction: A Survey. *International Journal of Social Robotics*, 5(2), pp.291–308.

Van Lente, H., Spitters, C. & Peine, A., 2013. Comparing technological hype cycles: Towards a theory. *Technological Forecasting and Social Change*, 80(8), pp.1615–1628.

Lepori, B., 2011. Coordination modes in public funding systems. *Research Policy*, 40(3), pp.355–367.

Lewes, G.H., 1874. *The Problems of Life and Mind*,

Lewison, G. & Dawson, G., 1998. The effect of funding on the outputs of biomedical research. *Scientometrics*, 41(1–2), pp.17–27.

Lewison, G. & Markusova, V., 2010. The evaluation of Russian cancer research. *Research Evaluation*, 19(2), pp.129–144.

Leydesdorff, L. & Rafols, I., 2011. Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of Informetrics*, 5(1), pp.87–100. Available at:

http://dx.doi.org/10.1016/j.joi.2010.09.002.
Liu, B., 2010. Sentiment Analysis and Subjectivity. *Handbook of Natural Language Processing*, pp.1–38. Available at: http://www.cs.uic.edu/~liub/FBS/NLP-handbook-sentiment-analysis.pdf%5Cnhttp://people.sabanciuniv.edu/berrin/proj102/1-BLiu-Sentiment Analysis and Subjectivity-NLPHandbook-2010.pdf.
Lo, A.C. et al., 2010. Robot-Assisted Therapy for Long-Term Upper-Limb Impairment after Stroke. *New England Journal of Medicine*, 362(19), pp.1772–1783.
Lok, C., 2010. Science funding: Science for the masses. *Nature*, 465(7297), pp.416–418.
Lucio-Arias, D. & Leydesdorff, L., 2009. An indicator of research front activity: Measuring intellectual organization as uncertainty reduction in document sets. *Journal of the American Society for Information Science and Technology*.
Luk, B.L. et al., 2005. Intelligent legged climbing service robot for remote maintenance applications in hazardous environments. *Robotics and Autonomous Systems*, 53(2), pp.142–152.
Luk, B.L. et al., 2006. Tele-operated climbing and mobile service robots for remote inspection and maintenance in nuclear industry. *The Industrial Robot*, 33(3), pp.194–204. Available at: http://search.proquest.com/docview/216994626?accountid=15300%5Cnhttp://sfx.cbuc.cat/upc?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&genre=article&sid=ProQ:ProQ%3Aengineeringjournals&atitle=Tele-operated+climbing+and+mobile+service+robots+for.
Lum, P.S. et al., 2002. Robot-assisted movement training compared with conventional therapy techniques for the rehabilitation of upper-limb motor function after stroke. *Archives of Physical Medicine and Rehabilitation*, 83(7), pp.952–9.
Lyall, C. et al., 2013. The role of funding agencies in creating interdisciplinary knowledge. *Science and Public Policy*, 40(1), pp.62–71.
Marescaux, J. et al., 2001. Transatlantic robot-assisted telesurgery. *Nature*.
Martin, B.R., 1995. Foresight in science and technology. *Technology Analysis & Strategic Management*, 7(2), pp.139–168. Available at: http://www.tandfonline.com/doi/abs/10.1080/09537329508524202.
McLurkin, J. et al., 2013. Using multi-robot systems for engineering education: Teaching and outreach with large numbers of an advanced, low-cost robot. *IEEE Transactions on Education*.
Medhat, W., Hassan, A. & Korashy, H., 2014. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), pp.1093–1113. Available at: http://dx.doi.org/10.1016/j.asej.2014.04.011.
Melton, N., Axsen, J. & Sperling, D., 2016. Moving beyond alternative fuel hype to decarbonized transportation. *Nature Energy*, 1(February), pp.1–29.
Meppem, T., 2000. The discursive community: Evolving institutional structures for planning sustainability. *Ecological Economics*, 34(1), pp.47–61.
Ministry of Economy Trade and Industry of Japan, 2015. *New Robot Strategy*, Available at: http://www.meti.go.jp/english/press/2015/pdf/0123_01b.pdf.
Mobjörk, M., 2010. Consulting versus participatory transdisciplinarity: A refined classification of transdisciplinary research. *Futures*, 42(8), pp.866–873.
Mohr, J.W. & Bogdanov, P., 2013. Introduction-Topic models: What they are and why they matter. *Poetics*, 41(6), pp.545–569.
Moraes, R., Valiati, J.F. & Gavião Neto, W.P., 2013. Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications*, 40(2), pp.621–633.
Muller, R., 1980. Innovation and Scientific Funding. *Science*, 209, pp.880–883.
Murphy, R.R., 2014. *Disaster Robotics*,

Muscio, A., Quaglione, D. & Vallanti, G., 2013. Does government funding complement or substitute private research funding to universities? *Research Policy*.

Musselin, C., 2014. Empowerment of French Universities by Funding and Evaluation Agencies. In pp. 51–76. Available at: http://www.emeraldinsight.com/doi/10.1108/S0733-558X20140000042002 [Accessed February 24, 2018].

Naito, O. et al., 2013. Actual State of the Robotic Study Changed by the Implementation in the Real World. *Journal of the Robotics Society of Japan*, 31(8), pp.804–815. Available at: http://jlc.jst.go.jp/DN/JST.JSTAGE/jrsj/31.804?lang=en&from=CrossRef&type=abstract.

Nakano, R., 2012. *Beyond the Western liberal order : Yanaihara Tadao and empire as society*, Palgrave Macmillan. Available at: http://apps.webofknowledge.com/full_record.do?product=WOS&search_mode=GeneralSearch&qid=14&SID=D1SodTAvB7GDhURjQnu&page=1&doc=3&cacheurlFromRightClick=no [Accessed April 7, 2018].

National Science Foundation, 2016. *A roadmap for US robotics*, Available at: http://jacobsschool.ucsd.edu/contextualrobotics/docs/rm3-final-rs.pdf.

Nelson, B.J., Kaliakatsos, I.K. & Abbott, J.J., 2010. Microrobots for Minimally Invasive Medicine. *Annual Review of Biomedical Engineering*, 12, pp.55–85.

Nelson, R.R. & Winter, S.G., 1982. *An Evolutionary Theory of Economic Change,*

Newman, D. et al., 2006. Analyzing entities and topics in news articles using statistical topic models. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3975 LNCS, pp.93–104.

Nowotny, H., 2005. Science and society. High- and low-cost realities for science and society. *Science (New York, N.Y.)*, 308(5725), pp.1117–1118.

Nowotny, H., 1999. The place of people in our knowledge. *European Review*, 7(2), pp.247–262. Available at: https://www.cambridge.org/core/product/identifier/S1062798700004026/type/journal_article [Accessed November 28, 2017].

Nowotny, H., 2015. The radical openness of science and innovation: Why uncertainty is inherent in the openness towards the future. *EMBO reports*, 16(12), pp.1601–1604.

Nowotny, H., Scott, P. & Gibbons, M., 2001. *Re-thinking science: Knowledge and the public in an age of uncertainty / Helga Nowotny, Peter Scott, Michael Gibbons*, Cambridge: Polity Press.

O'Malley, B.W. et al., 2006. Transoral Robotic Surgery (TORS) for Base of Tongue Neoplasms. *The Laryngoscope*.

Olson, E.M. et al., 2001. Patterns of cooperation during new product development among marketing, operations and R&D: Implications for project performance. *Journal of Product Innovation Management*, 18(4), pp.258–271. Available at: http://doi.wiley.com/10.1111/1540-5885.1840258 [Accessed November 28, 2017].

Parameshwaran, R., Praveen Kumar, S. & Saravanakumar, K., 2015. An integrated fuzzy MCDM based approach for robot selection considering objective and subjective criteria. *Applied Soft Computing Journal*, 26, pp.31–41.

Paul-Hus, A., Desrochers, N. & Costas, R., 2016. Characterization, description, and considerations for the use of funding acknowledgement data in Web of Science. *Scientometrics*, 108(1), pp.167–182.

Persson, O., 2010. Identifying research themes with weighted direct citation links. *Journal of Informetrics*, 4(3), pp.415–422. Available at: http://www.sciencedirect.com/science/article/pii/S1751157710000283%5Cnhttp://www.sciencedirect.com/science?_ob=MiamiImageURL&_cid=273584&_user=3425974&_pii=S1751157710000283&_check=y&_origin=article&_zone=toolbar&_coverDate=31-Jul-2010&view=c&originContentF.

Petersen, A., 2001. Biofantasies: Genetics and medicine in the print news media. *Social Science and Medicine*, 52(8), pp.1255–1268.
Porter, A.L. et al., 2002. Measuring national “emerging technology” capabilities. *Science and Public Policy*.
Radicchi, F., Fortunato, S. & Vespignani, A., 2012. Citation networks. *Understanding Complex Systems*, pp.233–257.
Rangel, S.J., Efron, B. & Moss, R.L., 2002. Recent Trends in National Institutes of Health Funding of Surgical Research. *Annals of Surgery*, 236(3), pp.277–287.
Rigby, J., 2011. Systematic grant and funding body acknowledgement data for publications: New dimensions and new controversies for research policy and evaluation. *Research Evaluation*, 20(5), pp.365–375.
Rosen-Zvi, M. et al., 2004. The author-topic model for authors and documents. *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, pp.487–494. Available at: http://portal.acm.org/citation.cfm?id=1036902.
Rosendahl, J. et al., 2015. Scientists’ situated knowledge: Strong objectivity in transdisciplinarity. *Futures*.
Rotolo, D., Hicks, D. & Martin, B.R., 2015. What is an emerging technology? *SPRU Working Paper Series*, 2015–6(10), pp.1–40. Available at: http://dx.doi.org/10.1016/j.respol.2015.06.006.
Ruef, A. & Markard, J., 2010. What happens after a hype? How changing expectations affected innovation activities in the case of stationary fuel cells. *Technology Analysis & Strategic Management*, 22(3), pp.317–338.
Rusu, R.B., Gerkey, B. & Beetz, M., 2008. Robots in the kitchen: Exploiting ubiquitous sensing and actuation. *Robotics and Autonomous Systems*, 56(10), pp.844–856.
Sakata, I. et al., 2013. Bibliometric analysis of service innovation research: Identifying knowledge domain and global network of knowledge. *Technological Forecasting and Social Change*, 80(6), pp.1085–1093. Available at: http://dx.doi.org/10.1016/j.techfore.2012.03.009.
Salter, A.J. & Martin, B.R., 2001. The economic benefits of publicly funded basic research: A critical review. *Research Policy*, 30(3), pp.509–532.
Sanders, D., 2009. Comparing speed to complete progressively more difficult mobile robot paths between human tele-operators and humans with sensor-systems to assist. *Assembly Automation*, 29(3), pp.230–248. Available at: http://apps.webofknowledge.com/full_record.do?product=WOS&search_mode=GeneralSearch&qid=1&SID=P2bIWgHMqlk79rH5ZPV&page=1&doc=1 [Accessed January 30, 2016].
Sanders, D. et al., 2011. Simple expert systems to improve an ultrasonic sensor-system for a tele-operated mobile-robot. *Sensor Review*, 31(3), pp.246–260.
Saranli, U., Buehler, M. & Koditschek, D.E., 2001. RHex: A Simple and Highly Mobile Hexapod Robot. *The International Journal of Robotics Research*.
Schaal, S., 1999. Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences*.
Schulz, D. et al., 2000. Web interfaces for mobile robots in public places. *Robotics Automation Magazine, IEEE*, 7(1), pp.48–56.
Serrano-Guerrero, J. et al., 2015. Sentiment analysis: A review and comparative analysis of web services. *Information Sciences*, 311, pp.18–38. Available at: http://dx.doi.org/10.1016/j.ins.2015.03.040.
Shen, C.-C. et al., 2016. Research impact of general and funded papers. *Online Information Review*, 40(4), pp.472–480.
Shepherd, R.F. et al., 2011. Multigait soft robot. *Proceedings of the National Academy of Sciences*.
Shibata, N. et al., 2009. Comparative Study on Methods of Detecting Research Fronts Using DifferentTypes of Citation. *Journal of the American Society for Information Science and Technology*, 60(3), pp.571–580.

Shibata, N. et al., 2008. Detecting emerging research fronts based on topological measures in citation networks of scientific publications. *Technovation*, 28(11), pp.758–775.

Shibata, N. et al., 2011. Detecting emerging research fronts in regenerative medicine by the citation network analysis of scientific publications. *Technological Forecasting and Social Change*, 78(2), pp.274–282. Available at: http://dx.doi.org/10.1016/j.techfore.2010.07.006.

Shimoga, K.B., 1996. Robot Grasp Synthesis Algorithms: A Survey. *The International Journal of Robotics Research*, 15(3), pp.230–266. Available at: http://journals.sagepub.com/doi/10.1177/027836499601500302 [Accessed November 21, 2017].

Sievert, C. & Shirley, K., 2014. LDAvis: A method for visualizing and interpreting topics. *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, pp.63–70. Available at: http://www.aclweb.org/anthology/W/W14/W14-3110.

Sismondo, S., 2009. Ghosts in the Machine: Publication Planning in the Medical Sciences. *Social Studies of Science*, 39(2), pp.171–198. Available at: http://journals.sagepub.com/doi/10.1177/0306312708101047 [Accessed February 24, 2018].

Slotine, J.-J.E. & Weiping Li, 1987. On the Adaptive Control of Robot Manipulators. *The International Journal of Robotics Research*.

Small, H., 1973. Co-citation in the Scientific Literature: A New Measure of the Relationship between Two Documents. *Journal of the American Society for Information Science*, 24(4), pp.265–269. Available at: http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=415edd28-79a4-4fa3-8b2c-cd145ef82a4e%40sessionmgr114&vid=4&hid=127.

de Solla Price, D.J., 1965. Networks of scientific papers. *Science*, 149(3683), p.510. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-0000432228&partnerID=tZOtx3y1.

SPARC, 2016. Robotics 2020 Multi-Annual Roadmap For Robotics in Europe. *SPARK The partnership for Robotics in Europe and the European Commission*. Available at: https://eu-robotics.net/sparc//wp-content/uploads/2014/05/H2020-Robotics-Multi-Annual-Roadmap-ICT-2016.pdf [Accessed March 30, 2017].

Sparrow, R., 2009. Predators or plowshares?: Arms control of robotic weapons. *IEEE Technology and Society Magazine*.

Šubelj, L., van Eck, N.J. & Waltman, L., 2015. Clustering scientific publications based on citation relations: A systematic comparison of different methods. *PLoS ONE*, pp.1–24. Available at: http://arxiv.org/abs/1512.09023.

Taboada, M. et al., 2011. Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics*, 37(2), pp.267–307.

Takano, Y., Kajikawa, Y. & Ando, M., 2017. Trends and typology of emerging antenna propagation technologies: Citation network analysis. *International Journal of Innovation and Technology Management*, 14(1), pp.2872–2881. Available at: http://www.worldscientific.com/doi/abs/10.1142/S0219877017400053.

Takano, Y., Mejia, C. & Kajikawa, Y., 2016. Dynamics of the research classification schema across Technologies: Case study of IoT-related technologies. In Y. Fei, ed. *The First International Conference of Innovation Studies*. Beijing, China: Tsinghua University, p. 15.

Takeda, Y. et al., 2009. Nanobiotechnology as an emerging research domain from nanotechnology: A bibliometric approach. *SCIENTOMETRICS*, 80(1), pp.23–38.

Tedeschi, F. & Carbone, G., 2015. Hexapod walking robot locomotion. In G. Carbone & F. Gomez-Bravo, eds. *Mechanisms and Machine Science*. Springer, pp. 439–468.

Tewari, A., Srivasatava, A. & Menon, M., 2003. A prospective comparison of radical retropubic and robot-assisted prostatectomy: Experience in one institution. *BJU International*.

The New York Times, 2016. Times Developer Network - Article Search API v2. Available at: http://developer.nytimes.com/docs/read/article_search_api_v2 [Accessed February 2, 2016].

Thelwall, M. et al., 2016. Alternative metric indicators for funding scheme evaluations. *Aslib Journal of Information Management*, 68(1), pp.2–18.

Thrun, S., 2001. Robust Monte Carlo localization for mobile robots. *Artificial Intelligence*.

Tress, G., Tress, B. & Fry, G., 2005. Clarifying Integrative Research Concepts in Landscape Ecology. *Landscape Ecology*, 20(4), pp.479–493. Available at: http://link.springer.com/10.1007/s10980-004-3290-4 [Accessed November 28, 2017].

Truffer, B., Voß, J.P. & Konrad, K., 2008. Mapping expectations for system transformations. Lessons from Sustainability Foresight in German utility sectors. *Technological Forecasting and Social Change*, 75(9), pp.1360–1372.

Urcola, P. & Montano, L., 2011. Adapting robot team behavior from interaction with a group of people. In *IEEE International Conference on Intelligent Robots and Systems*.

Wallach, H.M. et al., 2009. Evaluation methods for topic models. *Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09*, (4), pp.1–8. Available at: http://dl.acm.org/citation.cfm?id=1553374.1553515%5Cnhttps://www.era.lib.ed.ac.uk/bitstream/1842/4587/1/MurrayI_Evaluation Methods for.pdf.

Walter, A.I. et al., 2007. Measuring societal effects of transdisciplinary research projects: Design and application of an evaluation method. *Evaluation and Program Planning*, 30(4), pp.325–338. Available at: http://www.sciencedirect.com/science/article/pii/S0149718907000754?via%3Dihub [Accessed November 28, 2017].

Wang, J. & Shapira, P., 2011. Funding acknowledgement analysis: An enhanced tool to investigate research sponsorship impacts: the case of nanotechnology. *Scientometrics*, 87(3), pp.563–586.

Wang, J. & Shapira, P., 2015. Is there a relationship between research sponsorship and publication impact? An analysis of funding acknowledgments in nanotechnology papers. *PloS one*, 10(2), p.e0117727.

Web of Science, 2008. Funding Acknowledgements. (Online). *Clarivate Analytics*. Available at: http://wokinfo.com/products_tools/multidisciplinary/webofscience/fundingsearch/ [Accessed March 30, 2017].

Weinstein, G.S. et al., 2007. Transoral Robotic Surgery. *Archives of Otolaryngology–Head & Neck Surgery*.

Weinstein, G.S. et al., 2007. Transoral robotic surgery: supraglottic partial laryngectomy. *Ann Otol Rhinol Laryngol*.

Whitley, R., Gläser, J. & Engwall, L., 2010. *Reconfiguring knowledge production: Changing authority relationships in the sciences and their consequences for intellectual innovation*,

Wolcott, H.N. et al., 2016. Modeling time-dependent and -independent indicators to facilitate identification of breakthrough research papers. *Scientometrics*, 107(2), pp.807–817.

Woodburry, D., 1927. Dramatizing the "robot." *The New York Times*, p.87. Available at: https://timesmachine.nytimes.com/timesmachine/1927/11/06/95018768.html?pageNumber=81.

Woodman, R. et al., 2012. Building safer robots: Safety driven control. *The International Journal of Robotics Research*, 31(13), pp.1603–1626. Available at: http://ijr.sagepub.com/content/31/13/1603.short.

Yan, E., 2014. Research dynamics: Measuring the continuity and popularity of research topics. *Journal of Informetrics*, 8(1), pp.98–110.

Yegros-Yegros, A. & Costas, R., 2013. Analysis of the web of science funding acknowledgement information for the design of indicators on "external funding attraction." In J. Gorraiz, ed. *The 14th International Society of Scientometrics and Informetrics Conference*. Viena, Austria, pp. 84–95. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-84896874684&partnerID=40&md5=4f327d10e423a71fa0688fc1e04b6788.

Yim, M. et al., 2007. Modular self-reconfigurable robot systems - Challenges and opportunities for the future. *IEEE ROBOTICS & AUTOMATION MAGAZINE*, 14(1), pp.43–52.
Yoshikawa, T., 1985. Manipulability of Robotic Mechanisms. *The International Journal of Robotics Research*.
Zhao, D., 2010. Characteristics and impact of grant-funded research: A case study of the library and information science. *Scientometrics*, 84(2), pp.293–306.
Zoss, A.M. & Börner, K., 2012. Mapping interactions within the evolving science of science and innovation policy community. *Scientometrics*, 91(2), pp.631–644.

7 APPENDICES

7.1 Selected news for hypes and topics

Hype 1	"U.S. robot sales rose 71% in 1984" (1985-06-04) "sales rise for robots" (1984-06-27)
Hype 2	"Honda's Robot Advances, but to Where?" (2004-12-16) "A Robot For the Masses" (2004-11-28)
Hype 3	"Smart drones" (2013-03-17) "researchers put sense of touch in reach for robots" (2013-04-28)
Topic 5	"A Blow to NASA's Hubble Rescue" (2004-12-12) "Phoenix Lander Is Ready for Risky Descent to Mars" (2008-05-20)
Topic 13	"A Woodstock for Robots" (2013-12-22) "Bits Pics: Robot Competitions" (2010-07-21)
Topic 18	"When Robotic Surgery Leaves Just a Scratch" (2012-11-18) "Prepping Robots to Perform Surgery" (2008-05-04)

7.2 Highly ranked robotics research

7.2.1 Robotic research with high emergence score

The highest clusters of outlying nature, are described as exemplary of emergent technologies within the robotics domain.

Soft robotics

This sub-clusters gathers research on bio-inspired robotics. The materials that compose the elements of soft-robots are flexible or malleable, usually elastopolymers. Soft-robots are able to reach fluid motion and are deemed suitable for unstructured, unpredictable environments (Shepherd et al. 2011; Ilievski et al. 2011; Kim et al. 2013). This cluster is composed of 672 articles, being 2014 the average publication year. It belongs to the robot locomotion cluster (ID: 2-3-1-1).

Robotics as learning tool

This sub-cluster is about application of robotics in the classroom. The mission of these researchers is to develop robotics and programming skills in children by hands-on experience with robotic systems (Bilotta et al. 2009; Gabriele et al. 2017). This cluster contains 11 articles. It is part of the autonomous robot cluster (ID:3-7-3-7) Being the cluster with the highest coherence score.

Intelligent wheelchairs

Research in this cluster aims to improve the conditions for elderly care. These articles study mechanisms to design smart wheelchairs tackling on the issue of modular systems, and interconnectivity (Braga et al. 2011; Urcola & Montano 2011; Kobayashi et al. 2015). This is a small cluster of 10 articles, having the second largest coherence score (ID: 3-8-3-5).

Modeling robot behavior

In this sub-cluster it is found research for the design and robot learning of complex behaviours. Methods like non-linear dynamic system and reinforcement learning are introduce for the application in robotics. The objective of this research deviate from the hard-coded behaviours. Instead, a set of primitive instructions is given and then robots learn by imitation and repetition. (Calinon et al. 2007; Ijspeert et al. 2013; Andrew Bagnell 2014). This cluster contains 807 articles, published in 2012 on average (ID: 2-1-1-2)

7.2.2 Robotic research with high social attention

Virtual laboratories (ID: 3-3-4-3)

Similar to top emergent cluster previously identified on robots as learning tool, this sub-cluster also address the issue of better prepare new generations to tackle on the study in robotics. In this case, when robot equipment is not available in house. Research in this cluster is about design and case studies on experiences in virtual laboratories, or tele operated laboratories (Jara et al. 2011; Candelas et al. 2003).

Autonomous robots for teaching science (ID: 3-7-6-0)

In the same line of educational robotics. This cluster explores cases on hand-on learning orientated to higher level education (Beer et al. 1999; McLurkin et al. 2013).

7.2.3 Robotic research with high solution potentiality

Rescue robotics (ID: 3-2-2-2)

The development of rescue robots, that have to explore difficult environments, and interact with people in extreme situations is studied in this cluster (Burke et al. 2004; Murphy 2014).

Social robots (ID: 3-2-1-2)

This sub cluster explores theory and advancements on human factors for the design of socially assistive robotics. Particularly robots as partner. Both the manifestation of personality in robots, and views on the human user are studied to design robots that enter society smoother (Breazeal 2003b; Dautenhahn 2007; Duffy 2003).

7.3 Robotics and social issues

In this chapter we study the feasibility of robotics research to be linked to social issues. In a similar approach to chapter 3, clusters from a citation network are compared to the topic model of news articles. However, this time the dataset differs. For this case study, we use the sub-set of articles about social robotics; and from news articles, the case of news about Japan are analysed.

To extent the benefits of datamining to the analysis of society problems, there is the necessity to find the unit of measure to map across different social issues in different countries with the same versatility of papers in science and technology. When applying a topic model in a collection of articles about a country or region, the model outputs the different topics that have received attention in the press, some of them are expected to contain social issues.

Topic models and citation networks are well developed methodologies, and we want contribute in this stream of research by bringing them together for the use of knowledge discovery by exploiting the text similarity of topics and networks. Concretely, this chapter explores the possibility of finding solutions to social issues by exploiting the semantic relations between topic models from newspapers articles and citation networks of academic publications. The former representing society, and the later technology. The proposed methodology may help to scope over the large, and ever increasing, amount of technology research and extract from them plausible solutions to concerns in society.

We show our methodology by linking the topic model of newspaper articles about Japan to the network of academic papers on social robotics. Social robots are interactive partners that perform a social role (Hegel et al. 2009). They are also categorized as service robots, having as main characteristics a high level of complexity and autonomy (Haidegger et al. 2013). The idea of linking robotics to social issues is not new, Andrade et al (Andrade et al. 2014) reviewed robotic technologies for health care, and Ittipanuvat et al (Ittipanuvat et al. 2014) for the elderly society. Our research contributes on these explorations by assessing multiple social issues at once according to the necessities of the selected country.

The rest of the chapter is as follows. We describe the process of creating the topic model and citation network, and how to establish semantic connections between them. The methodology is then applied to newspaper articles on Japan, and academic articles of robotics and their connections are analyzed to figure out the robotics technologies that may bring a solution to a social concern. We conclude the chapter with a brief summary of the results and suggestions for future work.

7.3.1 Data and methods

The overview of this research is presented in Figure 7-1. It consists of 2 stages of data processing where social issues are extracted from newspaper articles, and robotic technologies are obtained from journal articles respectively. Finally, the connections between social issues and robotics are established by text similarity. Detailed explanation is provided in the following sections.

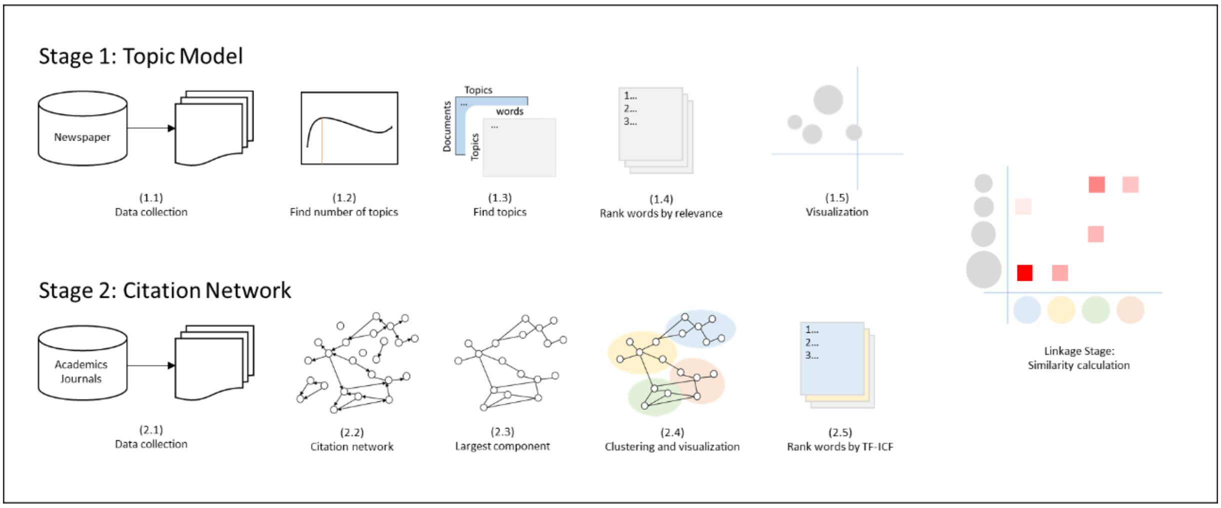

Figure 7-1 Overview of methodology in chapter 4

7.3.1.1 Topic modeling social issues

The first step is the data collection (1.1). The dataset of news articles related to a target country is extracted from a newspaper database. Next, a cleaning process is conducted on this dataset by removing stopwords and infrequent terms that appeared 5 times or less in the text corpus.

Once the news are obtained and cleaned, topic model is conducted to uncover the underlying themes in the corpus. Topic models are algorithms that analyze and classify the words in a corpus of unstructured text, like news articles, to find the underlying themes present in the set of documents (Blei et al. 2010). This themes, or topics, can be described as sets of words that co-occur repetitively across the corpus. More precisely, a topic is a multinomial distribution over the total vocabulary found in all the documents (Newman et al. 2006).

Similarly, documents are represented as multinomial distribution over topics. Simply put, documents are made of topics, and topics are made of words. In practice, the output of those algorithms are the document-topic matrix and the topic-word matrix filled with their respective conditional probabilities. Having that information, we are able to calculate the probability of a word w given a document d as follows:

$$\boldsymbol{p(w|d) = \sum_{t=1}^{T} p(w|t)p(t|d).} \qquad \text{Eq. 7-1}$$

Where t stands for each of all possible T topics.

In fact, the news data we collected describes the left side of the equation, and we have to get the topics on the right side. There are many algorithms to achieve this, in this chapter we use Gibbs sampling which is a Markov chain Monte Carlo algorithm that iterates over all words in the corpus to classify them into the desired T number of topics (Griffiths & Steyvers 2004). The Gibbs sampling algorithm is broadly used in topic modelling and has been applied on patents, journal articles, and newspaper corpus (Hu et al. 2014; Rosen-Zvi et al. 2004; DiMaggio et al. 2013).

To analyze news articles using Gibbs sampling we need to set the number of topics T beforehand (step 1.2). This is not a trivial task, because too few or too many topics compromise their human interpretability. So that, the number of topics is selected by running the topic model using several alternatives of T and choosing the one with the highest log-likelihood, which at the same time serves as evaluation method for the model (Wallach et al. 2009). The Maximum likelihood in topic models can be described as the number of topics T that most likely reproduce the observed dataset, the news, when using (1).

Once the number of topics is chosen, the document-topic matrix and the topic-word matrix are extracted (step 1.3). The topic model can be considered finished by simply ranking the words by their probability. However, this has been proven to be a suboptimal solution and a better method of ranking words in topics called relevance (step 1.4) was proposed by Sievert and Shirley (Sievert & Shirley 2014). Relevance is a sorting function driven by a parameter that can take a value between 0 and 1, it represents a tradeoff between how exclusive the word is to the topic (relevance = 0) and how frequent a word appears in that topic (relevance = 1). The authors found that an optimal solution lies in about 0.6, thus we used that value of relevance in this chapter.

The final step (1.5) is the visualization of the model. Distance between each pair of topics is calculated using the Jensen-Shannon divergence method, resulting in a T by T distance matrix. This is then reduced to T by 2 matrix using multidimensional reduction of the Principal Components. Once reduced, topics can be plot in 2-dimensional space.

7.3.1.2 Citation network of academic articles

Journal articles are different from news articles in several aspects, one key difference is the use of bibliographic references. Each academic paper is required to report previous works from where its results were built upon, creating a network of connections to previous papers. This citation networks, also called academic landscapes, have been used to study knowledge flows (Sakata et al. 2013) and research fronts(Shibata et al. 2011).

To create a citation network first we have to extract the bibliographic information of papers from a database of academic journals (step 2.1). Papers are treated as nodes, and each node gets connected to the papers they mention in their references (step 2.2). This is represented as arrows pointing backwards in time. In the following step, the directions of the arrows are removed and the largest connected component is extracted (step 2.3). We use the largest component because it captures the most information possible of the field we are analyzing.

In the next step (2.4), the network is clustered using a modularity maximization algorithm, which at the same time determines the best number of partitions for the network. As result, the citation network is divided in C number of clusters.

Finally, (step 2.5), the abstract of the papers in each cluster is analyzed and words ranked according to the term frequency – inverse cluster frequency TF-ICF weighting.

7.3.1.3 Semantic similarity between topics and clusters

The previous two stages outputted T topics and C clusters, for news and papers respectively. The words in topics are ranked by their relevance weighting, whereas those in clusters are ranked by TF-ICF weighting. In this part of the process we establish semantic relations between topics and clusters by using cosine similarity.

Cosine similarity measures the similarity of two vectors by calculating the cosine of the angle between them. It is used in text mining, where documents are represented as vectors of size V, being V the size of vocabulary in the corpus (i.e. the set of unique words). Any type of weighting can be applied to the words as long as they are transformed into normalized vectors, and all documents share the same vector space. Therefore, to compute the cosine similarity we have to go through two preparation steps: Obtain the total vocabulary, and normalize the vectors.

The total vocabulary is simply the union of the vocabulary found in the topics and the one found in the clusters: $V_{total} = V_{topics} \cup V_{clusters}$. Then, topics and clusters are expressed as vectors of size V_{total} which are sparse, filled with their respective weights. Following, the vectors are normalized.

Finally, cosine similarity, which is the dot product of normalized vectors, is computed over all possible pairs of topic and cluster vectors. The result is a T by C similarity matrix that can be visualized as a heatmap, from where the most similar topic and cluster can be observed.

7.3.1.4 Data sources

The methodology presented above is text driven. Given that the topics in newspaper articles are linked to academic papers in English, we chose an international newspaper to source the information of Japan. The New York Times was selected as database because is worldwide known, regarded as a quality provider of news, and several academic research has been done on its corpus (Melton et al. 2016), (Newman et al. 2006).

News articles about Japan were extracted from The New York Times by using the Application Programing Interface (API) they provide. The New York Times API covers a comprehensive archive of international news dating back to September 1851 (The New York Times 2016). Furthermore, they keep a controlled vocabulary, also called Times Tags or Times Topic Pages, to refer to specific people, events or places. As query, instead of searching the news that contains the term “Japan” we search for the Times Topic “Japan” in order to be sure we get the news articles *about* Japan. Results were filtered to get only those of blog, business, education, food, front-page, health, open, opinion, public editor, science, technology, times topic, topics, and world sections (i.e. those sections that are unlikely to represent a social issue as obituary, dining & wine were not taken into account). Because we are interested in recent social issues we retrieved articles from January 2010 to November 2015, resulting a total of 1528 news articles.

For the papers, the Web of Science core collection by Thomson Reuters was used to retrieve bibliographic data about robotic technologies. The Web of Science is known as a reliable source of journal articles, indexing more than 12000 journals in basically all disciplines since 1970. The query was as follow “(social* or service* or personal or consumer) NEAR/2 robot*” spanning trough all years available. 3908 records were extracted.

7.3.2 Results

1528 news articles about Japan were extracted from The New York Times, and 3908 journal articles about social robotics were retrieved from the Web of Science. In this section we present the characteristics of each of those datasets and how they are related. News articles were analyzed through topic models. We fitted the model using different amount of topics in a range from 2 to 50, and determined that 49 topics was the best solution for the news data based on the maximum log-likelihood as is shown in Figure 7-2.

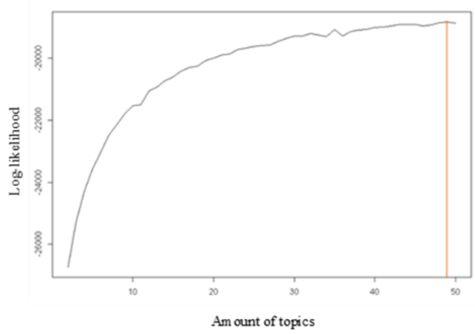

Figure 7-2 Estimation of the number of topics.
Based on maximum log-likelihood the highest value is reached at 49 topics.

The 49 topics about Japan are shown in Figure 7-3. Axes are the two principal components after multidimensional reduction representing relative distance among topics, those appearing closer each other are likely to be semantically related. Topics are ordered based on the size, which is an indication of how many words the topic collected in relation of all words in the corpus. Topics were named by human judgment based on their contents as shown in Table 7-1. Not all topics are about social issues. Because we are mining news about a country it is expected to get all kinds of topics related to that country. So that, topics about Japanese films, politics and so on can be spotted in the model. The largest topic 1, seems to be meaningless. It collects general words that cannot be allotted in other topics. These type of topics is called junk topics, and do not contribute to the analysis. They are also expected and are subject of study for its automatic detection and elimination (AlSumait et al. 2009).

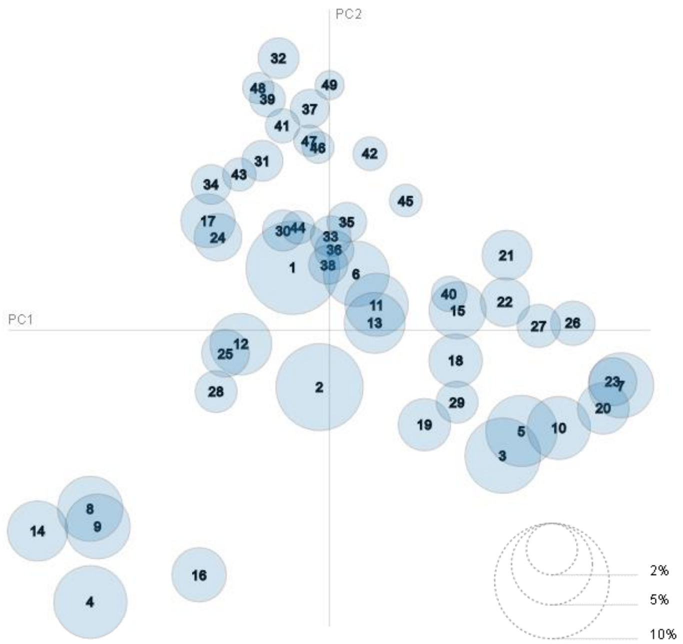

Figure 7-3 Topics about Japan obtained from news articles in The New York Times.
Axes are the two principal component after multidimensional reduction. They represent relative distance.

Table 7-1 Names of topics after manually inspecting their contents.

List of topics		
1. General (Junk Topic)	17. Local Migration	33. Journalism
2. Government	18. Airspace	34. Undersea Exploration
3. US Diplomatic Relations	19. Political Scandals	35. Whaling
4. Fukushima Daiichi	20. Military	36. Investment
5. Senkaku Islands	21. World War II	37. Overseas Education
6. Japanese and Foreigners	22. Historical Relations with China	38. Middle East
7. China		39. Culture and Traditions
8. Concerns on Radioactivity	23. War Apologies	40. North Korea
9. Nuclear Reactor Cooling	24. Damages on Infrastructure	41. Internet
10. Prime Minister	25. Earthquake and Tsunami	42. Law and Religion
11. Economics	26. South Korea	43. Climate Change
12. Accident and Disasters	27. United Nations	44. Academic Research
13. Politics	28. Nuclear Energy Policy	45. Film
14. Radiation Leaks	29. U.S. Military Bases	46. Social Demonstrations
15. International	30. Electronics Companies	47. Sports

16. Nuclear Plant Reactivation	31. International Business 32. Birthrate and Elderly Society	48. Research on Health 49. Literature and Arts

The bottom-left corner of Figure 7-3 groups 5 topics all referring to nuclear energy, particularly about the issues after the Fukushima incident. It is clear, according to the news in our dataset, that nuclear energy and its implications have received great attention in Japan, and by grouping them together is the main issue observable. Other topics that can be regarded more strictly as social issue are 17.- Local migrations and 32.- Birthrate and elderly society. Nevertheless, many others also contain a *social* component.

Table 7-2 shows the top words and news of selected topics. The junk topic collects words that do not convey meaning as a whole, and it is verified when looking at the news which seems to be unrelated. The other topics are coherent. Neglecting topic 1, topic model was able to map themes about Japan in a semantically meaningful way. Some of the topics may point to social issues, others are just an account of people, things or events. The next task is to link the topics to robotic technologies.

Table 7-2 Composition of selected topics from news articles.
Top 10 words and top 3 news are shown.

Topic 1: General (Junk Topic)	
Top words	Question, people, don't, time, make, things, sense, point, hard, ways…
Top News	Living with Mistakes Is Nuclear Power Simply Too 'Brittle'? Ghosts in a Secular Age
Topic 9: Nuclear Reactor Cooling	
Top Words	Fuel, reactor, rods, cooling, containment, water, vessel, spent, hydrogen, pools…
Top News	A Look at the Mechanics of a Partial Meltdown In Fuel-Cooling Pools, a Danger for the Longer Term Chemistry 201: Why Is Fukushima So Gassy?
Topic 32: Birthrate and Elderly Society	
Top Words	School, children, women, parents, son, mother, care, female, marriage, elderly…
Top News	On The Liberal Marriage Hypothesis A New Ratio for the Japanese Cram School Desperate Hunt for Day Care in Japan

An academic landscape was created by exploiting the citation relations of the articles on social robotics. The largest component of such citation network is shown in Figure 7-4. It consists of 1562 nodes, and 3079 edges, divided in 33 clusters. The figure also shows the top 5 clusters by number of nodes. In the network each cluster captures the research areas on social robotics. Table 7-3 shows the contents of selected clusters.

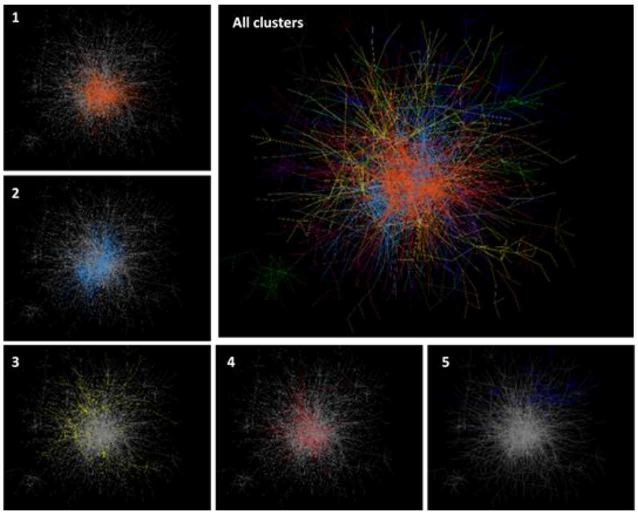

Figure 7-4 Citation network of social robotics.

The top 5 clusters are: 1. Robots for therapy and education, 2. Human-Robot interaction, 3. Mobility in closed space, 4. Robot interaction in the work place, 5. Robots for domestic tasks.

Table 7-3 Composition of selected clusters on social robotics.

Top 10 words and top 3 papers are shown

Cluster 1: Robots for therapy and education	
Top words	Child, autism, asd, social, disorder, emotion, therapy, expression, imitation, autistic
Top papers	A survey of socially interactive robots (Fong et al. 2003) Socially intelligent robots: dimensions of human-robot interaction (Dautenhahn 2007) Anthropomorphism and the social robot (Duffy 2003)
Cluster 2: Human Robot Interaction	
Top words	Older, social, child, acceptance, dementia, attitude, participant, assistive, paro, adult
Top papers	Interactive robots as social partners and peer tutors for children: A field trial (Kanda et al. 2004) Acceptance of Healthcare Robots for the Older Population: Review and Future Directions (Broadbent et al. 2009) Social Robots for Long-Term Interaction: A Survey (Leite et al. 2013)
Cluster 16: Robots for Hazardous Environments	
Top words	Unmanned, climbing, inspection, façade, vehicle, tele, wall, industry, adhesion, sea
Top papers	Intelligent legged climbing service robot for remote maintenance applications in hazardous environments (Luk et al. 2005) Tele-operated climbing and mobile service robots for remote inspection and maintenance in nuclear industry (Luk et al. 2006) Comparing speed to complete progressively more difficult mobile robot paths between human tele-operators and humans with sensor-systems (Sanders 2009)

Semantic connections were stablished among papers and news articles by calculating the similarity of terms that conforms each topic and cluster. Only the 100 most relevant words (forming a total vocabulary of 5559) were used to compute the similarity matrix that is presented in the form of a heatmap in Figure 7-5. The intensity of red imply more similarity. A higher similarity is an indication that they are semantically related. However, because the nature of language, it is possible that some pairs share words that are meaningless in the opposite context. The similarity matrix is just a tool to rank connected pairs, but they still have to be manually evaluated to assess the quality of their relation.

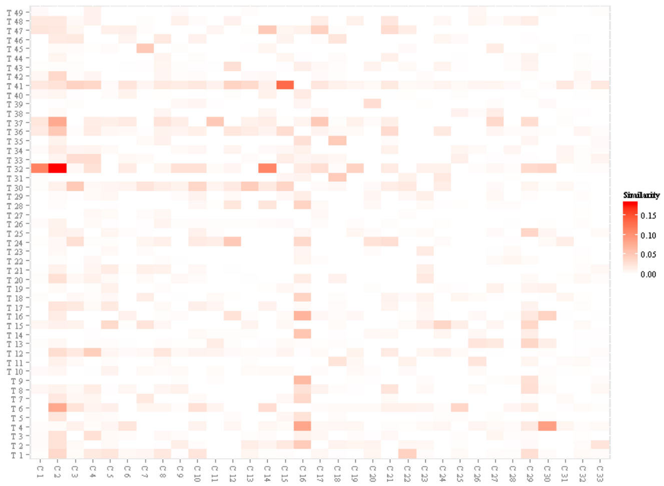

Figure 7-5 Similarity between topics of Japan and clusters of social robotics.
Y-axis and x-axis respectively

The 10 most similar topic and cluster pairs are shown in Figure 7-6. Three topics are associated to the clusters of human-robot interaction and robotics for hazardous environments. On the other direction 3 clusters point to the topic birthrate and elderly society. At first glance the linkage seems to have a good semantic relation, and a further analysis was conducted to verify their relatedness.

Birthrate and elderly society is a topic that collects several social issues. Some news were shown above in table 1, and some other headlines like “The Fertility Implosion”, “How in the World Will We Care for All the Elderly?”, “Japan’s Need for Women Workers”, appear in this topic along with some associated to children education. Each of them refers to known social concerns of Japan like fertility, elderly society, women empowerment, and others. This topic is linked to the clusters of human-robot interaction, robotics for therapy and education, and emotional attachment to robots, being a plausible connection between social issues and possible solutions through robotics. On the many possibilities of robotics for elderly care has already been reported by Ittipanuvat (Ittipanuvat et al. 2014), we can highlight the cases of robot Paro (Chang et al. 2013) and Keepon (Kozima et al. 2009) robots companions for the elderly and children, as examples of the many socially interactive robots that belong to the clusters. On table 3 we indicated more papers suitable for this linkages. For the emotional attachment to robots cluster we can mention the work of Bainbridge (Bainbridge et al. 2011) on the benefit of physical interaction with robots in games.

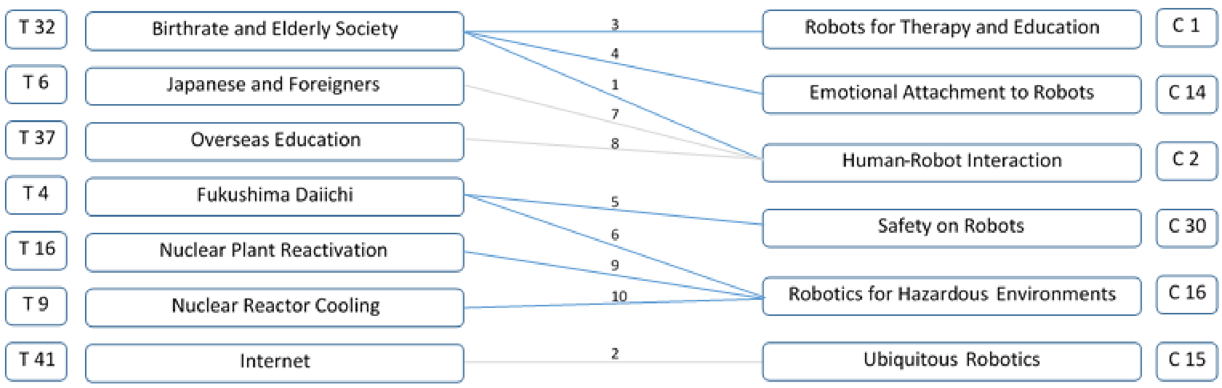

Figure 7-6 Pairs of topics about Japan and clusters of social robotics having the highest similarity.

The number in the connection is the ranking based on similarity, blue lines indicate a problem-solution relationship

Another group of linkages is the one connecting nuclear issues to robots for hazardous environments. Examples of headlines in these topics are: “Accident at Fukushima Daiichi Nuclear Plant”, “Japan Restarts Nuclear Reactor”, “Japan Plans Safety Assessments of Nuclear Plants” and those shown in table 1. A clear concern on nuclear energy is observed. For the issue on safety on nuclear energy production robotics may provide some solutions. Cluster 16 refers to tele operated robots able to move in unstructured terrain or climb vertical surfaces, some of them specifically designed for nuclear environments (Luk et al. 2006). Linkage to cluster 30 is also present, this contains indication on the design for safer robots (Woodman et al. 2012).

Not all connections discovered have a problem – solution relation. The second highest similar linkage joins the topic of internet and the cluster ubiquitous robotics. While it is true that both share many words, they belong to different semantic space. The topic internet, contains news related to the Internet as a whole having headlines of the type “Hackers Find Way to Outwit Tough Security at Banking Sites” or “Virus Infects Computers in Japan’s Parliament” that are unrelated to the ubiquitous robotics. The cluster refers to technologies connecting the robot to multiple sensors in the environment (Rusu et al. 2008) or to tele-operate robots through the Internet (Schulz et al. 2000).

Lastly, the topics of Japanese and foreigners, and overseas education where linked to the cluster of human-robot interaction. Both topics contains stories of foreigners living in japan, or Japanese living abroad. This kind of news use several words conveying emotional meaning (e.g. friends, feel) that happen to be present in the human-robot interaction cluster as well. Because the topics do not relate to specific issues or concerns, no solution can be associated from the cluster.

As shown above, linkage between topics and clusters by measuring the similarity of their text only, cannot be considered as evidence of a problem – solution relation for all connections. However, it is useful to recognize possible options and after inspection decide whether exist technologies that answers the problems observed

7.3.3 Summary

In this chapter we established semantic connections between topics about Japan found in an international newspaper and social robotic technologies published in academic articles. Semantic relations were helpful to connect local but broad issues of the selected country, to specific technologies in the robotics field. We used the robotics technologies as example, but other technology research fields may be used to connect the same topic model. Comparing the connections of multiple technologies to the topics of the same country might be useful for policy makers to spot research opportunities. By studying the topic model, important social issues (i.e larger topics) that don not show strong connection to technologies may serve as indicator of those opportunities. So that, the connection map either represented as heatmap or as the linkages in Figure 7-6, might be beneficial for analyzing both, stronger and weaker relationships. Such comparison deserve deeper analysis that will be conducted in future research.

Limitations were found in this research. Newspaper articles about Japan were extracted from an international newspaper. Even though this source is regarded as reliable, the use of domestic newspapers may be expected. Some Japanese newspapers also offer information in English but the availability of those databases was limited and could not be used. Nevertheless, the topic model described here reveals the themes on Japan that are relevant from an international point of view, which is at the same time a hint of relevancy, because topics that get great attention locally are prone to be displayed internationally as well. More research is needed to understand the usage of international press as proxy to evaluate local issues. On the other hand, the advantage of international outlets of news is that one source may be used to analyze several countries, or regions. These opportunities are yet to be explored.

Extracting text similarity served as indicator of relatedness, but not all the observed connections shown a solution to a problem. The tasks of separating the social issues from general topics, and identifying solutions from the connections still depends on human reasoning, and further work will be required to understand how completely automate them. However, the methodology is helpful to reduce the effort of selecting the best alternatives.

Further work is needed to evaluate the performance of topic models compared to other text mining methodologies in the domain mapping and research fronts identification tasks when analyzing academic articles. We will consider such questions in the future.