Lisbon School of Economics & Management
Universidade de Lisboa

MASTER

ECONOMICS AND MANAGEMENT OF SCIENCE TECHNOLOGY AND INNOVATION

MASTER´S FINAL WORK

DISSERTATION

INDUSTRIAL ESPIONAGE:

MAPPING MEANINGS AND ASSESSING TRENDS

ANA RITA ALMEIDA DO CARMO

JANUARY – 2023

Lisbon School of Economics & Management
Universidade de Lisboa

MASTER

ECONOMICS AND MANAGEMENT OF SCIENCE TECHNOLOGY AND INNOVATION

MASTER´S FINAL WORK

DISSERTATION

INDUSTRIAL ESPIONAGE:

MAPPING MEANINGS AND ASSESSING TRENDS

ANA RITA ALMEIDA DO CARMO

SUPERVISION:

PROFESSOR DOUTOR SANDRO MIGUEL FERREIRA MENDONÇA

PROFESSOR DOUTOR BRUNO MIGUEL PINTO DAMÁSIO

JANUARY – 2023

Acknowledgments

I would like to express my gratitude to Professors Sandro Mendonça and Bruno Damásio for all their understanding, and availability. Their advice was essential for my evolution.

To my family, thank you for being my biggest supporters and for always going above and beyond to make sure I have everything I needed.

Last but not least, I owe a big debt of gratitude to all my colleagues and friends who helped me stay motivated to accomplish this journey, which now that it is over, represents the conclusion of yet another cycle in my trajectory.

Abstract

Ever since the term “industrial espionage” exists, authors have been struggling with its definition. It is a multi-sided method used mainly for the purpose of obtaining trade secrets from competitors in the market, however, its methods may be numerous. Based on the existing research produced by the scientific community on the subject, we produced a bibliometric review regarding the same. Bibliometrics has an important role in the analysis of research production. Indicators regarding publication output and impact constitute an evidence base that helps to reveal the rhythm and direction of scientific agendas, technical concerns, and socio-economic/political debates. In this dissertation, we focus on a key topic that lies at the intersection of innovation studies and security studies and, we present a qualitative and quantitative analysis of the existing research regarding industrial espionage. For this analysis, publications in the Scopus database were used, analyzing, and processing them using the R studio analysis tool. The number of publications on industrial espionage is still relatively low, but there has been a positive trend in recent years. Through this analysis it was possible to verify the most productive authors and journals in this area, as well as to identify subtopics within industrial espionage that have the potential to be deepened in the future, and which are no longer so relevant for the scientific community. Through the Network analysis it was also concluded that the topic of industrial espionage is very broad, and that it is present in several other backgrounds and areas of study.

KEYWORDS

Industrial espionage. Economic espionage. Commercial espionage. Bibliometrics

Resumo

Desde que o termo "espionagem industrial" existe, autores têm-se debatido com a sua definição. É um método multifacetado utilizado principalmente com o objetivo de obter segredos de negócio dos seus concorrentes, no entanto, os seus métodos podem ser numerosos. Com base na investigação existente, produzida pela comunidade científica sobre o tema, produzimos uma revisão bibliométrica sobre o mesmo. A bibliometria tem um papel fundamental na análise da produção da investigação. Os indicadores relativos à produção e impacto da publicação constituem uma base de evidência que ajuda a revelar o ritmo e a direção das agendas científicas, preocupações técnicas, e debates socioeconómicos ou políticos. Nesta dissertação, concentramo-nos num tema chave que se situa na intersecção de estudos de inovação e estudos de segurança e, apresentamos uma análise qualitativa e quantitativa da pesquisa existente relativamente ao tema de espionagem industrial. Para esta análise, foram utilizadas publicações que se encontravam na base de dados Scopus, analisando e processando os mesmos através da ferramenta de análise R Studio. Verificou-se ainda um número relativamente baixo de publicações concerne espionagem industrial, no entanto há uma tendência positiva nos últimos anos. Através desta análise foi possível verificar os autores e revistas mais produtivos nesta área, bem como identificar subtópicos dentro de espionagem industrial que tem potencial e ser aprofundados no futuro, e quais já não são tão relevantes de análise para a comunidade científica.

Através da análise de Network conclui-se também que o tópico de espionagem industrial é muito abrangente, sendo que o mesmo está presente em várias outras backgrounds e áreas de estudo.

PALAVRAS-CHAVE

Espionagem industrial. Espionagem económica. Espionagem comercial. Bibliometria

Table of Contents

List of Figures

List of Tables

1. Introduction

With the significant economic development and increased competition (Hou & Wang, 2020), the last decades have seen a growth in industrial espionage among corporations (Thorleuchter & Van den Poel, 2013). However, there is still a concern among entities in admitting a breach violation, fearing loss of trust. As a result, the news of espionage are rarely made public, unless they are leaked by intelligence agencies (Søilen, 2016). Still, one of the biggest examples, and one of the most researched is the Cold War between Eastern and Western blocs. In the 1970s and 1980s, when the division between East and West was made, the competition between the two regions was substantial. For the communist regime on the East side, industrial espionage played a substantial role in keeping up with the Western capitalist side's increased productivity, and economic growth (Glitz & Meyersson, 2017).

Espionage practiced then, not only has not slowed down, but it has also evolved and adapted to the current times. Companies feel the pressure to keep up with the market while trying to stay protected from outsider attempts of intelligence gathering and/or espionage, conducting their own research on the subject. Recently, a survey conducted by CyberEdge Group (2021), in which 1200 IT security decision makers attended, claimed that 86,2% of the responding organizations experienced at least one successful cyberattack in the previous twelve months. Furthermore, 15,0% of those, experienced more than ten attacks in the same period. Due to the growth of attacks, companies are increasingly committed to protecting their clients and their data, not only through investing in security but also by training their employees through lessons, e-learnings, conferences, and fishing e-mail tests, to know how to detect a possible security breach and know how to approach them.

This paper aims to provide the research community with a qualitative and quantitative tool that identifies the amount of scientific research on industrial espionage. Consequently, we hope that, through this research, researchers and public entities will gain a deeper knowledge of the subject and understand how it can affect society, directly or indirectly. Additionally, we look forward to contributing to the scientific community with more research on the subject and suggest additional research points.

This dissertation addresses the question "*How has industrial espionage been studied in the scientific literature?*" in a systematic way from the relevant academic literature. We expect to contribute to the already existing literature that systematically analyzes the subject of industrial espionage, namely distinguishing it from competitive intelligence (Hou & Wang, 2020), through the analysis of existing scientific production and, be able to effectively perceive which authors and journals are most cited, therefore, more valid scientifically. The present study identifies what authors and journals have the greatest impact on citations and what are the topics on industrial espionage that have been studied with the highest frequency and are currently attracting more interest. Additionally, we present the evolution of published papers throughout the years.

We begin this research by defining espionage, identifying its different types, present a systematic literature review on industrial espionage and discuss some of history's most famous cases. Then, follows the identification of the data selected in our analysis and an explanation of how it was obtained. Finally, we present our results, by quantifying the key indicators of a bibliometric review (received citations, most cited articles, most cited authors and journals, and the distribution of publications) and propose some topics to further investigation.

2. Terms and Definitions

2.1 What is espionage?

The term *spy* originates from the old french *espie* (8th - 14th century), with Germanic origin in the word *espier*. According to the Oxford English Dictionary, *Spy* is "a person employed by a government or other organization to collect and report secret information on an enemy or competitor", "a person who observes others secretly" (*OED*, 2008, p. 1399). The term *espionage* has its origin in the French term *espionnage*, and it is "the practice of spying or of using spies" (*OED*, 2008, p. 487).

During the American Civil War, an instruction signed by US President Abraham Lincoln to the Union forces of the United States called "*Instructions for the Government of the Armies of the United States in the Field",* had the purpose of dictating how soldiers should conduct themselves during warfare. In the instructions published in General Orders No. 100, 24 April 1863, article 88:

> "A spy is a person who, secretly, in disguise or under false pretenses, seeks information with the intention of communicating it to the enemy. The spy is punishable with death by hanging by the neck, whether or not he succeed in obtaining the information or in conveying it to the enemy." (Lieber & Hartigan, 1863).

Espionage can be defined as the access to sensitive information by questionable means to use against the public interest (Crane, 2005). Nevertheless, it may be the action of providing information about the national defense, to someone who is not authorized to obtain it (Wagner, 2012). Espionage occurs to acquire a counterpart's private information in cases with asymmetric information (Ho, 2008). It is a strategic option in which the actor considers purchasing information on their opponents' decisions by studying the value of its private information. Which is delineated according to what this actor can attain by choosing this strategic option (Solan & Yariv, 2004). For the purpose of this dissertation, we are adopting Crane's definition of espionage as standard.

2.2 Different Spying Approaches

Espionage has been evolving through the decades, through times some terms such as 'economic espionage', 'industrial espionage', 'corporate espionage', and 'commercial espionage' have been emerging (Button, 2019). To simplify and focus on the last three terms mentioned, we will use Button's classification, and consider industrial espionage, commercial espionage, and corporate espionage as having the same meaning, referring only as industrial espionage from this point.

Economic espionage is the government's action of stealing information or trade secrets (Nasheri, 2005, p. 7) from domestic companies or governments to provide an advantage to a foreign state (Danielson, 2009). It is a new type of "white-collar crime" (Nasheri, 2005, p. xii). The US Economic Espionage Law (1996) defines economic espionage as the act of intentionally or knowingly benefiting any foreign government, foreign instrumentality, or foreign agent by stealing, copying, or receiving a trade secret, knowing the same to have been appropriated without authorization.

Economic espionage should not, however, be mistaken for government surveillance. During World War II, this surveillance method grew and increased its resources, both physical and human. For instance, increasing amounts of data regarding citizens and organizations - business transactions, cash flows, propaganda schemes, and tendencies in foreign media - were gathered between Denmark and Germany (Marklund, 2019). In

the United States, government surveillance was later legitimized and justified to American citizens, by the government, in 1971, as a necessity to be able to obtain information of any disturbance or threat to public security as long as it was not obtained in an intrusive or illegal way (Christie, 1972). Currently, US's intelligence services collect large amounts of information (mostly over the Internet) related to their citizens, justifying this act as not being a violation of privacy, since the information is not "accessed" (Macnish, 2016, p. 14). Of course, this has been going on for a while now. Since the creation of the US intelligence agency in 1952, the NSA - National Security Agency - has been illegally spying on multiple individuals (especially since the 1960s), additionally, it has also been collecting, storing, and analyzing data on a large proportion of the internet's user base since 2007 (Summer, 2016, p. 32), mainly due to 9/11 (Token, 2014). This was all disclosed (or confirmed) in 2013, when an NSA employee, Edward Snowden, took about 1.7 million documents from NSAs SharePoint, of which, up to 200 thousand were shared with journalists (Token, 2014; Summer, 2016, p. 22), to draw the public's attention to the NSAs acts of espionage on citizens (Token, 2014). Of all the documents shared by Snowden, two major surveillance programs stand out: Muscular and Prism. Muscular was operated mainly by intelligence agency GCHQ (Government Communications Headquarters, formed in 1919 in the UK) but with NSA's interference aimed to collect private, unencrypted public data extracted from communications sent using Yahoo and Google. Prism, also a program from both NSA and GCHQ, aimed to access encrypted online communications (such as video calls, photos, files, emails, instant messages, and others) under FISA (Foreign Intelligence Surveillance Act) authority. That meant that the information was obtained legally, under Section 702 of the FISA Amendments Act of 2008, forcing companies such as Google, Microsoft, Yahoo, or Facebook to disclose records of non-US citizens living outside the US, for up to one year (Summer, 2016, pp. 20-26).

In fact, in recent years, surveillance has been evolving to a new phase of marketing and advertising, named surveillance capitalism, in which the production of goods and services is subjected to the human experience, gathering information of the consumer, and consequentially using that data to predict the human behavior and therefore, increase sales and encourage consumerism (Zuboff, 2019, p. 8).

After World War II, the US intelligence agency, the CIA, created in 1947, took over and focused intelligence gathering through OSINT (Open Source Intelligence) and

HUMINT (Human Intelligence) collection analysis (Norton, 2011). The information these two intelligence analyses cannot provide is then provided by SIGINT (Signals Intelligence) (Badiru & Maloney, 2016). HUMINT, the gathering of intelligence through human means (Rose, 2009), and SIGINT, the gathering of intelligence through the analysis of other countries' communications (Diffie, 2006) are both used and applied by agencies such as the NSA or the GCHQ (Aid, 2003).

Military espionage, another espionage approach, was first applied in the context of war. In the early 1910s, a military spy was first indicated as a person who, in disguise or under false pretenses, seeks information from the enemy. Furthermore, the hiring of a spy was not considered an act against the laws of war, as long as he or she, was fully aware of the offense against the enemy and the consequent punishment, if caught (Halleck & Davis, 1911; Delupis, 1984). However, one who is captured and accused of espionage may not be judged as a military spy and therefore be treated as a prisoner of war, if there is an absence of a uniform, being then, considered a common spy (Delupis, 1984). A more recent term that emerged with World War II, and the advent of the atomic bomb, is nuclear espionage. During the Cold War between the US and the Soviet Union, KGB's spies were crucial for the Soviet Union for the gathering of information regarding the Manhattan Project. In fact, KGB's spies were more successful than previously expected by the US, they might have been even more so if they were not as "shortsighted, parochial, and bureaucratic as any of their Western counterparts" (Herken, 2009, p. 69).

With the emergence of more and more complex forms of espionage, governments have now made counterespionage one of their top concerns (Barrachina & Forner-carreras, 2022). And, despite the fact they are increasingly applying efforts on it, through legislative changes and developments, these remain insufficient and not only are scarce in the face of attacks but also, are not used as leverage to obtain new information of common interest (Oxnevad, 2019). So that counterespionage can be used more efficiently, two measures may be considered; 1) Protecting domestic corporations from economic espionage by foreign firms; 2) Extending the boundaries of what is defined as economic espionage and increase penalties for engaging in it (Barrachina & Forner-Carreras, 2022). Additionally, Grabiszewski & Minor (2018, pp. 276-277) found that applying a common counterespionage policy that covers all national companies may not be as optimal as expected, i.e., if a national company has high levels of R&D and there

is already any type of economic espionage protection prevailing, additional governmental protection will cause a reduction of the companies R&D intensity and consequentially, encourage foreign firms to increase espionage.

Finally, Personal Data Breach is also a term being increasingly used. Personal data may include identification card numbers, names, residences, dates of birth, credit card information, medical records, and contact information (Phua, 2009). According to the General Data Protection Regulation, a personal data breach is a "breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed" (Art. 4(12) GDPR).

After we have settled on the concept of espionage and approached some used practices, we are presenting some historical examples of espionage, to understand what it means and what its results may be in practice.

2.3 Historical Background

The Portuguese Discoveries

The 15th and 16th centuries were a very interesting period for the European countries, especially, the Portuguese. King Dom João II had shown an interest in the Nautic discoveries at an early age, and by his death, in 1495, his successor, Dom Manuel, had already ordered the explorer Vasco da Gama on the first nautic expedition to India. (Sérgio, 1977, p. 59). In the year 1500, after his return to Portugal, King Dom Manuel ordered Pedro Álvares Cabral to lead the Portuguese, with the biggest armada ever seen so far, on their second trip to trade spices with the Oriental countries. Instead, they came back with stories of contact with a new land - today known as Brazil - which in truth told, were not exceptionally surprising. According to Dom Manuel, the Portuguese were previously familiar with the American continent, even before Christopher Columbus. This discovery would not have been disclosed before that time for reasons of confidentiality and to prevent espionage by foreign nations (Sérgio, 1977, p. 64).

The Portuguese possessed nautical knowledge that rivals in the Netherlands, France, and England wanted. Information like charts, navigational instructions, the locations of logs, rocks, and reefs, as well as eddies and currents - knowledge that is today widely

shared - was fiercely held as secrets. All this knowledge gathered over time in all of the Portuguese experiments was essentially state secret and disclosing it to outsiders was punishable in court. (Hobhouse, 1986, pp. 120-121). Due to these discoveries, until the 17th century, Portugal maintained the biggest empire in the world, which Spanish revolutionaries from the Netherlands overthrew. The Portuguese court would later recover in the 18th century during Brazil's golden age (Vardi, 2001).

The Tea War

After the discovery of the sea route to India by Vasco da Gama, in 1497, trade between Europe and the Orient by sea began slowly (Hollins, 2020, pp. 17-18). Tea was first consumed in Europe by the Portuguese, around 1580. Only a few decades later, in 1652, it was introduced to the British as "Cha" (Hobhouse, 1986, pp. 117-118). And finally, in 1689 it was first imported from China to England (Mair & Hoh, 2009). Over the years, the consumption of tea became exceptionally engraved in Britain among all other European countries and, in the 18th century, tea drinking became a national custom of the British (Zhong, 2010).

It is estimated that by 1820, 30 million pounds of tea were already being consumed per year, in the UK alone. However, in almost two decades of tea consumption, only the Chinese had the knowledge of how tea was produced (Hobhouse, 1986, pp. 118-119). In the 19th century, the British East India Company sent the botanist Robert Fortune to China to obtain the tea processing technology and collections of tea trees (Fortune, 1847). Between 1843 and 1861, Fortune traveled to China several times to obtain plants from the Chinese (Yip, 1999). In his book *"Three years' wonderings in the Northern Provinces of China, including a visit to the tea, silk, and cotton countries: with an account on the agriculture and horticulture of the Chinese, new plants, etc."*, in 1847, he wrote:

> "We may safely conclude that the Chinese could easily supply any increased demand which we are likely to make upon them from this country, and supply it without ultimately advancing the prices of tea. (…) But were there any doubts of our being able to procure an almost unlimited supply of tea from China, let us turn to our own dominions in India. (…) All my experience of the tea districts in China goes, however, to show that the north-western districts of the mountains in India (…) are much better suited for this purpose than the more southern country of Assam. (…) The advantages which would result from the successful cultivation of the tea plant in India are immense. The vast population of our empire in the East

> would have a cheap and harmless beverage produced amongst themselves, and thousands of families would find a healthy and profitable employment in the cultivation and manufacture of tea. These results are altogether independent of the benefit which would be conferred upon our population at home." (Fortune, 1847, pp. 218-222).

This allowed Great Britain to grow tea in India, breaking China's control of the tea monopoly from the 19th century on.

The Cold War

In the 20th century, after the end of World War II and with the beginning of the Cold War, the world testified the division of Germany into two main sides, the Western, controlled mainly by the United States, and the Eastern, controlled by the Soviet Union, with this division, the West imposed a trade embargo on their Eastern Bloc counterparts, originally focusing on limiting the trade of gun and weapon technology. Over the years the Western control over technology transfer and exported goods with cutting-edge technologies and other commodities from the military and nuclear industries was only increasing. Therefore, East Germany started to rely upon, to a greater extent, on industrial espionage as the trade embargo, especially in the economic sector, which turned out to have significant benefits, allowing East Germany's economy to partially keep up with productivity increases in the 20th century. This could have facilitated East Germany to catch up to the capitalist side, as investments in espionage demonstrated economies of scale. However, after the unification of Germany in 1990, Western corporations had decades of experience conducting effective R&D, while Eastern enterprises lost their main sources of technological know-how (Glitz & Meyersson, 2017).

When the Cold War was over, thousands of discharged spies started looking for work in the private sector (Søilen, 2016), since the biggest organizations - such as NATO, the KGB, or the CIA - were redefining themselves, most spies began to engage in corporate espionage, adapting their skills to new responsibilities and duties. Governments had also begun redefining priorities, refocusing their policies to raise the economic standards, and recognizing that economic superiority had now equaled military superiority in importance (Nasheri, 2005, p. 19). Nowadays, as a consequence of the Cold War, most of the time those who engage in economic and industrial espionage employ a variety of collection techniques in a concerted effort that combines legal and illegal, conventional

and cutting-edge procedures (Nasheri, 2005, p. 82). Today, even though the relative advantages may be fewer than they were a few decades ago, mainly because of globalization, the costs of industrial espionage are also decreasing due to the development of cyber techniques, making this method of productivity as relevant as then (Glitz & Meyersson, 2017).

Now that we have introduced the most common types of espionage existing (such as economic, military, and nuclear espionage), as well as surveillance methods and data breaches, we are directing our attention to industrial espionage, which, unlike other terms already addressed, is not as straight forward.

2.4 Industrial Espionage

After identifying economic espionage as a government's efforts to gather information or trade secrets, industrial espionage aims to benefit a private company through the same means, however with no governmental direct influence (Nasheri, 2005; Danielson, 2009). According to Hou & Wang (2020), industrial espionage is an interdisciplinary object without a standard definition. According to Boulouard et al. (2016), industrial espionage is an object based on information outside the company and its methods of obtaining the data are illegal and unethical.

However, in a more complete approach, according to Button (2019), industrial espionage is a set of legal or illegal activities, used to obtain an advantage in the market in which the company is located. Also, industrial espionage can take many forms, from research in public sources to *Cyber-attacks*. Between these two forms, the human factor is an important aspect, since spies are essential, one may not even be aware that he or she is a source of intelligence (Sivanesan, 2011), being intelligence a strategy for obtaining, analyzing, interpreting, and disseminating value from data to assist in the decision-making (López-Robles et al., 2019). Crane (2005) admits that there is a line that should be drawn between the ways "ethical" and "unethical", meaning that ethical industrial espionage is competitive intelligence. Therefore, calling an action 'industrial espionage' implies that it is unethical, although not necessarily illegal, since the law may be incapable to set such a line, furthermore, espionage may be classified as "intelligence practices of questionable ethics" (Crane, 2005, p. 234). To be able to identify industrial espionage, Crane (2005) suggests three criteria: the tactics must be "questionable" (one may question if the information was obtained legally or ethically);

the information should be considered private or confidential; and the data gathered "is to be used against the public interest" (Crane, 2005, p. 236).

Industrial espionage is an increasingly recurrent theme these days, but it possibly has existed since the moment cavemen first saw the use of fire and tried to replicate it, thousands of years ago (Bergier, 1975). Much later, it was largely driven by the industrial revolution (Champion, 1998). And, in the 20th century, it took hold and was a very popular productivity method in the Soviet Union during the Cold War (1970 – 1989), where espionage was used to keep up with technological developments in capitalist organizations (Glitz & Meyersson, 2017). Obtaining information from other companies can bring benefits to the market, when more than one company has access to the same information, they are able to offer the same product to the market, thus avoiding monopolies and giving the consumer a choice, either by lower prices or by a differentiated product (Heims, 1982).

Alongside industrial espionage comes intellectual property and intellectual property theft (Waziri & Yerima, 2011). Intellectual property is any land, structure, and equipment that has the purpose to create or transform materials into finished manufactured products (Holmes, 1948 as cited in Ambrose, 1990, p. 356). Consequently, intellectual property theft is its appropriation, therefore the theft of tangible assets or operational information (for instance datasets, R&D, or planning of marketing strategies) (Waziri & Yerima, 2011). In 1883, with the goal of protecting industrial property, was signed The Paris Convention, it provides all the signatory nations the same industrial property protections that are provided to their own citizens and ensures the constant enforcement of industrial property rights. Though, one of its flaws is the tolerance and perpetuation of weak national laws (Arnam, 2001). Intellectual property is the protection and attribution of legal rights to intangible property, for example, an idea or an invention. Intellectual property is protected by four legal areas, those are patents, copyrights, trademarks, and trade secrets (Task Force on Intellectual Property, 2004). As a result, intellectual property theft is a violation of one of these four laws (Mclllwain, 2005). Even though these laws are created to protect intellectual property, some authors argue they are flawed, the patents system, for instance, may not always benefit the domestic economy, depending on the "political economy issues", also, patents are not correlated with the countries' innovative production (Grabiszewski & Minor 2018). Some companies, in order to protect

themselves against intellectual property theft, are being incentivized to publish scientific papers. By doing so, they are using it as a strategy for establishing legal property rights for their discoveries as well as managing and protecting their intellectual property portfolio (Rotolo et al., 2022).

Of course, the rise of technological assets through the years is visible. In the 1980s the intangible assets, represented about 40% of the total companies' value, in 2006, they represented about 75%. With the world much more digitalized, it is expected that with the increase in technology and the means, the incentives to practice industrial espionage increase as well (Podbregar, 2006). In addition, it is often physically easier to steal intangible property, since one may be able to do so through a laptop or computer (Nasheri, 2005, p. 52). A more recent question is Forced Technology Transfer, it is the pressure attributed to a foreign company to transfer technology to a domestic company if the foreign company is conducting or about to conduct business in the domestic company's country (Carbaugh & Wassell, 2019). In the last years, the most apparent case is China, the policies applied by the country that force the transfer of technology in or to the country attempt to oblige technology transfer through negative incentives (Prud'homme & Zedtwitz, 2019).

3. Research Methodology

In order to understand this dissertation, we should first find the definition of bibliometric review, and understand what it requires. Potter (1981, p. 5) defined bibliometrics as "(…) the study and measurement of the publication patterns of all forms of written communication and their authors". A bibliometric review aims to collect large amounts of bibliometric data to describe the status of the intellectual structure and emerging trends of a topic or field of research. A literature review is a more complete analysis, when compared to a systematic literature review, for it conducts a qualitative and quantitative analysis of the object under study, while the other, simply studies the object qualitatively (Donthu et al., 2021).

3.1 The Bibliometric approach

With the increasing number of scientific articles being published in the scientific community, bibliometrics becomes a result of all publications as a whole, with the

purpose of identifying top-performing researchers and articles in the entire community (Ball, 2018, pp. 8-10). Bibliometric data provide insight into some major macro-indicators. These could be the structure of academic activities (cognitive and/or social) and scientists' knowledge of developments in particular academic branches, at the national level; the academic productivity in each country and field; the impact of each country or region in specific fields of knowledge; the extent of international and regional collaboration and the use of formal channels of communication in each country and field; the scientific product and its influence on the community; and finally, institutional collaboration, for example, private versus public. The objects of this type of research are the creators of the publications (as individual authors or as teams), the journals and articles on their own, and finally, their descriptive characteristics and citation analyses (Moed et al., 1992, pp. 8-10; Jesus & Mendonça, 2018; Lyra et al., 2022; Mendonça et al., 2022).

In order to prepare a bibliometric analysis, one should choose the right indicators. A bibliometric indicator is a statistic that is quantitatively ascertainable, thus measurable, and makes a remark about an academic publication. In principle, the number of indications that may be used to make a remark about an academic paper is limitless. The decision of which statement to make is always critical for the selection and application of an indicator (Ball, 2018, pp. 17-18). There are two general classifications of bibliometric indicators: qualitative, for example, the impact factor of a publication or journal; and quantitative, such as the *h*-index (García-Villar & García-Santos, 2021). Even though an indicator can be either qualitative or quantitative, it must always be measurable, must provide an answer to a concrete question, and also, it must be possible to determine the indicator with a correlation between its effort and benefit (Ball, 2018, p. 18).

A bibliometric analysis contains two main techniques, performance analysis and science mapping (also known as bibliometric mapping (Cobo et al., 2011) or bibliometric method (Zupic & Cater, 2015) (Noyons et al., 1999; Donthu et al., 2021)). Evaluation of the research performance of nations, universities, departments, or individuals is done using performance analysis, which is based on publication output and received citations (Noyons et al., 1999). Since it is common practice for reviewers to provide the performance of various research components, performance analysis can be found in the majority of reviews, including those that do not engage in scientific mapping. There are

several performance analysis metrics. The most important metrics are the number of publications and citations made each year or for each study component. Publications are used to gauge productivity, while citations are used to gauge influence and impact (Donthu et al., 2021). A citation is defined as "a reference to another publication in an academic paper" (Ball, 2018, p. 81). Such citations show that the author referenced has done work that is relevant to the current research frontier and valuable to others striving to expand it (Diamond Jr, 1986). To assess the effectiveness of research constituents, other metrics are used, such as citations per publication or the *h*-index (number of publications that have at least *h* citations) that combine citations and publications (Donthu et al., 2021). Although, regarding the latter, there are a few factors that can have a negative impact on it (as well as other indicators): Inadequate consideration of self-citations and multi-authored articles; significant reliance on the specific database used to gather the original publication and citation information; failure to account for changes in age and experience (Egghe, 2010).

Science mapping shows the relationships between various fields, specializations, and individual papers or authors (Cobo et al., 2011; Donthu et al., 2021). There are five main techniques used in a science map: Citation analysis, Co-citation analysis, Bibliometric Coupling, Co-word analysis, and Co-authorship analysis (Donthu et al., 2021). Out of these, the most common are Co-citation and Co-word analysis (Cobo et al., 2011). Co-citation links publications, authors, or journals based on their shared appearances in reference lists. Co-word connects words that exist in the same title, abstract, or keyword list of different publications. Regarding the remaining three, Citation measures the impact of articles, authors, or journals using citation rates; Bibliometric Coupling relates publications, authors, or journals according to how many references they have in common; and Co-author connects authors that write the same papers (Zupic & Cater, 2015).

Since the first citation analysis made in the scientific community, in 1927, different methods have been developed and employed to help researchers to gather and interpret documents (Budd, 1988). Out of those, were created, what today we know as bibliometric laws. The firm groundwork for bibliometrics was built by these three essential laws (Mathankar, 2018): *Bradford's, Lotka's*, and *Zipf's* laws.

Bradford's law, suggested in 1934, aims to determine which journals are the most relevant and provide the most information on a certain topic (Chueke & Amatucci,

2015). Bradford described the pattern of literary dispersal in numerous journals. He discovered that if a large collection of articles is sorted in decreasing productivity of papers related to a specific topic, a standard division may be identified (Mathankar, 2018). “A standard classification must be adopted, so that references to the same subject would be brought together by the classification, irrespective of source or abstracting bureau, when, without an increase of labor, a complete index to scientific literature would be achieved” (Bradford, 1976, p. 103). Thus, journals in a certain topic area can be divided into three categories: Those that provide more than four references per year; those that produce two to four references per year; and those that carry one or fewer references each year (Mathankar, 2018).

Lotka’s law, approached in 1926, documents the relationship between authors and publications, concluding that publication output was inversely related to the number of scientists on a topic (Ball, 2018, p. 11). This correlation is now known as Lotka’s law, and its main purpose is to assess the impact of an author's production on a field of knowledge (Chueke & Amatucci, 2015). The invariability of Lotka’s law may allow it to be used to disseminate the production of certain groups of authors (Arsenora, 2013).

In 1932, Zipf argued that given a corpus of natural language utterances, the frequency of each word is inversely related to its rank in the frequency table, which became known as *Zipf's law* (Adel, 2021), in other words, if the words in a long text are sorted in decreasing order of frequency, the rank of any given word in the text will be inversely proportional to the frequency of occurrence of the term (Mathankar, 2018). Zipf's law is used to estimate the most recurrent themes related to a field of knowledge (Chueke & Amatucci, 2015).

There are several measures and indicators that may be employed in bibliometric investigations. Therefore, it is important to use a complete database in a bibliometric study. The amount of academic production by a person, institution, country or another group (aggregated on several levels) is the baseline parameter for a bibliometric output study (Ball, 2018, p. 18). Only types of output that are considered publications and can be documented should be considered in the bibliometric analysis. Although other outputs (such as seminars, hearings, exhibitions, commentaries, reviews, and other academic output) have recently been attempted in alternative metrics, “altmetrics”, which also covers webometrics (Ball, 2018, p. 19), these will not be acknowledged in the measuring logic of this bibliometric review, thus, only being considered scientific

articles and review articles. Nowadays, three primary databases are utilized in bibliometric investigations, each with a vast range of data and diverse metrics. These are Web of Science, Scopus, and, more recently, Google Scholar Metrics (Vieira & Wainer, 2013).

3.2 Systematic Literature Review

Before engaging in the quantitative review of industrial espionage, we must first do a qualitative analysis of the subject, by doing a systematic literature review. The benefits of this analysis are the gathering of evidence concerning a subject, identifying gaps in the area of study, and provide a background for new research activities (Kitchenham & Charters, 2007). There are three main steps to conducting a systematic literature review, that may be, then, subdivided (Kitchenham & Charters, 2007); Xiao & Watson, 2017):

1. Planning the review
    a. Formulate the problem
    b. Develop and validate the review problem
2. Conducting the review
    a. Search the literature
    b. Screen for inclusion
    c. Assess quality
    d. Extract data
    e. Analyze and synthesize data
3. Reporting the review
    a. Report findings

Considering, therefore, the stages above identified, we must then start with a systematic literature review on industrial espionage. Bearing in mind that this subject has already been systematically revised by Hou and Wang (2020), we aim to complement and deepen the research done before by understanding its causes, effects, and solutions.

As discussed in chapter 2, there is no standard definition of industrial espionage, it is an umbrella term, to which several authors add their own interpretations, this makes it difficult for researchers, companies, or the general public to understand what industrial espionage is, often mistaking it with economic espionage.

Competitive intelligence between companies is becoming even more and more competitive, there are a lot of advantages for a company being a pioneer in introducing a new product in the market. Competing organizations want to set the profile of corporate leaders by using techniques such as recruiting, tactical surveillance, corporate employee profiling, information assurance, and elicitation training. This is, of course, a great motivation for corporations to engage in competitive intelligence, some CEOs even insist that operating without competitive intelligence is a huge disadvantage from the beginning. Even though competitive intelligence is not industrial espionage, being the latter, clearly, the appropriation of information or material illegally, these two terms are not black and white. There is a huge gray area between the two that the law is not capable of acting on (Nasheri, 2005, pp. 74-77). Additionally, throughout time, this gray area is becoming bigger and bigger, that is partially because companies and entities have much more access to tools that are used in surveillance than they used to, and , governments have their role to play as well. Since they rely on domestic companies to increase economic growth, they frequently play a significant role in aiding or preventing "so-called "industrial" espionage" (Crane, 2005, p. 239).

There are several collection methods in competitive intelligence, Nasheri (2005, pp. 82-89) indicates some: Theft of Trade Secrets and Critical Technologies; Open-Source Collection; Unsolicited Requests for Information; Solicitation and Marketing Service; Acquisition of Export-Controlled Technologies, Joint Ventures, and Front Companies; Acquisition of Technology and Companies (often involving a third party, interested in the transaction); Exploitation of Visits to Companies, Commercial Markets, and Technology Transfers; Co-Opting of Former Employees and Cultural Commonalties; International exhibits, conventions, and seminars; Internet Activity (Cyberattack and Exploitation).

Of course, there are markets more vulnerable to industrial espionage than others. Any disruption in critical infrastructures such as electricity networks, transportation, and financial services maintained on huge networks linked to the internet might have an immediate and far-reaching impact, affecting citizens' well-being, safety, and security. There is an increasingly greater need for awareness among the public and public sectors to encourage them to collaborate to improve and expand standardized actions (both preventive and reactive) for a better response to a cyber incident (Parn & Edwards, 2019). Stewin and Bystrov (2013), provide us with a common issue of industrial

espionage. Cyber attackers are always looking for new ways to obtain information through stealthy attacks on companies' hardware, one of them is called DMA (Direct Memory Access) malware, which is malware that runs on dedicated hardware and uses DMA to undertake covert assaults against the host. The creation of decentralized blockchain technology may be the apex of cyber security research advances in cryptocurrency. Parn and Edwards (2019) defend that blockchain technology provides an innovative and safe means of storing information, executing data transactions, performing tasks, and establishing trust, making it suited for sensitive digital infrastructure data.

Thonnard et al. (2012), introduced the term "targeted attacks", which are a group of attacks based on e-mails containing malicious payloads, known in the field as "campaigns". These attacks are on a much smaller scale than other hostile, non-targeted activities carried out on a far wider scale. These campaigns can either target a single (type of) organization or a group of organizations with a shared purpose. Although successfully targeted assaults are still uncommon in comparison to traditional, profit-driven malware operations, they may be exceedingly disruptive. The victims of targeted assaults are intentionally chosen by the perpetrator. There are several factors behind the perpetrator's selection: the attacker may think that the attacked persons have access to high-value information; the compromised systems can be used to launch attacks against additional high-value systems or individuals; the malware differs from that used in non-targeted assaults and is generally more sophisticated. Thonnard et al. (2012), also observed that targeted attacks are divided into phases: incursion (tentative to penetrate a network), discovery (evaluation and spread through the network), capture (spread in additional computers), and data exfiltration (download of data and tools).

From an economic standpoint, the short-term consequences of many high-profile cyber-attacks are rather minimal when compared to the implications of lower-profile attacks with longer-term consequences, such as intellectual property theft (Andrijcic & Horowitz, 2006). We have already discussed some of the consequences when a company engages in industrial espionage, Andrijcic and Horowitz (2006) also approached which sectors would cause the greatest impact on the economy if they were to suffer an intellectual property theft. They first began by accessing which were the top ten sectors in the United States economy and assumed that all sectors have the same probability of suffering an intellectual property theft: Computer and electronic products

manufacturing; Motor vehicle, body, trailer, and parts manufacturing; Chemical manufacturing; Machinery manufacturing; Fabricated metal products manufacturing; Publishing, including software; Plastics and rubber products manufacturing; Primary metal manufacturing; Nonmetallic mineral products manufacturing. They first began by accessing which sectors would have the greatest consequences on the economy, concluding that Computer and electronic product manufacturing; Motor vehicle, body, trailer, and parts manufacturing; and Publishing, including software sectors would be the top three with the greatest impact on equity loss in the US economy. Secondly, they accessed which industries would suffer the greatest indirect impacts if they suffered an intellectual property theft, discovering that besides the sector in which the industry is allocated, the other sector may recur in indirect losses as well, as a result of the intellectual property theft. And finally, they estimated, out of a total of 59 major sectors (indicated in the US Bureau of Economic Analysis (BEA)), which ten sectors would experience the highest indirect revenue losses if any of the 59 sectors suffered an attack, concluding that the sector that would suffer the biggest amount of indirect losses, would be the Professional Scientific and Technical Services sector, followed by the Real Estate sector, and in third place would the Administrative and Support Services sector.

Crane (2005) recalls that, even though it may look like that as the markets evolve, there will always be incentives for companies to engage in industrial espionage, however, he suggests that instead of looking at competitors as enemies, companies should consider them as “allies” and therefore, be able to discern shared interests and interconnected flows of resources and rewards. Additionally, engaging in industrial espionage may be harmful to the company, in the long term, as stakeholders might lose trust and interest in the company.

3.3 Data Selection and Analysis

Bibliometric analysis is an excellent tool for summarizing and synthesizing literature. However researchers must consider these limitations when doing this analysis: bibliometric data from scientific databases are not developed solely for bibliometric analysis and may contain inaccuracies; the nature of the bibliometric methodology is in itself a limitation, because the bibliometric analysis is quantitative in nature, and the link between quantitative and qualitative outcomes is frequently ambiguous; bibliometric studies can only provide a short-term projection of the research field,

academics should avoid making overly optimistic claims about the research field and its long-term influence (Donthu et al., 2021).

A bibliometric review follows three steps (Santos et al., 2014):

1. Choice of the database and the main criteria to be used for the collection of data
2. Data collection in the database
3. Representation and analysis of data.

Out of the databases already mentioned in chapter 3.1 (Web of Science, Scopus, and Google Scholar Metrics), the data for this research was retrieved from the Scopus database, considering, for this dissertation, the one that is more complete and with has more valid data. As a research criterion, we use the keywords "*industrial espionage*", OR "*corporate espionage*", OR "*commercial espionage*" and look for them not only in titles but also in the abstract and keywords. As a methodological approach, there was a bibliometric search in articles and review articles only, all editorial materials, such as book chapters, articles in events, conferences, and others, were excluded from the survey. We carry this research out in the time frame between 1969 and 2021, analyzing documents that display the terms selected for this research.

After applying the research filters, the results presented a total of 660 publication records, 160 document results, and 500 secondary documents, which are publications that are cited by primary articles (citation within citation) (Poje & Groff, 2021). These are not included in this research; however, more research might be conducted by including those groups of records as well.

Subsequently, the 160 documents were analyzed and worked on in the Bibliometrix R-Studio analysis tool, allowing the elaboration of the thematic map of industrial espionage sub-topics; the analysis of the evolution of the terms throughout the years, through an evolution map; the analysis of the co-citation maps among papers, authors, and journals; and finally, the elaboration of an analysis confirming Lotka's law.

4. Results and discussion

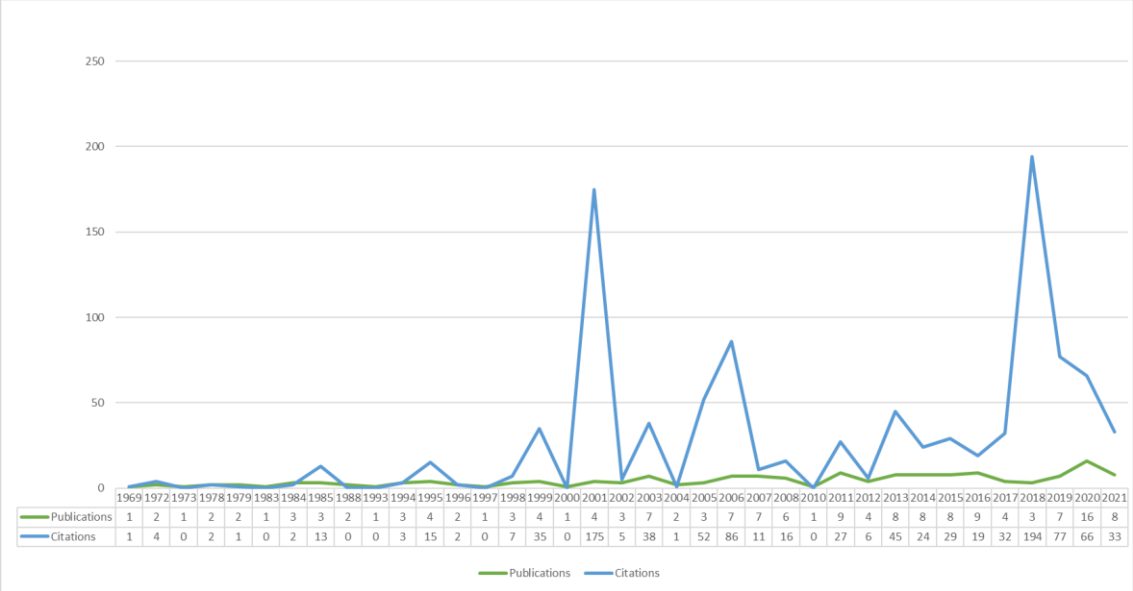
4.1 Descriptive Research

By focusing on a total of 160 documents we acknowledge that there is a small number of articles, yet there is an upward trend of published papers per year. The range of papers was published in a total of 139 journals, by 256 authors that possess affiliation to a total of 151 institutions in 35 countries. For the total of articles, 1021 references were considered, with an average of approximately 6 references per article. The results are presented in Table 1.

Table 1 *General bibliometric survey results (1969-2021)*

Bibliometric data	Quantity
Publications	160
Journals	131
Authors	256
Institutions affiliated	151
Countries	35
Citations	1021
Citations average	6.381 (1021/160)

The evolution of publications and, mainly, of citations made concerning the theme of industrial espionage, is not uniform. Until 2002, the publications on the theme did not exceed a total of four per year. As of 2003, there is a slight increase in the number of published papers. The year 2020 was, clearly, the period with the highest number of publications, with a total of 16 publications, followed by the years 2011 and 2015 with 9 publications each. This growth in the number of publications since 2002 can be attributed to the 9/11 event. Since that period there has been a greater awareness of industrial espionage. In addition, the Snowden case (mentioned above) since 2013 has also contributed to a greater study of this field.



	1969	1972	1973	1978	1979	1983	1984	1985	1988	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Publications	1	2	1	2	2	1	3	3	2	1	3	4	2	1	3	4	1	4	3	7	2	3	7	7	6	1	9	4	8	8	8	9	4	3	7	16	8
Citations	1	4	0	2	1	0	2	13	0	0	3	15	2	0	7	35	0	175	5	38	1	52	86	11	16	0	27	6	45	24	29	19	32	194	77	66	33

Citations are used to assess impact and influence (Donthu et al., 2021). It can be observed that there are a large number of citations for a small number of papers, for example, in the year 2018, the year with the highest number of citations, there were a total of 194. The second-highest number of citations was in 2001, with a total of 175.

Figure 1 *Annual distribution of publications on industrial espionage (1969-2021)*

As mentioned in the previous chapter, for this review, we are only considering the documents of type article and review. In the entire group of documents, the majority are article papers, representing 87% of the total 160 articles, the remaining are review articles (Figure 2). Regarding the language in which the article was published (Table 2), the English language is predominant, with a total of 90% of all publications, followed then by the German language. There were also publications in other languages, such as Italian and Russian, with two publications each, and finally Spanish, with only one publication.

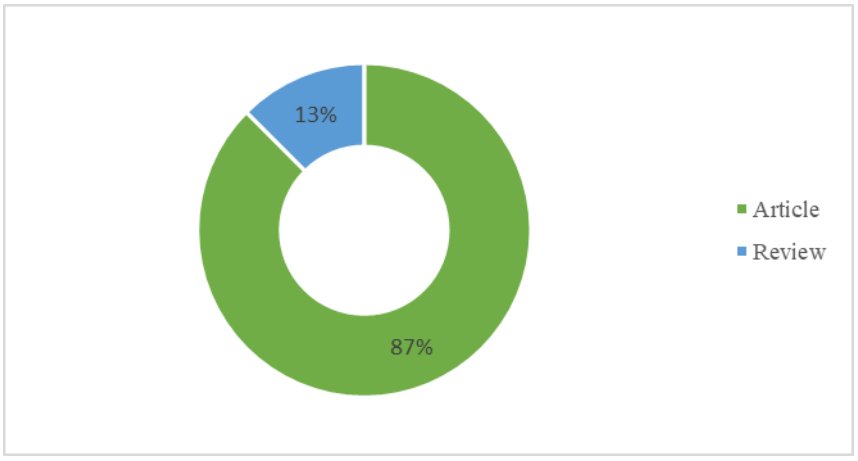


Figure 2 *Representation of the division of document type*

Table 2 *Language of original document distribution*

Language	Publications
English	144
English; German	3
German	7
Italian	2
Russian	2
Spanish	1
(blank)	1
Total	160

Table 3 identifies the most representative journals on the subject of industrial espionage. We analyzed 131 journals indexed in Scopus, taking into account the number of articles published, and the total number of citations in the database, we analyzed the top 20 journals. Of these 131 journals, we found that concerning the journal with the largest number of published articles in *Computers and Security*, with a total of 7 articles until 2021, has, however only 18 citations, which can denote a very low-impact journal, for example, compared to the *Journal of Manufacturing Systems*, which is only one paper obtained 194 citations, followed by *Lecture Notes in Computer Science*, with 151 citations in 3 journals, which can be considered to have a higher impact, with a citation/publication ratio of 50.33.

Table 3 *Journals with more citations on industrial espionage (1969-2021)*

Journal	Publications	Citations	Ratio (Cit./Publications)
Journal of Manufacturing Systems	1	194	194
Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intel	3	151	50.33333333
Engineering, Construction and Architectural Management	1	69	69
Business Horizons	1	47	47
Risk Analysis	1	42	42
Journal of Business Ethics	2	39	19.5
Sensors (Switzerland)	1	38	38
Journal of Economic Growth	1	26	26
Macroeconomic Dynamics	1	21	21
Computers and Security	7	18	2.571428571
Expert Systems with Applications	1	17	17
Water (Switzerland)	1	17	17
Journal of Workplace Learning	1	16	16
Journal of Legal, Ethical and Regulatory Issues	3	14	4.666666667
Computer Fraud and Security	5	13	2.6
Proceedings of the National Academy of Sciences of the United States of America	1	13	13
Ecologist	1	12	12
Industrial Archaeology Review	1	11	11
Games and Economic Behavior	1	10	10
International Journal of Critical Infrastructure Protection	1	10	10

Regarding the publishing authors, for a total of 160 documents, there were a total of 256 authors involved (Table 1). In Table 4 are found the 10 authors with more publications in the period in question. We find that Lee, with 3 publications, is the author with the highest number, followed by a series of authors with 2 publications each. In Table 5 are the 10 most cited authors. In this case, the author with more publications, Lee, is not on the list, having only 6 citations total. Both Hailes and Tuptuk (co-authors) are on the top, being that they have 2 publications and a total of 211 citations. These two co-authors are followed by other two pairs of co-authors with the next largest number of citations, published in 2001 and 2019.

Table 4 *Authors with more publications (1969-2021)*

Author	Publications	Year	Citations	Country
Lee C.-M.	3	2013;2014;2015	6	South Korea
Hailes S.	2	2018;2021	211	United Kingdom
Tuptuk N.	2	2018;2021	211	United Kingdom
Cozzi G.	2	2001;2006	47	Italy
Barrachina A.	2	2014;2021	14	Spain
Tauman Y.	2	2014;2021	14	Israel; United States
Urbano A.	2	2014;2021	14	Spain
Hassan H.	2	2020;2020	3	Malaysia
Jalil J.A.	2	2020;2020	3	Malaysia
Lee J.	2	2020;2021	3	South Korea

Table 5 *Most cited authors (1969-2021)*

Author	Publications	Year	Citations	Country
Hailes S.	2	2018;2021	211	United Kingdom
Tuptuk N.	2	2018;2021	211	United Kingdom
Jakobsson M.	1	2001	147	United States
Wetzel S.	1	2001	147	United States
Edwards D.	1	2019	69	United Kingdom
Parn E.A.	1	2019	69	United Kingdom
Cozzi G.	2	2001;2006	47	United Kingdom
Crane A.	1	2005	47	United Kingdom
Andrijcic E.	1	2006	42	United States
Horowitz B.	1	2006	42	United States

In Table 6 are the top 10 most cited articles, in the first position we find Tupnuk and Hailes's paper, with a total of 194 citations and a yearly average of times that it has been cited of 38.8. By comparing these results to Table 3, it is found that this paper allows the *Journal of Manufacturing Systems* to be the source with more citations. We can also assess that the paper in the second position, by Jakobsoon and Wetzel, has a total of 147 citations, allowing the journal *Lecture Notes in Computer Science* to be the second most cited journal, which has a total of 151 citations.

Table 6 *Most cited documents (1969 - 2021)*

Paper	Authors	Year	Journal	Total Citations	TC Per Year	Normalized TC
Security of smart manufacturing systems	Tuptuk N., Hailes S.	2018	Journal of Manufacturing Systems	194	38.8	3
Security weaknesses in bluetooth	Jakobsson M., Wetzel S.	2001	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in	147	6.68	3.36
Cyber threats confronting the digital built environment: Common data environment vulnerabilities and block chain deterrence	Parn E.A., Edwards D.	2019	Engineering, Construction and Architectural Management	69	17.25	6.27
In the company of spies: When competitive intelligence gathering becomes industrial espionage	Crane A.	2005	Business Horizons	47	2.61	2.71
A macro-economic framework for evaluation of cyber security risks related to protection of intellectual property	Andrijcic E., Horowitz B.	2006	Risk Analysis	42	2.47	3.42
Lightweight authentication protocol for M2M communications of resource-constrained devices in industrial internet of things	Lara E., Aguilar L., Sanchez M.A., García J.A.	2020	Sensors (Switzerland)	38	12.67	9.21
Inventing or spying? Implications for growth	Cozzi G.	2001	Journal of Economic Growth	26	1.18	0.59
Corporate Espionage and Workplace Trust/Distrust	Chan M.	2003	Journal of Business Ethics	23	1.15	4.24
Intellectual appropriability, product differentiation, and growth	Cozzi G., Spinesi L.	2006	Macroeconomic Dynamics	21	1.24	1.71
A systematic review of the state of cyber-security in water systems	Tuptuk N., Hazell P., Watson J., Hailes S.	2021	Water (Switzerland)	17	8.5	4.12

In order to visualize the representativeness of the countries of origin of the institutions of affiliation of the 256 authors of the 160 papers mapped in this bibliometric study, we identified the countries with scientific production on the subject of industrial espionage,

which can be seen in Figure 3. It can be confirmed in Table 7 that there is a large advantage for the United States in terms of the number of publications, with a total of 39 by 2021, followed by the United Kingdom with 19 publications. There are also 47 publications in which the origin country of the publication is not defined. Although the United States has more publications, by analyzing Table 5 the United Kingdom has more relevance and prominence in the community since it has a considerably higher number of citations than the United States.

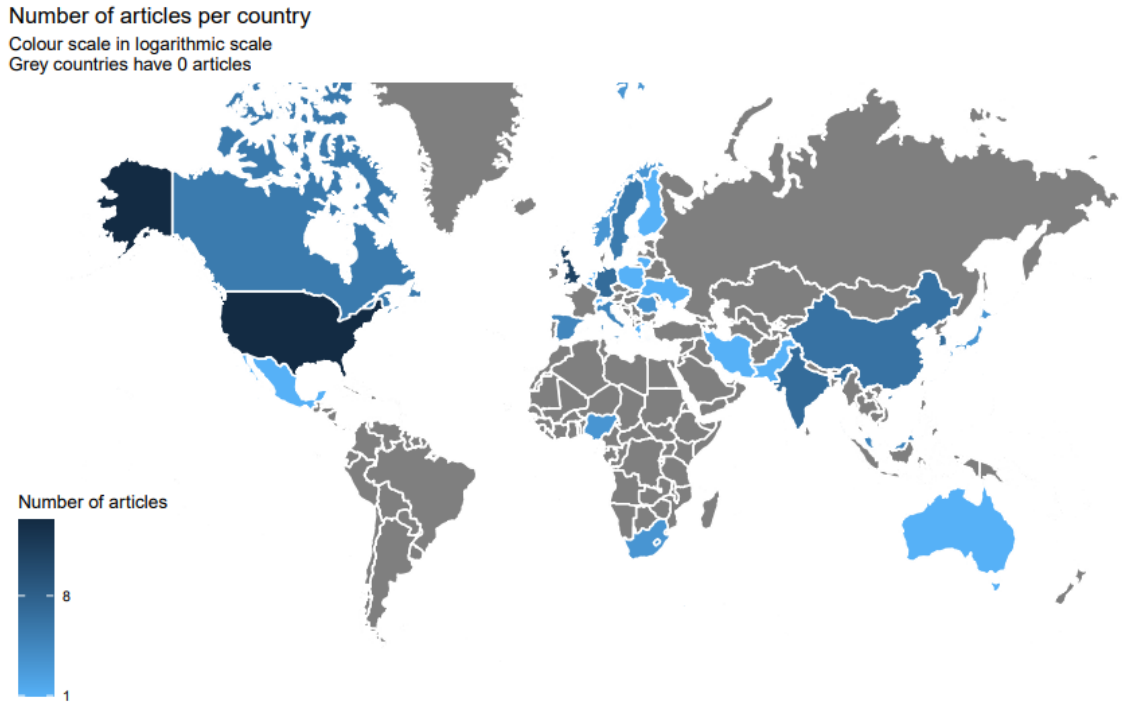


Figure 3 *Representation of publications per country (1969 - 2021)*

Table 7 *Number of publications per country of origin of the authors' institutions (1969 - 2021)*

Country	Number of publications	Country	Number of publications
United States	39	**Taiwan**	2
United Kingdom	19	**Australia**	1
Germany	12	**Cuba**	1
India	6	**Cyprus**	1
South Korea	6	**Finland**	1
Sweden	4	**Greece**	1
Canada	3	**Hong Kong**	1
China	3	**Iran**	1
Italy	3	**Lithuania**	1
Malaysia	3	**Mexico**	1
Spain	3	**Netherlands**	1
Belgium	2	**Pakistan**	1
Israel	2	**Poland**	1
Japan	2	**Russian Federation**	1
Nigeria	2	**Singapore**	1

Norway	2	**Switzerland**	1
Romania	2	**Ukraine**	1
South Africa	2	**Undefined**	47

About the research areas in which the papers under analysis fall, there are a total of 21 areas in which they can be found. Among all of them, in Figure 4 is shown that the Social Sciences area is the most predominant, with 59 associated papers, followed by the Computer Science area, with a total of 50 publications, and the Engineering area in third place, with 40 publications in all.

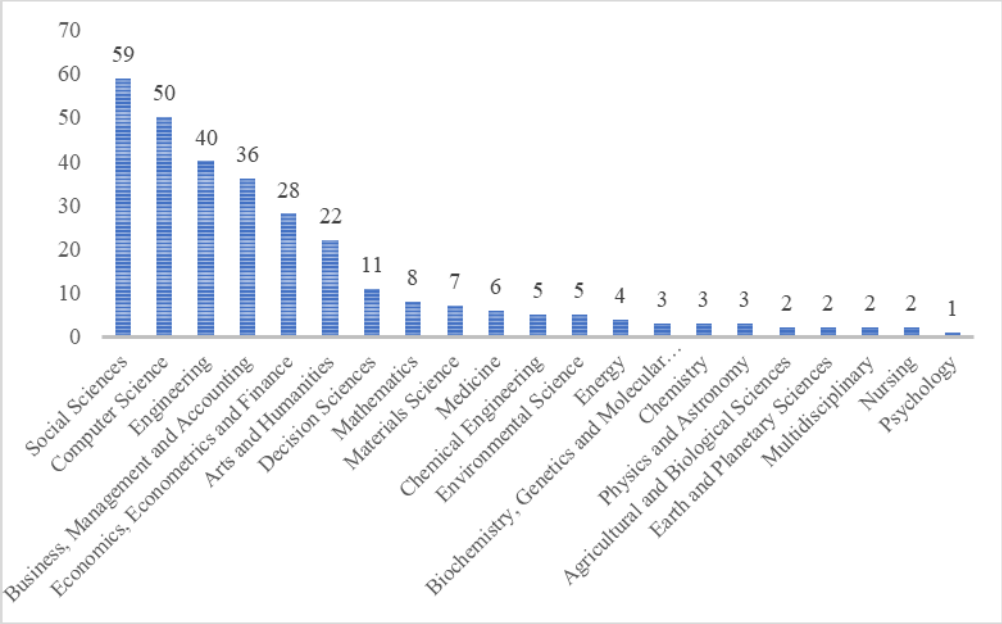


Figure 4 *Published documents per research area*

Analyzing the thematic map regarding industrial espionage (Figure 5), a few emerging topics can be extracted. A thematic map allows the visualization of four different types of cluster themes based on density (y-axis) and centrality (x-axis). Cluster categorization provides a more precise description of the condition of a particular network, as well as the position and degree of development of the clusters that comprise it (Callon et al., 1991). Starting then, by the cluster of type one, identified in the Figure as *Motor Themes,* some themes can be identified as central and with a great degree of development, indicating that cluster topics such as *Intellectual Property* or *Computer Crime* are "core" (Callon et al., 1991, 166) topics for researchers when conducting a study on industrial espionage. Following to the second quadrant, identified as *Basic Themes,* are the topics that are considered relevant, however, are not objects of

significant investigation yet. Although this quadrant is mostly empty (indicating that, currently there are no relevant and undeveloped themes), we can see the cluster *Personal Computing* and *Mobile Devices* appearing in this quadrant, which might suggest an interesting topic for further research. On the third quadrant, classified as *Niche Themes*, we find topics that are currently generating "less and less interest" (Callon et al., 1991, 166) than they were previously, and might have already been studied. In this quadrant, we find only one cluster, regarding *Artificial Intelligence*. On the fourth quadrant (identified as *Emerging or Declining Themes*), are the cluster topics that are not very relevant nor developed, representing the margins of the map. In this quadrant, we find the themes of *Competitive Advantage* and *Cyber-attacks*.

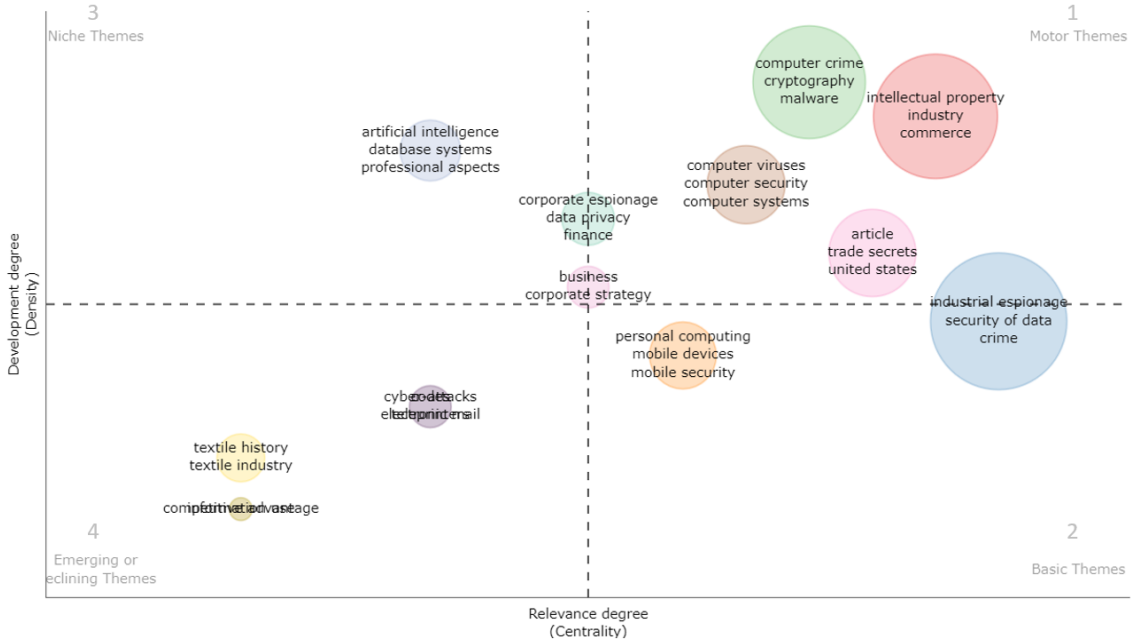


Figure 5 *Thematic map on industrial espionage*

4.2 Network Analysis

Being our research conducted over 6 decades, it is to be expected that terms and topics studied over time evolve or suffer some changes. To do that analysis, it is applied a thematic evolution map. An evolution map is a method used to evaluate the thematic progression of a study topic through time, dividing it by periods. Indicating the evolution of a thematic area over different subperiods (Cobo et al., 2011). In the period of analysis, from 1969 to 2021, there were identified two subperiods (Figure 6), from 1969 to 2011, and from 2012 onwards. The sizes of endpoints in each timeframe

reflect the number of papers within each subject. The thickness of connecting lines demonstrates the similarity of terms shared by themes across time (Cobo et al., 2011). So, by analyzing this Figure we acknowledge that until 2011, the terms *intellectual property* and *industrial espionage* represent the largest number of publications within that period, contrasting to the second period, the term *industrial espionage*, as well as competition, are the most prominent. Being the latest, along with *commerce* and *blockchain* the division of the term *intellectual property*, from the first period. Finally, it is worth referring that the terms *competition* and *industrial espionage* are the only terms that are present in both subperiods.

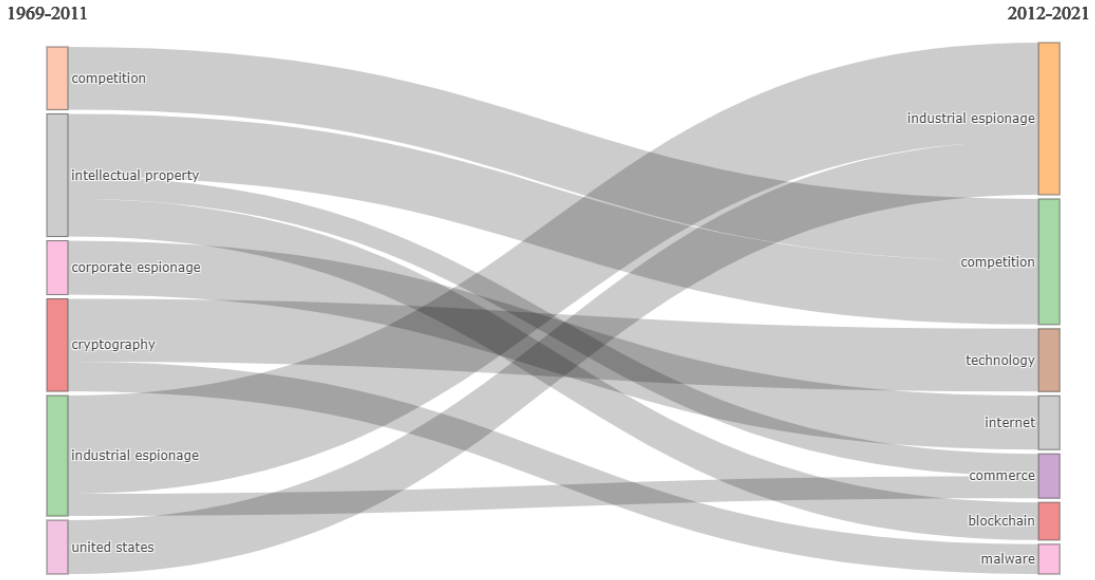


Figure 6 *Thematic evolution map (1969 – 2021)*

Figures 7, 8, and 9 demonstrate the analysis of the co-citation network. This is used to determine the connection between similar nodes, in this case, publications, authors, and journals respectively. Co-citation analysis aims to track pairs of nodes that are referenced together in the source articles. When the same pair of publications are referenced together research clusters emerge (Small, 1973; Surwase et al., 2011). The thickness of the lines connecting each node represents the strength of co-citation between each pair (Small, 1973). In our analysis in Figure 7, there were identified 33 papers and 7 clusters. Also, the minimum number of edges in this study is 1, meaning that if there is no line connecting a pair of publications, there is no co-citation at all between the pair. In this case, we identified 3 clusters connected between them,

indicating that there is co-citation in at least one pair of publications between the 3 clusters. The remaining 4 clusters are isolated, signifying that there is no correlation of references between them.

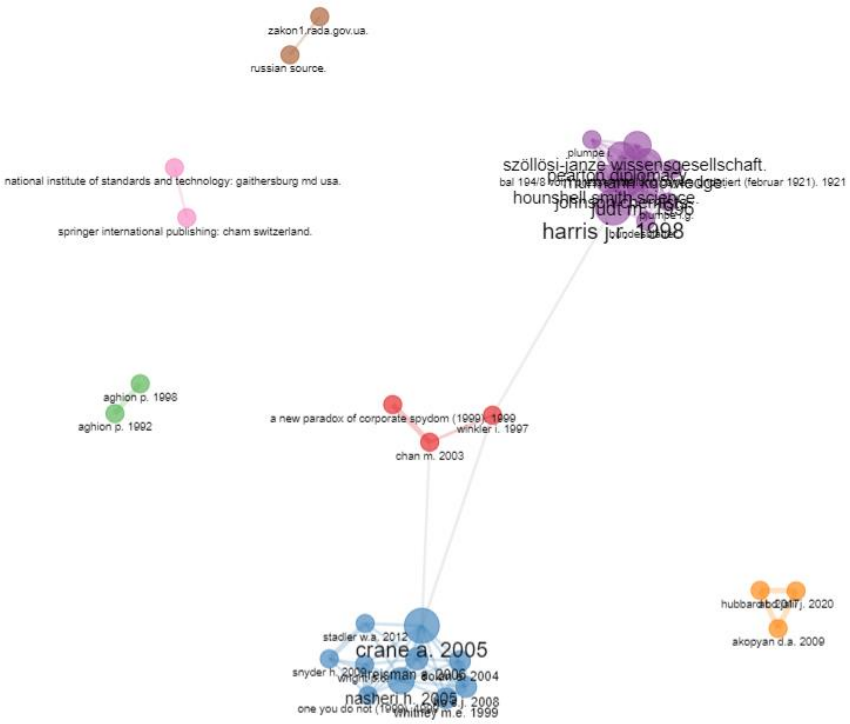


Figure 7 *Papers Co-citation network map*

In Figure 8, we analyze the author's co-citation network, identifying 47 nodes and 12 clusters. The 5 authors more co-cited are Nasheri, H. (United Kingdom); Harris, J. H. (United Kingdom); Lee, C.M. (South Korea); Jones, A (United Kingdom); and Crane, A (United Kingdom).

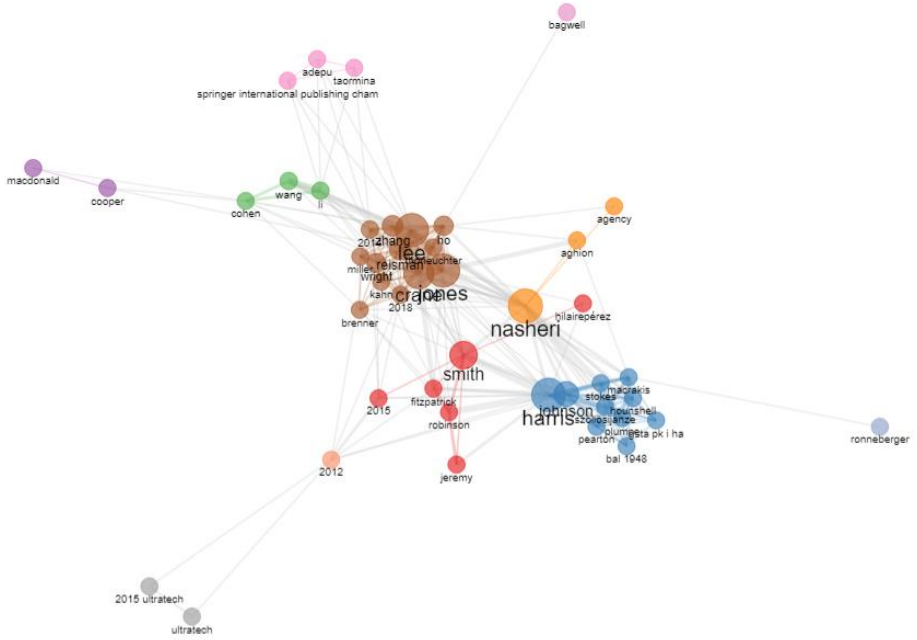


Figure 8 *Authors Co-citation network map*

Finally, in Figure 9, we find the analysis of the journal's co-citation map. Here we find that in the total of 5 clusters, all have at least one pair of journals that are cited together. In the journals co-citation network map, there are 44 journals co-cited, and the top 5 cited journals are Business Horizons, Communications of the ACM, Journal of Business Ethics, Econometrica, and Security Journal.

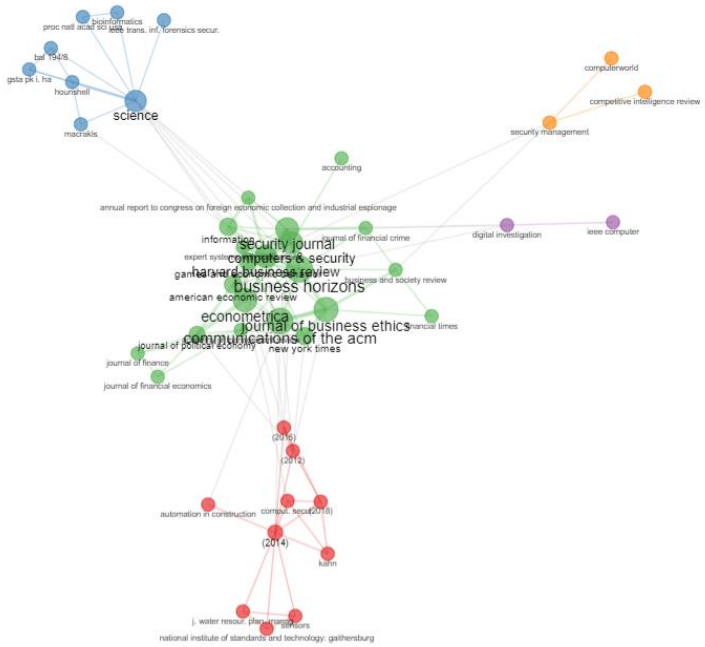


Figure 9 *Journals Co-citation network map*

By last, we did an analysis to confirm Lotka's Law (Figure 10), assessing that the number of publications is inversely proportionate to the number of authors. Concluding that 93% of the authors made only one publication, 5.9% made two publications, and the remaining made three or more publications concerning industrial espionage.

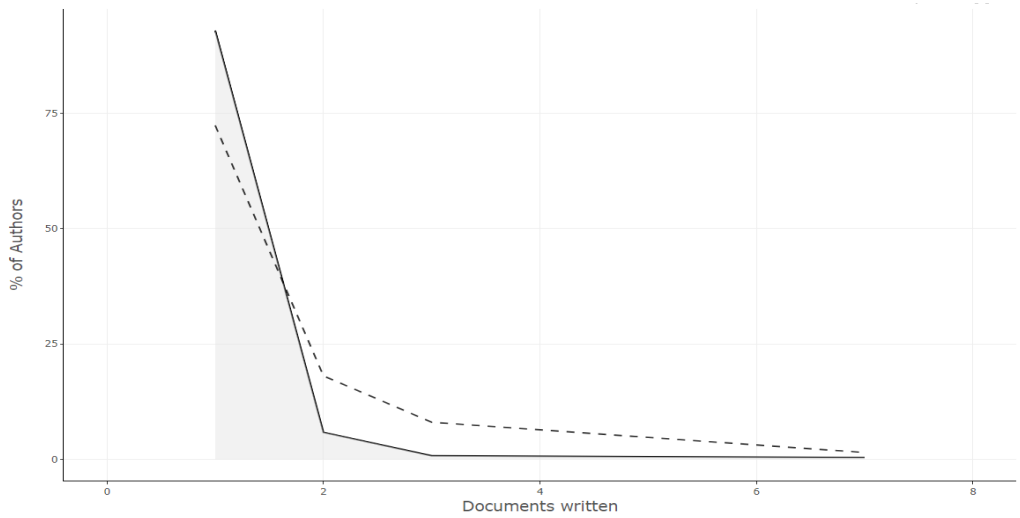


Figure 10 *Lotka's Law representation*

Conclusion

Papers have the ability to legitimize knowledge and have a significant impact on science regulation, appraisal of its development, resource allocation, and career opportunities (Vilaça, 2018). Bibliometrics takes place when doing a quantitative analysis of published papers. Bibliometric approaches are an essential component of research assessment methodology, particularly in scientific and practical subjects (Ellegaard & Wallin, 2015).

This research study has the goal of providing a bibliometric analysis of industrial espionage by answering the question "*How has industrial espionage been studied in the scientific literature?*". We first began by identifying several definitions of espionage and then focused on defining industrial espionage itself, finding that it is still, an umbrella definition, and is still under discussion in the scientific community. We then did a brief systematic literature review on the subject of industrial espionage. Bearing in mind that this issue was already approached by Hou and Wang (2020), our goal was to complement their work previously done, by quantifying the research on this topic. We finally analyzed and explained the existing scientific production on industrial espionage.

Our analysis was elaborated based on the Scopus database, being that, amongst the remaining, is the most complete. Our interpretation of the data was then made using mainly Bibliometrix R-tool software.

After our analysis, we acknowledge that there is still a small number of publications on the subject of industrial espionage, however, since 2012 there is a growing tendency in scientific production regarding this topic. It is found that during the analyzed period, between 1969 and 2021, the most productive journal was the *Computers and Security* journal, with a total of 7 published papers, however, the journal with the greatest impact is the *Journal of Manufacturing System*, with the biggest number of citations. As far as authors are concerned, the most published author, until 2021 is Lee, from South Korea, with a total of 3 papers published. We also analyzed the most productive countries, concluding that the United States is the most representative country as a total of papers published concerned. The United Kingdom comes in second place, as the most representative, and in the first place, as the country with the highest number of cited authors.

We found that the topic of industrial espionage is being published in a wide variety of journals, regarding diverse research topics, indicating that the authors may be from several different backgrounds, and, although they are related to the topic of industrial espionage, they are concerning scattered research areas, and there is no uniformity concerning the topics in the journals. This may also be reinforced by the fact that there is a low level of co-citation, especially among the various analyzed papers.

Even though the subject of industrial espionage is not new, there is still a big gap when it comes to its study. And even though we only considered published articles and review articles in our analysis, reducing our sample to a total of 160, the total documentation found on this topic is still very scarce. In this sense, we consider that, not only it would be beneficial to do a similar study on all documentation regarding industrial espionage, including, for instance, book chapters, and conferences, among others; we also believe that, with the increase of research arising on this subject, it would be interesting to do similar research in the future, with a more significant amount of data. Additionally, as we have identified on the thematic map (Figure 5), the subtopics of Personal Computing, Mobile Devices, and Mobile Computing are still understudied in the field of industrial espionage, which might suggest a relevant topic of research.

Bibliography

Adel, M. (2021). *Zipf's law applications in patent landscape analysis.* World Patent Information. Volume 64, March 2021, 102012

Aid, M. (2003). *All Glory is Fleeting: Sigint and the Fight Against International Terrorism*, Intelligence and National Security, Volume 18(4), Pages 72-120

Ambrose, B. (1990). *An Analysis of the Factors Affecting Light Industrial Property Valuation.* Journal of Real Estate Research, Volume 5(3), Pages 355-370

Andrijcic, E. & Horowitz, B. (2006). *A macro-economic framework for evaluation of cyber security risks related to protection of intellectual property*. Risk Analysis. Volume 26(4), Pages 907-923

Arsenova, I. (2013). *The 2nd International Conference on Integrated Information. New application of bibliometrics.* Procedia - Social and Behavioral Sciences, 73(2013), Pages 678-682

Badiru, A. B. & Maloney, A. E. (2016). *A Conceptual Framework for the Application of Systems Approach to Intelligence Operations: Using HUMINT to Augment SIGINT*. American Intelligence Journal, Volume 33(2), Pages 41-46

Ball R. (2018). *An Introduction to Bibliometrics: New Development and Trends*. United States: Chandos Publishing (2018)

Barrachinas, A. & Forner-Carreras, T. (2022). *Market must be defended: The role of counter-espionage policy in protecting domestic market welfare.* Information Economics and Policy. Volume 58(2022), 100964

Bergier, J. (1975). *Secret Armies: The Growth of Corporate and Industrial Espionage*, Indianapolis: Bobbs-Merrill Co., Inc.

Boulouard Z, Koutti L, Chouati N, Haddadi A, Dousset B, Haddadi A, Bouhafer F. (2016) *Visualizing large Figures out of unstructured data for competitive intelligence purposes.* In: Proceedings of SAI Intelligent Systems Conference (IntelliSys). Springer; 2018. Pages 605-626

Bradford, S. C. (1976). *Sources of Information on Specific Subjects*. Collection Management, Volume 1(3-4), Pages 95-104

Budd, J. (1988). *A bibliometric analysis of higher education literature*. Agathon Press, Inc. Volume 28(2), Pages 180-190

Button, M. (2019). *Editorial: economic and industrial espionage*. Security Journal (2020) Volume 33, Pages 1-5

Callon, M.; Courtial J. P. & Laville, F. (1991). *Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry.* Scientometrics, Volume 22(1), Pages 155-205

Carbaugh, B. & Wassell, C. (2019). *Forced technology transfer and China.* Economic Affairs, Volume 39(3), Pages 306-319

Christie, G.C. (1972). *Government surveillance and individual freedom: A proposed statutory response to Laird v. Tatum and the broader problem of government surveillance of the individual.* N.Y.U. Law Review. Volume 47, Pages 871-902

Chueke, G. & Amatucci, M. (2015). *O que é bibliometria? Uma introdução ao Fórum.* Internext, Volume 10(2), Pages 1-5

Cobo, M.J., López-Herrera, A.G., Herrera-Viedma, E. & Herrera, F., (2011). *An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field.* Journal of Informetrics, Volume 5 (2011), Pages 146-166

Crane, A. (2005). *In the company of spies: When competitive intelligence gathering becomes industrial espionage*. Business Horizons, Volume 48(3), Pages 233-240

Danielson, M. (2009). *Economic Espionage: A Framework for a Workable Solution*. J.L. SCI. & TECH. 2009; Volume 10(2), Pages 503-548

Delupis, I. (1984). *Foreign Warships and Immunity for Espionage.* The American Journal of International Law, Volume 78(1), Pages 53-75

Diamond Jr, A. (1986). *What is a Citation Worth?.* The Journal of Human Resources, Volume 21(2), Pages 200-215

Diffie, W. (2006). *Chattering about SIGINT.* IEEE Security & Privacy, January/February 2006, Volume 4, Page 9

Donthu N., Kumar S., Mukherjee D., Pandey N., Lim W. *How to conduct a bibliometric analysis: An overview and guidelines*. Journal of Business Research, Volume 133(2021), Pages 285-296

Egghe, L. (2010). *The Hirsch index and related impact measures*. Annual Review of Information Science and Technology, Volume 44(1), Pages 65-114

Ellegaard, O. & Wallin, J.A. (2015). The bibliometric analysis of scholarly production: How great is the impact? Scientometrics, Volume 105, Pages 1809-1831

Fortune, R. (1847). *Three years' wonderings in the Northern Provinces of China, including a visit to the tea, silk, and cotton countries: with an account on the agriculture and horticulture of the Chinese, new plants, etc.* (2nd edition). London: John Murray, Albemarie Street

García-Villar, C. & García-Santos, J. (2021). *Bibliometric indicators to evaluate scientific activity*. Radiología (English Edition) Volume 63(3), Pages 228-235

Glitz, A. & Meyersson, E. (2017). *Industrial Espionage and Productivity*. Research Policy, IZA DP, Volume 10816, Pages 2-10

Grabiszewski, K. & Minor, D. (2018). *Economic Espionage*. Defence and Peace Economics. Volume 30(3), Pages 269-277

Halleck, H. W. & Davis, G. D. (1911). *Military Espionage.* The American Journal of International Law, Volume 5(3), Pages 590–603

Heims, P. A. (1982). *Countering industrial espionage*. England: 20th Century Security Education Ltd

Herken, G. (2009). *Target Enormoz: Soviet Nuclear Espionage on the West Coast of the United States, 1942–1950*. Journal of Cold War Studies, Volume 11(3), Pages 68-90

Ho, S. J. (2008). *Extracting the information: espionage with double crossing*. Journal of Economics, Volume 93(1), Pages 31-58

Hobhouse, H. (1986). *Seeds of Change - Six Plants That Transformed Mankind* (1st edition). New York: Harper & Row

Hollins, S. C. (2020). *A Dark History of Tea* (1st edition). Great Britain: Pen & Sword Books Ltd

Hou, T. & Wang, V. (2020). *Industrial espionage – A systematic literature review (SLR)*. Computers & Security. Volume 98, November 2020, 102019

Kitchenham B. & Charters S. (2007). *Guidelines for performing systematic literature reviews in software engineering.* School of Computer Science and Mathematics - Keele University, UK

Jesus, A. & Mendonça, S. (2018). *Lost in Transition? Drivers and Barriers in the Eco-innovation Road to the Circular Economy*. Ecological Economics. Volume 145, Pages 75-89

Lieber F. & Hartigan, R. S. (1863). *Lieber's Code and the law of war.* Chicago: Precedent

López-Robles, J.R., Otegi-Olaso, J.R., Gómez, I. & Cobo, M.J. (2019). *30 years of intelligence models in management and business: A bibliometric review.* International Journal of Information Management, Volume 48, 22-38

Lyra, M.S., Damásio, B., Pinheiro, F.L. & Bacao, F. (2022). *Fraud, corruption, and collusion in public procurement activities, a systematic literature review on data-driven methods.* Applied Network Science, Volume 7(83) (2022)

Macnish, K. (2016). *Government Surveillance and Why Defining Privacy Matters in a Post-Snowden World*. Journal of Applied Philosophy, Volume 35(2), Pages 417-432

Mair, V. H. & Hoh, E. (2009). *The True History of Tea (*1st edition). London: Thames and Hudson

Marklund, A. (2019). *Trawling the Wires: Mass Surveillance of Border-crossing Communication in Denmark during World War II*. Technology and Culture, Volume 60(3), Pages 770-794

Mathankar, A. (2018). *Bibliometrics: an overview.* International Journal of Library & Information Science (IJLIS) Volume 7(3), Pages 9-15

Mclllwain, J. S. (2005). *Intellectual property Theft and Organized Crime: The Case of Film Piracy. Trends in Organized Crime.* Volume 8(4), Pages 15-39

Mendonça, S., Damásio, B. Freitas, L.C., Oliveira, L., Cichy, M. & Nicita, A. (2022). The rise of 5G technologies and systems: A quantitative analysis of knowledge production. Telecommunications Policy. Volume 46(4), 102327

Moed, H., Bruin, R., Hederhof, A., Van Raan, A. & Tijssen, R., (1992). *State of the Art Bibliometric Macro-Indicators. An Overview of Demand and Supply*. Commission of the European Communities. Luxembourg: Office for Official Publications of the European Communities, 1992

Nasheri, H. (2005). *Economic espionage and industrial spying.* Cambridge: Cambridge University Press.

Norton, R. A. (2011). *Guide to Open Source Intelligence. A Growing Window into the World.* The Intelligencer, Volume 18(2), Pages 65-67

Noyons, E., Moed, H. & Luwel, M. (1999). *Combining mapping and citation analysis for evaluative bibliometric purposes: A bibliometric study.* Journal of the American Society for Information Science, Volume 50(2), Pages 115-131

Oxnevad, I. (2019). *Corporate Privateering and Economic Counter Espionage in U.S. Great Power Competition.* Orbis, Volume 63(3), Pages 391-405

Pan, X., Yan, E., Cui, M., Hua, W. (2018) *Examining the usage, citation, and diffusion patterns of bibliometric mapping software: A comparative study of three tools.* Journal of Informetrics, Volume 12(2), Pages 481-493

Parn, E. & Edwards, D. (2019). *Cyber threats confronting the digital built environment. Common data environment vulnerabilities and block chain deterrence*. Engineering, Construction and Architectural Management. Volume 26(2), Pages 245-266

Phua, C. (2009). *Protecting organizations from personal data breaches*. Computer Fraud & Security, Volume 2009(1), Pages 13-18

Podbregar, I., (2006). *Some Patterns of Industrial Espionage*. Journal of Criminal Justice and Security ISSN 1580-0253, 323-331

Poje, T. & Groff, M. Z. (2021). *Mapping Ethics Education in Accounting Research: A Bibliometric Analysis.* Journal of Business Ethics (2022) Volume 179, Pages 451-472

Potter, W. G. (1981). Introduction to Library Trends. Graduate School of Library and Information Science. University of Illinois at Urbana-Champaign, Volume 30(1). Pages 5-8

Prud'homme, D. & Zedtwitz, M. (2019). *Managing "forced" technology transfer in emerging markets: The case of China*. Journal of International Management Volume 25(3), 100670

Rose, M. (2019). *Commanding Spies*. Orbis, Volume 63(2), Pages 258-280

Rotolo, D., Camerani, R., Grassano, N. & Martin, B. R. (2022). *Why do firms publish? A systematic literature review and a conceptual framework.* Research Policy. Volume 51(10)

Santos, J., Kalsing, M. & Hansen, P. (2014) *Redes de cooperação interorganizacional: uma análise sistemática da produção científica na Web of Science de 1981-2013.* SEMEAD - Seminários em Administração, 17., São Paulo: FEAUSP, 2014, Pages 1-15

Sérgio, A. (1977). Breve Interpretação da História de Portugal (7th edition). Lisbon: Livraria Sá da Costa Editora

Sivanesan, G. (2011). *The human factor in espionage*. Computer Fraud & Security, Volume 2011(2), February 2011, Pages 15-16

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, Volume 24(4), Pages 265-269

Søilen, K. S. (2016). *Economic and industrial espionage at the start of the 21st century – Status quaestionis*. Journal of Intelligence Studies in Business Volume 6(3), Pages 51-64

Solan, E & Yariv, L (2004). *Games with espionage.* Games and Economic Behavior, Volume 47(1), Pages 172-199

Stewin, P. & Bystrov, I (2013). *Understanding DMA malware*. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Volume 7591 LNCS, Pages 21-41

Summer, S. (2016). *You: for Sale || The Snowden Revelations*. Protecting Your Personal Data and Privacy Online. Pages 17-48

Surwase, G., Sagar, A., Kademani, B.S. & Bhanumurthy, K. (2011). *Co-citation Analysis: An Overview.* Beyond Librarianship. Creativity, Innovation and Discovery (BOSLA National Conference proceedings). September 2011. Mumbai

Thonnard, O., Bilge, L., O'Gorman, G., Kiernan, S. & Lee, M. (2012). *Industrial espionage and targeted attacks: Understanding the characteristics of an escalating threat*. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Volume 7462 LNCS, Pages 64-85

Thorleuchter, D. & Van den Poel, D. (2013). *Protecting research and technology from espionage*. Expert systems with applications, Volume: 40 (9). Pages: 3432-3440

Toxen, B. (2014). *The NSA and Snowden*. Communications of the ACM, Volume 57(5), Pages 44-51

Vardi, L. (2001). *Book Reviews: Industrial Espionage and Technological Transfer: Britain and France in the Eighteenth Century. By J. R. Harris*. Business History Review, Volume 75(2), Pages 432-435

Vieira, P. & Wainer, J. (2013). *Correlações entre a contagem de citações de pesquisadores brasileiros, usando o Web of Science, Scopus e Scholar*. Perspectivas em Ciência da Informação, 18(3), Pages 45-60

Vilaça, M.M. (2018). *A publicação como obsessão, a pressão como efeito e a integridade como discurso/desafio: uma análise crítico-provocativa da cientometria vigente*. Motrivivência, Florianópolis, Volume 30(54), Pages 51-73

Wagner, R. (2012). *Bailouts and the Potential for Distortion of Federal Criminal Law: Industrial Espionage and Beyond.* Tulane Law Review, Volume 86, Pages 1017-1054

Waziri, K. M. & Yerima, T. F. (2011). *Industrial Espionage and Intellectual property Rights Protection: How Legal is it Legal*. International Journal of Sustainable Development, Volume 4(3)

Xiao, Y. & Watson, M. (2017). Guidance on Conducting a Systematic Literature Review. Journal of Planning Education and Research. 2019, Volume 39(1), Pages 93-112

Yip, K. L. (1999). Pleuridium japonicum newly reported from China. Cryptogamie, Bryol. Volume 20(4), Pages 255-256

Zhong, W. (2010). *The roles of tea and opium in early economic globalization: A perspective on China's crisis in the 19th century*. Front. Hist. China, Volume 5(1), Pages 86-105

Zuboff, S. (2019). *The age of surveillance capitalism. The fight for a human future at the new frontier of power* (Ed 1). New York. PublicAffairs

Zupic, I. & Cater, T. (2015). *Bibliometric Methods in Management and Organization.* Organizational Research Methods, Volume 18(3), Pages 429-472

References

CyberEdge Group. 2021 Cyberthreat Defense Report. Retrieved from https://www.isc2.org/-/media/ISC2/Research/Cyberthreat-Defense-Report/2021/CyberEdge-2021-CDR-Report-v10--ISC2-Edition.ashx [accessed 05.01.22]

Espionage. (2008). In *Oxford English Dictionary* (11th ed., p. 487). Oxford: Oxford University Press.

European Commission (2020). General Data Protection Regulation (GDPR). Programme of the European Union. Retrieved from https://gdpr.eu/article-4-definitions/ [accessed 10.10.22]

Spy. (2008). In *Oxford English Dictionary* (11th ed., p. 1399). Oxford: Oxford University Press.

Task Force on Intellectual property, 2004. *Report of the Department of Justice's Task Force on Intellectual property*. Washington, D.C.: United States Department of Justice. Retrieved from https://www.justice.gov/sites/default/files/olp/docs/ip_task_force_report.pdf [accessed 11.02.22]