POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

EXECUTIVE SUMMARY OF THE THESIS

COVID-19 and R&D Trends: Evidence from Patent Data.

TESI MAGISTRALE IN MANAGEMENT ENGINEERING – INGEGNERIA GESTIONALE

AUTHOR: FRANCESCO PERRONE

ADVISOR: CHIARA FRANZONI

ACADEMIC YEAR: 2021-2022

1. Introduction

Patents and patent analytics are widely considered to be instruments used by scholars and practitioners to have a proxy of innovation activities in technological fields and predict technological change (Hall et al., 2005; Pavitt, 1985). In this research, I will analyze patent data in the most innovative technological fields to study the R&D trends that have characterized these in the last 10 years (2012 – 2022). I focus especially on evidence of changes in the same trajectories that might have emerged during the two year period marked by the COVID-19 pandemic (2020 – 2022). The motivation behind this work, is to unravel the main technological trends and look at preliminary evidence of changes in which the pandemic could have played a role. The aim of the work is to provide evidence-based hypotheses on R&D efforts trajectory changes and possible reasons behind the different behaviors of these trajectories.

The hypotheses proposed in this thesis mainly concern the three different ways in which COVID-19 might have influenced innovation and patenting activity.

The first hypothesis deals with the patenting trends in which data shows a not-significant variations in their trajectory during the pandemic period (2020 – 2022).

The second hypothesis concerns patenting trends in which data shows acceleration in the trajectory during the pandemic period (2020 – 2022).

The third hypothesis deals with patenting trends in which data shows deceleration in the trajectory during the pandemic period (2020 – 2022).

In order to choose the set of emerging technologies to study, I have drawn from the *MIT Technology Review*, which every year lists the 10 breakthrough technologies that are most likely to impact businesses and consumers in the short-medium term. In this way, I analyzed a total number of 110 breakthrough technologies classified in 21 different technological fields.

This research highlights the discontinuities that the pandemic has created between previous and future necessities. In this regard, my work could be helpful on a practical perspective as it provides data analysis on research trends that could hint at those technological fields that are likely to receive greater interest form the market in the future. Secondly, in this work I provide a set of new hypotheses on the reasons why COVID-19 may have or not impacted R&D activity in certain technological fields. This could support innovation research since it supplies new hypotheses of technological change that may be relevant for future innovation studies.

2. Data and Methods

In this chapter, I will describe the main processes used to source, extract, encode, clean and re-classify the data gathered for my research. The final results of this procedure will be the technological field patent trends that will be analyzed in the *Results* chapter.

2.1. Data sourcing

To detect the patenting trends over the last 10 years to measure the effects of the COVID-19 pandemic, I decided to consult *MIT Technology Review* to identify radically innovative technologies for my work. Every year, the *MIT Technology Review* (MIT, 2022) lists 10 breakthroughs in technology that are likely to impact business and citizens daily life in the near future. I collected each of the 110 technologies mentioned in the *Review* from the year 2012 to the year 2022, providing the name of each technology and a brief description of it. I then dedicated the second step of the sourcing activity to seek patent information for each of the technology using the open-source patent research tool *Patentscope*, that allows advanced searches through the use of queries that support own field codes limiting the research to restricted fields and the use of Boolean and wildcards operators (

Table 1). In Equation (2.1. 1) an example of one of the research queries.

Table 1. Tables of used field codes.

Field Codes	Description
FP:	Search text or numeric information on the front page of a patent document.
EN_TI:	Search text or numeric information on the title of a patent document only in English language.
AD:	Allows to search for the application date in patent document.
DP: (…AND…)	Allows to search information during an indicated period of time marked by two dates (it has to be used concurrently with the Boolean operator AND).
IC:	Searches through indicated codes from the IPC classification.
CPC:	Searches through indicated codes from the CPC classification.

FP:(EN_TI: KEYWORD OR KEYW* OR "KEYWORD")
AND AD:(DP:(01.01.2012 TO 29.09.2022))

(2.1. 1)

2.2. Data extraction and encoding

I conducted the activity of data extraction concurrently to the activity of data encoding, using data from *Patentscope* and using *Microsoft Excel* as a database manager.

I collected four types of patent data in the preliminary patent search:

- *Patentscope* research: I recorded the query used to draw out the preliminary patent screening on a specific technology.
- IPC relevant codes: I recorded the first 4 digits of the *International Patent Cooperation* (IPC) codes that appeared in the patent documents resulting from the search, together with the amount of times they appear for each one of them. In particular, I only collected the ones that I deemed to be relevant and reflecting the description

of the specific technology, thus avoiding the collection of codes that were not connected to the technology description. The first 4 digits indicate (WIPO, 2022):

- First letter: Section (e.g., H = all inventions related to ELECTRICITY)
- Second and Third Number: Class (e.g., H01 = all inventions related to BASIC ELECTRIC ELEMENTS)
- Fourth letter: Subclass (e.g., H01S = all inventions related to DEVICES USING THE PROCESS OF LIGHT AMPLIFICATION BY STIMULATED EMISSION OF RADIATION [LASER] TO AMPLIFY OR GENERATE LIGHT)

- Year of first publication: I recorded the year in which the first patent publication appeared.
- Total number of patents found.

These data were easily collected thanks to the *analysis* interface available in *Patentscope*.

2.3. Data cleaning

After having extracted and encoded the patent data, I proceeded to refine the research and obtain a more fine-grained analysis from which then draw the patent trends. Fine-tuning the research was particularly useful for two reasons:

- Eliminate patent–outliers in a technology-specific patent landscape.
- Detect missing patents in a technology-specific patent landscape.

For this purpose, the process used to clean data collected was the one of data saturation (Faulkner & Trotter, 2017).
The first phase of the search aimed at minimizing omittances, by including all the potentially relevant technologies. By doing this, I aimed at reducing *type II* errors in the patent search, given that the research query in the extraction phase was excessively narrow and might have left out some "*false negatives*" patent documents from the end results. The search resulted in a very broad set of outcomes, that includes several technologies patented in a span of more than 10 years. The second phase of research aimed at minimizing the inclusion of non-pertinent patents, by means of a careful data cleaning process. In this phase, I aimed at reducing instead *type I* errors in the patent search, gradually offsetting the broadness of the research to exclude "*false positives*" patent documents. I cleared out data until saturation by gradually adding in the research query keyword terms that could identify the specific technology at hand (Equation (2.3. 1)).

AD:(DP: DP:(01.01.2012 TO 29.09.2022)
AND IC:("X12Y" OR ...)
AND FP:("KEYWORD" OR KEYWORD OR KEY*)

(2.3. 1)

The last step of the data cleaning process entailed a recursive round of the fine-grain refinements of the previous research. In this phase, I aimed at obtaining the highest level of saturation possible for each technology, obtaining in all cases a level of saturation of at least 95% from the first phase.

Once obtained all the end results for each technology, I recorded new information on the table collecting all the data:

- *Patentscope* final refined research: I recorded the query used to draw out the final patent screening on a specific technology.
- IPC relevant codes: I recorded the first 4 digits of the *International Patent Cooperation* (IPC) codes that appeared in the patent documents resulting from the final search, together with the amount of times they appear for each one of them.
- Level of saturation
- Year of first publication: I recorded the year in which the first patent publication appeared.
- Number of patents obtained after the final search.

2.4. Technology classification process

When the final sample was completed, I dedicated a last phase for re-aggregating the singular IPC codes obtained from the previous phases. This process proved especially useful in drawing out the macro trends in the Results, given that the aggregation of singular IPC codes in macro-categories better represents a trending scenario in technological fields. In the first phase of the classification process, the first activity I did was grouping all the relevant IPC codes found in the refined research into specific technological sectors. I followed the approach of the *IPC technological concordance*, a paper commissioned by WIPO that explains the current concordance table for 35 different fields of technology. Once every IPC code was assigned to the pertinent technological field, I employed a classification matrix to understand how the technologies of the research can be identified inside these categories. I used a share classification system based on the ratio of the number of times a specific IPC codes appears in a patent search for a specific technology and the total number of times the technology relevant IPC codes appear in a patent search for a certain technology.

$$Contribution\ share\% = \\ = \frac{n°singular\ IPC\ code\ appears\ in\ a\ patent\ search}{total\ n°\ relevant\ IPC\ codes\ appear\ in\ a\ patent\ search} \tag{2.4. 1}$$

The last phase of the classification process entailed the retrieving of comprehensive patent publication data from the previously search for every focal technology. This process, coupled with the classification, will be the fundamental input for the subsequent analysis.

2.5. Method for obtaining the patent trends

In this sub-section, I will illustrate how the final patenting trends have been found, building on the groundwork provided in the methodology chapter. The data necessary to plot the trends came from the classification matrixes and from the count of the applications published every year. For each yearly list of 10 breakthrough technologies from 2012 to 2022, I created a new table that aimed at gathering the count of applications published every year for each IPC technological concordance field. The data points for each year were calculated multiplying the contribution shares of the 10 technologies in each technological field by the count of publications of a certain technology in a specific year. I then repeated the same process with the count of publications in the following year to obtain the next data point, until 2021 (Equation (2.5. 1)).

$$a_{n,m} = \\ = contribution\ share_{tech1n} \times count_{tech1m} \\ + contribution\ share_{tech2n} \\ \times count_{tech2m} + \ldots + contribution\ share_{technn} \\ \times count_{techmm} \tag{2.5. 1}$$

Once all data points were available, I aggregated them to build a table that gathered the general patenting trends per each technological field from the year 2013 (first year of publications) to 2021. Data points in this table were represented by the summation of the values obtained from the previous calculations, considering all the technologies in each year of analysis from 2012 to 2022. In some cases, the analysis did not return enough data to conduct a consistent analysis in some fields, therefore at the end I have analyzed patenting trend in 21 out of 35 technological fields.

Table 2. Generic appearance of the general trend table with aggregate data from every research year.

Year	Technological field
2013	$\sum_{2012}^{2022} a_{1,2013,2012} + a_{1,2013,2013} + \ldots + a_{1,2013,2022}$
2014	$\sum_{2012}^{2022} a_{1,2014,2012} + a_{1,2014,2013} + \ldots + a_{1,2014,2022}$
…	…
…. (until 2021)	…

3. Results

In this chapter, I will illustrate the results of the research conducted in *Patenscope* and the analyses on the patenting trends that arose from it. The chapter aim is to reflect on the co-occurrence between patents trends and the pandemic event worldwide. I will proceed to introduce three different sub-sections:

The first sub-section will be devoted on the patenting trends that have remained unchanged during the COVID-19 pandemic (2020 – 2022), leading to hypothesize that the exogenous shock brought by the virus might have not deviated an already existing R&D trend trajectory.

The second sub-section will deal with the trends that instead have experienced an acceleration during the pandemic (2020 – 2022), leading to hypothesize a promotion of the R&D activity in certain technological domains that proved to be useful in fighting COVID-19.

Lastly, the third sub-section will outline the patenting trends in those technological domains that have shown a deceleration concurrently with the pandemic (2020 – 2022), leading to hypothesize that scientists and inventors might have de-emphasized the relevance of innovations in field that were less likely to help in coping with the pandemic.

3.1. Trends remained unchanged during the pandemic

Several technological fields have been experiencing different patterns of growth in patents in the last 10 years. The ones examined in this sub-section, were characterized by unchanged trends trajectory until 2021 and a trend forecast that did not deviate from the initial trajectory in 2022. This led me to hypothesize that the pandemic event (2020 – 2022) has not significantly influenced research focus and innovation activities in these technological fields, leading to suppose an overall independence. The fields identified were the following: *Computer technology; Audio-visual technologies; Optics*; *Control*; *Measurement*; *Handling*; *Machine tools; Other special machines; Materials, metallurgy; Transport; Food chemistry* and *Environmental technology.* For what concerns *Computer technology*, *Audio-visual technologies* and *Optics*, I found it plausible that COVID-19 might not have interfered in their general growth in patenting activity, given that the technologies included in these categories mostly are radically innovative digital technologies that have experienced great diffusion over the last decade, therefore they were greatly backed up by R&D efforts over the whole examination period, also considering the growing pervasiveness of digital in every aspect of the daily life both for businesses and consumers. In the case of *Control* and *Measurement* technological field, I hypothesized that the pandemic might not have been influential in their decelerating trend patterns since the technologies included in these categories are mostly related to instruments and methods used to steer and allow communication between machines, especially autonomous vehicles, which already encountered performance maturity and are now moving towards market application, overcoming the phase of R&D development. In the technological fields *Handling*, *Machine tools*, *Other special machines* and *Transport*, I notice that in the field evidencing always increasing growth (*Handling*), there has been a concurrent increase in demand that was reflected in the innovation activity over the whole examination time span. Conversely, in fields like *Machine tools*, *Other special machines* and *Transport*, data evidences slower patterns of growth that may indicate either a still low market applicability of the technologies or a focus in different steps of the innovation process may hint at steady decelerations in these fields. *Food chemistry* field experienced exponential growth patterns starting from 2019 thanks to the advent of meat-free burgers, which might have contributed to the entrance of new-comers and innovative start-ups in the market that, thanks to their patenting activity, shifted the trends trajectory to an exponential one. This is suggesting that the pandemic might not have played a role in the growth patterns of this field. In the case of *Environmental technologies*, I hypothesized that research efforts have been constantly growing over the years, suggesting that scientists interest in developing this kind of breakthroughs was mostly influenced by the urgency of facing climate-change challenges that could have not be related to COVID-19.

3.2. Trends accelerated during the pandemic

The trend analyzed in this section experienced accelerations in their trajectories during the years of the pandemic (2020-2022). This led me to hypothesize that COVID-19 might have impacted positively R&D efforts in these technological fields so as to suppose a potential concurrency. These fields were: *Telecommunication*, *Digital communications*, *Analysis of biological material*, *Medical technology*, *Biotechnology* and

Pharmaceuticals. For what concerns *Telecommunication* and *Digital communications*, I suggest that COVID-19 may have positively influenced research efforts in these categories as, during forced social distancing and closure of most productive activities, the demand for products belonging to this field increased drastically, therefore motivating scientists and inventors in increasing patent activity towards this kind of technologies. *Analysis of biological material* and *Medical technology* encountered rebound accelerations during the pandemic. Thus, I hypothesized COVID-19 may have revitalized the research effort in these technological categories, suggesting that innovation activity refocused on diagnostics and medical monitoring concurrently to the spread of the virus. For *Biotechnology* and *Pharmaceuticals*, I suggested that the urgency of addressing a worldwide challenge such as COVID-19, might have played a role in increasing innovativeness in these fields, directing majorly R&D efforts into finding new solution to dab the spread of the pandemic.

3.3. Trends decelerated during the pandemic

The trend analyzed in this section have experienced decelerations during the pandemic time (2020-2022). This led me to hypothesize that COVID-19 might have impacted negatively R&D efforts in these technological fields so as to suppose a potential concurrency. These fields were: *Electrical machinery, energy and apparatus*; *Semiconductors*; *Thermal process apparatus*. In all three of these categories, I hypothesized that the technologies included in these categories reached high level of maturity in terms of development in performance that might have hindered research efforts in these fields to focus R&D activities to more crisis-critical sector directly addressing the pandemic.

4. Conclusions

The core objective of the thesis was to reflecting around patterns of change in the worldwide R&D activity before, during and after the COVID-19 pandemic, with the aim of formulating hypotheses on whether the exogenous shock of the pandemic could have played a role in changing the innovation landscape.

It is important to stress that, due to the methods used, this thesis is not aimed at drawing causality links between the trends presented and the pandemic, but rather to generate evidence-based hypotheses that reflect how patent activity might have changed in consequence or in response to the pandemic in the most radically innovative technologies in the market.

As a matter of fact, the pandemic has caused discontinuity between previous and future necessities, such as a greater need to access technologies enabling remote working and healthcare or the urgency to commercialize vaccines to fight the spread of the virus. In this regard, my research may serve as a base point to further explore future innovation activities in several technological fields to understand whether COVID-19 has shifted R&D efforts directing technological change to specific areas. In particular, I believe that the thesis has implications for innovation research in two different areas. First of all, the methodology used in this dissertation can be useful for the identification of the technological research trends within the data. This could be helpful in understanding what differences subsist between consumer habits and new research trajectories. In fact, identifying R&D trends within patent data may be beneficial for spotting availability of new technological ideas in the future and target the ones that are likely to receive greater interest from the market with additional research. On a managerial perspective, this could help businesses and research hubs to steer their investment decisions or direct initiatives of corporate venture capital. Secondly, my research points out several hypotheses on the reasons why the pandemic may or may not have impacted patent activity in certain technological fields. This adds up to an already wide literature of patent analytics approaches to preview technological developments, contributing with an analysis on a highly disrupting factor such as COVID-19. The thesis can also be helpful in innovation research since it supplies new hypotheses in technological change that may be relevant for future innovation studies. In conclusion, I invite future research to investigate the impact of the pandemic in different technological fields and in a more detailed fashion, possibly outlining causality links between the outbreak of the virus and patenting activity, using different sets of data and more advanced forecasting methods in order to provide an even more accurate overview of future technological development.

References

Faulkner, S. L., & Trotter, S. P. (2017). Data Saturation. In *The International Encyclopedia of Communication Research Methods* (pp. 1–2). Wiley. https://doi.org/10.1002/9781118901731.iecrm0060

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market Value and Patent Citations. *The RAND Journal of Economics*, *36*, 16–38.

MIT. (2022). *MIT Technology Review*.

Pavitt, K. (1985). Patent statistics as indicators of innovative activities: Possibilities and problems. *Scientometrics*, *7*(1–2), 77–99. https://doi.org/10.1007/BF02020142

WIPO. (2022). *Guide to the International Patent Classification (2022)*. https://www.wipo.int/classifications/ipc/

5. Acknowledgements

This thesis represents the end of a two-year cycle dense with sacrifice and commitment but equally filled with personal satisfaction and pride.

I feel I must thank first of all Professor Chiara Franzoni for guiding me with constancy and great mentorship during the development of thesis, providing always inspiring thoughts on the research standpoints and helping me organize the work in the best possible way.

I want to thank my family for supporting me throughout this journey. To my father, who with his immense wealth of knowledge, wisdom and love has always been a stronghold that has supported me in every step and decision of my life. To my mother, whose unceasing emotional support and unconditional love has always given me the strength to hold on by always standing by my side. To my sister Paola, who through her honesty and maturity has always been a point of reference in my life as well as a source of pride and love. I thank you from the bottom of my heart, I am aware that without your invaluable support I would not have reached this milestone.

I want to thank Massimo, the person with whom I have shared every moment these past two years, who through his loyalty and friendship has significantly enriched my life and will be always the best travelling companion I could possibly wish for.

To Angela, the person to whom I owe much of the credit for what I have achieved. Thank you for sharing by my side every moment of your everyday life and making my life special with your affection and boundless love.

To Giò, Manfredi and Tareq, the best flat mates and brothers I have ever had. Thank you for letting me experience some of the best times of my life and turning a roommate house into a real family.

To my lifelong friends and the ones I have met along the way, I thank you for having supported me these past two years, without you all this would not have been possible.