MDSAA

Master Degree Program in
Data Science and Advanced Analytics

A SCIENTOMETRIC ANALYSIS TO STUDY THE TRENDS ON NUCLEAR WEAPONS-RELATED RESEARCH

Ana Luísa Silva Cruz Guerreiro Mestre

Dissertation

presented as partial requirement for obtaining the Master Degree Program in Data Science and Advanced Analytics

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

A SCIENTOMETRIC ANALYSIS TO STUDY THE TRENDS ON NUCLEAR WEAPONS-RELATED RESEARCH

by

Ana Luísa Mestre

Dissertation presented as partial requirement for obtaining the Master's degree in Data Science and Advanced Analytics, with a Specialization in Business Analytics

Supervisor / Co Supervisor: Bruno Damásio

Co Supervisor*:* Sandro Mendonça

November 2022

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledge the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Ana Luísa Mestre

Lisboa, November 2022

ABSTRACT

From their very first use in the Bombings of Hiroshima and Nagasaki at the end of the World War II, nuclear weapons became a serious threat to humanity and the world. Over the years, many protests and treatises have been developed to promote disarmament and non-proliferation of these weapons. Scientometrics is defined as a branch of the general "science of science" that covers quantitative methods to analyze research process. Although research on nuclear weapons is increasing rapidly, there has been little use of scientometric approaches on the referred topic. Within this framework, this study seeks to address the global trends on 5038 nuclear weapons-related articles from 1946 to 2021, extracted from the Web of Science (WoS) Core Collection database and complemented with data obtained through Scopus APIs, by drawing on scientometric analysis and methods. The analysis was performed on a main authors' level and their affiliation at the time of publication of the article. The study of the worldwide patterns of academic literature related to nuclear weapons can assist researchers of various knowledge areas in identifying the main trends in the field and possibly deepen the study of some of them. In particular, the findings identify 1990 as the year when the production of scientific literature on nuclear weapons begins to grow more, the United States as the most contributing country in the field, Social Sciences as the most popular category, as well as provide other useful insights on the quantity, quality, and impact of research "elements" like authors, affiliations, sources, and keywords.

KEYWORDS

Scientometrics; Python; Scopus APIs; WoS Core Collection database; Nuclear weapons; Atomic bombs; Hydrogen bombs

INDEX

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Nuclear weapons represent one of the most serious threats to mankind with severe health and environmental impacts. Containing and preventing nuclear war remains one of the greatest challenges facing world leaders today, and disarmament and non-proliferation remain fundamental tools to help create and promote a secure environment favorable to ensuring human development (1). Today, nine countries possess nuclear weapons, and nearly all the major nuclear powers, including the United States, Russia, and China, are now significantly increasing their nuclear arsenals in size and capability (2). Additionally, their command-and-control systems are becoming more vulnerable to cyber-attacks which raises the risk of a nuclear war (3). With the aim of minimizing this threat, it becomes fundamental to increase people's awareness of the danger of nuclear weapons being used intentionally, accidentally, or by miscalculation by nuclear-armed countries, as well as of their immediate and long-term effects on people, society, heath care systems, and the environment.

Scientometrics is a field of study that has become more and more popular in the last years across the spectrum of scientific fields (4) defined as a branch of the general "science of science" that covers all quantitative methods to analyze science and research process (5). Over time, a lot of studies covering the application of scientometric tools on different topics were carried out, such as the contribution of specific countries to the research output in different fields (6; 7; 8); on forest health and tree diseases (4); on rural depopulation (9); and on social media and big data (10; 11; 12; 13), 5G technologies and systems (14), and finance and economy-related fields (15; 16; 17).

Having this in mind, and the fact that an overview about nuclear weapons-related research is not known yet, this study investigates the knowledge production related to nuclear weapons to provide the worldwide trends on the referred topic. For this purpose, a scientometric analysis was carried out using Python software, Scopus APIs, and Bibliometrix R-tool, and applied to a set of 5038 documents retrieved from the WoS Core Collection database. The results obtained may be of interest to researchers from different knowledge areas, as well as useful in the development of future research. They consider different levels of analysis, namely who writes about nuclear weapons (authors), which institutions are associated with whom writes about the topic (affiliations), who publishes on the topic (sources), in which disciplinary fields the articles are inserted (categories), what was written (documents), and what keywords are commonly used (author keywords and keywords plus).

More specifically, the study focuses on providing answers to the following questions:

- Is there any relation between the time when there is growth in the number of articles published with historical events on nuclear weapons? Can these events influence the scientific production on nuclear weapons?
- What are the most contributing regions on the referred topic, considering the main authors and their affiliations?
- Which sources publish more on nuclear weapons-related articles and how is their production over time?
- What are the most common categories and their distribution over the years?
- Who are the most important authors in the field considering articles published, citations, and H-Index?
- What are the most popular articles, and what are some of their characteristics?

- What keywords are commonly used in nuclear weapons-related articles and with which other keywords they usually appear?

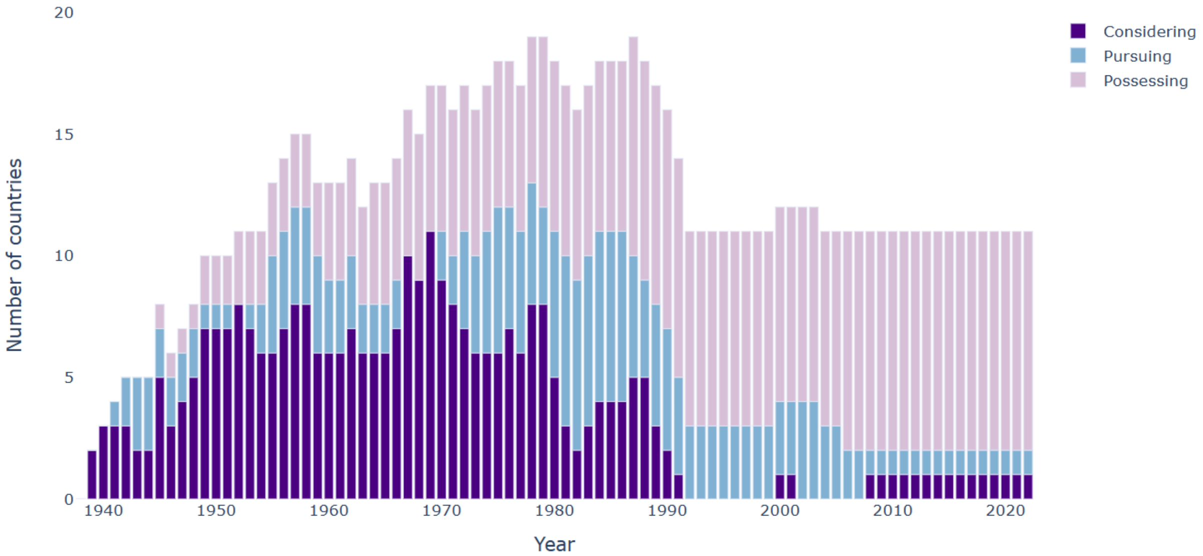
The reminder of this article is structured as follows: firstly, Section 2 outlines a theoretical perspective on nuclear weapons evolution and history; Section 3 describes the methodological set-up; Section 4 proceeds with the articulation of the main decisions taken to clean and to prepare the data for the analysis; this is followed by Section 5 that presents and debates the principal results and patterns obtained; finally, Section 6 covers the final remarks adding the main conclusions, the description of some limitations encountered, and the suggestions of future research.

2. NUCLEAR WEAPONS: A BRIEF SUMMARY

2.1. INTRODUCTION TO NUCLEAR WEAPONS

Nuclear weapons are the most destructive weapons ever created, representing a serious threat to humanity and the world, with the capacity to kill millions directly and billions through their impact on agriculture (18) by suddenly releasing vast amounts of energy (19). The radioactive pollution on the atmospheric, aquatic, and underground environments due to nuclear weapons tests (20) increases rates of leukemia and other cancers among populations who were downwind of these tests. For example, the impacts of the hundreds of nuclear tests carried out from 1946 by the United States, Britain, and France in the Pacific, on colonized lands and oceans, from Australia to the Marshall Islands, Kiribati to French Polynesia, are still being felt today. Some of them occurred between 1952 and 1957 in South Australia, where nuclear tests conducted by Britain caused health problems for local Aboriginal people who were at the highest risk of radiation (21).

Figure 1. Number of countries that consider, pursue, or possess nuclear weapons, 1938 – 2022. Source: own elaboration on Our World in Data (22).

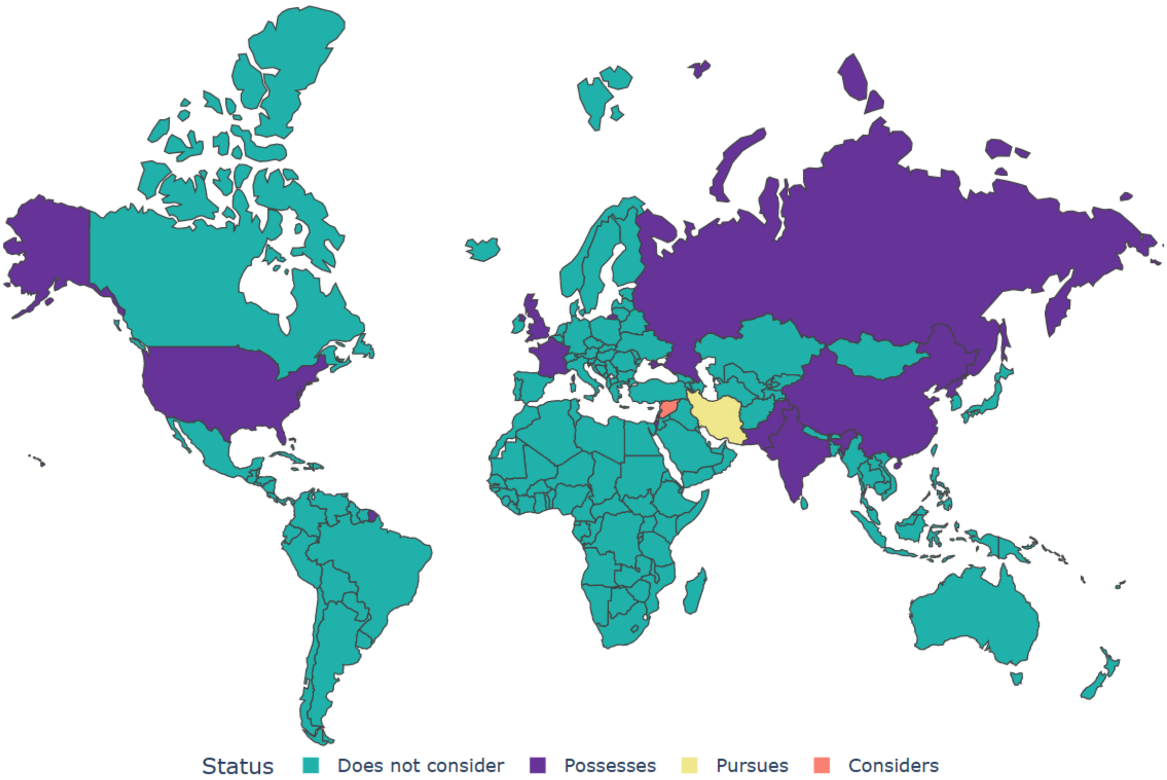


The history of nuclear weapons is traced back to 1930s, when the technology behind them begins to be developed. In particular, the discovery of nuclear fission in 1938 opened the possibility of nuclear technologies (23) applicable in agriculture and food, medicine, space exploration, water desalination, including weapons (24). This paved the way for their very first use in the catastrophe of Hiroshima and Nagasaki few years later at the end of World War II. Since then, controlling and preventing the proliferation of nuclear weapons to additional countries has been an important issue in international relations (25). Unfortunately, despite the efforts of different leaders and organizations, and although the number of nuclear weapons in the world has decreased since its peak in the 1980s (26), more countries are exploring and pursuing nuclear weapons (18), as evidenced on Figure 1.

2.2. THE NUCLEAR WORLD, NUCLEAR POWERS AND PROLIFERATION

When thinking of the nuclear world, what is happening nowadays is that while most countries support a nuclear weapon-free world, which has resulted in nuclear weapons being banned in areas on the earth, a smaller number of countries still possess nuclear arsenals because of the security, prestige, and power associated with their possession. As such, nine countries are known to possess nuclear weapons today, including Russia, the United States, China, France, India, Israel, North Korea, Pakistan, and the United Kingdom (Figure 2) and approximately ninety percent of all nuclear warheads belong to the United States and Russia combined (25). Most of the weapons of these countries are strategic weapons, which cause large-scale damage and are capable of intercontinental delivery; others are tactical weapons which are small nuclear warheads and delivery systems meant to carry out a limited strike in a smaller area (27).

Figure 2. Country position on nuclear weapons, 2022. Source: own elaboration on Our World in Data (22).



However, the nuclear world is not just where nuclear weapons exist (26). Many countries without nuclear weapons want to protect themselves against nuclear threats without developing nuclear weapons. To ensure this, a security guarantee must exist, which extends the perceived security a nuclear arsenal provides to allied countries that do not possess nuclear weapons. For this reason, this guarantee provided by a nuclear-armed country is also known as a "nuclear umbrella". The United States nuclear umbrella, for example, covers the North Atlantic Treaty Organization (NATO) member countries that do not have nuclear weapons, Japan, Australia, and South Korea.

According to Levy (19), regarding the forces that have been taking throughout the years to stop nuclear actions, in fact, there is no comprehensive treaty that bans the use or mandates the destruction of nuclear weapons, but rather a series of overlapping, withdrawn treatises, such as:

- the Non-Proliferation of Nuclear Weapons Treaty, a multilateral treaty which opened for signature in 1968, entered into force in 1970, and was extended indefinitely in 1995 (28);
- the Anti-Ballistic Missile Treaty ratified in 1972 between the United States and the Soviet Union, from which the United States withdrew in 2002 in the wake of the September 11, 2001 terrorist attack on the World Trade Center and the Pentagon (29);
- the Comprehensive Nuclear Test Ban Treaty, which opened for signature in 1996 and has not entered into force yet (30).

During the Cold War – a period defined by tensions between the United States and the Soviet Union – , nuclear annihilation was a persistent threat because each country possessed enough nuclear weapons to destroy the other (26). Another alarming situation is the conflict between India and Pakistan (31) which, although neighboring countries, have fought several conventional wars over the years (mainly due to an unresolved boundary dispute) (32). On the other hand, the conflict in Ukraine since February 24th, 2022, has also increasing the possibility of a nuclear war (33).

2.3. The history of nuclear weapons

Now let's deep dive in some of the critical actions and decisions that led to today's distribution of those weapons and the world's non-proliferation regime.

2.3.1. The beginning of the atomic age

Scientific discoveries in the late 1930s made nuclear weapons a possibility for the first time in history (23). In **1942**, in response to fears that German scientists had been working on a weapon using nuclear technology, President Franklin D. Roosevelt authorized the formation of the Manhattan Project – a secret research effort led by the United States government to develop the first nuclear weapons by bringing together various scientists and military officials working on nuclear research (34). On August 6th, 1945, the United States dropped the first atomic bomb used in combat on the city of Hiroshima, Japan, killing thousands of people. When the Japanese did not immediately surrender, the United States dropped a second atomic bomb three days later, on the city of Nagasaki, Japan, which resulted in other thousands of victims. On August 15th, 1945, Japan surrender, ending World War II. This is the only country to have been attacked by nuclear weapons and where anti-nuclear sentiment runs particularly deep (35).

The end of World War II marked the beginning of the atomic age. Between 1945 and 1964, in the context of the Cold War, geopolitics and the lack of effective international disarmament policies, countries like the United States, the USSR, the United Kingdom, France, and China became nuclear powers, and many nuclear tests were conducted in all global environments – atmosphere, underground, and underwater. The severe environmental damage caused by these nuclear tests, as well as the general context of global nuclear weapons tests, have created the premises of the first instance of large-scale international cooperation to eliminate nuclear weapons testing (20).

2.3.2. Antinuclear movements

In **1957**, the International Atomic Energy Agency (IAEA) was created with the mission of promoting and overseeing the peaceful use of nuclear technology (34). The antinuclear movement emerged as a social movement in **1961** at the height of the Cold War. During Women Strike for Peace – a women's peace activist group in the United States – protests on November 1, 1961, thousands of women marched in various cities in the United States to demonstrate against nuclear weapons. Antinuclear movements captured national attention again in the 1970s and 1980s with protests against nuclear reactors after the Three Mile Island accident – a nuclear meltdown at a Pennsylvania power plant in 1979 (36). Additionally, in **1982**, one of the largest political protests in the United States history took place in New York City, where around one million people marched protesting nuclear weapons and urging an end to the Cold War nuclear arms race (20).

2.3.3. The Cuban missile crisis

At the end of **1962**, the Soviet Union installed nuclear-armed missiles in Cuba which resulted in a thirteen-day military and political standoff known as the Cuban Missile Crisis. This crisis remains the closest the world has ever come to a nuclear war. The year of **1963** marked a milestone with the entry into force of the Limited Test Ban Treaty (LTBT) – a treaty stipulating a ban on nuclear weapons tests in all global environments, except for the underground (37). The LTBT had a genuine impact in limiting radioactive isotopes in the atmosphere in the two hemispheres (38).

2.3.4. The non-proliferation of nuclear weapons treaty

In **1968**, the United States, the Soviet Union, and some sixty other countries signed the Non-Proliferation Treaty (NPT) to seek the end of the nuclear arms race and promote disarmament (35). The NPT went into effect in 1970, was extended indefinitely in 1995, and not only bars nuclear weapons states from propagating weapons to other states, but also prohibits states without nuclear weapons to develop or acquire nuclear arsenal. It permits the use of nuclear energy for peaceful purposes (35) and it has been the cornerstone of the international effort to prevent the spread of nuclear weapons (2). Although the provisions of the treaty were to be implemented with the help of the IAEA (39), geopolitical experiences of the last four decades have shown that, outside the scope of the NPT, a different category of nuclear states had emerged, including India, South Africa, Pakistan, North Korea, and apparently Israel (unconfirmed nuclear status) (40), with some of these countries conducting nuclear weapons tests of their own. However, after the Soviet Union collapsed in 1991 and the Cold War ended, real progress was made to strengthen the NPT, get former Soviet countries to join the treaty, and stop further use of nuclear weapons (34).

2.3.5. The comprehensive nuclear test ban treaty

A new phase in stopping all types of nuclear tests began in **1996** with the United Nations (UN) adoption of the Comprehensive Nuclear Test Ban Treaty (CTBT) (36). This treaty banned nuclear explosions of any kind, including weapons tests (41). In **1998**, India and Pakistan conducted several nuclear tests in response to one another which dismayed the international community (31). World leaders admonished the two long-time adversaries in breaking the CTBT. The United States imposed strict economic sanctions against both countries and lobbied for the World Bank, International Monetary Fund, and other countries to do the same. The sanctions were lifted in 2001 when the United States needed Pakistan and India's support to fight Al Qaeda and other terrorist cells in Afghanistan (35).

2.3.6. North Korea: a big nuclear threat

North Korea remains one of the most unstable nuclear powers today. Although the country is believed to have a small nuclear arsenal, its continued nuclear tests have raised widespread concerns over possible nuclear conflict. The country withdrew from the NPT and all the other international agreements on non-proliferation in 2003 and conducted its first nuclear test in **2006** (36). Since then, it has not only threatened to use its nuclear weapons against the United States, Japan, and South Korea (26), but also openly tested nuclear weapons, drawing sanctions from various nations and international bodies. In **2017**, it also tested two long-range intercontinental ballistic missiles (42).

In **2002**, the American President George W. Bush named Iran, Iraq, and North Korea as the "Axis of Evil", in part due to United States suspicions of those countries having weapons of mass destruction (WMD). Later that year, intelligence reports indicated that North Korea possessed crucial equipment in the production of nuclear weapons (35). In 2003, the United States invaded Iraq as a legitimate, defensive response to the threat of WMD, which the Saddam Hussein government allegedly had in its arsenal, according to the view of the American government and their allies. The UN investigated the claims that Iraq was stockpiling WMD, and after months of investigation, they could not find any Iraqi WMD, which proved US intelligence was faulty (43). In the same year, the IAEA announced that it had found traces of highly enriched uranium at a nuclear power plant in Iran (44).

In April **2009**, negotiations among China, Japan, North Korea, Russia, South Korea, and the United States to find a peaceful resolution to North Korea's nuclear weapons program fell apart after the UN Security Council condemned a North Korean test launch of a rocket, which it had disguised as part of its civilian space program. The negotiations, known as the Six Party Talks, had lasted six years but failed to reach a resolution (34).

2.3.7. The joint comprehensive plan of action

In **2015**, the United States, China, France, Germany, Russia, the United Kingdom, and the European Union reached a nuclear agreement with Iran after years of negotiation, known as the Joint Comprehensive Plan of Action (JCPOA) (34). This plan intends to contribute to regional and international peace and security, marking a fundamental shift in the countries' approach to nuclear issues (45).

2.3.8. Treaty on the prohibition of nuclear weapons

Finally, the Treaty on the Prohibition of Nuclear Weapons (TPNW) was adopted on 7th July, 2017 by the majority of UN member states, opened for signature on 20th September, 2017, and entered into force on 22nd, January, 2021 (46). International organizations and civil society groups also supported the negotiations of the TPNW (47). This is the first comprehensive and universal prohibition of nuclear weapons, prohibiting all state parties from developing, testing, producing, acquiring, stockpiling, using, or threatening to use nuclear weapons. State parties are also prohibited from assisting any other state in conducting such activities.

3. METHODOLOGY

Pautasso (4) refers to Scientometrics as the quantitative study of patterns in the scientific literature. According to Hood (48) the term Scientometrics was coined by Vassily V. Nalimov in the 1960s and refers to the science of measuring and analyzing of science, such as a discipline's structure, growth, change, and interrelations. In this sense, Nalimov (49) describes Scientometrics as the quantitative study of various kinds of intelligence process in the development of science which has formed a stable field and obtained broad acceptance from the journal Scientometrics, created in 1978. Finally, Vilanov (50) adds that the definition of Scientometrics focused on the study of scientific information is that it analyzes the quantitative aspects of the generation, propagation, and utilization of scientific information to contribute to a better understanding of the mechanism of scientific research activities.

This approach may have different applications such as the following ones described by Pautasso (4): the examination of temporal trends in the number of publications on a certain topic; the investigation of the spatial distribution of the research attention given to a certain topic; the study of the relative number of papers dealing with various research areas; the exploration of the productivity of researchers and its likely quality; the comparison of the research performance and the diversity of institutions and countries; the documentation of patterns and determinants of the length of peer review processes; or the study of interdisciplinary collaboration activities.

Moreover, this is an important area of study because today different rankings use scientometric indicators for university rankings and universities themselves use them as Key Performance Indicators (KPIs) (51; 52; 53). On the other hand, bibliometric indicators are also widely used by governments and researchers in the evaluation of research output. The analysis of huge volumes of information generated every year by institutions and organizations can be facilitated by Scientometrics, whose current growth in different contexts has a positive impact on the way of analyzing science information (51). Additionally, it provides scientific basis for scientific decision making and management (54).

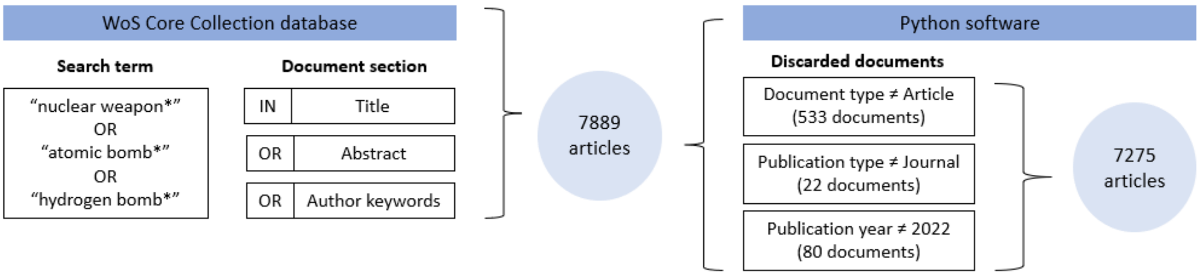
The methodology used in this paper is based on scientometric techniques with the objective of identifying research patterns in a particular subject, in this case nuclear weapons-related scientific articles. More specifically, the proposed methodology aims to perform a robust analysis on the publications of said topic considering several levels of analysis such as authors, affiliations, sources of information, categories, and keywords, for example, in an objective and reliable way. Further details regarding authors and affiliations, and categories and keywords will be provided in sections 4.4.2 and 4.4.3, respectively.

As such, and knowing that this methodology depends on the quantity analysis of all publications in the chosen field based on specific words, the first step was to extract the necessary raw data, which includes publication (bibliometric) and content (textmetric) data, from the online version of the WoS Core Collection database – a publisher-independent global database often used by researchers (9). The data search was carried out by filtering the articles[1] which are published in journals (more information on Appendix, Table 12) and which contain at least one of the following terms "nuclear weapon*", "atomic bomb*", or "hydrogen bomb*" in at least one of the articles' elements: title, abstract, or

[1] Article – Reports of research on new and original works that are considered citable and that were published in a journal and/or presented at a symposium or conference. An article usually contains abstract, graphs/tables, and cited references (77).

author keywords. The wildcard (*) and the Boolean operator (OR) were used to amplify and to improve the quality of the data extraction. In the case of the wildcard, it already includes the plural of the words and may also include other varieties of them (e.g.: atomic bombing, hydrogen bombardment). Additionally, the year 2022 was excluded because at the time of this study the year had not finished. From this preliminary selection resulted 7275 articles as indicated on Figure 3. Of these 7275 articles, 2350 contemplate the search term "nuclear bomb*", 1066 the search term "atomic bomb*", and 55 the search term "hydrogen bomb*" in the title, abstract, or author keywords.

Figure 3. Data extraction and filtration in WoS and Python. Source: own compilation.



One limitation of using WoS Core Collection database for extracting data is that it only allows the exportation of a maximum of 500 documents at a time. Thus, the documents were downloaded as several separated text files and then merged on a zip file so that they could be completely uploaded to the Bibliometrix R-Tool – a recent R-package which facilitates a more complete bibliometric analysis employing specific tools for both bibliometric and scientometric quantitative research (9) – that was used in the beginning of the study to get a first and quicker understanding of the data. From here, the data was exported as a Microsoft Excel spreadsheet and finally imported to Python software, where data cleaning and preprocessing, and data analysis and exploration were performed.

Finally, the Scopus database – another database commonly used by researchers – had also a fundamental role in this study since it enabled to retrieve important information related with authors, affiliations, categories, and keywords through different Elsevier Scopus Application Programming Interfaces (APIs) after creating an API Key on Scopus website and then importing the pybliometrics Python library, which accesses Scopus's APIs from Python on a large scale (55). In the end, the Bibliometrix R-Package was used again, but this time as a tool to complement the keywords' analysis, namely, to understand the relationships between the keywords used in nuclear weapons-related articles.

4. DATA PREPROCESSING

When importing the raw data to a Jupyter Notebook, some preprocessing was primarily needed, and some decisions had to be taken to improve the quality of the results in the analysis and visualizations. Some problems encountered in the data were duplicated articles, missing data, articles without authors, authors without affiliations, articles not found in Scopus, and many combinations of categories which makes it challenging to group. There were 7275 articles in the beginning of this phase and 5038 articles in the end.

4.1. UNNECESSARY VARIABLES

The dataset had 7275 articles and 71 variables, with each row of the dataset representing an article and each column representing an attribute of the article. Since not all the variables (columns) are going to be used in the analysis, 36 of the 71 were dropped. Most of the eliminated variables contained information related with proceeding papers[2], early access[3], or book chapters[4] which are document types that were not studied in this research. Other variables contained detailed information about the articles that was also not considered such as their number of pages or volume.

4.2. DUPLICATED DATA

Regarding duplicated data, articles with the same document title were considered duplicated (56). As such, 38 articles were identified as duplicated. From the duplicated articles, another decision was taken: to select only the ones with cited references. However, four situations were identified (Table 1), and in the end 21 articles were dropped in this step.

Table 1. Situations identified and decisions taken in relation to duplicated data. Source: own compilation.

Situations identified	Decisions taken
Duplicated articles where one had references and the other did not.	The article with references was selected (6 in total were dropped).
Duplicated articles where all had references with different number of references.	The article with the highest number of references was selected (5 in total were dropped).
Duplicated articles where all had references with same number of references.	An article entry was selected arbitrarily (2 in total were dropped).
Duplicated articles where none had references.	An article entry was selected arbitrarily (8 in total were dropped).

4.3. MISSING DATA

Of the 35 remaining variables that now exist in the dataset, 20 had missing values, however, for the purpose of this study, only the variable DI[5] was considered regarding missing data (Appendix, Table

[2] Proceeding paper – Full papers in a wide range of disciplines that were or will be presented at a symposium or meeting. Generally published in a book of conference proceedings (77).

[3] Early access – An article that has been electronically published by a journal before it has been assigned to a specific volume and issue (77).

[4] Book chapter – A monograph or publication written on a specific topic within a main division in a book (77).

[5] DI – Alphanumeric string used to uniquely identify articles published electronically (76).

12). This was a very important variable for the analysis because it was used as input of the Abstract Retrieval API (57), which initialized the process of retrieving significant, complementary information for the study. There were 1287 articles without DI, so they were eliminated in this phase.

4.4. Extracted data with Scopus APIs

In this phase additional information related with the authors, affiliations, categories, and keywords was retrieved from the Scopus database using three different Scopus APIs – Abstract Retrieval, Author Retrieval, and Affiliation Retrieval – accessible through pybliometrics Python library. In the end new variables were created with this information. However, during this process, other incoherencies in the data were encountered, such as articles that did not exist in Scopus, articles without authors, and authors without affiliations.

4.4.1. Articles that do not exist in Scopus

As previously mentioned, the process of retrieving complementary information for the analysis started by interacting with the Abstract Retrieval API through the usage of the variable DI as its input. However, 467 articles were identified as not existing in Scopus database in this step. This means that the entity (value of DI of each article) that was used to search the article's information in Scopus through the API does not exist, according to the documentation of the pybliometrics Python library on the error messages that may occur when accessing Scopus database via its API (58). Thus, in the end 467 articles were discarded from the study.

4.4.2. Selection of authors and affiliations

The next step was to apply the property "authors" of the Abstract Retrieval API to the remaining 5500 articles to retrieve the authors' and the affiliations' identifiers (ids). However, when applying this property to the apparently already prepared data (5500 articles), 4 articles were identified as not having an author associated, and 458 articles as not having affiliations associated with the main author of each article. Thus, these articles (4 + 458) were also discarded from the analysis.

The final data (5038 articles) had 3210 articles authored by more than one author and 498 articles where the main author had more than one affiliation. In this phase of the analysis, two important decisions were taken: one in the selection of the articles' authors and the other in the selection of the authors' affiliations.

Regarding the authors of an article, it was decided that only the first author of each article would be considered in the analysis since it is the one with higher contribution to the article. As for the main authors with multiple affiliations, only the first affiliation of the main author was selected. This is because when authors have multiple affiliations, for any given author, affiliations are often listed in descending order of importance. Typically, there is a "primary" affiliation which comes first, and the order of remaining affiliations is of lesser importance (59).

The author affiliation identifies where the author worked or studied when the research was conducted (60). To complement the decision described above, it was also concluded in this section that the ids of the affiliations of the main authors represented the ids of the institutions where they worked at the time they developed the study (and not their current affiliations[6]). For each article, the affiliation is associated taking into consideration the combination of the author and the article (and not only the author). Consequently, even after applying this criterion of selection (select the first affiliation of the main author for each article), there are authors that continue to have more than one unique affiliation in the dataset, which corresponds to the different locations that these main authors had at each moment of elaboration of each article.

The authors' and affiliations' information thus retrieved was provided as input to the Author Retrieval (61) and Affiliation Retrieval (62) APIs, respectively, as further shown in Figure 4 and described in Section 4.5.

4.4.3. Categories grouping

The data contained 712 unique combinations of WoS categories (Appendix, Table 12). To facilitate the grouping, the Scopus subject areas were used instead of the hundreds retrieved from the WoS Core Collection database. This is because titles in Scopus are classified under four broad subject areas (life sciences, physical sciences, health sciences and social sciences and humanities), which are further divided into 27 major subject areas – the ones used in the analysis – and more than 300 minor subject areas. Additionally, titles may belong to more than one subject area (major and minor) (63) which explains the big number of categories' combinations in the data. The abbreviations for each one of the 27 major subject areas were thus retrieved using the property "subject_areas" of the Abstract Retrieval API, and then the correspondence with its respective name was carried out based on the information provided in Garcia (64). Appendix, Table 13 provides additional information on the subject areas considered in the analysis.

4.4.4. Author keywords and keywords plus

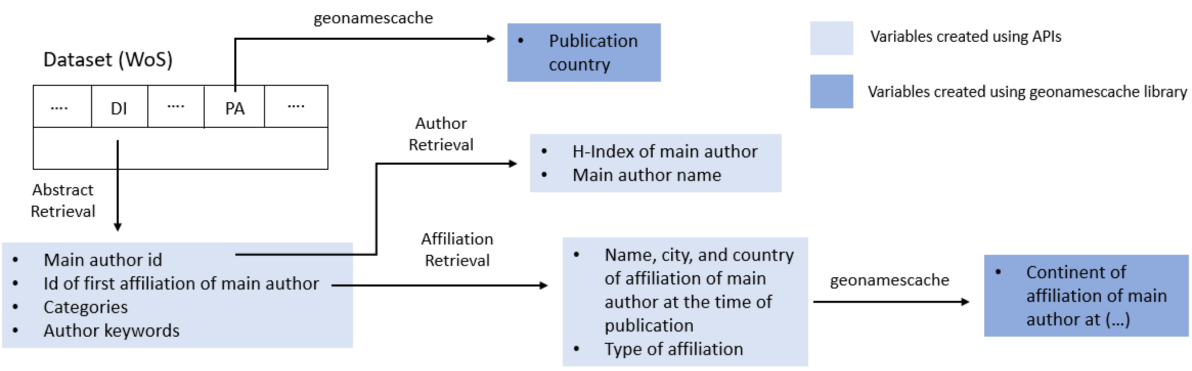
There are two types of keywords in the data: author keywords and keywords plus (Appendix, Table 12). The first ones are defined by the author(s) of the article; the second ones are words or phrases that frequently appear in the titles of an article's references, but do not appear in the title of the article itself. Keywords plus are thus independent supplement for author keywords (13) generated from cited titles. Additionally, they are based upon a special algorithm that is unique to Clarivate databases (65), enhancing the power of cited-references searching by searching across disciplines for all the articles that have cited references in common (65). In the end, since there were more articles without author keywords in the data extracted from the WoS database than in the data obtained with the Abstract Retrieval API, the data of the second option was considered for the analysis (by applying the property "authkeywords" of the API to the data). As for the keywords plus, since there is no property of the Scopus APIs to retrieve this information, the data of the WoS database was considered in this case. Again, Section 4.5 summarizes the creation process of additional variables for the analysis including the ones with information on the categories and author keywords.

[6] An analysis on the current affiliations of the main author using the property "affiliation_current" of the Author Retrieval API (61) was considered but at the end discarded because there was the risk of associating articles with the current affiliation of the main author, and not with the affiliation at the time of development of the article.

4.5. NEW VARIABLES CREATED

This section describes the chain creation process of several variables throughout the different steps mentioned in Section 4.4, with complementary data extracted through the Scopus APIs and represented in lighter blue on Figure 4, as well as two other variables represented in darker blue on the same figure. In this sense, Figure 4 presents a brief schema on the creation of these new variables, informing that they were obtained using two main strategies: the Scopus APIs (accessed through the pybliometrics Python library), and the geonamescache Python library.

Figure 4. Diagram with information contained in the variables created. Source: own compilation.



4.5.1. Using Scopus APIs

Most of the variables (represented in lighter blue on Figure 4) were created by applying different properties of the Scopus APIs or from variables that resulted from this "application". As stated before, this chained process of retrieving additional information for the study started with the usage of the variable DI of the WoS Core Collection database as input of the Abstract Retrieval API. From here resulted variables containing information on the authors' and affiliations' ids, as well as one variable with information on the articles' abbreviation categories and another one on the authors' keywords. At a later stage, the main author id (obtained previously) was used as input of the Author Retrieval API, from where resulted two more variables containing information on the main author of each article (name and H-Index). Finally, the id of the first affiliation of the main author (obtained in the first step also) was used as input of the Affiliation Retrieval API, from where resulted variables containing information on the first affiliation of the main author (name, city, and country), and one variable that characterizes the type of affiliation. For this variable it was considered that the affiliations' ids which start with a 6 (6XXXXXXX) are academic affiliations and the ones which start with a 1 (1XXXXXXX) are non-academic affiliations, according to the documentation of the pybliometrics Python library (55). More detailed information on the process of retrieving additional information for the study through Scopus APIs and the consequent creation of new variables with this data is available for consultation on Appendix, Table 14.

4.5.2. Using geonamescache Python library

The geonamescache library was used to create two more variables for the analysis (represented in darker blue on Figure 4). This is a Python library that provides functions to retrieve names, ISO and FIPS codes of continents, countries as well as US states and counties as Python dictionaries (66). The library was used in these cases as a solution to get the names of all existing countries from their respective codes, and to then compare them with the information contained in the variables of the publication address and country of affiliation of main author. It is important to note here that in the creation of the publication country variable, the publication address variable was used instead of the publication city variable (both retrieved from the WoS database, Appendix, Table 12) because there are cities with the same name in different countries (e.g.: there is a city named York in the USA and in England), and when comparing with a list of all the world countries, these cities were going to be associated with more than one country.

5. RESULTS AND DISCUSSIONS

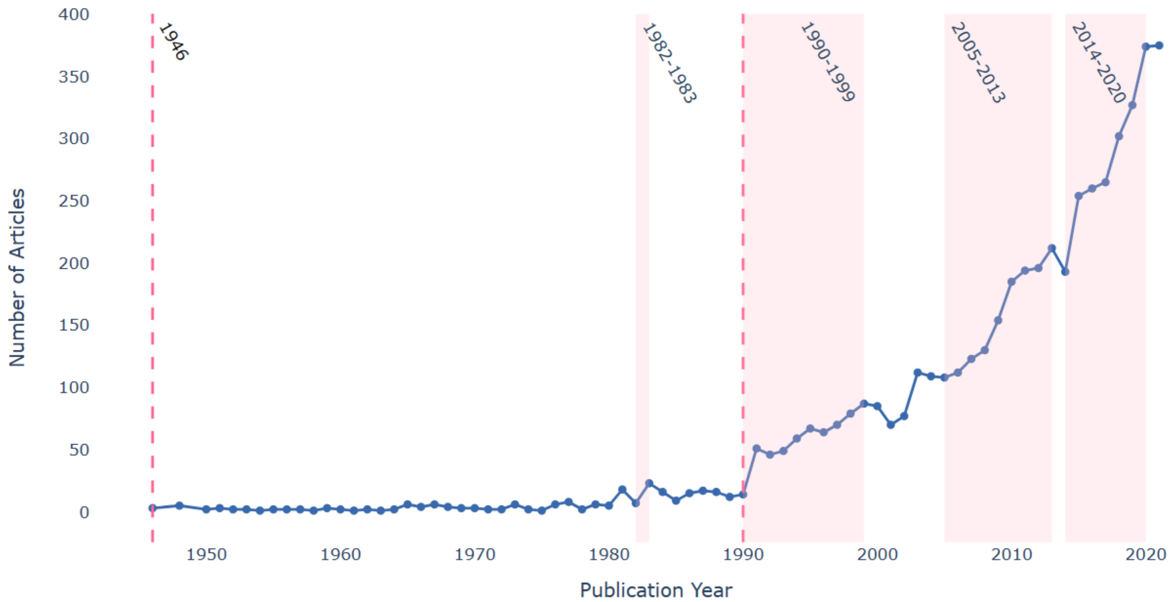
In this section, 5038 articles that focus on or refer to nuclear weapons, atomic bombs, or hydrogen bombs-related words in the title, abstract, or author keywords, and published between 1946 and 2021 were studied. They are authored by 3966 individuals (considering the main author of each article) linked to affiliations in 98 different countries (considering the first affiliation of the main author at the time of publication of each article). The articles are written in 18 different languages, being English the most predominant one (around 97 percent). Moreover, they are published in 1772 journals, considering the unique International Standard Serial Numbers [7] (ISSNs), spread across 47 countries.

5.1. THE EVOLUTION OF NUCLEAR WEAPONS-RELATED RESEARCH: AN OVERVIEW

5.1.1. Relating data behavior and characteristics with historical events

The number of articles published on nuclear weapons-related research by year between 1946 and 2021 is represented on Figure 5. In general, the quantity of publications increases over the years considered, particularly since 1990. On average, the number of articles published every year is 68 with a standard deviation of approximately 98. This is a high standard deviation, which indicates that the data is more spread out in relation to the mean, as it can be concluded on Figure 5. To find connections between the content of the articles and History throughout time, in some of the years when it was identified growth in the number of published articles, specific words were searched according to nuclear weapons historical events that happened in different moments in time. For this purpose, the following years were considered: year 1946; from 1982 to 1983; from 1990 to 1999; from 2005 to 2013; and from 2014 to 2020.

Figure 5. Number of published articles on nuclear-weapons-related research by year, 1946-2021. Source: own compilation.



[7] An ISSN uniquely identifies all types of print and electronic media, such as scientific journals, magazines, and newspapers (75). In the data studied, all the articles are published in journals (Appendix, Table 12).

First year 1946: right after bombings of Hiroshima and Nagasaki

The data tell us that 1946 is the first year with articles published on the topic, with only three 3 articles. These were authored by Harkins, William Draper; Keller, Paul D., and Stephenson, Reginald J., published in the United States, and have 0, 21, and 2 cited references respectively. Additionally, these articles were published approximately – it is not possible to have information about the exact publication date because this variable is missing for these articles – one year after the first atomic bombs were dropped in Hiroshima and Nagasaki in August 1945. Although the variables with information on the keywords and on the abstract are also missing in these articles, it can be concluded by their titles below that they reflect studies on the atomic bombs:

- The Neutron, the Intermediate or Compound Nucleus, and the Atomic Bomb
- A Clinical Syndrome following Exposure to Atomic Bomb Explosions
- A Brief Account of the Physics of the Atomic Bomb

From 1946 to 1980 the number of articles did not vary too much, with an average of around 3 articles published per year, and a maximum of 8 articles published in 1977.

1982 – 1983: antinuclear movements

There is a higher increase in the numbers from 1982 (7 articles) to 1983 (23 articles). This was during the time when the antinuclear movement was capturing attention again through the realization of many protests of nuclear reactors in response of the Three Mile Island accident in 1979 (67). In 1982 there was one of the largest political protests in the United States History, with around a million people protesting against nuclear weapons and urging an end to the Cold War nuclear arms race (23; 36). By analyzing the titles of the articles (author keywords could have been also used but this variable is missing for these articles) published in 1982, 3 of them mention antinuclear-related topics, more specifically:

- The Movement for a Nuclear-Weapon-Free Zone in the Nordic Region-Impact of the Karlskrona Incident
- The Nuclear Non-Proliferation Treaty and No 1st Use of Nuclear-Weapons
- Verification of a Nuclear-Weapons Freeze

In 1983, the publication of articles related to this topic continues, with 6 out of 23 articles containing the word "free" in the title:

- Submarines, Territorial Integrity and a Nordic Nuclear Weapon-Free Zone
- Freeze on Nuclear-Weapons Development and Deployment – Pro
- A Nuclear Weapon-Free Zone in the Nordic Area – Conditions and Options – A Norwegian View
- Freeze on Nuclear-Weapons Development and Deployment – Con
- Nuclear Disengagement in Europe – Problems of Nuclear Weapon-Free Zones
- A Nordic Nuclear Weapon-Free Zone as a (Partial) Depletion Zone

1990 – 1999: collapse of the Soviet Union and end of the Cold War

Between 1990 and 1999, the number of articles published keeps increasing in general (from 14 to 87 articles), which coincides with the collapse of the Soviet Union and the end of the Cold War in the early 1990s, and many other events related with improvements in the nonproliferation efforts. This could have been because during the Cold War the threat of nuclear war was always present, but after the Soviet Union collapsed and the Cold War ended, real progress was made to change this reality (25). During this decade, more specifically, Belarus, Kazakhstan, and Ukraine gave up nuclear weapons, the Comprehensive Nuclear Test Ban Treaty – a treaty that banned nuclear explosions of any kind, including for weapons tests – opened for signatures, and the International Atomic Energy Agency's model additional protocol – a model that gave the agency further access to information and nuclear sites to better monitor whether countries were using nuclear technology for peaceful purposes – is introduced (34). As an example, 4 of the articles published between 1990 and 1999 contain the words "Cold" or "Soviet" in the title or author keywords (Table 2).

Table 2. Information on articles published between 1990 and 1999 with the words "Cold" or "Soviet" in the title or author keywords. Source: own compilation.

Title	Author keywords	Publication year	Publication country
Nationalism, Nuclear Policy, and Children in Cold War America	Children; Cold War; Nationalism; Nuclear Policy; US	1997	United States
Framing National-Security Elite Public Discourse on Nuclear-Weapons During the Cold-War	Frames; Nuclear Weapons; Peace Movements; Public Discourse; Social Movements	1995	United States
Sociology and Nuclear-Weapons – A View from Outside	Sociology; Nuclear Weapons; Nuclear War; Cold War	1992	United States
Tactical Nuclear-Weapons, Soviet Command Authority, and the Cuban-Missile-Crisis – A Note	It does not have author keywords.	1993	Canada

2005 – 2013 and 2014 – 2020: North Korea's withdrawal from Six Party Talks, Fukushima accident, nuclear powers agreement with Iran, UN adoption of Nuclear Weapons Ban Treaty

Two more significant growths occurred between 2005 and 2013 (+ 104 articles) and between 2014 and 2020 (+ 181 articles), being 2020 the year with the highest number of articles published (374 articles). Some historical events that may have influenced these numbers occurred in 2006, when North Korea announced that it had successfully conducted its first nuclear test (36); in 2009, when the same country walked out of Six Party Talks – the name given to the negotiations among China, Japan, North Korea, Russia, South Korea, and the United States to find a peaceful resolution to North Korea's nuclear weapons program (68); in 2011, when there was a nuclear accident at the Fukushima Daiichi Nuclear Power Plant caused by an earthquake and tsunami natural disaster (69) becoming the second worst nuclear disaster in the history of nuclear power generation, after the Chernobyl accident in 1986 (70); in 2015, when World Powers (the United States, China, France, Germany, Russia, the United Kingdom, and the European Union) reached a nuclear agreement with Iran after years of negotiation (JCPOA); and in 2017, when the United Nations adopted the Nuclear Weapons Ban Treaty/Treaty on the Prohibition of Nuclear Weapons – the first legally binding treaty for nuclear disarmament in 20 years (34). In the data sample, 31 out of the 930 articles published between 2005 and 2010 contain the word “North Korea” in the title, abstract or author keywords, and 4 articles contain the words “nuclear test” and “North Korea” in the title, abstract or author keywords. Additionally, all the 18 articles published between 2011 and 2013 contain the word “Fukushima” in the title, abstract, or author keywords. As for the second range of years (2014-2020), 8 out of the 2027 articles published between 2014 and 2020 contain the acronym “JCPOA” in the title, abstract or keywords, and 24 out of the 1300 articles published between 2017 and 2020 contain the word “Ban Treaty” in the title, abstract or author keywords.

5.2. The contribution of different regions: continents, countries, and affiliations

5.2.1. Analyzing the geographical distribution of authors

As previously mentioned, the first author of the list of authors in each article is the one who contributes the most for the research, being the one considered in the analysis, and the first affiliation of the main author of each article represents the institution where the author worked at the time the study was performed.

Table 3. Information on geography of authorship (considering main author and his/her first affiliation at the time of publication of the article). Source: own compilation.

First author's continent	Unique main author's countries	Unique main authors	Percentage unique main authors	Country with highest number of unique main authors	Highest number of unique main authors per country
North America	5	1832	0.457	United States	1705
Europe	34	1135	0.283	United Kingdom	350
Asia	28	765	0.191	Viet Nam	376
Oceania	5	161	0.040	New Zealand	134
Africa	18	61	0.015	Zimbabwe	32
South America	8	51	0.013	Venezuela	32

Consequently, by plotting the aggregated authorship information on Table 3 it is possible to picture the international distribution of knowledge production. In total, 98 different countries had a contribution in the articles of this study, spread across the 6 continents. Most authors (1832 authors) are established in North America in only 5 different countries. Europe occupies the second place with approximately 30 percent of the population of authors in 34 countries, followed by Asia in the third position with around 2 percent of the total number of authors, in 28 countries. Looking at the countries of each continent with the highest number of authors, the top three is composed by the United States in the first place again with 1705 authors (only the remaining 127 are spread across the other 4 countries), Vietnam in the second position now, with 376 authors, and the United Kingdom in the third, with 350 authors.

5.2.2. Comparing the research output with publication country and main author's country

Another way of exploring the contribution that different regions have in the articles studied is by comparing the quantity of research output considering the publication country with the quantity of research output considering the country where the main author worked at the time. This is because an article may have been written by an author who worked in a different country than the one where the article was published.

Figure 6. Top 15 – Quantity of research output per publication country and per country of main author's affiliation at the time. Source: own compilation.

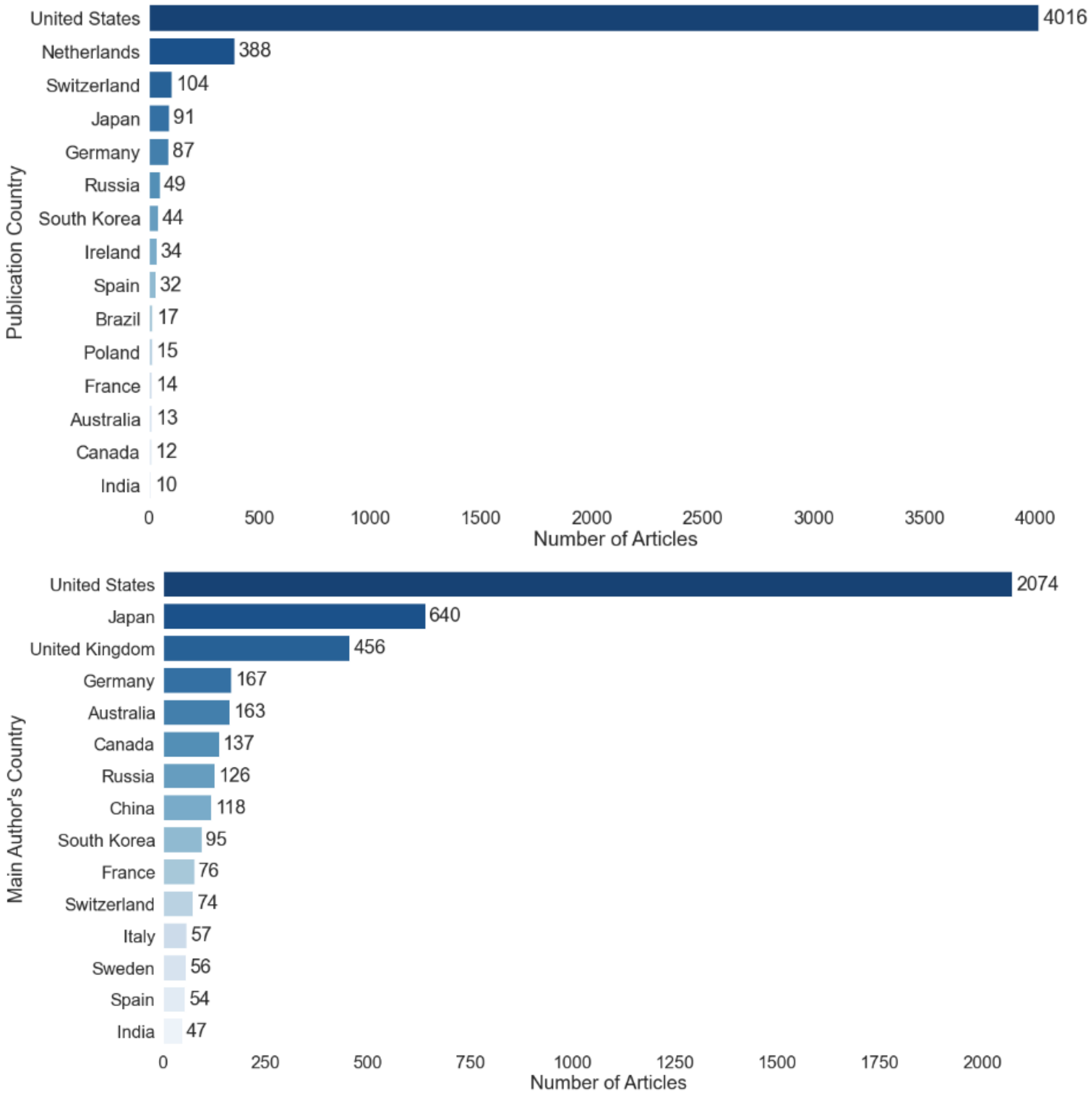
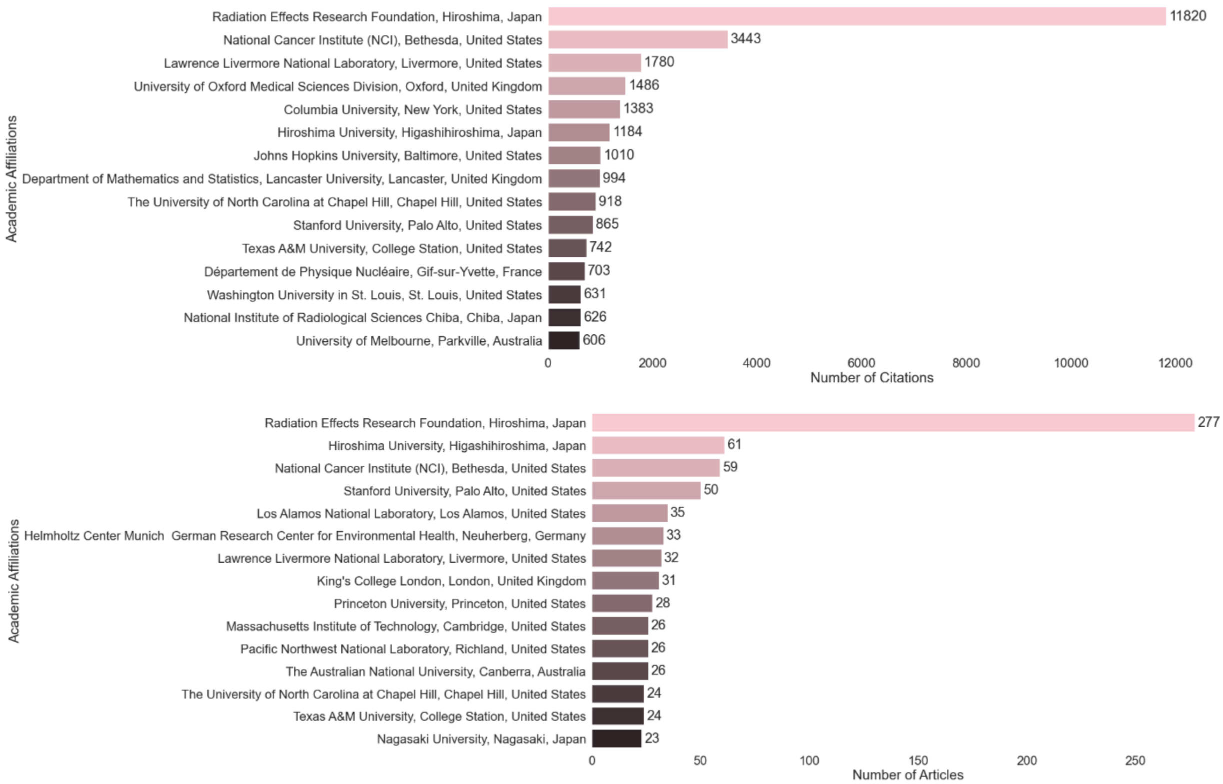


Figure 6 represents the quantity of nuclear weapons-related research per publication country (first graph) and per country of main author's affiliation (second graph). In the first case the United States is the country with the highest number of articles published (4016), representing around 80 percent of the total number of articles studied. The second one is the Netherlands, with only 388 articles, a big difference of 3628 articles less than the United States. Then there is Switzerland, Japan, and Germany with similar numbers of articles published (an average of 94 articles each). Regarding the number of articles per main author's country at the time of publication of the article, although the United States continues to be the country with the highest number of articles, now Japan occupies the second position, meaning that 640 of the articles studied were written by authors whose affiliation at the time was in Japan (these articles may have been published in Japan or not). The United Kingdom occupies the third position, with 456 articles whose author's affiliation was in the United Kingdom at the time. The next countries are Germany, Australia, and Canada, with an average of 156 articles per each one of the three.

5.2.3. Evaluating affiliations activity through number of articles and citations

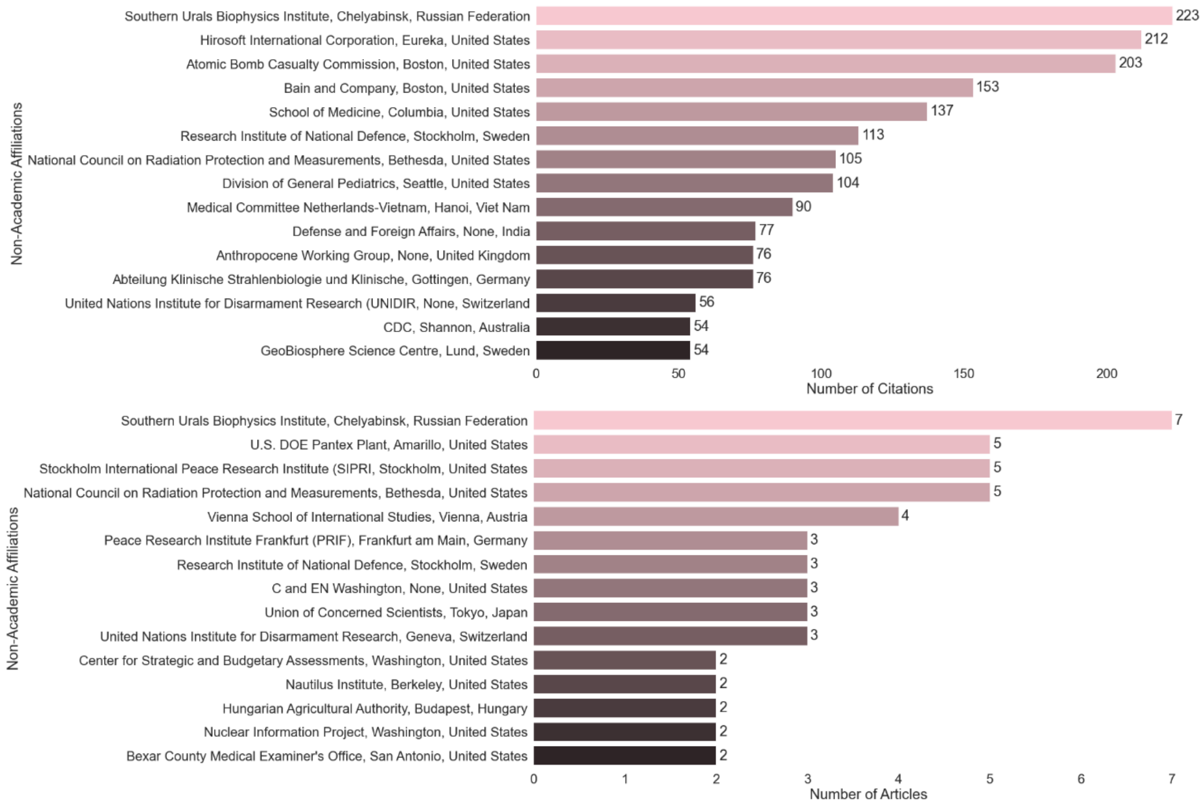
There are 4568 articles written by authors linked to academic affiliations at the time, and only 470 articles written by authors associated with non-academic affiliations. Figure 7 and Figure 8 represent the most relevant academic and non-academic affiliations, respectively, in terms of quantity of research output and number of citations. For this, the 1696 unique academic affiliations were considered in Figure 7, and the 387 unique non-academic affiliations in Figure 8.

Figure 7. Top 15 – Academic affiliations in quantity of research output and citations. Source: own compilation.



Regarding academic affiliations (Figure 7), the "Radiation Effects Research Foundation" is the one with the highest number of articles (277). The second position is occupied by "Hiroshima University" with 61 articles, and the following ones by the "National Cancer Institute (NCI)", "Stanford University", and "Los Alamos National Laboratory" in the United States, with 59, 50, and 35 articles, respectively. As for the top academic affiliations in terms of number of citations, the first one is again the same Japanese affiliation as in the first case with a total of 11820 citations. This means that the articles whose authors are affiliated to this institution (277) have a total of 11820 citations. In the second place is the "National Cancer Institute (NCI)" with 3443 citations, followed by "Lawrence Livermore National Laboratory" in the United States, “University of Oxford Medical Sciences Division” in the United Kingdom, and “Columbia University” in the United States with 1780, 1486, and 1383 citations respectively.

Figure 8. Top 15 – Non-academic affiliations in quantity of research output and citations. Source: own compilation.



As for the top non-academic affiliations (Figure 8) in terms of number of citations, “Southern Urals Biophysics Institute” in Russia, and “Hirosoft International Corporation” and “Atomic Bomb Casualty Commission” in the United States occupy the first three positions with a total of 223, 212, and 203 citations respectively. The first position is occupied by the same institution in terms of number of articles (with only 7 articles associated), being the next two ones also North American institutions – “Stockholm International Peace Research Institute” and “U.S. DOE Pantex Plant” – with 5 articles each.

5.3. Publishing outlets and thematic profile

5.3.1. Most relevant sources

It is also interesting to examine the journals and respective country which publish more articles on the topic. As already indicated on Figure 6, the United States is the country with the highest number of articles published (4016), which can be again confirmed by Figure 9, where from the 15 journals that occupy the top positions in terms of quantity of research output, only three (“Journal of Radioanalytical and Nuclear Chemistry”, “Science of the Total Environment”, “Journal of Radiation Research”) are not North American. As such, “Radiation Research” is the top 1 journal with 198 articles published, followed by “Health Physics” and “Bulletin of the Atomic Scientists”, with 177 and 158 articles published on the topic, respectively.

Figure 9. Top 15 – Number of articles published per journal and respective country. Source: own compilation.

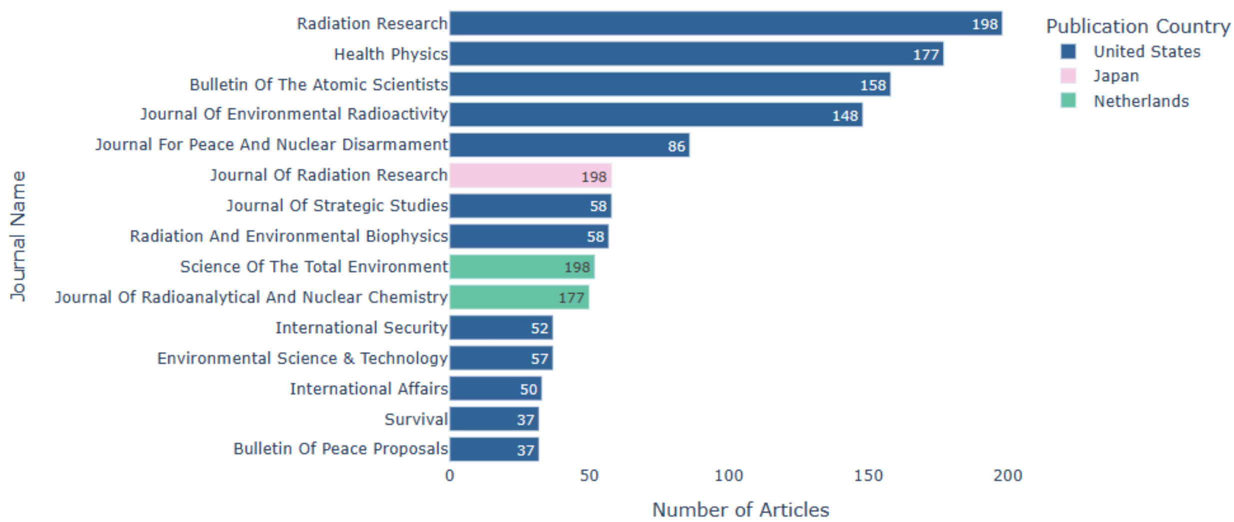
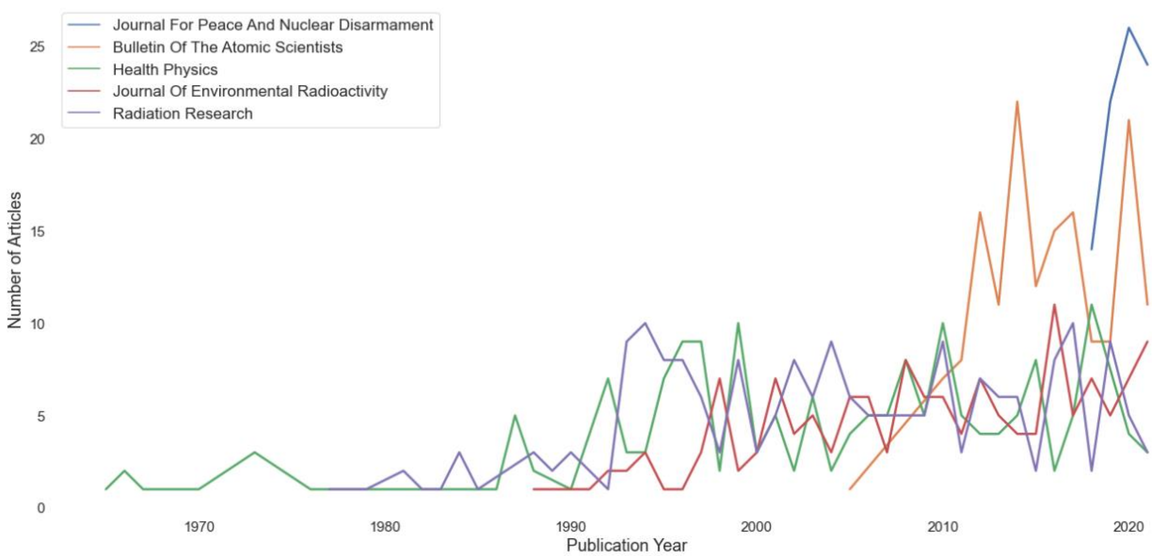


Figure 10 plots the evolution of the five most productive sources on nuclear weapons-related articles throughout the years. The first journal with publications is the "Health Physics", and it surges in the 1960s. At the end of the 1970s, the "Radiation Research" appears. Besides these journals are the first ones to publish articles on the referred topic, they are also the ones with the highest number of articles published, as already identified on Figure 9. The "Journal of Environmental Radioactivity" is the third one to surge, at the end of the 1980s. Only many years later, in around 2005, the" Bulletin of the Atomic Scientists" (the third one with the highest number of articles published) begins to publish articles on nuclear weapons. Considering the first five most productive journals, the "Journal for Peace and Nuclear Disarmament" is the most recent one publishing articles on the topic, only since around 2019. For this graph, only the most five productive sources on the referred topic were considered, meaning that there may be journals with less articles published which started publishing earlier on nuclear weapons.

Figure 10. Top 5 – Journals productivity over time, 1965-2021. Source: own compilation.



5.3.2. Main knowledge categories

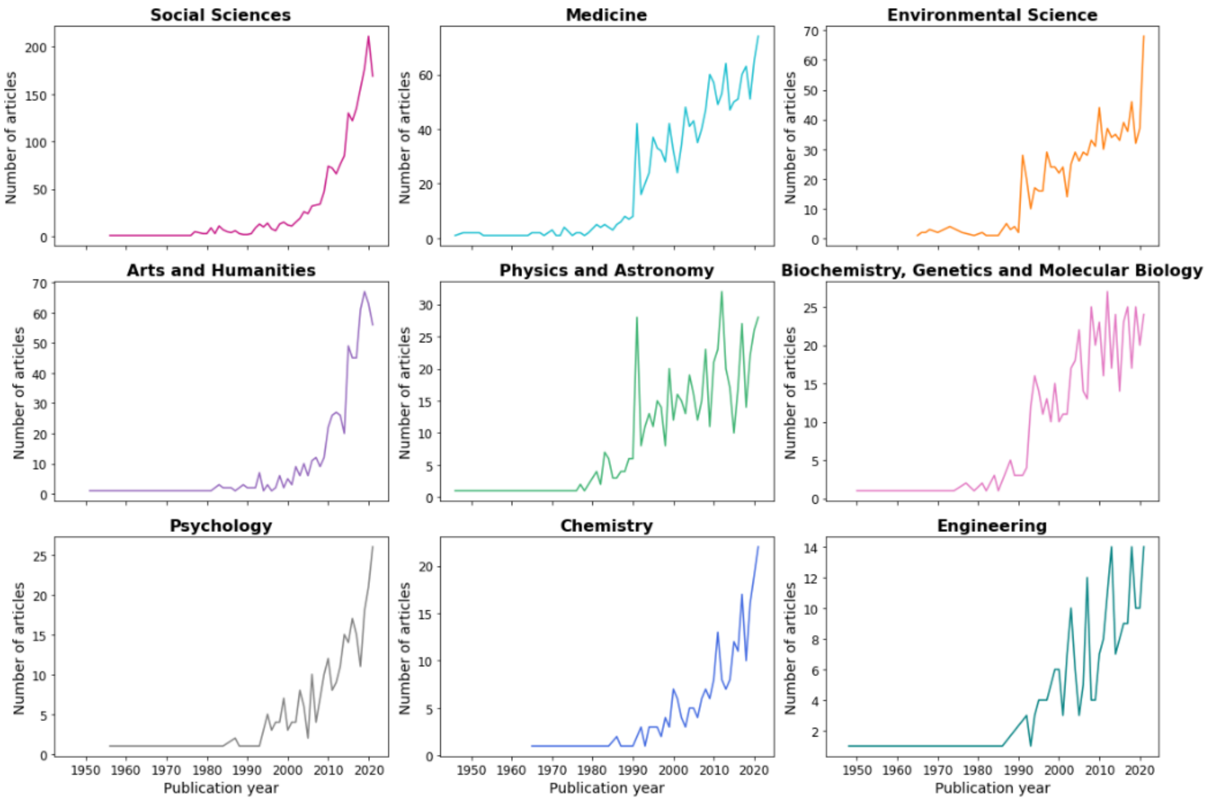
To develop a broader categories' classification scheme for the articles, 27 major disciplinary categories obtained from the Scopus Abstract Retrieval API were considered in this step (Appendix, Table 13). It is important to remember that articles can be classified into more than one field, which generates overlap: in the data studied, around 46% of the articles belong to more than one of the 27 Scopus major categories. Table 4 provides information on the quantity and quality of the 10 (out of 27) most relevant categories ordered by number of articles. Around 23% of the total nuclear weapons-related research is classified as belonging to the knowledgebase of "Social Sciences", followed by "Medicine" with approximately 17%, and "Environmental Science" with 11.42%. Regarding quality, "Medicine" is the area which contains the highest number of citations with a total of 44484 citations, then "Biochemistry, Genetics and Molecular Biology" with 22378 citations, and "Physics and Astronomy" with 18102, considering the first 10 categories with more articles. In terms of averages, "Biochemistry, Genetics and Molecular Biology" is the category with the highest average of citations (around 41 citations), whereas "Medicine" occupies the first position with an average of 674 citations per year.

Table 4. Top 10 – Information on categories ordered by number of articles. Source: own compilation.

Category	Number of articles	Percentage of articles	Number of citations	Average of citations	Average of citations per year
Social Sciences	1887	22.55%	18059	9.57	354.10
Medicine	1458	17.42%	44484	30.51	674.00
Environmental Science	956	11.42%	16238	16.99	331.39
Arts and Humanities	634	7.58%	4767	7.52	113.50
Physics and Astronomy	592	7.07%	18102	30.58	369.43
Biochemistry, Genetics and Molecular Biology	542	6.48%	22378	41.29	486.48
Psychology	270	3.22%	6372	23.60	177.00
Chemistry	237	2.83%	4981	21.02	150.94
Engineering	230	2.75%	2727	11.86	77.91
Earth and Planetary Sciences	216	2.58%	5446	25.21	155.60

The time path of the 9 categories containing more nuclear weapons-related research is represented on Figure 11. Most of these "core" fields begin to grow more in the decade between 1980 and 1990. Additionally, "Medicine", "Physics and Astronomy", "Biochemistry, Genetics and Molecular Biology", "Engineering", and "Arts and Humanities" encompass the "oldest" articles.

Figure 11. Top 9 – Thematic distribution of articles, 1946-2021. Source: own compilation.



Note: The scale of each subplot is adjusted to the number of articles in each category.

5.4. AUTHORS' IMPACT AND PERFORMANCE DYNAMICS

5.4.1. Measuring author's impact

Another useful result that can be retrieved from this study is the impact of the different authors present in the articles analyzed.

When comparing the authors' names with the authors' ids, there were less 14 names than ids. In particular, there were 14 names which corresponded to 2 different authors each. This means that there were 28 authors with equal names (considering the first and the last name of each author). In this way, besides the author's name, it was also necessary to include the id of each author in the evaluation of the authors' impact to prevent the situations where an author name is being interpreted as one person, when it may be associated to two different people.

Considering number of articles and citations

The authors' impact can be evaluated not only in terms of total number of articles that the author has published, but also in terms of total number of citations that his/her articles have (number of articles citing his/her work). Thus, for each article it was considered the main author and his/her affiliation at the time the work was performed, and then aggregated each pair (main author, affiliation) in each situation presented below: quantity of research output on Table 5 and quantity of citations on Table 6.

Table 5. Top 15 – Number of articles per main author and his/her first affiliation at the time. Source: own compilation.

Main author name	Affiliation of main author at the time of publication	Number of articles
Kristensen, Hans M.	Federation of American Scientists, Washington, D.C., United States	13
Kusunoki, Yoichiro	Radiation Effects Research Foundation, Hiroshima, Japan	13
Cologne, John B.	Radiation Effects Research Foundation, Hiroshima, Japan	11
Richardson, David B.	The University of North Carolina at Chapel Hill, Chapel Hill, United States	11
Shizuma, Kiyoshi	Hiroshima University, Higashihiroshima, Japan	11
Yamada, Michiko	Radiation Effects Research Foundation, Hiroshima, Japan	10
Reed, Bruce Cameron	Alma College, Alma, United States	10
Arbatov, Alexei G.	Primakov National Research Institute of World Economy and International Relations, Russian Academy of Sciences, Moscow, Russian Federation	9
Yoshida, Kengo	Radiation Effects Research Foundation, Hiroshima, Japan	8
Fuhrmann, Matthew	Texas A&M University, College Station, United States	8
Nakashima, Eiji	Radiation Effects Research Foundation, Hiroshima, Japan	8
Ron, Elaíne	National Cancer Institute (NCI), Bethesda, United States	7
Kroenig, Matthew	Georgetown University, Washington, D.C., United States	7
Land, Charles E.	National Cancer Institute (NCI), Bethesda, United States	7
Greenberg, Michael R.	Edward J. Bloustein School of Planning and Public Policy, New Brunswick, United States	7

The 15 most productive authors (in terms of number of articles) are represented on Table 5. The ones who occupy the first 5 positions are Kristensen, Hans M.; Kusunoki, Yoichiro; Cologne, John B.; Richardson, David B., and Shizuma, Kiyoshi, linked to the Radiation Effects Research Foundation and the Hiroshima University both in Japan, and to the Federation of American Scientists and the University of North Carolina, both in the United States, with 13 and 11 articles published.

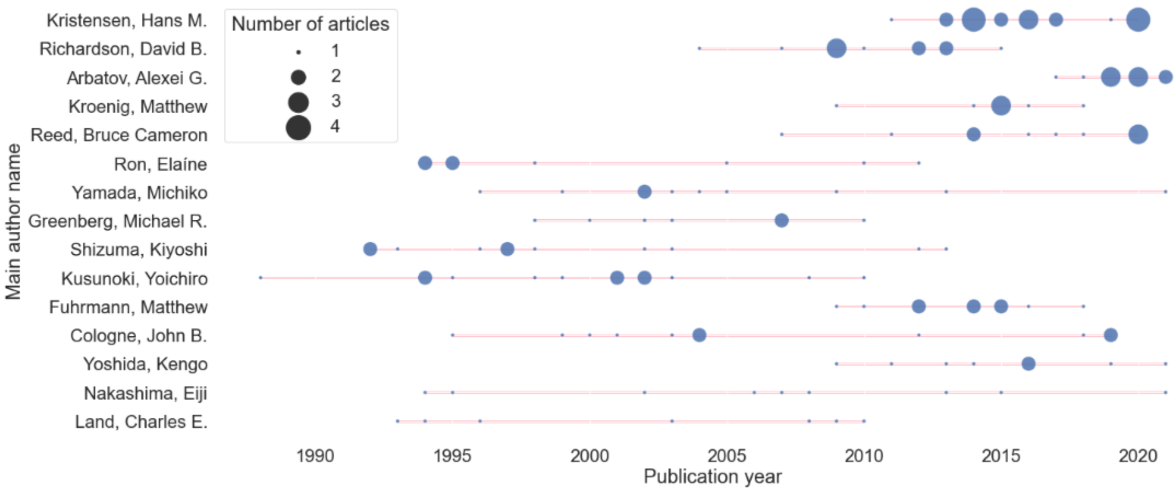
Table 6. Top 15 – Number of citations per main author and his/her first affiliation at the time. Source: own compilation.

Main author name	Affiliation of main author at the time of publication	Number of citations
Preston, Dale L.	Radiation Effects Research Foundation, Hiroshima, Japan	2061
Pierce, Donald Alan	Radiation Effects Research Foundation, Hiroshima, Japan	1552
Berrington de González, Amy	University of Oxford Medical Sciences Division, Oxford, United Kingdom	1283
Ron, Elaíne	National Cancer Institute (NCI), Bethesda, United States	1229
Diggle, Peter J.	Department of Mathematics and Statistics, Lancaster University, Lancaster, United Kingdom	994
Campbell, Jacquelyn C.	Johns Hopkins University, Baltimore, United States	867

Brenner, David J.	Columbia University, New York, United States	851
Kersting, Annie B.	Lawrence Livermore National Laboratory, Livermore, United States	832
Mueller, Th A.	Département de Physique Nucléaire, Gif-sur-Yvette, France	703
Yamada, Michiko	Radiation Effects Research Foundation, Hiroshima, Japan	635
Payne, B. Keith	Washington University in St. Louis, St. Louis, United States	631
Ozasa, Kotarou	Radiation Effects Research Foundation, Hiroshima, Japan	561
Shimizu, Yukiko	Radiation Effects Research Foundation, Hiroshima, Japan	495
Lam, Shujie	University of Melbourne, Parkville, Australia	454
Land, Charles E.	National Cancer Institute (NCI), Bethesda, United States	453

As for the most influential authors (in terms of number of citations), the results are a little different. In this case the author with the highest number of citations is Preston, Dale L., affiliated to the Radiation Effects Research Foundation in Japan, with 2061 citations. This author is followed by Pierce, Donald Alan affiliated to the same institution, with 1552 citations, and Berrington de González, Amy affiliated to the University of Oxford Medical Sciences Division in the United Kingdom, with 1283 citations (Table 6). In both situations, the most predominant institution in which the authors were affiliated at the time they developed their work is the Radiation Effects Research Foundation in Japan (linked to 5 of the authors in each scenario), as it had already been verified by Figure 7, where the top academic affiliations were analyzed.

Figure 12. Top 15 - Production (first) and influence (second) of authors over time, 1967-2021. Source: own compilation.



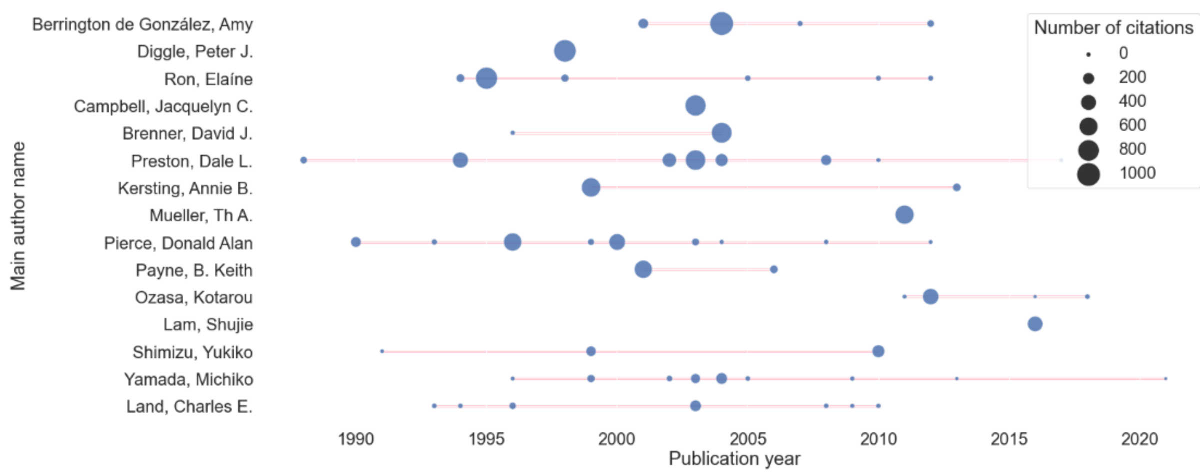


Figure 12 identifies the 15 most productive and influence authors considering their name and their id over the years. Contrarily to the analyzes carried out on Table 5 and on Table 6, for this one it was not consider the authors' affiliations, but instead the total number of articles and citations per author over the years independently of their affiliations at the time. In terms of quantity of articles, Kusunoki, Yoichiro, with a total of 13 articles published, is the first author publishing nuclear weapons-related articles, at the 1960s. On the other hand, Kristensen, Hans M., with a total of 19 articles (from which 13 were developed when the author was affiliated to the Federation of American Scientists, as indicated on Table 5), only begins to publish articles at around 2011. Regarding the number of citations, Preston, Dale L. is the first author with citations, at the end of the 1980s, followed by Pierce, Donald Alan, who begins to have them in 1990.

Considering H-Index

Besides the number of articles and citations, the authors' impact can also be measured through the H-Index, which is believed to be a reliable indicator of the authors' scholarly achievement (71). This index was proposed by J.E. Hirsch in 2005 and it is a quantitative metric based on analysis of publication data using publications and citations to provide an estimate of the importance, significance, and broad impact of a scientist's cumulative research contributions (72). The H-Index refers to the highest number 'h' such that the author has h publications, each with at least h citations.

Table 7. Articles authored by Boice, John Dunning. Source: own compilation.

Article	Main author name	Main author ID	Affiliation of main author at the time of publication	Main author H-Index
1	Boice, John Dunning	7103070062	Vanderbilt-Ingram Cancer Center, Nashville, United States	94
2	Boice, John Dunning	7103070062	National Council on Radiation Protection and Measurements, Bethesda, United States	94
3	Boice, John Dunning	7103070062	National Council on Radiation Protection and Measurements, Bethesda, United States	94
4	Boice, John Dunning	7103070062	International Epidemiology Institute, Rockville, United Sates	94
5	Boice, John Dunning	7103070062	National Council on Radiation Protection and Measurements, Bethesda, United States	94
6	Boice, John Dunning	7103070062	National Council on Radiation Protection and Measurements, Bethesda, United States	94
7	Boice, John Dunning	7103070062	National Cancer Institute (NCI), Bethesda, United States	94
8	Boice, John Dunning	7103070062	National Council on Radiation Protection and Measurements, Bethesda, United States	94

Note: The H-Index is the same although the author's affiliations change.

Contrarily to what happened in the previous situations with the number of articles and the number of citations (Table 5 and Table 6), the H-Index of an author does not depend on his/her affiliation at the time of development of an article. To illustrate this, on Table 7 it is represented the case of author

Boice, John Dunning, who has eight articles published in the data studied, associated with four unique affiliations (articles published in four different moments in time). As it can be verified, the H-Index of the author always assumes the same value, independently of the author's affiliations.

Table 8. Articles authored by Kato, Kazuo. Source: own compilation.

Article	Main author name	Main author ID	Affiliation of main author at the time of publication	Main author H-Index
1	Kato, Kazuo	57225954804	Radiation Effects Research Foundation, Hiroshima, Japan	3
2	Kato, Kazuo	57225954804	Radiation Effects Research Foundation, Hiroshima, Japan	3
3	Kato, Kazuo	57225954804	Radiation Effects Research Foundation, Hiroshima, Japan	3
4	Kato, Kazuo	57225954804	Radiation Effects Research Foundation, Hiroshima, Japan	3
5	Kato, Kazuo	57225954804	Radiation Effects Research Foundation, Hiroshima, Japan	3
6	Kato, Kazuo	7406870302	Hiroshima University, Higashihiroshima, Japan	8
7	Kato, Kazuo	7406870302	Hiroshima University, Higashihiroshima, Japan	8

Note: The H-Index varies because the authors' ids are different.

In the same way, in the cases where two authors have the same name (considering the first and the last name) as represented on Table 8, the H-Index assumes two different values because they refer to two different people (the identifiers are different).

Table 9. Top 15 – H-Index per main author name, ID, and number of affiliations. Source: own compilation.

Main author name	Main author ID	Number of unique affiliations of main author	Main author H-Index
Key, Timothy J.	7102740999	1	136
Degenhardt, Louisa	7004040033	1	118
Blot, William James	7102114492	1	114
Gale, Robert Peter	56480006900	1	111
Breslau, Naomi	35421498600	1	103
Lund, E.	55423514100	1	101
Pearce, Neil	56491440700	1	101
Linet, Martha S.	7005032041	1	95
Allen, Naomi E.	7202442006	1	95
Boice, John Dunning	7103070062	4	94
Yates, John T.Y.	57203637727	1	94
Goodman, Marc T.	7401997536	1	93
Jensen, Mark P.	7401786989	1	93
Slovic, Paul	7007111080	1	92
Shapiro, Steven D.	7402734374	1	92

In this context, Table 9 lists the first 15 authors with the highest H-Indexes, their respective ids, and number of unique affiliations. Key, Timothy J. occupies the first position with an H-Index of 136, meaning that this author has 136 articles published, each with at least 136 citations. The next positions are occupied by Degenhardt, Louisa, and Blot, William James, with H-Indexes of 118 and 114, respectively. To confirm what was already stated above, for the author Bouce, John Dunning with four unique affiliations, the H-Index is always 94 independently of the affiliations.

It is important to emphasize here that the results of this study were obtained with only a sample of data. This means that only some of the articles of the authors considered were analyzed (the ones containing the search terms mentioned in the beginning of this work in the title, abstract, or author keywords). The H-Index – calculated with the property "h_index" of the Author Retrieval API (Appendix, Table 14) – considers all the articles that an author has published independently of the affiliations in different moments in time (and not only the articles present in the dataset used in this research). Another point to remember is that the H-Index considers all the articles in which an author contributed to, meaning that he/she may have been the first author in the list of authors (and the one who contributed the most), or occupied other positions in the list of authors of each article. In this study it is only being considered the main author of each article and the articles related to nuclear weapons, but the calculation of the H-Index considers other articles where the author does not appear in the first position of the list of authors (71) and/or articles which are related to other fields.

Table 10. Articles authored by Lund, E. (name, ID, affiliation at the time, number of citations, and H-Index). Source: own compilation.

Article	Main author name	Main author ID	Affiliation of main author at the time of publication	Number of citations	H-Index
1	Lund, E.	55423514100	UiT Norges Arktiske Universitet, Tromso, Norway	29	101
2	Lund, E.	55423514100	UiT Norges Arktiske Universitet, Tromso, Norway	4	101

For example, in the data studied author Lund, E. (Table 10) has only two articles published in which he is the main author: one with 29 citations and the other with 4 citations. However, the author has an H-Index of 101, meaning that he has 101 articles, each with at least 101 citations. As such, there are clearly more articles for which the author had a contribution being considered in the calculation of this index: articles where the author was the main author (and were not used in this research), and/or others in which the author was not the main author (and may have been used in this study and/or not).

The H-Index is thus independent of the affiliations of the authors, and it is calculated considering all their articles although the institutions where they were working when they conducted different articles changed throughout the time.

5.4.2. Highly cited papers

Table 11 provides information on the 20 most cited articles sorted by number of citations. In this group of articles, 16 unique authors, 15 unique publication years, 14 unique academic affiliations, and 6 countries were identified. Additionally, the Radiation Effects Research Foundation in Japan is the most present affiliation (6 articles), while the United States is the affiliation country with the highest number of articles associated (8 articles). Preston, Dale L. is the author with the highest number of articles (3 articles). Of all the articles studied, the most cited article was published in 2004, has the title "Risk of Cancer from Diagnostic X-Rays: Estimates for The UK and 14 Other Countries", 1095 citations, and it is authored by Berrington de González A. whose affiliation at the time was the University of Oxford Medical Sciences Division in the United Kingdom. This article aims to estimate the extent of the diagnostic X-rays based on the annual number of diagnostic X-rays undertaken in the United Kingdom and in fourteen developed countries. The second position is occupied by the article with the title "Model-Based Geostatistics", published in 1998, with 994 citations, and authored by Diggle, Peter J. whose affiliation at the time was the Department of Mathematics and Statistics, Lancaster University in the United Kingdom.

Table 11. Top 20 – Details on highly cited articles. Source: own compilation.

Title	Number of authors	Main author name	Publication year	Number of citations	Affiliation of main author at the time of publication	Affiliation type	Main author affiliation country	Number of affiliations of main author
Risk of Cancer from Diagnostic X-Rays: Estimates for The UK and 14 Other Countries	2	Berrington de González, Amy	2004	1095	University of Oxford Medical Sciences Division	Academic	United Kingdom	1
Model-Based Geostatistics	3	Diggle, Peter J.	1998	994	Department of Mathematics and Statistics, Lancaster University	Academic	United Kingdom	1
Thyroid-Cancer After Exposure to External Radiation – A Pooled Analysis of 7 Studies	9	Ron, Elaíne	1995	950	National Cancer Institute (NCI)	Academic	United States	1
Risk Factors for Femicide in Abusive Relationships: Results from A Multisite Case Control Study	18	Campbel, Jacquelyn C.	2003	867	Johns Hopkins University	Academic	United States	1
Studies of Mortality of Atomic Bomb Survivors. Report 13: Solid Cancer and Noncancer Disease Mortality: 1950-1997	5	Preston, Dale L.	2003	806	Radiation Effects Research Foundation	Academic	Japan	1
Migration of Plutonium in Ground Water at The Nevada Test Site	6	Kersting, Annie B.	1999	721	Lawrence Livermore National Laboratory	Academic	United States	1
Improved Predictions of Reactor Antineutrino Spectra	12	Mueller, Th A.	2011	703	Département de Physique Nucléaire	Academic	France	1
Studies of the Mortality of Atomic Bomb Survivors. Report 12 .1. Cancer: 1950-1990	5	Pierce, Donald Alan	1996	641	Radiation Effects Research Foundation	Academic	Japan	1

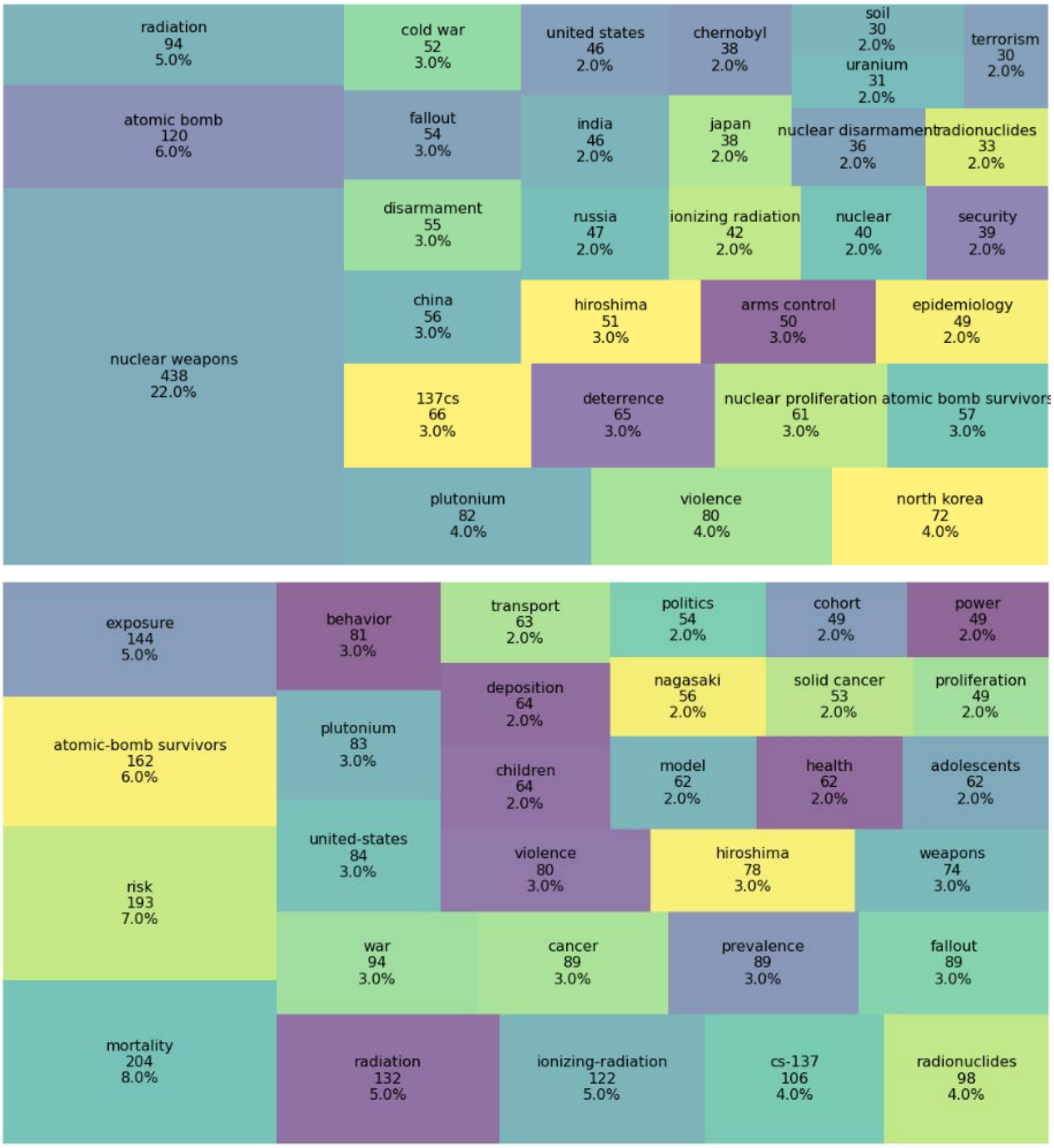
Prejudice and Perception: The Role of Automatic and Controlled Processes in Misperceiving a Weapon	1	Payne, B. Keith	2001	631	Washington University in St. Louis	Academic	United States	1
Radiation-Related Cancer Risks at Low Doses Among Atomic Bomb Survivors	2	Pierce, Donald Alan	2000	524	Radiation Effects Research Foundation	Academic	Japan	1
Studies of the Mortality of Atomic Bomb Survivors, Report 14, 1950-2003: An Overview of Cancer and Noncancer Diseases	9	Ozasa, Kotarou	2012	508	Radiation Effects Research Foundation	Academic	Japan	1
Estimated Radiation Risks Potentially Associated with Full-Body CT Screening	2	Brenner, David J.	2004	505	Columbia University	Academic	United States	1
Cancer Incidence in Atomic-Bomb Survivors 3. Leukemia, Lymphoma and Multiple-Myeloma, 1950-1987	13	Preston, Dale L.	1994	487	Radiation Effects Research Foundation	Academic	Japan	1
Combating Multidrug-Resistant Gram-Negative Bacteria with Structurally Nanoengineered Antimicrobial Peptide Polymers	11	Lam, Shujie	2016	454	University of Melbourne	Academic	Australia	1
Marked Longevity of Human Lung Parenchymal Elastic Fibers Deduced from Prevalence of D-Aspartate and Nuclear-Weapons Related Radiocarbon	5	Shapiro, Steven D.	1991	422	University of Utah Health Sciences	Academic	United States	1
Retrospective Birth Dating of Cells in Humans	5	Spalding, Kirsty L.	2005	395	Karolinska Institutet	Academic	Sweden	1
Radiation Effects on Breast Cancer Risk: a Pooled Anslysis of Eight Cohorts	6	Preston, Dale L.	2002	386	Radiation Effects Research Foundation	Academic	Japan	1
Effects of Low Doses and Low-Dose Rates of External Ionizing-Radiation-Cancer Mortality Among Nuclear Industry Workers in 3 Countries	17	Cardis, E.	1995	377	International Agency for Research on Cancer	Academic	France	1
Perceived Risk, Trust, and the Politics of Nuclear Waste	3	Slovic, Paul	1991	348	Decision Research, Eugene, United States	Academic	United States	2
Radiation Risks Potentially Associated with Low-Dose ET Screening of Adult Smokers for Lung Cancer	1	Brenner, David J.	2004	325	Columbia University, New York, United States	Academic	United States	1

5.5. The most common keywords

5.5.1. Most popular keywords

Author keywords are the keywords defined by the author(s) of an article, whereas keywords plus are words that frequently appear in the titles of an article's references, but do not appear in the title of the article itself. Figure 13 indicates the frequency – number of documents that contain a specific keyword – and respective percentage of some of the most popular author keywords and keywords plus used in nuclear weapons-related articles. Regarding author keywords represented in the first tree map, "nuclear weapons", "atomic bomb", and "radiation" are the top three keywords known in the studied literature, present in 438, 120, and 94 articles, respectively. As for the keywords plus, "mortality", "risk", and "atomic-bomb survivors" occupy the first three positions, occurring in 204, 193, and 162 articles, respectively.

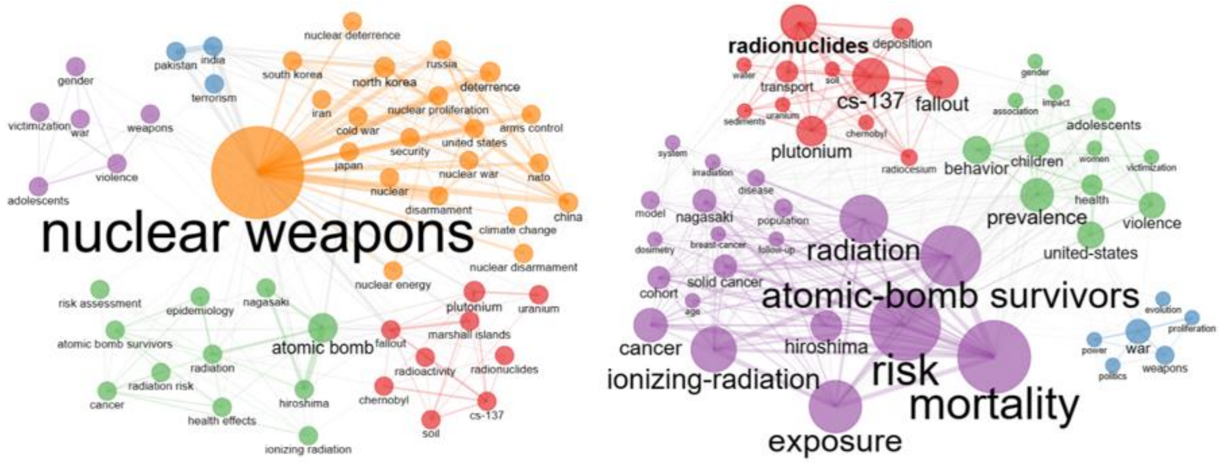
Figure 13. Top 30 – Author keywords (first) and keywords plus (second) used in the articles studied. Source: own compilation.



5.5.2. Keywords co-occurrence network

In order to complement the analysis and to understand the relationships between the different keywords used in nuclear weapons-related articles, a co-occurrence network was plotted in the Bibliometrix R Package after uploading the excel file containing the previously preprocessed data in Python software (5038 articles). Applied to the author keywords and to the keywords plus in this case, a keywords co-occurrence network (KCN) focuses on understanding the knowledge components and knowledge structure of a scientific/technical field by examining the links between keywords in the literature (10).

Figure 14. Keywords co-occurrence network composed by 5 clusters for the author keywords (left) and 4 clusters for the keywords plus (right). Source: bibliometrix R-package.



In this context, Figure 14 focuses on the analysis methods based on KCNs, which have been used in theoretical and empirical studies to explore research topics and their relationships in selecting scientific fields. In each one of the networks, 50 keywords were considered, and isolated keywords (keywords that only appear in only one document) were eliminated. The minimum edge is 2, meaning that at least two keywords must appear together in 2 documents to be considered in the networks (73). As such, the network on the left is composed by 5 clusters of author keywords, and the network on the right is composed by 4 clusters of keywords plus. If the keywords are grouped into the same cluster, they are more likely to reflect identical topics. Regarding the first network, each cluster has different numbers of author keywords, being the blue and the purple clusters the ones with smaller number of keywords. For example, in the orange cluster (the one with the highest number of keywords), the word "nuclear weapons" contemplates the highest number of articles (represented by the biggest circle), and usually appears with words like "disarmament", "nuclear proliferation", or "nuclear energy". On the other hand, in the green cluster the word "atomic bomb" is the one which appears in the highest number of articles, and it is commonly used with words like "atomic-bomb survivors", "Nagasaki" or "radiation". In relation to the second network, the blue cluster is the one with less keywords plus and the purple cluster the one which contains the highest number of keywords. In the purple cluster, the words "atomic-bomb survivors", "risk", and "mortality" are represented by the biggest circles (the three most popular keywords plus, as identified before on Figure 13) and tend to appear with words like "ionizing radiation", "disease", or "irradiation".

6. FINAL REMARKS

6.1. CONCLUSIONS

The aim of this study was to identify the worldwide trends in nuclear weapons scientific production considering different levels of analysis namely the relation between periods of growth in the number of articles published and relevant historical events; the contribution of different regions and sources; the main knowledge areas in which the articles are inserted; the evaluation of the authors' impact; the citation analysis of the research output; and the analysis of the most common keywords. For this purpose, a scientometric analysis using Python software, Scopus APIs, Bibliometrix R-package, and 5038 articles retrieved from the WoS Core Collection database in the period from 1946 to 2021 was performed.

The results obtained in this research are based on two important decisions made on an authors and affiliations' level: the selection of the main author and the selection of the first affiliation of the main author of each article. Additionally, the authors' affiliation selected in each article represents the main institution where the authors worked at the time they conducted the study. As such, even after the selection of the first affiliation of the main author, it is possible that an author has more than one affiliation, meaning the institutions that he/she had in different moments in time.

Regarding the authors' H-Index, computed with the Scopus APIs, it was concluded that in its calculation it considers all the publications that an author has and not only his/her publications in the set of 5038 articles analyzed. Besides this, in the articles to which the authors contributed they may occupy the first position in the list of authors of each article, or another one. Thus, there are situations where an author has, for example, two articles published where he surges as the main author, with 29 and 4 citations each, and an H-Index of 101. This means that among all his/her articles, 101 have received at least 101 citations. However, only two of all the articles in which the author participated are "visible" for us, not only because just nuclear weapons-related articles were considered for the purpose of this study, but also because the analysis was made on a main author's level.

According to the results obtained, the following are the answers to the questions referred in the beginning of the study:

- **Is there any relation between the time when there is growth in the number of articles published with historical events on nuclear weapons? Can these events influence the scientific production on nuclear weapons?**

The results suggest that the interest in research on the topic of nuclear weapons has been increasing in general in the last decades, particularly since 1990. Additionally, some of the periods when the production of articles registered higher growth coincided and may have been influenced by significant historical events, since the findings also reveal the presence of specific keywords (related with such events) in the articles published during these years (Section 5.1.1).

- **What are the most contributing regions on the referred topic, considering the main authors and their affiliations?**

North America and Europe are the most contributing continents to the research output. From these continents, the United States and the United Kingdom are the countries with more authors (Section 5.2.1). The analysis resulted from the comparison of the quantity of research output considering the main author's country and the publication country (Section 5.2.2) identified the United States as the country with highest number of articles in both situations, which makes sense since it is the one where most authors resided/worked at the time they wrote the articles. Japan occupies the second position considering the main author's country, but the Netherlands is the second country considering the publication country, which suggests that authors may be affiliated to a certain country and their articles are published in a different one. Finally, the results of the affiliations' analysis (Section 5.2.3) revealed that the most productive and relevant academic affiliation is the "Radiation Effects Research Foundation" in Japan; as for the non-academic affiliation, the most contributing one is the "Southern Urals Biophysics Institute" in Russia, in both scenarios too.

- **Which sources publish more on nuclear weapons-related articles and how is their production over time?**

The sources analysis identified that the American journal the "Radiation Research" leads the production with 198 articles published. Additionally, from the 15 most productive journals, 12 are American. The "Health Physics" is the first journal to publish articles on nuclear weapons, whereas the "Journal for Peace and Nuclear Disarmament" is the last one, considering the 5 most productive journals (Section 5.3).

- **What are the most common categories and their distribution over the years?**

"Social Sciences", "Medicine", and "Environmental Science" are the categories with higher number of articles on the topic. The decade between 1980 and 1990 is when the majority of the top 9 most popular categories verify an increase in the number of articles (Section 5.3.2).

- **Who are the most important authors in the field considering articles published, citations, and H-Index?**

Kusunoki, Yoichiro and Kristensen, Hans M. are the authors with more articles published on the topic, whereas Preston, Dale L. is the author with more citations. Kusunoki, Yoichiro and Preston, Dale L. were affiliated to the Japanese institution the "Radiation Effects Research Foundation" at the time the work was conducted, whereas Preston, Dale L. was affiliated to the "Federation of American Scientists". These authors may occupy different positions in the results if they were affiliated to other institutions at the time they wrote other articles because the analysis is dependent on the author's affiliation. The H-Index, on the other hand, does not depend on the author's affiliations. According to the results, Key, Timothy J. is the author with the highest H-Index. The author's impact is thus different dependently on the level considered. For example, the author with more citations is not the same as the one with more articles published on the topic. However, of the three levels considered, the H-Index is the most reliable one since it measures the quality (citations of all the author's articles) and the quantity (total number of articles to which the author contributed) at the same time (Section 5.4.1).

- **What are the most popular articles, and what are some of their characteristics?**

The citation analysis of the research output revealed that the top three most cited articles were published in 2004, 1995, and 1998. Their authors worked at the "University of Oxford Medical Sciences Division" in the United Kingdom, at the "National Cancer Institute (NCI)" in the United States, and at the "Department of Mathematics and Statistics, Lancaster University" in the United Kingdom at the time they wrote these articles. Moreover, all the top twenty most cited articles are authored by authors who were affiliated to academic affiliations at the time (Section 5.4.2).

- **What keywords are commonly used in nuclear weapons-related articles and with which other keywords they usually appear?**

The most common author keywords are "nuclear weapons", "atomic bomb", and "radiation". As for the keywords plus, "mortality", "risk", and "atomic-bomb survivors" occupy the first three positions (Section 5.5.1). The keywords co-occurrence network suggests that the author keywords "nuclear weapons" are often used with words like "nuclear war", "nato", or "nuclear disarmament"; the word "atomic bomb" is usually used with words like "ionizing radiation", "radiation risk", or "health effects". Regarding keywords plus, "mortality", "risk", and "atomic-bomb survivors" usually appear with words like "disease", "solid cancer", or "ionizing radiation" (Section 5.5.2).

6.2. LIMITATIONS AND FUTURE WORK

The results of a scientometric analysis regarding an emergent theme like nuclear weapons can, in a few years or in shorter periods of time, be subject to substantial variations because the data used to perform this research relies on databases being constantly updated with the increase or decrease in the number of indexed journals. Another limitation of the present article is that it explores a particular theme using two different databases, and although they represent two of the most influential and commonly used databases, the global perspective could be improved with the inclusion of other ones. During the data preprocessing section, a lot of articles were discarded from the study because they did not exist in Scopus (only in WoS), which originated loss of important information for the study, constituting another limitation encountered. As previously noted, in the WoS Core Collection database it is only possible to retrieve 500 articles at a time, which is not practical and consumes a lot of time when there are thousands of articles to extract. Finally, the data extracted from the WoS Core Collection database is raw data which required robust data preprocessing before the analysis.

As for future developments, since this study only explores the articles' citations by other articles, an analysis on the articles' references may be also performed. Another possibility for future research is to select a category (e.g.: Social Sciences) or a country (e.g.: the United States) to deepen the analysis and get insights on specific groups of articles. It may be also interesting to explore the factors that contribute to some countries not being big contributors in the production of articles related with nuclear weapons, or to explore if in some of these countries the numbers are changing. Finally, besides the co-occurrence network, other network analysis may be carried out to understand, for example, which authors (author collaboration network) and countries (country collaboration network) usually work together.

Additional points to consider regarding the H-Index are that it may characterize the authors in different levels of importance (successful, outstanding, truly unique) according to the scientific activity of the author. For example, an H-Index of 40, after 20 years of scientific activity, characterizes outstanding scientists, likely to be found at the top universities or major research laboratories, but an H-Index of 10 after the same time of scientific activity represents authors with much less impact. Since the scientific activity of the authors was not considered in this study, it is a possibility for further studies and developments of the present article to detail the characterization of the impact of different authors including the scientific activity. Moreover, there are "paradoxical situations" for authors who have the same number of publications and the same H-Index, but different citation counts. As an example, author A has 8 publications which have been cited a total of 338 times and author B has 8 publications which have been cited a total of 28 times. Both authors have the same H-Index of 5 but author A has a higher citation rate than author B. Again, these situations were not explored in this study because the analysis is based on a specific topic and on a main author's level, and the H-Index is a metric to assess the entire body of scholarly output by an author, not intended for a specific timeframe or topic (72).

In the context of the Climate Crisis, which is a closer and very worrying matter today, and since nuclear weapons have severe ecological impacts, it may be useful for politicians and economists to explore the insights obtained in this article to investigate on politics and measures that promote disarmament and non-proliferation of these weapons to adopt in their countries. Scientists, physicians, doctors, and sociologists may also want to know more about the production of scientific literature on nuclear weapons and use the results of this article as input of other studies for different purposes.

To sum up, the outcome of this study can help researchers to dive deeper into specific areas discussed in the article. The findings provide researchers with a panorama of worldwide nuclear weapons research and suggest direction for further study to the field, offering the possibility for researchers to expand the databases used, the level of analysis according to different objectives, or even the possibility of using additional software tools that facilitate the analysis.

7. REFERENCES

1. **Geneva Environment Network.** Nuclear Disarmament and Ecological Impacts of Nuclear Weapons. *genevaenvironmentnetwork.* [Online] 2021. [Citação: 22 de October de 2022.] https://www.genevaenvironmentnetwork.org/resources/updates/nuclear-disarmament-and-ecological-impacts-of-nuclear-weapons/.

2. **Union of Concerned Scientists.** Nuclear Weapons Justice. *ucsusa.* [Online] 2022. [Citação: 23 de October de 2022.] https://www.ucsusa.org/nuclear-weapons/justice.

3. **ICRC International Committee of the Red Cross.** Nuclear weapons - an intolerable threat to humanity. 7 de August de 2018.

4. *Scientometrics of Forest Health and Tree Diseases: An Overview.* **Pautasso, Marco.** 17, 2016, Forests, Vol. 7, p. 12.

5. *Data Mining in Scientometrics: usage analysis for academic publications.* **Mryglod, Olesya, Holovatch, Yurij e Kenna, Ralph.** 9 de July de 2018, Institute of Electrical and Electronics Engineers (IEEE), p. 6.

6. *A big picture: bibliometric study of academic publications from post-Soviet countries.* **Chankseliani, Maia, Lovakov, Andrey e Pislyakov, Vladimir.** 18 de August de 2021, Scientometrics, p. 30.

7. *Impact and visibility of Norwegian, Finnish and Spanish journals in the fields of humanities.* **Sanz-Casado, Elías, et al.** 2021, Scientometrics, p. 19.

8. *Obesity researches in youth: A scientometrics study in Middle East countries.* **Atoof, Mehrdad Kazemzadeh, et al.** 2021, Journal of Research in Medical Sciences: the Official Journal of Isfahan University of Medical Sciences, p. 7.

9. *Worldwide trends in the scientific production on rural depopulation, a bibliometric analysis using bibliometrix R-tool.* **Rodríguez-Soler, Rocío, Uribe-Toril, Juan e Valenciano, Jaime De Pablo.** 2020, Land Use Policy, p. 20.

10. *Social media and e-commerce: A scientometrics analysis.* **Javid, E. Taiebi, Nazari, M. e Ghaeli, M. R.** 2019, International Journal of Data and Network Science, p. 22.

11. *The impact of social media in business growth and performance: A scientometrics analysis.* **Pourkhani, A., et al.** 2019, International Journal of Data and Network Science, p. 22.

12. *Big data and social media: A scientometrics analysis.* **Esfahani, Hossein Jelvehgaran, Tavasolia, Keyvan e Jabbarzadeh, Armin.** 2019, International Journal of Data and Network Science, p. 20.

13. *A bibliometric approach to tracking big data research trends.* **Kalantari, Ali, et al.** 2017, Journal of Big Data, p. 18.

14. *The rise of 5G technologies and systems: A quantitive analysis of knowledge production.* **Mendonça, Sandro, et al.** 2022, Telecommunications Policy, p. 39.

15. *Finance for the Environment: A Scientometrics Analysis of Green Finance.* **Cai, Rui e Guo, Jianluan.** 1 de July de 2021, Mathematics, p. 15.

16. *Eco-innovation in the transition to a circular economy: An analytical literature review.* **Jesus, Ana de, et al.** 2018, Journal of Cleaner Production.

17. *Eco-Innovation Diversity in a Circular Economy: Towards Circular Innovation Studies.* **de Jesus, A., et al.** 2021, Sustainability.

18. *Nuclear Weapons.* **Roser, Max, Herre, Bastian e Hasell, Joe.** 2013, Our World in Data.

19. **Levy, B.S. e Sidel, V.W.** Terrorism. *International Encyclopedia of Public Health.* Oxford : Academic Press, 2008, p. 6.

20. *Nuclear Weapons Tests and Environmental Consequences: A Global Perspective.* **Prăvălie, Remus.** 2014, AMBIO, p. 15.

21. *315 nuclear bombs and ongoing suffering: the shameful history of nuclear testing in Australia and the Pacific.* **Ruff, Tilman e Hawkins, Dimity.** 2020, The Conversation.

22. **Our World in Data.** Nuclear Weapons. [Online] 2022. https://ourworldindata.org/nuclear-weapons.

23. **Cirincione, Joseph.** *Bomb scare: the history and future of nuclear weapons.* New York : Columbia University Press, 2007.

24. **Emirates Nuclear Energy Corporation.** Other Uses of Nuclear Technology. *enec.* [Online] 2022. [Citação: 8 de October de 2022.] https://www.enec.gov.ae/discover/fueling-the-barakah-plant/other-uses-of-nuclear-technology/.

25. *Nuclear Weapons.* **Roser, Max, Herre, Bastian e Hasell, Joe.** 2013, Our World in Data.

26. **World101 Global Era Issues.** The Nuclear World. *world101.* [Online] 8 de October de 2017-2022. https://world101.cfr.org/global-era-issues/nuclear-proliferation/nuclear-world.

27. *With Putin's nuclear threat, what are Russia's 'strategic' and 'tactical' nuclear weapons?* **Singh, Rishika.** 2022, The Indian Express.

28. **United Nations Office for Disarmament Affairs.** Treaty on the Non-Proliferation of Nuclear Weapons (NPT). *un.org.* [Online] 2022. https://www.un.org/disarmament/wmd/nuclear/npt/.

29. *U. S. withdraws from ABM Treaty; global response muted.* **Boese, Wade.** 2002, Arms Control Today.

30. **Arms Control Association.** The Status of the Comprehensive Test Ban Treaty: Signatories and Ratifiers. *armscontrol.org.* [Online] August de 2022. https://www.armscontrol.org/factsheets/ctbtsig.

31. *International Nuclear Relations After the Indian and Pakistani Test Explosions.* **Walker, William.** 1998, International Affairs .

32. **Clerk, Aadarsh.** India-Pakistan Relations – Terrorism, Kashmir, and Recent Issues. 8 de March de 2019.

33. **Kimball, Daryl G.** The Nuclear Risk Dimension of the War On Ukraine. *armscontrol.* [Online] 2022. [Citação: 8 de October de 2022.] https://www.armscontrol.org/events/2022-06/nuclear-risk-dimension-war-ukraine.

34. **World101 Global Era Issues.** The History of Nuclear Proliferation. *World101.* [Online] 2017-2022. [Citação: 17 de September de 2022.] https://world101.cfr.org/global-era-issues/nuclear-proliferation/history-nuclear-proliferation.

35. **Asia Society.** *A Brief History of Nuclear Weapons States.* 2022.

36. **ICAN.** The road to a world free of nuclear weapons. *ican.* [Online] 2017. [Citação: 9 de October de 2022.] https://www.icanw.org/nuclear_weapons_history.

37. *"Crucified on a Cross of Atoms": Scientists, Politics, and the Test Ban Treaty.* **Rubinson, Paul.** 2011, Diplomatic History, p. 36.

38. *Observations and modelling of the global distribution and long-term trend of atmospheric 14CO2.* **Levin, Ingeborg, et al.** 2010, Tellus B: Chemical and Physical Meteorology, p. 20.

39. *Nuclear safety, nuclear security: Whither the IAEA?* **Weitz, R.** 2011, World Affs., p. 10.

40. *Nuclear pursuits: Non-P-5 nuclear-armed states.* **McDonnell, T.** 2013, Bulletin of the Atomic Scientists, p. 8.

41. **United Nations Office for Disarmament Affairs.** Comprehensive Nuclear-Test-Ban Treaty (CTBT). [Online] 2021. https://www.un.org/disarmament/wmd/nuclear/ctbt/.

42. **Atomic Bomb History.** History. *history.* [Online] 2020. [Citação: 8 de October de 2022.] https://www.history.com/topics/world-war-ii/atomic-bomb-history.

43. *The Reasons Behind U.S. Invasion of Iraq.* **Danju, Ipek, Maasoglu, Yasar e Maasoglu, Nahide.** 2013, Procedia - Social and Behavioral Sciences.

44. **Kaur, Harmeet, Kim, Allen e Sherman, Ivory.** The US-Iran conflict: A timeline of how we got here. 11 de January de 2020.

45. **European Union External Action.** Joint Comprehensive Plan of Action. *eeas.europa.eu.* [Online] 14 de July de 2015. https://www.eeas.europa.eu/eeas/joint-comprehensive-plan-action_en.

46. **United Nations Office for Disarmament Affairs.** Treaty on the prohibition of nuclear weapons. [Online] 2021. https://www.un.org/disarmament/wmd/nuclear/tpnw/.

47. **Knox, Jennifer.** A Historic Nuclear Ban Treaty Is Set to Become International Law: Here's What That Means. 2020.

48. *The Literature of Bibliometrics, Scientometrics, and Informetrics.* **Hood, William W. e Wilson, Concepción S.** 2001, Scientometrics, p. 23.

49. **Nalimov, V. V. e Mulchenko, Z. M.** *Measurement of Science. Study of the Development of Science as an Information Process.* United States : U.S. Air Force Systems Command, Foreign Technology Division, 1971.

50. **Vitanov, Nikolay K.** *Science Dynamics and Research Production.* Switzerland : Springer Cham, 2016.

51. *Modeling Scientometric Indicators Using A SDMX Ontology.* **Lopez-Rodriguez, Victor e Ceballos, Hector G.** 2022, Journal of Big Data.

52. *Do papers (really) match journals' "aims and scope"? A computational assessment of innovation studies.* **Santos, A.T., Mendonça, S.** 2022, Scientometrics.

53. *The small world of innovation studies: an "editormetrics" perspective.* **Santos, A.T., Mendonça, S.** 2022, Scientometrics.

54. **Qiu, J., et al.** *Informetrics: Theory, methods and applications.* s.l. : Springer, 2017.

55. **pybliometrics.** Navigation. *pybliometrics.* [Online] 2017-2021. [Citação: 15 de September de 2022.] https://pybliometrics.readthedocs.io.

56. **Goutsmedt, Aurélien.** Extracting and Cleaning Bibliometric Data with R (1). *Aurélien Goutsmedt.* [Online] 2022. [Citação: 15 de October de 2022.] https://aurelien-goutsmedt.com/post/extracting-biblio-data-1/.

57. **pybliometrics.** pybliometrics.scopus.AbstractRetrieval. *pybliometrics.* [Online] 2017-2021. [Citação: 24 de September de 2022.] https://pybliometrics.readthedocs.io/en/stable/classes/AbstractRetrieval.html.

58. —. Access. *pybliometrics.* [Online] 2017-2021. [Citação: 24 de September de 2022.] https://pybliometrics.readthedocs.io/en/stable/access.html.

59. **Academia.** Stack Exchange. *Academia.* [Online] 12 de July de 2015. [Citação: 14 de September de 2022.] https://academia.stackexchange.com/.

60. **International Science Editing.** Which author affiliation do I use? *internationalscienceediting.* [Online] 2022. [Citação: 24 de September de 2022.] https://www.internationalscienceediting.com/which-author-affiliation-do-i-use/.

61. **pybliometrics.** pybliometrics.scopus.AuthorRetrieval. *pybliometrics.* [Online] 2017-2021. https://pybliometrics.readthedocs.io/en/stable/classes/AuthorRetrieval.html#.

62. —. pybliometrics.scopus.AffiliationRetrieval. *pybliometrics.* [Online] 2017-2021. https://pybliometrics.readthedocs.io/en/stable/classes/AffiliationRetrieval.html.

63. **Elsevier.** Coverage you can count on. *elsevier.* [Online] 2022. [Citação: 8 de November de 2022.] https://www.elsevier.com/solutions/scopus/how-scopus-works/content.

64. *Ranking of the subject areas of Scopus.* **Garcia, Jose A., Fdez-Valdivia, J. e Rodriguez-Sánchez, Rosa.** 2011, Journal of the American Scociety for Information Science and Technology, p. 10.

65. **Clarivate TM.** KeyWords Plus generation, creation, and changes. *support.clarivate.com.* [Online] July de 2022. https://support.clarivate.com/ScientificandAcademicResearch/s/article/KeyWords-Plus-generation-creation-and-changes?language=en_US.

66. **geonamescache 1.5.0.** Python Software Foundation. *geonamescache.* [Online] 2022. [Citação: 14 de September de 2022.]

67. *Intra-Industry Effects of the Accident at Three Mile Island.* **Bowen, Robert M. and Castanias, Richard P. and Daley, Lane A.** 1983, Journal of Financial and Quantitative Analysis, p. 24.

68. **Buszynski, L.** *Negotiating with North Korea: The Six Party Talks and the Nuclear Issue.* London : s.n., 2013.

69. *2011 Fukushima Dai-ichi nuclear power plant accident: summary of regional radioactive deposition monitoring results.* **Hirose, Katsumi.** 2012, Journal of Environmental Radioactivity, p. 5.

70. **Britannica.** Fukushima accident. *britannica.* [Online] 2022. [Citação: 18 de September de 2022.] https://www.britannica.com/topic/accident.

71. **Towards Data Science.** Fastest way to calculate h-index of publications. *towardsdatascience.* [Online] 2021. [Citação: 28 de September de 2022.] https://towardsdatascience.com/fastest-way-to-calculate-h-index-of-publications.

72. **Bernard Becker Medical Library.** Tools for Authors: What is the h index? *beckerguides.* [Online] 2022. [Citação: 22 de October de 2022.] https://beckerguides.wustl.edu/authors/hindex#_ednref2.

73. **Dimensions API Lab.** Building a concepts co-occurence network. [Online] 2020. https://api-lab.dimensions.ai/cookbooks/2-publications/Concepts-network-graph.html.

74. *The 1963 Nuclear Test Ban Treaty: A Missed Opportunity for Détente?* **Mastny, Vojtech.** 2008, Journal of Cold War Studies, p. 22.

75. **Predatory Publishing.** What is an ISSN (International Standard Serial Number)? *predatory-publishing.* [Online] 2022. [Citação: 1 de October de 2022.] https://predatory-publishing.com/what-is-an-issn-international-standard-serial-number/.

76. **APA Style.** What is a digital object identifier, or DOI? *apastyle.apa.org.* [Online] 2020. https://apastyle.apa.org/learn/faqs/what-is-doi.

8. APPENDIX

Table 12. Web of science core collection metadata. Source: Web of science core collection help.

Variable	Description	Variable	Description
FN	File name	OI	ORCID identifier (Open Researcher and Contributor ID)
VR	Version number	FU	Funding agency and grant number
PT	Publication type (J=Journal; B=Book; S=Series; P=Patent)	FX	Funding text
AU	Authors	CR	Cited references
AF	Author full name	NR	Cited reference count
BA	Book authors	TC	Web of science core collection times cited count
BF	Book authors full name	Z9	Total times cited count (web of science core collection, BIOSIS citation index, chinese science citation database, data citation index, russian science citation index, SciELO citation index)
CA	Group authors	U1	Usage count (last 180 days)
GP	Book group authors	U2	Usage count (since 2013)
BE	Editors	PU	Publisher
TI	Document title	PI	Publisher city
SO	Publication name	PA	Publisher address
SE	Book series title	SN	International standard serial number (ISSN)
BS	Book series subtitle	EI	Electronic international standard serial number (eISSN)
LA	Language	BN	International standard book number (ISBN)
DT	Document type	J9	29-character source abbreviation
CT	Conference title	JI	ISO source abbreviation
CY	Conference date	PD	Publication date
CL	Conference location	PY	Year published
SP	Conference sponsors	VL	Volume
HO	Conference host	IS	Issue
DE	Author keywords	SI	Special issue
ID	Keywords plus	PN	Part number
AB	Abstract	SU	Supplement
C1	Author address	MA	Meeting abstract
RP	Reprint address	BP	Beginning page
EM	E-mail address	EP	Ending page
RI	Researcher ID number	AR	Article number
DI	Digital object identifier (DOI)	GA	Document delivery number
D2	Book digital object identifier (DOI)	PM	PubMed ID
EA	Early access date	UT	Accession number
EY	Early access year	OA	Open access indicator
PG	Page count	HP	ESI hot paper. Note that this field is valued only for ESI subscribers.

P2	Chapter count (book citation index)	DA	Date this report was generated
WC	Web of science categories	ER	End of record
SC	Research areas	EF	End of file

Table 13. Subject areas covered by Scopus database.

Scopus abbreviation	Subject area	Broad subject area
MULT	Multidisciplinary	
AGRI	Agricultural and Biological Sciences	Life sciences
ARTS	Arts and Humanities	Social sciences
BIOC	Biochemistry, Genetics and Molecular Biology	Life sciences
BUSI	Business, Management, and Accounting	Social sciences
CENG	Chemical Engineering	Physical sciences
CHEM	Chemistry	Physical sciences
COMP	Computer Science	Physical sciences
DECI	Decision Sciences	Social sciences
EART	Earth and Planetary Sciences	Physical sciences
ECON	Economics, Econometrics and Finance	Social sciences
ENER	Energy	Physical sciences
ENGI	Engineering	Physical sciences
ENVI	Environmental Science	Physical sciences
IMMU	Immunology and Microbiology	Life sciences
MATE	Material Science	Physical sciences
MATH	Mathematics	Physical sciences
MEDI	Medicine	Health sciences
NEUR	Neuroscience	Life sciences
NURS	Nursing	Health sciences
PHAR	Pharmacology, Toxicology, and Pharmaceutics	Life sciences
PHYS	Physics and Astronomy	Physical sciences
PSYC	Psychology	Social sciences
SOCI	Social Sciences	Social sciences
VETE	Veterinary	Health sciences
DENT	Dentistry	Health sciences
HEAL	Health Professions	Health sciences

Table 14. New variables created using Scopus APIs, respective description, and creation process. Source: own compilation.

Variables	Description	Creation procedure	Query used
article_info	Interaction with the Abstract Retrieval API.	Calling the Abstract Retrieval API by using the variable DI as its input.	AbstractRetrieval(DI)
authors_info	Information on all the authors of an article (author id, name, and affiliation id).	Applying the property authors of the Abstract Retrieval API.	article_info.authors
authors_ids	Author ids of all the authors of an article.	Getting the auid element of the list of namedtuples retrieved by the property authors.	authors_info.auid
authors_aff_ids	Affiliation ids of all the authors' affiliations of an article.	Getting the affiliation element of the list of namedtuples retrieved by the property authors.	authors_info.affiliation
first_author_id	Main author id of an article.	Selecting the first element of the list of all the authors' ids of an article.	authors_ids[0]
first_author_affs_ids	Main author affiliations' ids of an article.	Selecting the first element of the list of all the authors affiliations' ids of an article.	authors_aff_ids[0]
first_author_aff_id	Main author affiliation id of an article.	Selecting the first element of the list of all the first author's affiliations ids of an article.	first_author_affs_ids[0]
subject_areas_info	Information on the articles' categories (area, abbreviation, and code).	Applying the property subject_areas of the Abstract Retrieval API.	article_info.subject_areas
cat_abbreviations	Categories' abbreviations of an article.	Getting the abbreviation element of the list of namedtuples retrieved by the property subject_areas.	subject_areas_info.abbreviati on
akeywords	Author keywords of an article.	Applying the property authkeywords of the Abstract Retrieval API.	article_info.authkeywords
first_author_info	Interaction with the Author Retrieval API.	Calling the Author Retrieval API by using the variable first_author_id as its input.	AuthorRetrieval(first_author_ id)
main_author_hindex	H-Index of the main author of an article.	Applying the property h_index of the Author Retrieval API.	first_author_info.h_index
given_name	First name of the main author of an article.	Applying the property given_name of the Author Retrieval API.	first_author_info.given_name
surname	Last name of the main author of an article.	Applying the property surname of the Author Retrieval API.	first_author_info.surname
main_author_name	First and last name of the main author of an article.	Joining the variables given_name and surname.	given_name + surname
info_first_aff_of_first_auth or	Interaction with the Affiliation Retrieval API.	Calling the Affiliation Retrieval API by using the variable first_author_aff_id as its input.	AffiliationRetrieval(first_auth or_aff_id)
country_first_aff_of_first_a uthor	Country of the affiliation of the main author of an article.	Applying the property country of the Affiliation Retrieval API.	info_first_aff_of_first_author .country
city_first_aff_of_first_auth or	City of the first affiliation of the main author of an article.	Applying the property city of the Affiliation Retrieval API.	info_first_aff_of_first_author .city

name_first_aff_of_first_author	Name of the first affiliation of the main author of an article.	Applying the property affiliation_name of the Affiliation Retrieval API.	info_first_aff_of_first_author.affiliation_name
aff_of_main_author_at_the_time_of_publication	Name, city, and country of the author affiliation at the time the article was published.	Joining the variables name_first_aff_of_first_author, city_first_aff_of_first_author, and country_first_aff_of_first_author.	name_first_aff_of_first_author + city_first_aff_of_first_author + country_first_aff_of_first_author

Note: The pink rows refer to the variables where the calls of the three APIs were saved.

NOVA
IMS
Information
Management
School

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa