The Lab's Quarterly
2025/ a. XXVII / n. 0 – ISSN 2035-5548
Online First: 10 June 2025

ALGORITHMIC FEEDBACK LOOPS IN SOFT SCIENCE DISCIPLINES.

An application of the systematic literature review on the evolution of definitions from 2000 to 2023

di Gabriella Punziano, Giuseppe Michele Padricelli†, Antonio Vettori‡*

Abstract

In today's digital society, consumer cultures and practices have been reshaped by digital platforms. Cultural entertainment consumption, such as movies and music, is now largely mediated by platforms like Netflix, Spotify, and YouTube, which use AI-driven algorithms to recommend, filter, and rank content dynamically. This article presents a longitudinal study of scientific literature to examine how the concept of feedback loops has been addressed. It explores how this recursive process—where outputs influence new inputs—has evolved and been interpreted differently across soft and hard sciences over time.

Keywords

Feedback loop, Prisma SLR, algorithm role

* GABRIELLA PUNZIANO, Associate Professor of Sociology and Methodology of social research at the University of Naples Federico II, Department of Social Sciences. Among her research interests: the new analytical frontiers and the challenges introduced by new data, mixed and integrated and digital perspectives; social policies and welfare regimes; the analysis of public, institutional and political communication phenomena through innovative content analysis techniques; risk communication analysis on social and digital platforms.

Email: gabriella.punziano@unina.it

† GIUSEPPE MICHELE PADRICELLI, Post-Doc Fellow at the University of Naples Federico II and Adjunct professor in Methodology of social research at University of Campania L. Vanvitelli. His main research interests concern the methodological challenges in digital realm addressed by the study of collective action and political communication.

Email: giuseppemichele.padricelli@unima.it

‡ ANTONIO VETTORI, MD in 'Public, Social and Political Communication' at the University of Naples Federico II. His thesis work aims to understand user-algorithm interactions by focusing on the methodological aspects of the study.

Email: a.vettori@studenti.unina.it

DOI: 10.13131/unipi/zzr3-mm11

1. INTRODUCTION: MULTIPLE WAYS TO CONCERN SCIENCE

Scientific inquiry has long been centered on fundamental questions about what science can achieve and how its mechanisms operate. Scholars and intellectuals have historically debated these issues, leading to the emergence of multiple paradigmatic perspectives that have shaped the evolution of scientific thought. Despite their differences, these perspectives converge on the idea that science is characterized by a structured system of codes, techniques, and methods—essential tools that bridge theoretical frameworks with empirical observation (Kuhn, 1962; Popper, 1959). This methodological foundation serves as a key criterion for distinguishing scientific knowledge from other forms of understanding.

Over time, different conceptualizations of science have emerged, reflecting diverse epistemological orientations and disciplinary applications. The classification of scientific approaches has evolved in response to shifts in reflective and experimental fields, accommodating new ways of interpreting and engaging with reality (Latour, 1987; Feyerabend, 1975). These evolving frameworks highlight the dynamic nature of scientific knowledge, illustrating how methodologies and paradigms adapt to new discoveries and societal needs.

In this regard, the epistemological perspective, through Merton's (1973) contribution to the sociology of science, outlines a fundamental path to follow in systematizing the distinction between "sciences". The ability to measure objects, referents, and units of analysis using standardized measurement units has always been the only discriminating element useful in distinguishing, or rather dividing, sciences into few broad families. However, over the years, new systematizations and increasingly sophisticated classifications have emerged, directly linked to the role of the researcher and the mechanisms they can follow based on the objects they observe, deconstruct, and analyze. Following Wu et al. (2022), we now distinguish, for example, between *hot and cold sciences*, as well as *hard and soft* sciences.

The first distinction, based on the temperature metaphor, categorizes sciences according to their state of discovery advancement and the empirical evidence reported in studies. The advancement of evidence is directly proportional to temperature, which serves as an indicator of theoretical solidity. Within this distinction, Latour (1987) systematizes the concept of the "*black box*", that is, a portion of cold, closed knowledge that the researcher is responsible for opening and discovering to *heat it up*. The opening of black boxes and the change in temperature of theoretical frameworks within them contribute to the accumulation process that Kuhn (1962) already identified as the driving force behind scientific revolutions. In these terms, the process of destruction does not necessarily imply a paradigm shift but rather intervenes in clarifying the ambiguity or making explicit a theory (or parts of the concepts underlying it) through consensus-building and debate within the scientific community.

On the other hand, the distinction between hard and soft sciences, according to Nelson (2002), stems from the unequal evolution of scientific knowledge. The distinction between hard and soft sciences aligns with the longstanding debate between natural and social sciences. According to Storer's work *The Hard Sciences and the Soft* and further discussed by Shapin (2022), soft sciences are often seen as adhering to non-scientific criteria. Dang (2018) notes that hard sciences tend to follow a single paradigm that scholars adhere to. In contrast, soft sciences do not always share a single paradigm, leading to multiple, sometimes conflicting, approaches. However, this division does not fully capture the complexity and knowledge structures within various disciplines, though it serves as a helpful starting point for understanding the diversity of academic fields.

A key aspect of the common discourse surrounding this distinction is a qualitative evaluation of the sciences. Shapin (2022) points out that hard sciences have traditionally been regarded more positively for their methodological rigor, while soft sciences have often been seen as quasi-scientific. This perception often relates to individual views on what constitutes science. Despite the persistence of this view, it is gradually fading, especially as the future of the sciences becomes increasingly interdisciplinary.

The so-called digital turn and the pervasiveness of objects, devices, codes, and languages in everyday social life can, to some extent, be considered one of the most recent scientific revolutions (Knell, 2021), capable of reshaping theoretical frameworks in communication, economics, and society as a whole.

In this sense, interdisciplinarity emerges as an essential ingredient in approaching the theme of digitalization *tout court*, both from a technical and a social perspective. In this way, research heritages and connections between hard and soft sciences themselves contribute to narrowing the gap between them, especially in addressing processes and objects currently confined within so-called black boxes, such as feedback loops and, by extension, algorithms and recommendation systems.

Thoroughly studying these processes, with particular attention to the coexistence of these two perspectives, is crucial for understanding the mechanisms that connect hard and soft sciences. Through a longitudinal analysis based on a Systematic Literature Review this study covering scientific productions on the topic of feedback loops from 2000 to 2023, aims to shed light on the transfer of concepts, definitions, models, and interpretative capacities to soft sciences.

2. ADDRESS FEEDBACK LOOPS: THE INNOVATIVE PRACTICES OF CULTURAL CONSUMPTION

Since the so-called "digital turn", cultures and consumption practices have been in the middle of a significant remediation process for several years. The increasingly pervasive presence of platforms dedicated to the consumption of audiovisual cultural products, such as Netflix, Spotify, YouTube, and others, has undoubtedly, according to Poell et al. (2021), challenged the habits of cultural consumption in cinema, music listening, and TV series viewing. In this sense, social research has kept up, focusing on the opportunities for access to and consumption of audiovisual products by questioning the processes of searching, enjoying, and recommending entertainment products.

One of the most interesting aspects of the reflection on this topic can be summarized in the definition of the feedback loop process. According to Airoldi, Beraldo and Gandini (2016), these processes are characterized by an interaction between users and the recommendation systems of these platforms, generally never made public. Operating in the context of what Beer (2022) defines as a "recursive society", it is necessary to aim at understanding the possibilities of permeating reality through algorithmic logic. This is done while considering that data and outputs are not only the results of a direct action produced by the user (ibid) but also a product shaped by platform logic and thus by the digital traces left online by users that is, by previous feedback (intentional or not) that has shaped previous consumption actions in a potentially infinite loop.

Feedback loops are evident in these contexts. For example, on TikTok, videos that receive a high number of views, likes, and shares tend to be further promoted by the platform's algorithm, thus reaching a wider audience. This feedback encourages users to create similar content to gain visibility and engagement, fuelling a cycle in which the most popular content becomes increasingly visible and dominant.

Feedback loops on platforms like YouTube and Netflix are often tied to personalized recommendations. The algorithms analyse users' viewing behaviour to suggest new content. However, there are also more critical aspects of feedback loop application. For instance, if a user has specific and limited preferences, the algorithm might further narrow the variety of recommended content, creating a filter that limits the diversity of cultural experiences.

In summary, feedback loops play a key role in the consumption of cultural products on social and streaming platforms. They shape the content presented to users and influence their consumption choices. These cycles can significantly affect cultural diversity, user opinions, and social dynamics online.

The fluidity of cultural consumption practices today is associated with the increasingly pervasive presence of users on streaming platforms and the use of search and consumption tools facilitated by digital mediatization (Airoldi and Rokka, 2022). The materiality and ownership of cultural products in physical formats (DVDs, Blu-rays, CD-ROMs, etc.) are being replaced by services tailored to users' preferences: an adaptation of supply to demand, facilitated by the role of the algorithm.

According to Airoldi and Rokka (2022), every cultural product involves what is known as articulation, a moment when the perspectives of those who produce a product and those who consume it must meet. Today, this articulation also involves the so-called algorithmic training: a non-human intermediary that operates between the production and consumption process in a non-neutral, authoritative, and recursive way. However, this should be considered a dialectical process because, on the one hand, the algorithm controls and constrains the user based on its outputs, while on the other, the algorithm itself must adapt based on the user's behaviour something that can be understood through the concept of the feedback loop.

Nowadays, TikTok and YouTube are two significant platforms for this new way of consuming cultural products (Denicolai and Farinacci, 2020; Rico Kuntag and Sijabat, 2023).

The first is a Chinese social network, known as Douyin in the Asian country, which originally launched as Musical.ly in 2016 and was later rebranded as TikTok in 2018 by ByteDance. This company acquired it from founders Aleix Zhu and Luyu Yang in 2017 for approximately 750 million euros. It has become one of the most used social platforms, especially among teenagers, with 1.2 billion monthly active users worldwide. Its primary purpose is the sharing and viewing of short videos, up to a maximum of ten minutes. The innovative feature introduced by TikTok, which was quickly adopted by other social networks, is the “For You” page. This is the main screen displayed when the app is opened, showcasing multimedia content suggested by the algorithm based on the user's preferences through an analysis of viewing times, comments, shares, and likes. In this sense, almost all of the user's experience and video consumption is based on what the algorithm deems close to the consumer's preferences. Other sections, such as the profiles followed, are comparatively less influential. Nevertheless, it is important to highlight that despite the algorithm's pervasive role in shaping the user experience, the diversity of the "For You" page between different users, even those in close proximity, demonstrates that it is primarily the algorithm that adapts to the user, rather than the other way around.

YouTube was founded in 2005. Today, according to data published by Google and commented on by Forbes, it is one of the crown jewels of the tech giant and is managed by Alphabet. Nearly 10% of its revenue comes directly from this platform. Although it is now an integral part of the Google ecosystem, YouTube was initially created by three young men, Chad Hurley, Steve Chen, and Jawed Karim, who were working for PayPal at the time. The primary function of YouTube is to share and view multimedia content uploaded by other users. These include video blogs, movie trailers, reviews, gaming, and music. Indeed, several studies show that the sharing of music videos is predominant on the platform (Airoldi et al., 2016). The interface is not solely based on content users search for or channels they subscribe to; it also suggests recommended content on the homepage and the video being played. These recommended videos are generated by an algorithm that considers numerous factors but mainly creates a pattern of shared viewing (Airoldi et al., 2016). In other words, when examining a specific video that the user is watching, the platform suggests other videos that other users subsequently watched with a similar viewing pattern.

An illustrative study was conducted by Airoldi, Beraldo and Gandini in 2016, titled *Follow the Algorithm*, which investigates the dynamics behind related and recommended videos. These are determined not only through a computational process but also by the collective behaviour of users who contribute to categorizing audiovisual products. Although the algorithm of the Google-owned platform is not public, this categorization is based on the frequency of co-viewing between the currently playing video and those recommended in the sidebar. The study aims, through a mixed methods approach and a combination of network analysis with content analysis, to form clusters of videos, demonstrating that on the one hand, these clusters follow genre-based logic—songs of the same genre are grouped together while other clusters are defined by contexts or listening moments, such as music for meditation, relaxation, or children's music.

As previously mentioned, what emerges is an environment where cultural consumption changes not only in its form of fruition but also in how it is re-presented, while still adhering to historical and cultural categorizations like musical genres. In this sense, feedback loops become clearer, representing a phenomenon that connects human actors with network actants, such as algorithms. It is a balancing phenomenon that, on the one hand, mitigates the influence of the algorithm—at least considering the media narrative that emerges—but, on the other, allows users greater ease in content consumption, no longer chosen independently but based on what aligns with previously detected tastes.

What has been clarified so far briefly outlines the foundations of the feedback loop concept, centralizing the social sciences debate on the behavioural and algorithmic dynamics related to the consumption of cultural products online.

What is the starting point for a critical reflection on this topic? How has the debate evolved before reaching its current state of advancement? How have the scientific trajectories, spanning both soft and hard sciences, been influenced over the years by reflections on the user-algorithm relationship?

3. THE ALGORITHMIC FEEDBACK: A SYSTEMATIC LITERATURE REVIEW

The Systematic Literature Review (SLR) considered those studies published on the topic of algorithmic feedback over the past twenty-three years, covering the period from 2000 to 2023. The aim was to provide an overview of the evolution of publications across all fields, with a particular focus on those related to the humanities or soft sciences, which were then subjected to content analysis. Describing how this topic has been addressed in the research field over the years meets two main knowledge needs. The first is aimed at a transdisciplinary understanding of the definition and framing of the subject matter across various disciplines, from hard sciences to soft sciences; the second is to delve deeply into the field of social sciences. This approach is further supported by the fact that the concept of feedback originated in cybernetics, and its evolution as an interdisciplinary subject has faced a challenging path toward institutionalization.

The peculiarity of the SLR lies in treating the literature as a scientific process. Lamé (2019) states that by applying concepts from empirical research, the SLR becomes a more transparent and replicable process, consequently reducing the distorting effects that researchers might face.

The term "systematic" here refers to a step-by-step protocol to ensure that the objectivity of the research is protected from potential researcher biases (Sharif et al., 2019). This protocol consists of specific steps to be followed:

- definition of the question (it must be useful and significant, considering variables such as previous studies or the real need to use the systematic literature review approach);
- determination of study types (involves selecting the studies to be included in the review, which can be determined through a hierarchy of studies that follow a more structured approach or by choosing the study type most appropriate to the research question);
- literature search (can be done through databases such as Web of Science, Scopus, PubMed, etc., which often serve as the starting point for an SLR and may vary depending on the research discipline);
- results checking (presence of biases can compromise the integrity of the results, and thus these aspects need to be monitored. Various tools should be used to ensure proper control, such as an initial screening to eliminate duplicates and generally check the relevance of the selected studies);
- critical evaluation (an essential phase, without which the study would be unreliable, serving to systematize the study rather than to engage in criticism for its own sake);
- synthesis and interpretation of results (to synthesize and analyze the extracted data, this can be done using statistical methods (meta-analysis) where appropriate, or through a qualitative approach and narrative synthesis when the studies under review are more heterogeneous);
- dissemination of results (carried out through drafting a report and subsequent publication. This phase must be considered from the outset, not only at the end of data synthesis and interpretation. This step is crucial because it ensures that the research is as accessible as possible to the intended audience).

Lamé (2019) further notes that this approach also presents challenges, such as its limited adoption by researchers, as it partially hinders the ability to synthesize and accumulate results.

The SLR is a process that can consider a large amount of information and studies. One of the most significant aspects is its ability to distinguish and understand the difference between actual knowledge and what we perceive as knowledge. This discernment is made possible by analyzing and comparing studies with similar populations and research topics. This approach is more suited to answering specific questions and testing hypotheses than traditional literature reviews. In fact, the SLR is not comparable to a mere discussion of the literature but is a valid scientific tool; systematic and traditional reviews aim to satisfy different knowledge needs and should not be seen as alternatives but rather as approaches that can also complement each other. This method may also involve a statistical study approach, such as meta-analysis, which uses specific statistical techniques to synthesize the results of numerous studies into a single quantitative estimate. A systematic review should be conducted when it is necessary to provide a comprehensive overview of the evidence in a research area to facilitate future advancements on the topic; when the field of inquiry is still emerging and there are not yet enough registered and published studies; or when, as in this case, the field of study is well-established with a large number of published studies. In this study, the SLR can be used to summarize and simplify the already abundant amount of data produced (Petticrew and Roberts, 2005).

3.1 The PRISMA approach

The application of the SLR in this study followed the typical characteristics of the PRISMA model.

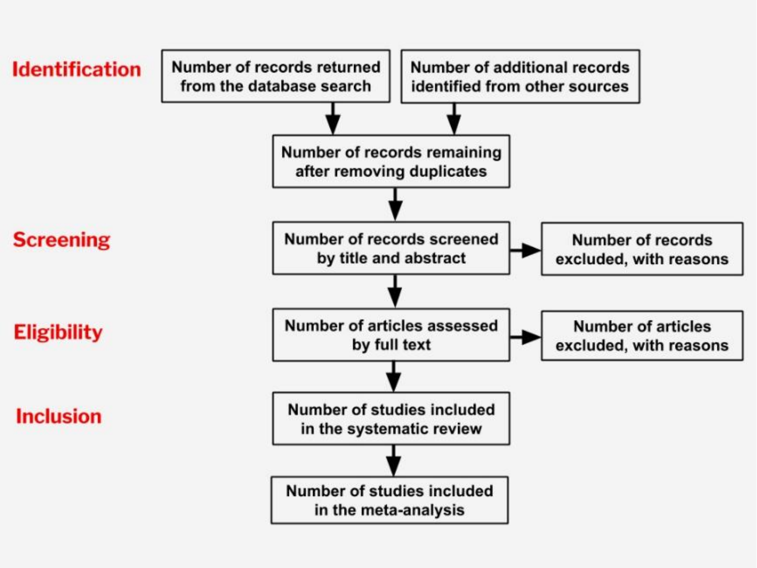
The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) is a reporting checklist designed to guide the conduct of a Systematic Literature Review through a series of predefined steps. It was introduced in 2009 and initially published by an international network of researchers primarily from medical disciplines.

Page, McKenzie et al. (2021) developed the PRISMA checklist that consists of 27 items, which are considered essential steps for successfully conducting a systematic review. These items can be divided into different sections (Table 1): *title, abstract, introduction, methods, results, discussions, and other information.*

Table 1 – SLR phases in PRISMA approach

Title
The first crucial point is to identify the report from the title with the term systematic review, so that the research can be more easily found and categorized.
Abstract
The abstract should consist of a detailed summary of what has been discussed within the review. It should specify the research objectives and the questions the review aims to answer, clearly outline the methods used, such as eligibility criteria, where the information was collected, including the number of studies in the report, any potential risks of bias, and a general synthesis and interpretation of the results.
Introduction
The two key points are the motivations and objectives. The first phase must consider the existing knowledge on the topic of investigation, while the objectives relate to the questions the review aims to answer.
Methods
In this section, the process followed for the SLR should be clearly outlined. This section includes the eligibility criteria, the sources from which the information is derived, the search strategies, the data selection process, risks, synthesis methods, and the evaluation of risks and results. One of the tools used in this step is the PRISMA flowchart (Figure 1), which provides a clearer view of the inclusion and exclusion process of articles at all its stages.
Results
The results are presented both through a written description and the inclusion of tables and graphs. The selection of studies and their characteristics, the risk of bias in the selection, the evidence emerging from the selected publications, and the syntheses are clearly outlined.
Discussion
The discussion provides a general interpretation of the results, considering any limitations and the implications of the findings for practice and future studies.

Figure 1. Prisma Flow Diagram (from www.edanz.com/blog/prisma-flow-diagram)



The data collection was conducted through the selection of academic works on the topics of feedback loops, algorithms, and artificial intelligence. More specific terms that are still related to the central theme, such as cybernetics, recommendation systems, and machine learning, were also included.

The consideration of these search terms is due to the fact that, as mentioned, the origin of the concept of feedback can be traced back to cybernetics. As an interdisciplinary concept, cybernetics became institutionalized following extensive debates between approaches aligned with the social sciences and those rooted in engineering (Carradore, 2013).

The data collection procedure was carried out using the Web of Science (WOS) portal, a citation database that allows researchers to gather indexed and categorized scientific articles based on a search query. As previously mentioned, the time span ranges from 2000 to 2023. This time frame was chosen to ensure comprehensive coverage of the evolutionary process of the web, from the emergence of what is known as Web 2.0 up to 2023, the last year for which the source provides complete coverage of published works.

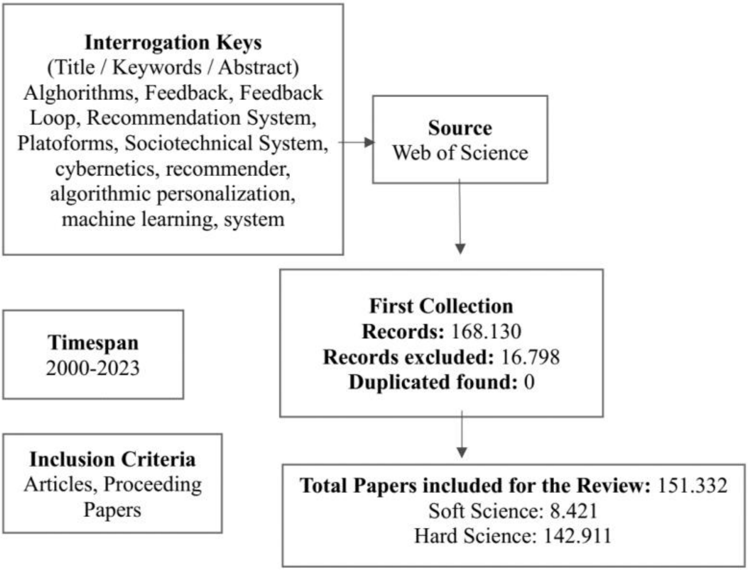
After running the query, 168,130 records were identified. This dataset was then refined by excluding publications outside the 2000-2023 time frame, reducing the dataset to 156,556 records. An additional inclusion criterion was applied based on the publication type, further narrowing the dataset to 151,332 records, focusing solely on peer reviewed articles and proceeding papers from any discipline. The choice to select these two types of publications was driven by their higher number, allowing for a more substantial sample, as well as the fact that both types typically undergo a certain degree of peer review.

A particular type of scientific article and proceeding paper holds significant importance in the context of academic conferences. These publications serve as a key means of disseminating research findings and new ideas within the scientific community, offering researchers in the social sciences and other fields an opportunity to share their results.

The final dataset (the processing of which is outlined in Figure 2) includes all articles that contained one or more terms—either isolated or combined—specified in the search query in their title, abstract, or keywords.

Each article collected in the final dataset is not exclusively attributable to a single keyword from the query but rather to a combination of these words. The selection of each individual keyword in the query is based on a direct connection to the central theme. The process began with a review of the relevant literature to identify useful terms, followed by test queries on the Web of Science database. These tests involved multiple attempts to combine each selected word with the term "feedback loop". The final query resulted from the set of combinations that yielded a substantial number of peer-reviewed articles, which were subsequently included in the dataset. Some keywords initially identified through the literature review, such as "Artificial Intelligence", did not return peer-reviewed scientific outputs that fell within the time frame defined for this study. In fact, the connection between "feedback loop" and "artificial intelligence" primarily led to more recent publications that exceeded the established temporal limit. The closure of this time frame in 2023 corresponds to the threshold within which Web of Science database ensures full coverage and record updates.

Figure 2. Data collection process (our elaboration)

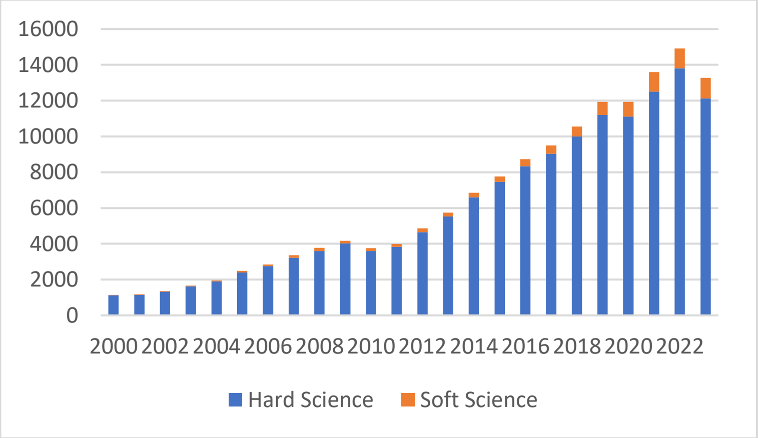


4. THE TEMPORAL TREND BETWEEN HARD AND SOFT SCIENCE

The empirical part of this study aims to understand how social sciences have addressed this topic and how their treatment has evolved. To achieve this, it is useful to apply typical sociological research methods, which often include diachronic analytical approaches to better understand changes in a phenomenon over time. This approach aligns with longitudinal studies, which are "based on the classic strategy of identifying specific units that are observed, surveyed, or exposed to the same stimuli repeatedly over time" (Caputo, Felaco, Punziano, 2017: 25). For this purpose, the dataset was segmented according to different disciplines using the categorization provided by the source, known as "WOS Categories". The disciplines were then grouped into "hard sciences" and "soft sciences". The soft sciences subset includes fields such as: *Anthropology; Archaeology; Art; Behavioural Sciences; Business and Business Finance; Communication; Cultural Studies; Development Studies; Economics; Education and Educational; Research; Ethics; History and Philosophy of Science; Hospitality, Leisure, Sport, and Tourism Humanities Multidisciplinary; Information Science and Library Science; Language and Linguistics; Law; Management;*

Philosophy; Political Science Psychology (Applied, Clinical, Experimental, Mathematical, Multidisciplinary, Social); Public Administration; Public, Environmental, and Occupational Health; Social Issues; Social Sciences (Biomedical, Interdisciplinary, Mathematical Methods); Sociology. This categorization resulted in two subsets: one containing publications from disciplines classified as hard sciences (142,911 records) and another from disciplines classified as soft sciences (8,421 records). A temporal analysis of publication trends was then conducted, with an initial summary in Figure 3. This figure displays a stacked column chart showing how the total number of publications has generally increased each year since 2000, except for the periods between 2009-2010 and 2022-2023. The division between publications in hard sciences and soft sciences is highlighted by different colors in the graph, revealing a substantial predominance of publications from hard sciences. The goal is to determine whether, in recent years, soft sciences have overtaken hard sciences in their focus on this topic, and to see if similar or different trends emerge between the two approaches. The initial hypothesis suggests that the topics of feedback loops and algorithms have gained more prominence in soft sciences in recent years. Figure 3 illustrates how, for both approaches, the number of publications generally surpassed the previous year's output almost every year. However, this visualization does not account for the initial disparity in attention given to each discipline.

Figure 3. Annual publication in hard and soft science (made by Microsoft excel)



To explore this last point in greater depth, the same descriptive inspection is replicated, focusing exclusively on the scientific production related to soft sciences.

Given the high sensitivity of the research area variable and the large number of categories in which it is operationalized, a reclassification of categories was carried out to avoid an overly fragmented analytical synthesis. In this process, the disciplines were grouped into five new categories:

- Human and Social Studies, which includes publications in the fields of anthropology, sociology, education, history, and linguistics.
- Psychology
- Communication and Cultural Studies
- Business and Economics, which includes publications in the fields of economics, finance, and business organization.
- Government and Law, which includes publications in the fields of law and political science.

Looking at the figure 4, in addition to observing a continuous upward trend in the number of publications per year, it is possible to discern the consistent proportionality in production across the different disciplines. Furthermore, another interesting aspect emerges, which becomes even more evident in Figure 5, where the count is standardized for each research area relative to the total number of works published each year. This aspect concerns the identification of the years in which each research area first appeared.

At the beginning of the timeline, total production was exclusively dominated by human and social studies, as well as business and economics. However, as the timeline progresses, psychology first and then studies on communication processes gradually claim a share of research production on the topic. It is only from 2007 onward that a decline in production within the social studies area is recorded. In the same period, there is a growing focus on communication studies, along with the emergence and consolidation (although still modest compared to other research areas) of research on the topic within the field of Government and Law.

Regarding this last aspect, it is not difficult to formulate research hypotheses linking the increase in scientific production on these topics to the rising attention to regulation and legal aspects, which have been central to much of the public and political debate. Isolating and conducting an in-depth qualitative analysis of articles in the Government and Law area published in these years presents a valuable opportunity for future research developments. This would allow for verifying the initial hypothesis and further exploring the role of algorithms and artificial intelligence in public and political decision-making processes, as well as delving into the sensitive issue of data protection and identity security in the digital space.

Figure 4. Annual publication in soft science

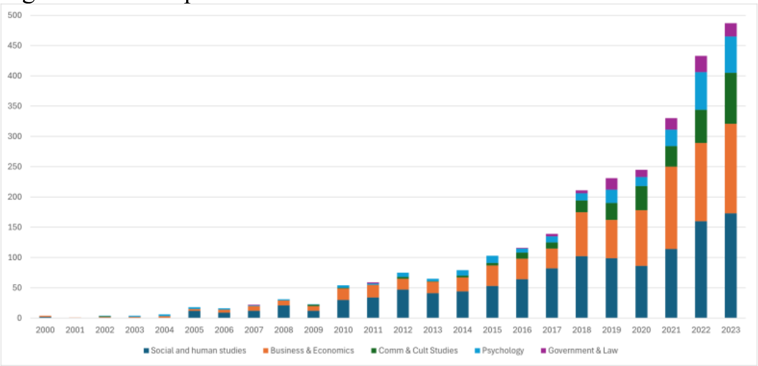
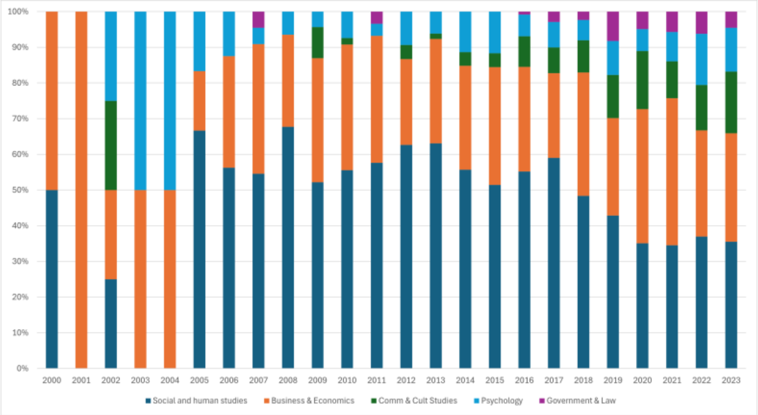


Figure 5. Annual publication per Research Area



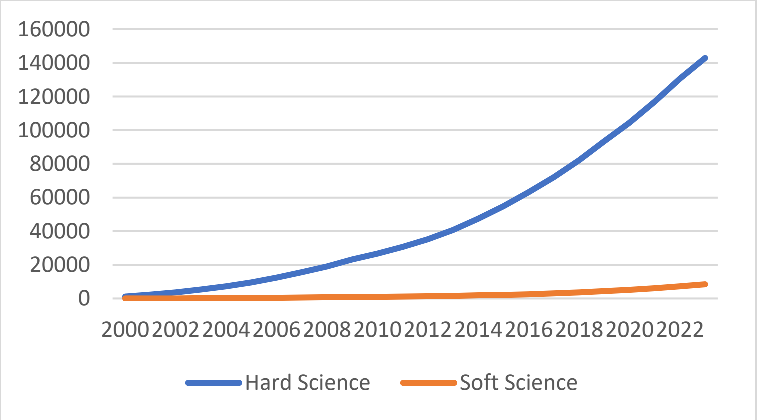
In light of this, another way to better address the limitations of absolute growth comparisons, is the employ of index numbers, a statistical measure used to express relative changes in a data series with respect to a reference or base value. The base value can be fixed—meaning changes are compared to a specific fixed time—or moving, where the comparison is made with the immediately preceding time period. Index numbers serve as a method for comparing different magnitudes over time or across distinct data sets, normalizing variations relative to a common reference point (Predetti, 2006).

To calculate an index number, the value of a variable at a given time is divided by its value at a reference time, multiplied by 100 to obtain a percentage value. The interpretation of an index number depends on its relation to 100:

- an index number greater than 100 indicates an increase relative to the base period;
- an index number less than 100 indicates a decrease relative to the base period;
- an index number equal to 100 signifies no change from the base period.

In the case of publications across hard sciences and soft sciences, the year 2000 was selected as the starting point for data collection. This year is significant for technological advancement, both from a technical perspective and a social one. Using 2000 as the reference year, Figure 6 displays the growth of publications starting from that year, illustrating the predominance of hard science disciplines in absolute numbers.

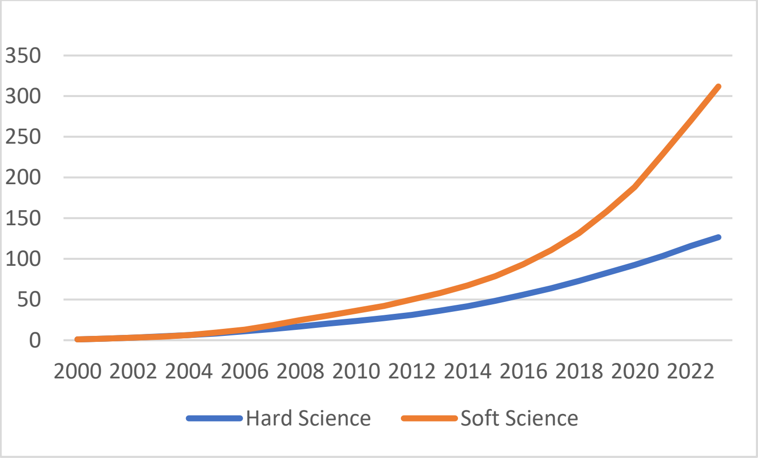
Figure 6. Absolute growth of publication from 2000 (made by Microsoft Excel)



By relating the growth to the starting point, and thus using fixed-base index numbers where the base is the number of publications in the year

2000 for both hard sciences and soft sciences, it is possible to see in Figure 7 the greater growth in publications within soft science disciplines.

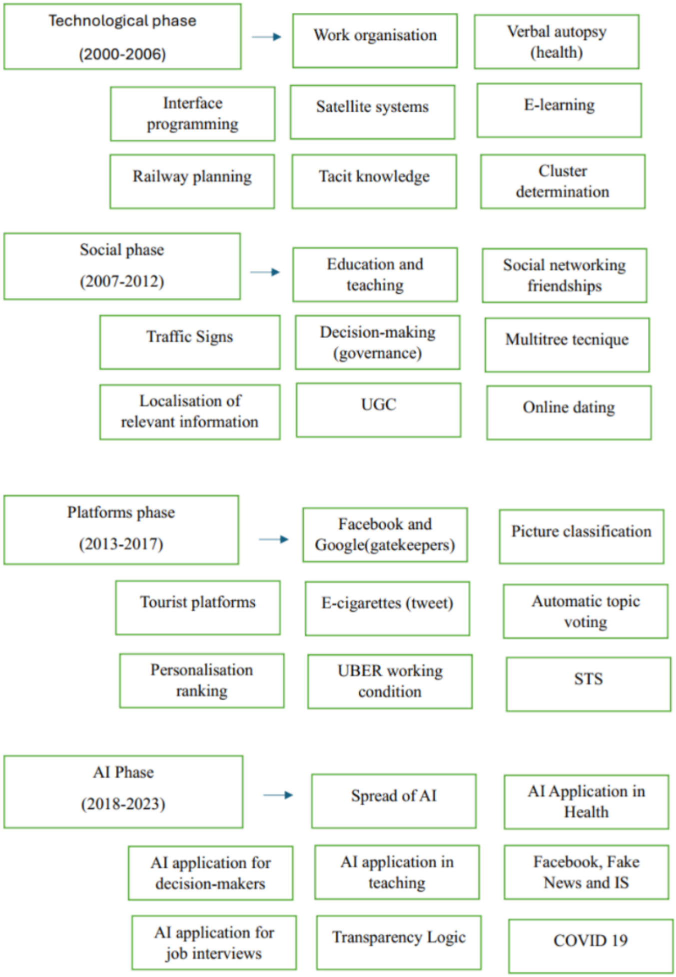
Figure 7 – Relative growth of publication from 2000 (made by Microsoft Excel)



As of 2023, this growth has been approximately 2.5 times higher for soft sciences than hard sciences. This number confirms our initial hypothesis that there has been a growing interest in the research topics within disciplines associated with the soft science approach.

This growth increased more significantly starting in 2007, with another surge in 2018. To understand the reasons for this growth, we reviewed the titles and abstracts of 120 papers—five articles categorized under soft sciences with the highest number of citations for each year (2000–2023). An interpretive analysis emerged from this, which is schematically represented (Figure 8), where temporal phases are associated with keywords from the articles published during those years.

Figure 8. Temporal phases of the most cited publications



Four chronological phases have been identified: the technological phase (2000-2006), the social phase (2007-2012), the platform phase (2013-2017), and the artificial intelligence phase (2018-2023).

- In the first phase, referred to as the technological phase, a large number of papers emerge that, in addition to addressing soft science disciplines, also touch on topics related to hard science fields such as engineering, computer science, and medicine. This phase is characterized by highly technical subjects, focusing on the design of devices capable of autonomously performing tasks. Examples include the scheduling of train timetables, the design of techniques to establish appropriate clusters in scientific studies, and a range of studies related to medical disciplines. The social-human factor is less prominent but not entirely absent in this phase, as some studies address topics such as product selection in online shopping, aided by automatic ranking systems—an anticipation of the next phase.
- In the second phase, referred to as the social phase, there is increased focus on topics more closely related to social aspects, such as education, decision-making, online product sales, and social networks. Themes such as user-generated content (UCG) emerge, which place human activity at the center of algorithm design and automated recommendation systems. There is a notable rise in studies that explore online environments more closely, particularly those of a collaborative nature, where social activity is essential to complete the network underpinning these contexts.
- In the third phase, referred to as the platform phase, the trend from the previous phase continues, with a stronger focus on social topics, more aligned with soft science, compared to hard science. During this phase, numerous studies emerge focusing on digital platforms, which, through the interplay of technology and human elements, become a dominant force in the online environment. These studies examine the role of gatekeepers like Facebook and Google, the influence of platforms within the information system, and platforms such as Yelp and Uber, addressing themes such as user reviews and the working conditions of platform drivers. In addition to being a central topic, platforms have also become essential for data collection. This phase sees the emergence of studies using user-generated data from social platforms to conduct research through digital methods, such as using tweets to understand the phenomenon of electronic cigarettes.
- In the fourth and final phase, referred to as the artificial intelligence phase, numerous articles emerge where the application of AI in various fields takes center stage, from healthcare to Industry 4.0, to the creation of collaborative filters, and to the issue of the information system on platforms like Facebook, particularly how

issues like misinformation and fake news are addressed. Algorithms are viewed as actors in society, fully integrated into the digital environment. AI is also associated with topics such as healthcare, policymakers, job interviews, and education. Moreover, the social component is predominant in studies questioning algorithmic awareness, tied to issues like the digital divide (Gran, Booth, and Bucher, 2020).

In conclusion, the increased attention given by soft science disciplines to algorithmic feedback and recommendation systems is linked to a greater focus on the social issues these technologies entail. The two moments of greatest growth are closely connected to the widespread diffusion of online environments, where users play a major role in activities, starting from the social phase of our study, and to the spread of AI technologies. These not only spark technological interest but also raise social concerns regarding their implications for society, such as ethical questions about certain applications or the opportunities they present.

5. THE EVOLUTION OF DEFINITIONS IN THE SOFT SCIENCES THROUGH THE APPLICATION OF LDA AND ACL

Recognizing the significant rise in interest in this subject in recent years—particularly within the soft sciences compared to the hard sciences—the next step in this research is to analyze how academic engagement with this topic has evolved. Using content analysis, we aim to highlight the changes that have occurred over time and across disciplines, determining whether different soft science disciplines interpret the same phenomenon differently from hard science disciplines. Content analysis is a set of techniques that enables the breakdown of information (e.g., a text) into its simplest constituent components. This analytical approach can partially be placed within the framework known as the Bag-of-Words method (Bolasco, 2013), a coding technique that disregards the order of words and their roles in the text, aiming to simplify the computational processing of a text.

Every day, documentary traces are produced, and with the advent of the internet, these traces have multiplied exponentially, forming what is known as Big Data. The substantial increase in documents has made content analysis a fundamental approach in contemporary research methods. The main purpose of using documents is to understand the worldview or interpretation of a particular actor (users, researchers, political figures, organizations, etc.) who produced the trace. It is important to recognize that documents are social products that convey the viewpoint of the person who produced them. In other words, documents are not neutral, as they stem from specific cultures, have particular purposes, and are thus influenced by context and individual characteristics (Amaturo and Punziano, 2013).

An essential aspect of content analysis is the data collection phase, which precedes the coding and classification of the unit of analysis and the analysis itself, culminating in the presentation of results. In the data collection phase, the researcher focuses on the following:

- data access: The researcher must determine whether the documents are institutional (and thus whether there is an archive and whether it is accessible) or private. In the latter case, the researcher must figure out how to obtain access, whether through direct contact, payment, or the use of an archive;
- data selection: The researcher must also decide how much and which part of the documentation to analyze, as well as consider the originality of the document—although this concern has become obsolete in the case of digital documents.

These two steps were carried out following the assumptions of content analysis as an inquiry (Losito, 1993), a type of content analysis that, in this case, falls into the third category (Rositi, 2000). Similar to the construction of a survey, the researcher interrogates the content, where, instead of sequentially posing questionnaire questions, a series of entries related to the collected content is considered. These entries «are nothing more than the variables to be inserted into the data matrix for analysis purposes» (Amaturo and Punziano, 2013: 144).

The operational definition phase, which underpins the process of constructing the empirical basis, involved dividing the content into 3 domains each containing a set of variables (Table 2):

1. general information;
2. publication information;
3. content information

The first domain includes the operationalization of variables related to *the language, the year and the geographical area of publication*. The second domain includes variables related to the *number of citations and the research area*. The third domain, finally, includes variables related to the *title, the abstract and the keywords* of each record.

Table 2 – The building of the empirical base

General Information			Publication Information		Content Information		
Language	Year of Publication	Geographical Area	Citation	Research Area	Title	Abstract	Keywords
Nominal variable	Ordinal variable	Nominal variable	Interval variable	Nominal variable	Text variables		

In the present research, after data collection, a corpus was constructed by concatenating the variables of the third domain, creating a single textual variable. This corpus was cleaned of natural noise resulting from the automatic data collection process. For instance, one publication was found to contain Spanish and Chinese text, despite being marked as English, and therefore the non-English sections were removed. Additionally, it was found that the acronym "AI" (which in our context refers to Artificial Intelligence) was used inconsistently, so studies using the acronym in an unexpected sense were manually excluded. This was achieved by observing the usage contexts of the acronym.

The corpus, generated using T-LAB software, was then processed using the Latent Dirichlet Allocation (LDA) algorithm (Blei et al., 2003), a bottom-up Bayesian procedure for topic modeling developed in 2003. This process is characterized by interpreting each document as a collection of topics, which in turn are determined by specific terms. These terms combine to form a topic, with the key feature of this process being the interchangeability of terms and topics within a document. In other words, LDA allows us to learn the distribution of topics from a collection of documents, enabling us to predict the topic distribution in a new document based on the terms it contains (Blei et al., 2003).

Subsequently, we applied Lexical Correspondence Analysis (LCA), a multidimensional analysis technique in which the minimal unit of meaning is the individual word or graphic form (Amaturo and Punziano, 2013). Once applied to a text, this procedure seeks and visualizes latent linguistic structures that express prevalent concepts or themes. The graphic representation produced by this approach is displayed on a factorial plane formed by two independent dimensions of meaning—factors—each representing a latent aspect of the association of the observed data. The dispersion of terms around the axes' origin shows the strength of the association. The farther a point (a word) is from the origin of an axis, the greater its contribution to that axis. The closer the points, the greater the interdependence between the categories represented by those points. Additionally, the value of the coordinates on the graph suggests the importance of a point relative to the axis. The proximity of words or word categories on the factorial plane indicates their combination or association within the text. In fact, the closeness of two terms on the factorial plane shows their similar use within the observed document. In contrast, the proximity of two documents on the plane indicates the presence of a similar vocabulary in both.

Next, a Cluster Analysis was conducted. This term refers to a set of data analysis procedures that group elements of a set into clusters based on their similarity and homogeneity. Clustering involves comparing objects based on their characteristics. The clusters adhere to two criteria: cohesion, meaning the statistical units in each group share similar characteristics, and separability, meaning different clusters are as distinct as possible.

The collected corpus consists of 8,051 documents, with each document containing the title and abstract of a specific publication. Afterward, the corpus underwent preprocessing, performed automatically by T-LAB software. This process reduced the heterogeneity of the corpus through lemmatization (i.e., reducing morphological heterogeneity), which involved removing numerical characters, punctuation, and stop words. Simultaneously, the software also cleaned lexical, reducing the text to minimal sense units and consolidating terms that shared the same semantic field.

Following this process, the corpus' size, denoted as N, was reduced to 1,674,150 tokens. Identical tokens were stacked, forming the vocabulary (V), which contains 42,526 types, of which 18,129 are hapaxes. Each type in the vocabulary is associated with a count of occurrences (i.e., how many times the token appears in the corpus). The type-token ratio (TTR) of the entire corpus indicates language diversity, with a value of 2.5%, below the 20% threshold, suggesting a collection sufficiently large to capture linguistic richness. The linguistic sophistication, indicated by the ratio between the number of hapaxes (words that appear only once) and the vocabulary size, is 42.6%, lower than the 50% threshold, and therefore considered acceptable.

In conclusion, the topic modeling returned the following themes:

1. Artificial intelligence (13%)
2. Algorithmic development (11%)
3. Algorithmic consuptiion (6%)
4. Data treatment (9%)
5. Health (7%)

6. Work and economy (10%)
7. Social Platform (9%)
8. Machine Learning (9%)
9. Recomendation and personalization (13%)
10. Eduation (12%)

5.1 Lexical Correspondence Analysis and Cluster Analysis

The themes that emerged constitute different perspectives on the same subject. To understand the relationships between topics and research areas, multidimensional analysis techniques are required, starting with Lexical Correspondence Analysis (ACL) (Benzecri, 1973; Lebart et al., 1998) and followed by Cluster Analysis (Lebart, 1994).

The results obtained through ACL represent two latent dimensions, called factors. These factors are crossed and represented on a factorial plane showing the topics previously described, the most representative terms of each topic, the active variables used for the analysis, and illustrative variables (Figure 11), which do not influence the construction of the factors. The active variables are the number of citations of the publication, the research area, and the topics outlined through topic modeling; the illustrative variables include the year of publication and the publication region. The total inertia extracted by the two factors is 18.21% (9.18% from the x-axis and 9.03% from the y-axis).

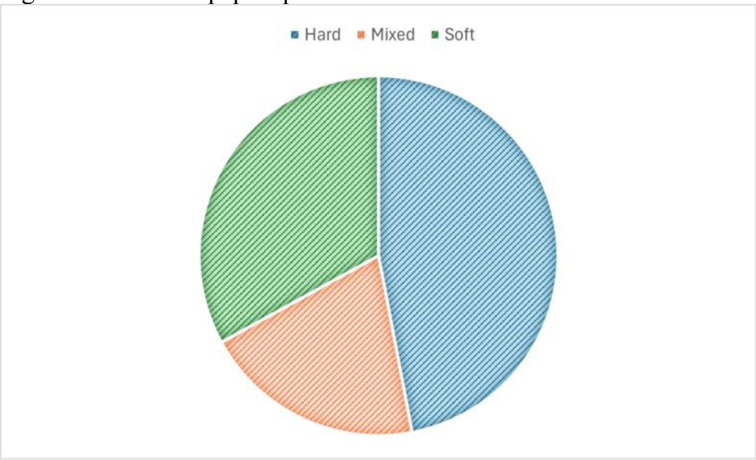
Specifically, regarding the active and illustrative variables, the number of citations (Table 3) is divided into quartile classes, each containing 25% of the entire corpus.

Table 3 – Number of citations in quartile classes.

Quartile Label	**number of citations**
FirstCit	0-1
SecondCit	2-4
ThirdCit	5-14
FourthCit	15+

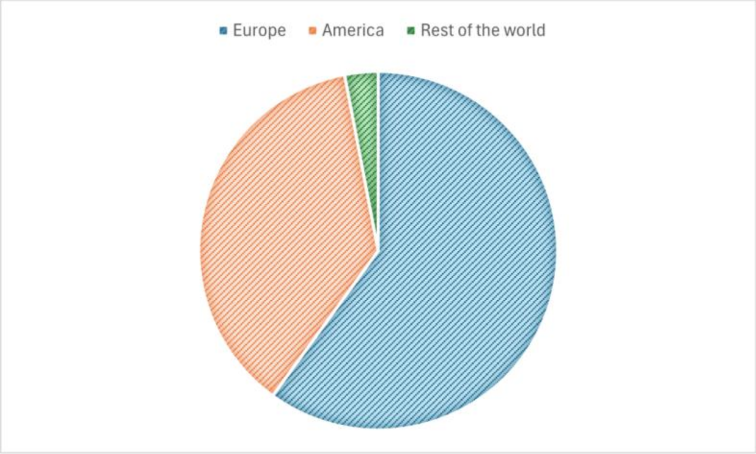
The research area (Figure 9) is divided into three categories: soft, mixed, and hard. “Soft” refers to a publication focused on a single discipline categorized as a soft science; “mixed” refers to a publication that spans two or more soft science disciplines; and “hard” refers to a publication that, besides covering at least one soft science discipline, also connects to at least one hard science discipline.

Figure 9- Scientific papers per research area



The region of publication was collected from WOS as the city of publication. To reduce data dispersion, cities were grouped by continent. Additionally, due to lower occurrences, Asia, Oceania, and Africa were grouped into a single category labeled "Rest of the World" (Figure 10).

Figure 10 – Scientific papers per geographical area



As shown in Figure 11, the first factor obtained relates to the social-technical sphere. On the positive semi-axis, publications focus on technical aspects of algorithms, recommendation systems, and feedback loops. On the negative semi-axis, publications discuss the social implications of adopting these technologies. The positioning of the research area confirms this factor points: soft science and mixed publications are shifted toward the negative semi-axis, while hard science publications are shifted toward the positive semi-axis. The social semi-axis is characterized by topics such as social platforms, algorithmic consumption, labor and economy, education, and artificial intelligence. These topics focus on the social consequences of applying certain technologies, such as changes in human relationships, social network dynamics, the spread of fake news, and the alteration of consumption patterns due to recommendation systems. Ethical and regulatory issues also arise from the increasing pervasiveness of AI in society. On the positive semi-axis, terms like "treatment," "cancer," "graph," and "recommendation" describe more technical applications. Conversely, the negative semi-axis includes terms that represent the social nature of the phenomenon: "influencer", "medium", "brand", "ethical", and "political".

The second dimension represents subjects on the positive semi-axis, reflecting the individual and personal use of algorithms and recommendation systems, with a sense of flexibility. The education topic characterizes this semi-axis, as publications discuss personalized curricula tailored to the specific needs of students. Key terms include "student," "teacher," "tutor," "project," "feedback," and "personalize." When discussing specific user demographics, the negative semi-axis concerns sectors, both disciplinary areas and target groups. The health topic characterizes this semi-axis, as publications focus on standardized treatment or diagnostic methods. Another significant topic is algorithmic consumption, reflecting the standardization of cultural consumption and the view of individuals not as unique but as part of a target group. Key terms include various diseases such as HPV, cancer, and diabetes, along with the names of popular social media platforms like YouTube and Twitter.

Figures 11- Factoral plan with active and illustrative variable illustrated (made by T-Lab)

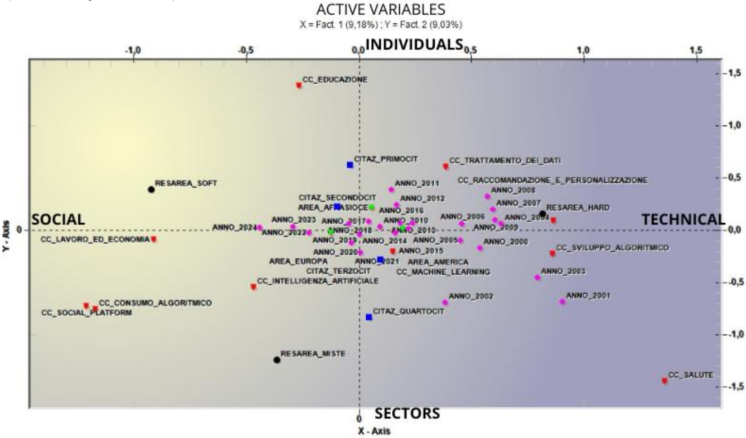
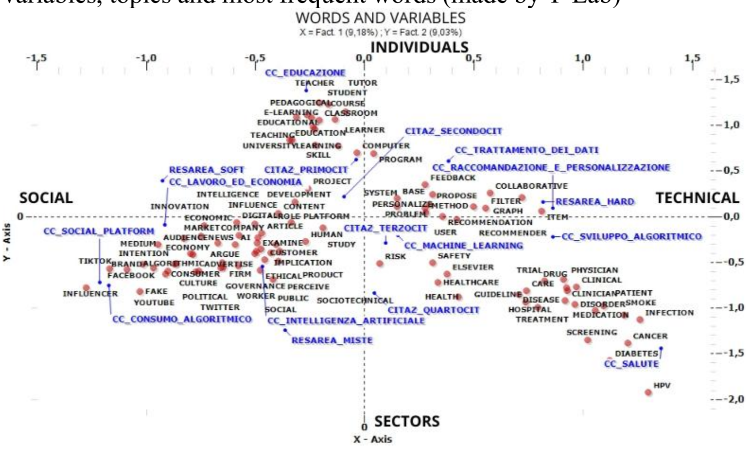


Figure 12 – ACL factorial plan with active variables, illustrative variables, topics and most frequent words (made by T-Lab)

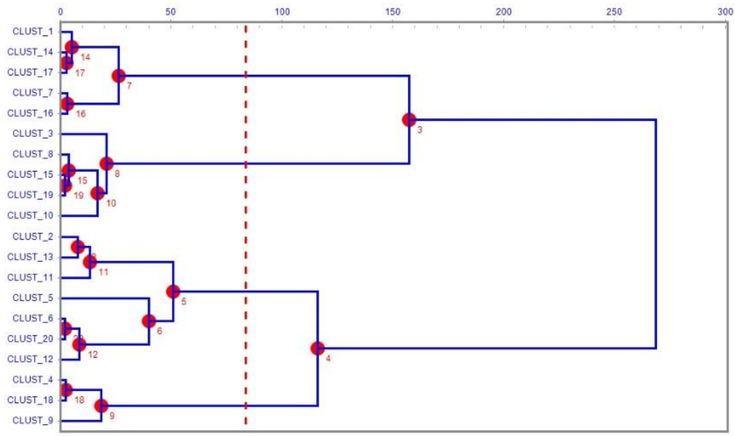


Having defined the latent dimensions of our plane, the objective is to systematize the findings, including the different interpretations and definitions of the same topic.

The reading of the key insights projected onto the graph is guided by the stretching of the temporal dimension on the graph. Observing the graph, it is possible to note how the four previously described temporal phases extend from the right side (which intersects individuals and sectors with the technical plane) toward the left side (which intersects individuals and sectors with the social plane).

To delve further into the analysis and further synthesize the information contained in the dataset, a cluster analysis was performed. A hierarchical clustering method was adopted, meaning that the number of clusters to analyse was not predetermined; instead, the optimal partition was chosen by cutting the dendrogram (Figure 13).

Figure 13- Dendrogram graph with three cluster partition (made by T-Lab)



Four clusters emerged (Figure 14), aligning with the quadrants of the factorial plane. Each cluster can thus be interpreted according to the semi-axes on which they are positioned, representing a conceptual sphere of our topic of investigation. These clusters reflect different interpretations and visions based on the context and the period when the analyzed documents were published.

1. *Cluster 1 (blue) - "Instrument in Design"***:** This cluster is characterized by the technical and subject semi-axes. The most representative terms include "user," "item," "recommendation," "propose," "filter," "model," "collaborative," "dataset," "method," and "retrieval." This cluster represents a design phase, with publications focusing on algorithmic development and recommendation and personalization systems. The technological applications are still in an experimental and investigative stage. The term "instrument in design" refers to both the technical and

conceptual phases, as the research topic is in an embryonic stage, with an ongoing search for appropriate scientific methods. The feedback loop is seen as a concept with unclear boundaries, still seeking optimal application. Algorithms are regarded as tools in the design phase.

2. *Cluster 2 (light blue) - "Technical Instrument":* This cluster is characterized by the technical and structure semi-axes and is strongly influenced by medical terms. The most frequent words include "patient," "health," "treatment," "screening," "clinical," "cancer," and "risk." Publications in this cluster are often highly cited and are positioned in the higher quartiles of citation distribution. This cluster illustrates how some disciplines, particularly hard sciences, have greater scientific prominence. The term "standardization" defines this cluster, as it reflects the search for a universal solution to a general problem, such as a disease. Algorithms are regarded as technical instruments for maximizing the results of a codified process, viewed as a technical tool for standardized processes.
3. *Cluster 3 (red) - "Social Instrument"* **:** This cluster is characterized by the social and subject semi-axes, predominantly related to education. The most common terms include "student," "learning," "course," "teaching," "program," "university," and "skill." The term "individualization" can summarize this cluster, as publications focus on algorithmic applications for customizing education to individual student needs. The feedback loop is treated from a social perspective but still considers individual needs. Algorithms are seen as technical tools applied in social contexts, though the cooperation between the user and the algorithm remains separate and not dynamic.
4. *Cluster 4 (pink) - "Social Actor"*: This cluster is characterized by the social and structure semi-axes. The defining terms include "medium," "consumer," "AI," "news," "social," "economy," and "digital." The term "massification" summarizes this cluster, where consumers are considered target groups. Feedback loops in social networks and platforms are central here, reinforcing the concept of filter bubbles (Pariser, 2011). Algorithms are regarded as social actors in digital environments, participating in processes of socialization, cultural consumption, and information dissemination. The interaction between users and algorithms becomes nearly instantaneous, as seen on platforms like TikTok, where content adapts based on minimal user interactions.

Following this reading, we see how scientific production progressively moves from a purely technical approach to the topic (concerning the development and structure of algorithms and the underlying processes) toward discussions related to the application contexts of algorithms and their resulting effects.

In this sense, the information contained in the dataset is synthesized into a progressive perspective that centers on the relationship between individuals and algorithms. Consequently, it cannot exclude a discussion on the topic of agency and the (non) neutrality of algorithms.

Proceeding in an orderly manner, it is possible to virtually overlay on the plane the three constitutive elements of the embedded process related to human agency (Klinger and Svensson, 2018), originally elaborated by Emirbayer and Mische (1998): *iteration, practical evaluation, and projectivity* (Figure 14).

Starting from the technical side, where the first two temporal phases are located, all aspects related to the connections between agency and the past (iteration) and agency and the present of individuals (practical evaluation) emerge. These aspects are evident in scientific productions focused on algorithmic innovation in terms of development and its possible implications for individual statuses and conditions over time (data processing, recommendations, and personalizations). This type of production aligns with the need to discuss patterns and selective models which, according to Emirbayer and Mische (1998), concern the opportunities to "sustain identities, interactions, and institutions over time" (p. 971). Initially, these discussions revolve around "algorithmic calculations, how they are designed to select, recognize types, locate categories, sort and rank big/thick/trace data from the past" (Klinger and Svensson, 2018, p. 4660).

If, on this side of the process, the discussion on algorithms focuses on their static and purely structural aspects (primarily addressed by fields related to hard sciences) the discourse shifts as we follow the temporal dimension and move toward the social side of the plane, where the last two phases are located.

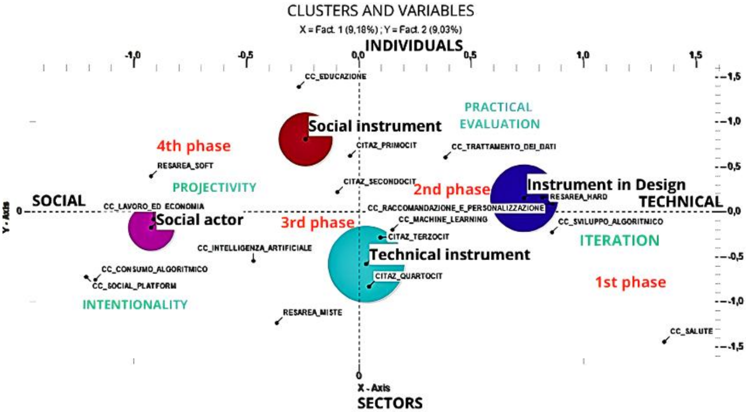
On this side of the plane, as mentioned, we find all elements of scientific production predominantly related to the field of soft sciences, which only partially concern structural aspects (such as artificial intelligence). Instead, they focus mainly on algorithmic applications in terms of context (work and economy, social platforms).

From this, a reading emerges that aligns with the third element of projectivity, corresponding to all possibilities of forecasting and commercial strategy applied even to social platforms, based on digital traces left online and the profiling and recommendations generated through the aforementioned processes of iteration and practical evaluation.

In light of the dataset's synthesis, there arises a need to further focus on the non-neutrality of the algorithm, understood as a series of processes capable of autonomously making decisions. If we also consider the fourth element of *intentionality* (Mitcham, 2014) it becomes necessary to interpret algorithmic logic as a complex process connected both to its application context (fields of intervention, media logics, etc.) and to human intervention. A process that, following Latour's (1992) contribution, possesses human, non-human, unhuman, and inhuman characteristics.

In conclusion, one crucial conclusion emerges clearly from this analsysis: today, it is essential to approach the study of this topic without neglecting the theoretical perspective of Actor-Network Theory (ANT) (Latour, 1992). The principles of ANT allow for a focus on the human-nonhuman relationship, enabling contemporary social researchers not only to explore digital technologies and devices but also to shed light on the role of algorithms and recommendation processes. This endeavour becomes possible by formulating the right research questions and drawing on the extensive body of literature produced on this topic over the past decades in the field of soft sciences, as empirically demonstrated in this work. These studies enable an appropriate categorization of processes and concepts linked to nonhumans and their agency, considering them as "Mediators" rather than mere "Intermediaries" (Sayes, 2014, p. 138). This role suggests moving beyond intellectual reductionism regarding their direct replacement of human actors, instead conceiving them as "transcendental conditions for our collectives, nor are they merely a black box that lines up other actors, nor are they merely placeholders for a human actor" (*ibidem*).

Figures 14 – Cluster superimposed on ACL factoral plan (made by T-Lab)



6. CONCLUSIONS

This work has explored how definitions and concepts related to the algorithmic feedback loop have evolved over time. Specifically, the changes in the field of soft sciences from 2000 to 2023 were analyzed. A Systematic Literature Review (SLR) was conducted to highlight the increasing interest in the topic and the resulting proliferation of publications, revealing that soft sciences experienced proportionally greater growth than hard sciences.

The temporal analysis identified four main phases: the technological phase (2000–2006), the social phase (2007–2012), the platform phase (2013–2017), and the artificial intelligence phase (2018–2023). These phases laid the groundwork for understanding the different perspectives that the same topic could take over time and across various disciplines.

Subsequently, ten topics were defined to systematize the various focal points of the publications, ranging from artificial intelligence to the application of algorithms in educational and healthcare contexts. Finally, the ACL and cluster analysis confirmed a division between a more technical scientific debate and a social vision, while also acknowledging the growing presence of interdisciplinary approaches that combine hard and soft sciences.

The interpretation of the cluster analysis highlighted different perspectives and views on the concept of the algorithmic feedback loop within our study sample. Four clusters were defined, bringing to the forefront the concept of an "actant," defined as "any entity that acts, regardless of its ontological status (human or non-human, concrete or abstract), size, scale (individual or aggregated), or specific characteristics" (Magaudda and Neresini, 2020). Feedback loops are processes that characterise the interaction between users and recommender systems (Airoldi, Beraldo and Gandini, 2016) and this leads to understand how the algorithms increasingly take on the role of active agents in society, embedded in a network of entities that, through their relationships, determine the observed effects. From this perspective, algorithms acquire a form of agency—the capacity to influence processes and actively participate in shaping them. This highlights how data and outputs are not merely generated by users but are also molded by the underlying logic of platforms. (Beer, 2022). This social role held by the algorithmic feedback loop, as explained in the social actor cluster, characterized by the most recent publications, strengthens the hypothesis that soft sciences are paying increasing attention to the topic, surpassing hard sciences in proportional growth from their starting point.

These initial conclusions, in turn, open up several other issues and challenges to address. Foremost among them is the methodological question: on the one hand, the paradigmatic directions to consider, and on the other, the set of research techniques and tools necessary and most helpful in pursuing the knowledge objectives set forth.

The study should not be considered conclusive; instead, it leads to other open questions and queries that will need to be investigated further. Notably, if the concept of algorithmic feedback loops has changed so drastically in recent years, it is reasonable to expect that it can change again, offering opportunities for further reflection and exploration. Attention should be drawn to the longstanding rivalry between the soft and hard sciences. Given the profound societal impact of algorithms (particularly on cultural diversity, user opinions, and online social dynamics) can reflections on this issue continue to follow separate disciplinary paths? Or is it now imperative that discussions bridge both paradigms to address the topic comprehensively? Ultimately, the results show how topics such as algorithms, recommendation systems, and feedback loops have undergone a conceptual shift. There has been a move from a strictly technical focus to increasingly emphasizing social and ethical aspects.

REFERENCES

AIROLDI, M. (2021). The techno-social reproduction of taste boundaries on digital platforms: The case of music on YouTube. *Poetics*. *89*: 101563. https://doi.org/10.1016/j.poetic.2021.101563

AIROLDI M., BERALDO D., GANDINI A. (2016). Follow the Algorithm: An Exploratory Investigation of Music on YouTube. *Poetics*. 57: 1–13.

AIROLDI, M., ROKKA, J. (2022). Algorithmic consumer culture. *Consumption Markets & Culture*. *25*(5): 411–428. https://doi.org/10.1080/10253866.2022.2084726

AMATURO, E., PUNZIANO, G. (2013). *Content analysis tra comunicazione e politica.* Roma: Carocci.

BEER, D. (2022). The problem of researching a recursive society: Algorithms, data coils and the looping of the social. *Big Data & Society*. *9*(2): 205395172211049. https://doi.org/10.1177/20539517221104997

BENZECRI, J. P. (1973). *L'analyse des donnees.* Dunod

BLEI D.M., NG A.Y. JORDAN M.I. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research. 3: 993-1022

BOLASCO, S. (2013). *L'analisi automatica dei testi. Fare ricerca con il text mining*. Roma: Carocci.

CAPUTO, A., FELACO, C., PUNZIANO, G. (2017). *La ricerca trasversale e longitudinale nelle scienze sociali*. Milano: Franco Angeli.

CARRADORE, R. (2013). Cibernetica e ordine sociale. Modelli e immagini di società. *Scienza & Politica*. 25(48): 149–173. https://doi.org/10.6092/issn.1825-9618/3899

DANG, T. N. Y. (2018). The nature of vocabulary in academic speech of hard and soft sciences. *English for Specific Purposes*. *51*:69–83. https://doi.org/10.1016/j.esp.2018.03.004.

DENICOLAI, L., FARINACCI, E. (2020). Te lo dico con un video. I linguaggi audiovisivi del quotidiano social. *L'avventura, International Journal of Italian Film and Media Landscapes.* Fascicolo Speciale/Dicembre: 145-165.

EMIRBAYER M., MISCHE, A. (1998). What is agency? *American Journal of Sociology.* 103(4): 962–1023.

FEYERABEND, P. (1975). *Against Method*. London: New Left Books.

GRAN, A. B., BOOTH, P., BUCHER, T. (2020). To be or not to be algorithm aware: a question of a new digital divide? *Information, Communication & Society*. *24*(12): 1779–1796. https://doi.org/10.1080/1369118X.2020.1736124.

KLINGER, U., SVENSSON, J. (2018). The end of media logics? On algorithms and agency. *New Media & Society*. 20(12): 4653-4670. https://doi.org/10.1177/1461444818779750.

KNELL, M. (2021). The digital revolution and digitalized network society, *Review of Evolutionary Political Economy*. 2: 9-25. https://doi.org/10.1007/s43253-021-00037-4.

KUHN, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

KUNTAG, J. R., SIJABAT, S. (2023). Pragmatics element in product promotion content through short video platform tik tok, instagram reels, and youtube shorts. *Jurnal Ekonomi*. 12(04): 2156–2165.

LAMÉ, G. (2019). Systematic Literature Reviews: An Introduction. *Proceedings of the Design Society: International Conference on Engineering Design*. 1(1): 1633–1642. https://doi.org/10.1017/dsi.2019.169

LATOUR, B. (1987). *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge: Harvard University Press.

LATOUR, B. (1992). Where Are the Missing Masses? In W. Bijker, J. Law (Eds.), *Shaping Technology/Building Society: Studies in Sociotechnical Change* (pp. 225–258). Cambridge: MIT Press

LEBART, L. (1994). *Complementary use of correspondence analysis and cluster analysis.*

LEBART, L., SALEM, A., BERRY, L. (1998). *Exploring textual data*. Dordrecht: Kluwer.

LOSITO, G. (1993). *L'analisi del contenuto nella ricerca sociale*. Milano: FrancoAngeli.

MITCHAM, C. (2014) Agency in humans and in artifacts: a contested discourse. In P. Kroes & P.P. Verbeek (Eds), The Moral Status of Technical Artefacts (pp. 11-29). Berlin: Springer.

MAGAUDDA, P., NERESINI, F. (2020). *Gli studi sociali sulla scienza e la tecnologia*. Bologna: il Mulino.

PAGE, M.J., MCKENZIE, J.E., BOSSUYT, P.M., BOUTRON, I., HOFFMANN, T.C., MULROW, C.D., SHAMSEER, L., TETZLAFF, J.M., MOHER, D. (2021). Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. *Journal of Clinical Epidemiology*. 134: 103-112. https://doi.org/10.1016/j.jclinepi.2021.02.003

MERTON, R.K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations.* Chicago: University of Chicago Press.

NELSON, R.R. (2002). On the uneven evolution of human know-how, *Research policy*. 32(6): 909-922. https://doi.org/10.1016/S0048-7333(02)00093-8

PARISER, E. (2011). *The Filter Bubble: What The Internet Is Hiding From You*. New York: Penguin press.

PETTICREW, M., ROBERTS, H. (2005). *Systematic reviews in the Social Sciences: A Practical Guide*. Oxford: Wiley-Blackwell.

POELL T., NIEBORG, D., DUFFY, B. (2021). Platforms and cultural production. Cambridge: Polity Press.

POPPER, K. (1959). *The Logic of Scientific Discovery*. London: Hutchinson.

PREDETTI, A. (2006). *I numeri indici. Teoria e pratica*, Giuffrè, Milano.

ROSITI, F. (2000). Analisi del contenuto. In F. Rositi, M. Livolsi (Eds.), *La ricerca sull'industria culturale*. *Roma: La nuova Italia Scientifica* (pp. 59-94). Roma: Carocci editore.

SAYES, E. (2014). Actor–Network Theory and methodology: Just what does it mean to say that nonhumans have agency?. *Social studies of science*, 44(1), 134-149. https://doi.org/10.1177/0306312713511867

SHAPIN, S. (2022). Hard science, soft science: A political history of a disciplinary array. *History of Science*. 60(3): 287–328. https://doi.org/10.1177/00732753221094739

SHARIF, S.P., MURA, P., WIJESINGHE, S.N.R. (2019). A systematic review of systematic reviews in tourism. *Journal of Hospitality and tourism management*. 39: 158-165.

WU, L., KITTUR, A., YOUN, A., MILOJEVIĆ, S., LEAHEY, E., FIORE, S.M., AHN, Y. (2022). Metrics and mechanisms: Measuring the unmeasurable in the science of science. *Journal of Infometrics*. 16(2): 1-30. https://doi.org/10.1016/j.joi.2022.101290.