



Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Beyond artificial intelligence controversies

What are algorithms doing in the scientific literature?

Munk, Anders Kristian; Jacomy, Mathieu; Ficozzi, Matilde; Jensen, Torben Elgaard

Published in:
Big Data and Society

DOI (link to publication from Publisher):
[10.1177/20539517241255107](https://doi.org/10.1177/20539517241255107)

Creative Commons License
CC BY-NC 4.0

Publication date:
2024

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Munk, A. K., Jacomy, M., Ficozzi, M., & Jensen, T. E. (2024). Beyond artificial intelligence controversies: What are algorithms doing in the scientific literature? *Big Data and Society*, 11(3).
<https://doi.org/10.1177/20539517241255107>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Beyond artificial intelligence controversies: What are algorithms doing in the scientific literature?

Big Data & Society
July–September: 1–20
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20539517241255107
journals.sagepub.com/home/bds



Anders Kristian Munk¹ , Mathieu Jacomy² ,
Matilde Ficozzi² and Torben Elgaard Jensen²

Abstract

Mounting critique of the way AI is framed in mainstream media calls for less sensationalist coverage, be it jubilant or apocalyptic, and more attention to the concrete situations in which AI becomes controversial in different ways. This is supposedly achieved by making coverage more expert-informed. We therefore explore how experts contribute to the issuefication of AI through the scientific literature. We provide a semantic, visual network analysis of a corpus of 1M scientific abstracts about machine learning algorithms and artificial intelligence. Through a systematic quali-quantitative exploration of 235 co-word clusters and a subsequent structured search for 18 issue-specific queries, for which we devise a novel method with a custom-built datascap, we explore how algorithms have agency. We find that scientific discourse is highly situated and rarely about AI in general. It overwhelmingly charges algorithms with the capacity to solve problems and these problems are rarely about algorithms in their origin. Conversely, it rarely charges algorithms with the capacity to cause problems and when it does, other algorithms are typically charged with the capacity to solve them. Based on these findings, we argue that while a more expert-informed coverage of AI is likely to be less sensationalist and show greater attention to the specific situations where algorithms make a difference, it is unlikely to stage AI as particularly controversial. Consequently, we suggest conceptualising AI as a political situation rather than something inherently controversial.

Keywords

Controversy mapping, issue mapping, artificial intelligence, machine learning, visual network analysis, semantic analysis

This article is a part of special theme on *Analysing Artificial Intelligence Controversies*. To see a full list of all articles in this special theme, please click here: <https://journals.sagepub.com/page/bds/collections/analysingartificialintelligencecontroversies>

The past decade has seen a proliferation of artificial intelligence controversies. As of February 2024, the *AI, Algorithmic, and Automation Incidents and Controversies* (AIAAIC) database, which is an independent public interest initiative, had recorded more than 1300 cases where AI posed a risk, caused observable harm, or otherwise generated enough public concern to prompt a response (AIAAIC, n.d.). In 2023, a report from the Stanford Institute for Human Centered AI estimated that the rate of such incidents had increased 26-fold since 2012 (Maslej et al., 2023). Somewhat paradoxically, this proliferation

of controversies stands in contrast to the way AI is often portrayed in mainstream media. A recent study in Canada

¹Department of Technology, Management & Economics, Section for Human-Centered Innovation, DTU, Lyngby, Denmark

²Department of Culture & Learning, TANTlab, Aalborg University, Copenhagen, Denmark

Corresponding author:

Anders Kristian Munk, Department of Technology, Management & Economics, Section for Human-Centered Innovation, DTU, Lyngby, Denmark.
Email: ankm@dtu.dk



showed that AI coverage prioritises potential benefits over risks (Dandurand et al., 2023), contributing to the formation of AI as what Michel Callon (1998) calls a “cold situation” in which the definition of problems and solutions are not in dispute. Similarly, a UK study looking at the coverage of AI in healthcare found that the dominant frame is how the technology outperforms its human counterparts (Bunz and Braghieri, 2022), efficiency being taken for granted here as the relevant problem to solve. Others posit that the picture is more blurry but still highlight the importance of mainstream media in shaping public perception of AI and its responsibility towards the formation of AI as a properly unfolded (i.e., less cold) matter of public concern. A study in the U.S. thus found that while risks tend to be dealt with in more detail, most coverage simply focuses on new opportunities in business and technology (Chuan et al., 2019) and a study of how ChatGPT has been covered in the UK found that it is framed both as an imminent systemic danger and as a ground-breaking solution to a range of existing problems (Roe and Perkins, 2023). According to the authors of the latter, it “serves as a timely reminder of the need for critical evaluation of media coverage, a shift away from sensationalism, and a movement towards comprehensive, expert-informed reporting on AI technologies” (ibid:8). The call for a less sensationalist coverage is also present in Lucy Suchman’s commentary for this special issue where she argues against the tendency to stage AI as if it was one (controversial or uncontroversial) thing and attend instead to some more basic questions such as “what is the problem for which these technologies are a solution? According to whom?” (Suchman, 2023). The fact that media coverage of controversy can also be sensationalist and contribute to AI hype is for example demonstrated in a recent study of how critical AI discourse is structured in French mainstream media. Here, Crépel et al. (2021) find a clear opposition between “a prophetic, even apocalyptic discourse” (ibid:86) and a more grounded attention to issues like filtering, censorship, surveillance, privacy, discrimination or disinformation. In short, the mounting critique of how AI is framed in mainstream media is not only about a lack of attention to AI controversy in general but about a lack of situated attention to the different problems in relation to which different AI technologies come to matter in different ways. The call for a more expert-informed coverage is first and foremost a call for specificity. In this paper we therefore explore how experts are themselves framing AI in situated ways by asking what algorithms are doing in the scientific literature.

The role of the public expert is of course not reducible to that of the scientist, and experts can certainly derive their authority from other places than academic institutions or scientific peer review, yet scientific discourse plays a significant role in what Noortje Marres (2012) calls “issuefication”. Controversy, she reminds us, is not only about “actors taking positions but a process of the problematisation of objects, by which they become charged with various social,

economic, political problematics or issues” (ibid.). When we are talking about emerging AI technologies, the capacity to charge them with such problematics or issues arguably rests in no small part with the experts who first develop and apply them, conceptualise their risks and potentials and observe their consequences in different situations. We are thus interested in the way algorithms gain agency in and through the scientific literature and we define agency in the actor-network theoretical sense typically adopted in controversy mapping (Venturini and Munk, 2021), namely as something that makes an observable difference to a situation. This capacity for action is closely associated with the way specific media contribute to the enactment of controversy (Marres, 2015) and we take peer-reviewed journal publications to be not just an institution that reports on, and therefore allows us to observe, what algorithms are doing in science, but an institution that contributes actively to the issuefication of algorithms in the sense that the inscription of algorithmic capacities in scientific discourse adds to their agency. It provides possible ways in which the public might come to grips with what algorithms can be and do – for example what problems they are a solution to, to follow Suchman’s question – and some of those ways would necessarily be part and parcel of a more expert-informed media coverage of AI.

To map the agency of algorithms in the scientific literature we conduct a semantic network analysis of 1 M abstracts from Scopus published between 2011 and 2021. We then use the semantic network as the basis for a three-step qualitative analysis. *First*, we annotate the topical structure of the network and ask broadly what algorithms are doing in different topical contexts in the scientific literature. *Second*, we query the corpus specifically for 18 issues where AI technologies are known to be controversial. This allows us to analyse the distribution of each issue across the topical structure of the corpus and assess the kind of agencies algorithms have in relation to the issue in question. Are they for example framed as a cause of the problem or as a solution to it? *Third*, we select one of these issues, namely the problem of fakes, for in-depth exploration.

This approach follows similar work in digital controversy mapping, for example analysing the scientific literature on ecosystems (Tancoigne et al., 2014) or obesogenic environments (Elgaard Jensen et al., 2019), meeting notes from the international climate negotiations (Venturini et al., 2014) or, indeed, critical AI discourse in legacy media (Crépel et al., 2021). The way we use the quantitative mapping to produce questions for further qualitative analysis also follows established quali-quantitative practices (Latour et al., 2012; Munk, 2019) although we develop our own systematic strategy for qualitative annotation and custom-build a datascape to support further qualitative exploration.

The approach differs somewhat from existing work in that we do not immediately focus on situations where AI

is framed as a problem. Instead, we construct the corpus in a way that allows us to explore all the ways in which algorithms have agency in the scientific literature. AI technologies have been the subject of a wide range of fruitful controversy analyses, all the way back to the early developments of neural networks in the 1950s (Olazaran, 1996), to recent debates about intentionality in large language models (Bender et al., 2021) or fairness in AI systems (Benbouzid, 2023). Much of it has profited from the disruptive force, or “ontological disturbances” (Whatmore, 2009), of events that render knowledge claims available for scrutiny, such as the failure of Google Flu Trends (Lazer et al., 2014), the crash of the self-driving Tesla Model S (Stilgoe, 2018) or the botched introduction of AI-enhanced photography in Huawei Moon Mode (Zhang, 2022). However, focusing directly on artificial intelligence controversies would obscure the ways in which AI technologies have agency in relation to non-AI problems and prevent us from appreciating broadly what a more expert-informed coverage of AI might indeed look like. Based on our findings, and to understand the way scientific discourse contributes to the issuefication of AI, we posit that attention is needed to the following set of facts:

- Scientific discourse is highly focused on specific algorithms in specific situations and rarely about AI in general.
- It overwhelmingly charges algorithms with the capacity to solve problems and these problems are rarely related to algorithms in their origin.
- Conversely, scientific discourse rarely charges algorithms with the capacity to cause problems and when it does, other algorithms are typically charged with the capacity to solve those problems.

To us, this suggests that the lack of controversiality in mainstream media coverage of AI is not about a lack of expert-informed attention to specific problems and situations, but a deeper question of how AI is issuefied in the first place. Speculating from the way algorithms are issuefied in scientific discourse, a more expert-informed media coverage would indeed be more situated and avoid sensationalist frames, but it is not obvious how this kind of detailed attention to the specific problems in relation to which AI technologies make a difference would also contribute to staging the controversiality of AI. In the discussion we therefore draw on Andrew Barry’s (2012) notion of political situations to understand how AI matters to a broader range of problems, and therefore also implicates a bigger hinterland of non-AI issues, than what can be captured with a focus on the situations where AI is explicitly controversial.

Corpus construction and semantic analysis

The two databases typically used for scientometric analysis are Thomson Reuters’ Web of Science and Elsevier’s Scopus. We chose to harvest our dataset through the latter. A recent study found that Scopus covered 39,758 scientific journals whereas Web of Science covered 13,610 in comparison, 99.11% of which were covered by Scopus as well (Singh et al., 2021). Scopus also had better coverage in the social sciences and humanities than Web of Science, although both areas were underrepresented compared to the physical, medical and life sciences in both databases. It should be noted that the Dimensions database is currently gaining recognition as a possible scientometric alternative to Scopus and Web of Science with much wider coverage in general and significantly better coverage for the social sciences and humanities (ibid.). Future iterations of the analysis presented here could therefore consider using Dimensions as an alternative data source.

We searched Scopus through the API for papers published in English language journals containing either of the search terms *algorithm**, *“machine learning”*, *AI*, or *“artificial intelligence”* in the abstract, title or keywords. We exported all bibliometric data including the abstract for each article published after 2010. As we conducted the search in October 2021, this left us with a corpus of 1,004,003 abstracts published between 2011 and 2021 and with the 10 complete years from January 1st, 2011, to December 31st, 2020, comprising 852,316 of those abstracts.

This was a key 10-year period for AI. DeepMind was established in 2010, Google Brain in 2011, the same year IBM Watson beat the reigning Jeopardy champions and managed to train the first neural network capable of recognizing cats from unlabelled images. Between 2011 and 2012, scientific papers about AI, algorithms and machine learning doubled in volume. Towards the end of the decade, machine learning models were making cancer diagnoses and AI-generated images were selling as fine art at auctions. As shown in Figure 1, our corpus more than triples in size over that time. When we consider the annual growth in papers on Scopus in general, which increases from 1.5 M in 2011 to 2.5 M in 2020 (Thelwall and Sud, 2022), the growth in volume of papers about AI, algorithms or machine learning is still noteworthy. More than 5% of all papers in Scopus published in 2020 are in our corpus, which is a two-and-a-half-fold increase compared to a decade earlier.

We used the SpaCy library in Python (Honnibal and Montani, 2017) to perform part of speech (PoS) tagging and named entity recognition (NER) on the abstract texts. This gave us a list of nouns and a list of named entities for each abstract. From the Scopus metadata we also had a list of author keywords and a list of scientific disciplines for each article. We combined the four lists to produce a co-word network using the NetworkX library in Python. Each node in

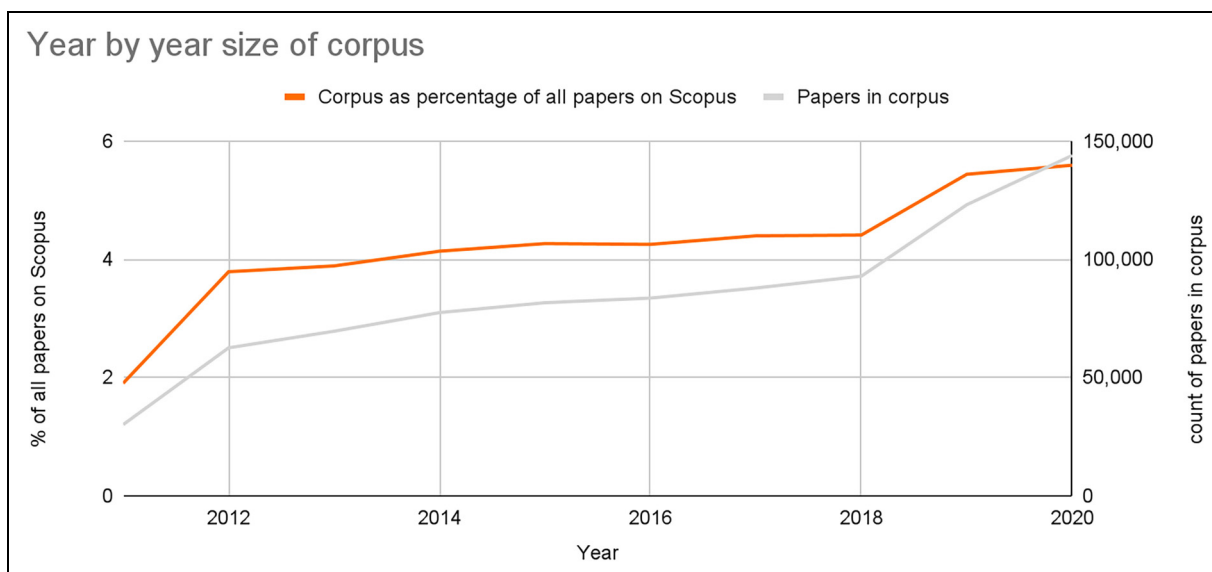


Figure 1. Annual growth in abstracts about algorithms, machine learning and artificial intelligence on Scopus.

the network is an n -gram (either a named entity, noun, author keyword or scientific discipline) and the edges represent co-occurrence of the n -grams in the same articles.

To filter the network, we ranked the edges by pointwise mutual information (PMI) (Bouma, 2009). Co-word networks are known to be noisy and produce hairballs (one-cluster structures). By filtering out the lower PMI edges we obtained a network of 7562 nodes, 85,215 edges and good topical clustering. We used thresholds ensuring that the network maintained a good coverage of the entire corpus, resulting in 86.56% of the abstracts having at least one n -gram featured in the final network.

To annotate the co-word clusters, we designed a qualitative process (see Figure 2) whereby we first computed all the k -cliques (Palla et al., 2005) where $k = 7$, then automatically constructed queries to retrieve a sample of abstracts for each clique, and subsequently read those abstracts to manually annotate the topical context. We had 166 cliques to annotate in total. The queries were designed so they returned papers that generated a triad in a clique. The triad would consist of the most central node in the clique, its closest neighbour in the clique (the node with the highest PMI) and one of the other nodes in the clique that were neighbouring the most central node. If multiple queries could be constructed with less than 10% node overlap for a clique, we allowed it. This resulted in a total of 235 queries for the 166 cliques.

Figure 3 shows a section of our qualitative annotation of the corpus. The bottom layer with cyan dots and labels visualises the co-word network after removal of the low PMI edges. The nodes were placed by a force-driven layout algorithm, designed to visually place groups of nodes (cyan dots) that are strongly interconnected closer to each other (Jacomy et al., 2014). The edges are not

shown to prevent visual clutter. Some of the original node labels, i.e., the nouns, named entities, author keywords or scientific fields mapped from the article abstracts, are also shown in cyan to provide a sense of the underlying data. Our qualitative annotation of the k -cliques is shown in red. Note that the compositing of the red layer makes the cyan layer look black when they overlap. If the layout algorithm has managed to make the clique discernible as a visual cluster, i.e., if the clique can be unambiguously located to a point on the map, the qualitative annotation is pinned to the clique as a *landmark* (visualised as a white dot with a label). If the clique has been stretched out because of the dimensionality reduction inherent in visual network layouts (Jacomy, 2021), the qualitative annotation is marked up as a *bridge* connecting other landmarks (visualised as a label along a line with arrows pointing away from the label on either side). Finally, assessing the annotated landmarks and bridges in each *area* of the map, we provide another layer of annotation of the overall theme in purple. In the following we will use the terms *landmarks* and *bridges* to refer to the annotated cliques and the term *areas* to refer to the overall annotated themes.

After qualitative annotation, the section of the network shown in Figure 3 turns out to be broadly in an area we have chosen to label “social science” with landmarks like “text classification”, “social network analysis”, or “socially responsible AI”. A smaller area is specifically related to economics with landmarks like “fintech”, “marketing” or “macroeconomics”. There are bridges connecting the landmarks, like “Facebook advertising” connecting “digital marketing” and “social network analysis”, or “language” connecting “optical character recognition” and “music genre recognition”. Two landmarks are annotated as artefacts

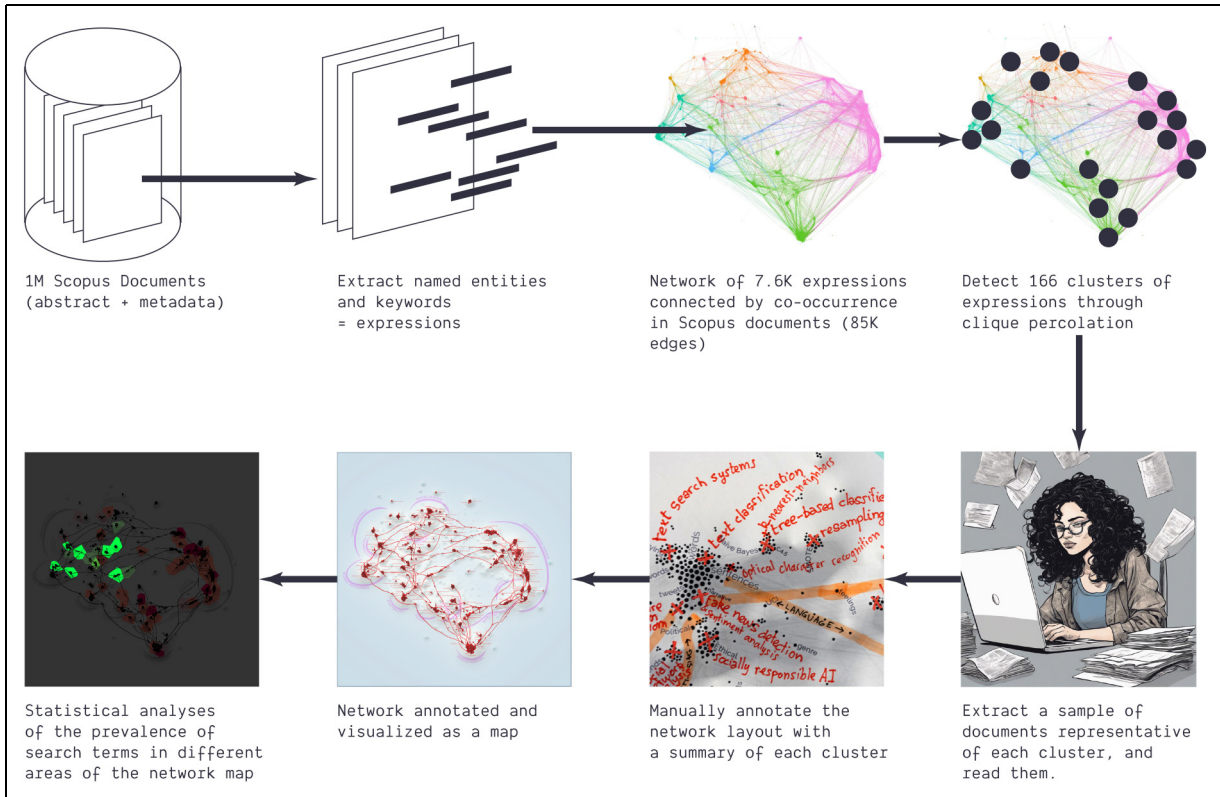


Figure 2. Our quali-quantitative process for semantic analysis and annotation of the corpus.

because they represent discourse specific to open access publications and publications from Springer respectively, i.e., an artefact of the dataset.

To explore the annotated corpus, which we now refer to as the base map because our queries can be located on it for analytical context, we built a datascape (last step in Figure 2). The datascape is a query system that allows us to search for subsets of abstracts containing specific search terms, published by specific authors, or in specific journals, fields or time periods, and visualise their representation in the different areas of the base map. Figure 4 shows the way we partitioned the basemap into cells derived from centroids in a Voronoi diagram designed to match the clusters of the map. We first applied a k-means on the network layout with $k=126$ so that the average cell would contain 60 nodes, a factor 2 above the commonly accepted size of 30 data points for simple statistical tests. To prevent the Voronoi cells from extending to empty areas, we added filler centroids with the same density using Poisson disc sampling.

For a given query, the datascape can now compute a score for each node based on how many documents mention the n-gram represented by that node. We normalise the score by the total number of documents where this expression is found to obtain a weight in the $[0,1]$ range. For each cell, we perform a statistical analysis of the node weights. If the cell contains less than 30 nodes, we do not

consider the test significant. We compute the average node weight in the cell, and we compute its normalised deviation from the expected average for a random distribution. We also perform a one-sample t -test to assess whether the deviation is statistically significant. We combine the p -value and the normalised mean deviation into a colour scale that displays whether the expressions of a cell are over- or underrepresented, but only if the test is statistically significant (Figure 5 Left). These colours have been chosen to provide context while highlighting (in bright green) the areas where expressions are overrepresented in the queried documents. Figure 5 (right) shows how the query “fake news” is overrepresented only in the social science area while it is underrepresented almost everywhere else.

The way we queried Scopus means that our corpus includes abstracts that mention *at least one* of three things: algorithms, machine learning or AI (in their title, author keywords or abstract). Accordingly, in the corpus we find abstracts that talk about algorithms or algorithmic phenomena in decidedly non-machine learning ways, for example analogue algorithms for diagnosing a patient. Whereas authors in the scientific literature might use general terms like machine learning or artificial intelligence when they write abstracts, they will often use an evolving variety of more precise technical terms instead. Before designing the query, we knew that when authors write about deep neural networks, large language models,

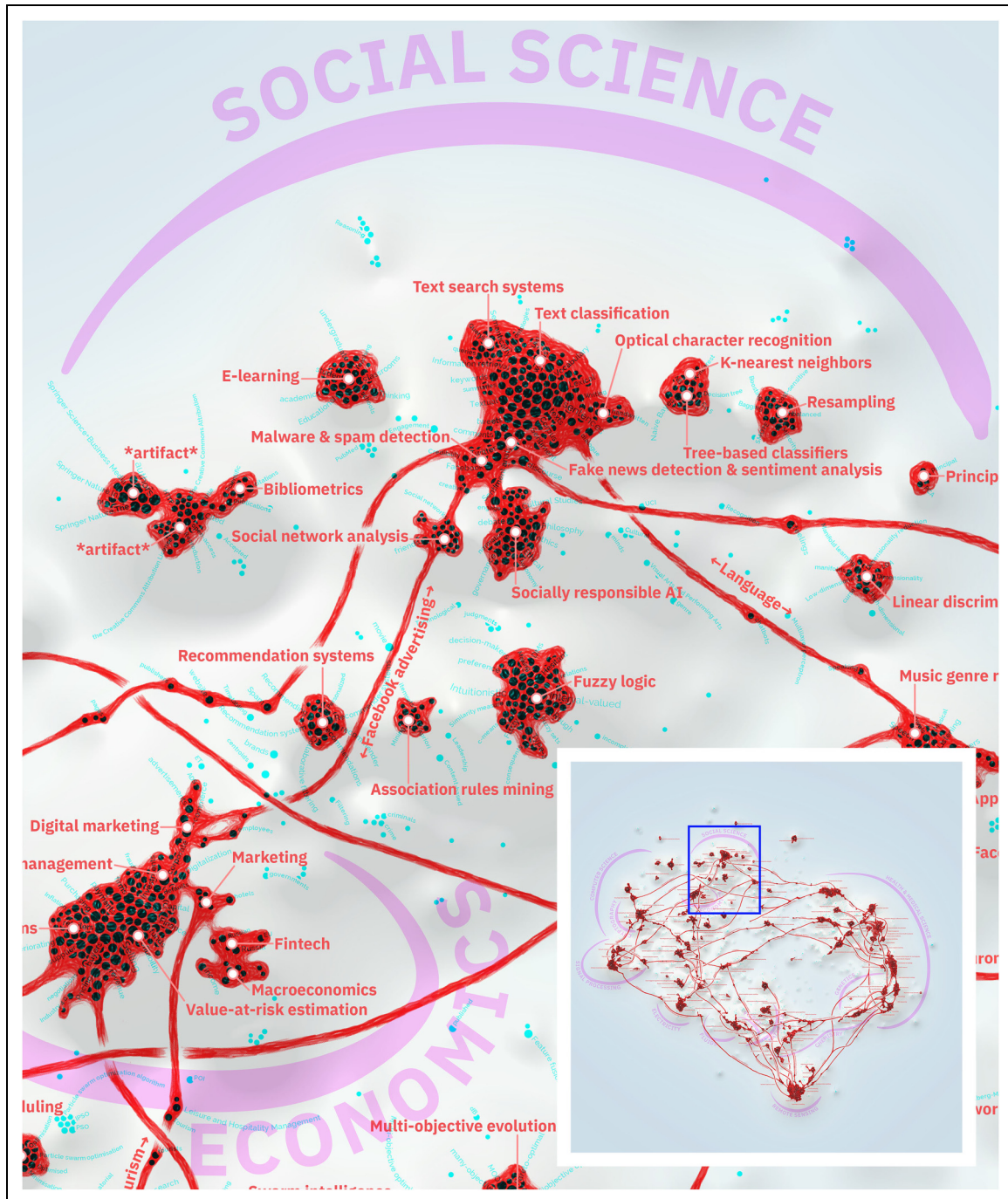


Figure 3. Section of qualitative annotations for the social science area of the semantic network. Underlying semantic network in cyan (and black when covered by red annotation layer); qualitative annotations of landmarks and bridges in red and of thematic areas in purple. The red annotations for bridges and landmarks are based on the network topology (k -cliques) while the purple annotations for areas are just a visual help and left intentionally vague. A full high-resolution version is available online.⁶

semi-supervised learning or natural language processing they are effectively writing about machine learning. We also knew, from our own fields, the names and acronyms of specific algorithmic techniques like GAN (Generative Adversarial Networks), NER (Named Entity Recognition), PCA (Principal Component Analysis) or KNN (K-Nearest

Neighbours), but it seemed prudent to assume that we did not have a full overview of what those algorithmic techniques would be (and even less what they would be called) across other scientific disciplines and contexts. A query composed of search terms specific to individual machine learning algorithms or algorithmic techniques would therefore likely

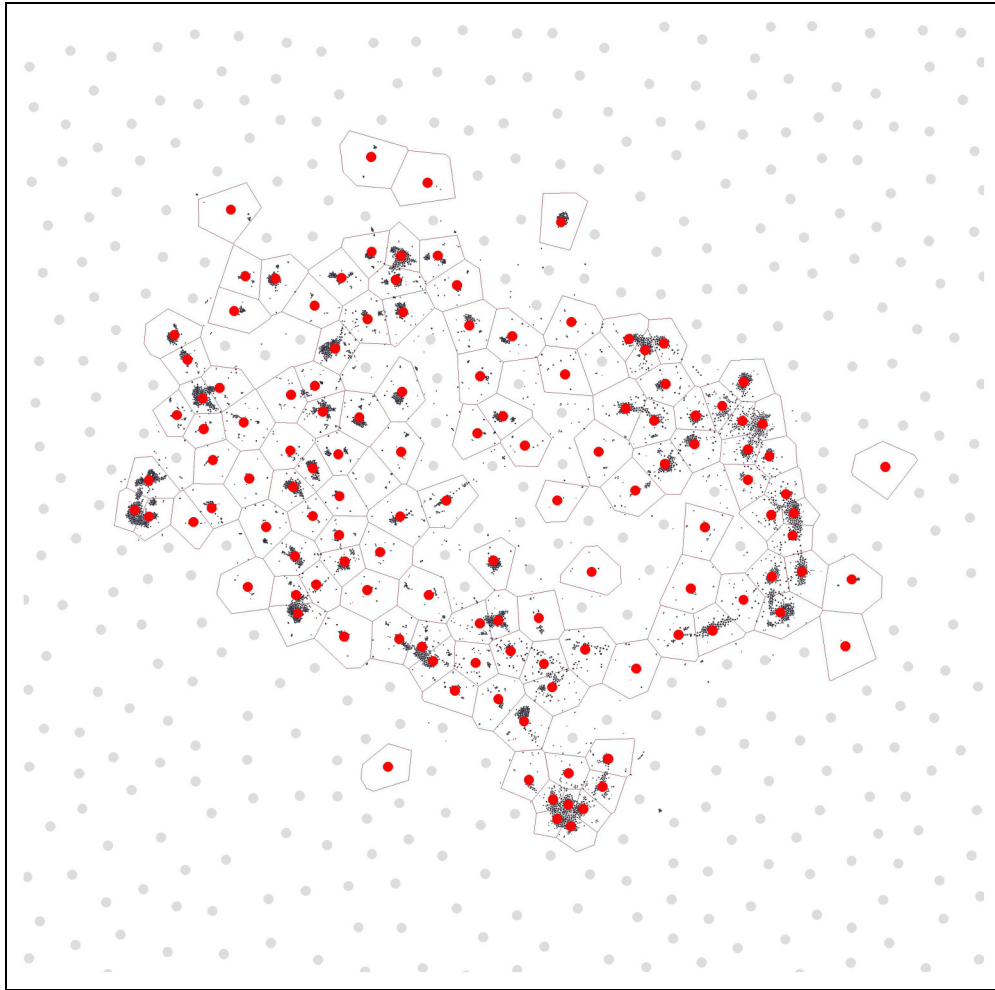


Figure 4. Centroids used to partition the basemap. In black, the nodes of the semantic network (essentially the same as the cyan layer in Figure 3). The red dots are the main centroids, computed with k-means to match the network clusters. The grey dots are filler centroids obtained via Poisson-disc sampling. The lines are the borders of the Voronoi cells around the main centroids. The cells around the red dots are used as a grid to partition the space. This strategy improves on a literal grid because the demarcations are mostly set between the clusters, preventing them from being split over multiple cells.

produce a corpus that was heavily biased towards the fields that we were already somewhat familiar with. The choice of *algorithm** as a search term is a deliberate way to be explorative; an attempt to catch all things algorithmic in the scientific literature without prejudicing the search towards specific types or fields of application.

As shown in Figure 6, the subsets of the corpus that mention machine learning or artificial intelligence are significantly smaller than the subset that mentions algorithms (a factor of 9 for machine learning and 27 for artificial intelligence). This does not mean that most of the corpus is not about machine learning or AI. Except for a few instances where the algorithms are analogue rules for making decisions, most of the papers do in fact deal with machine learning algorithms but reference them by their specific names rather than use the generic language of machine learning or artificial intelligence.

When we look at the annual proportion of each subset and compare it to the annual proportion of the total corpus (Figure 7) it becomes particularly clear that the generic language of artificial intelligence and machine learning is over-represented in 2019 and 2020. This could indicate that authors in the scientific literature perceive an increasing need to connect specific algorithmic techniques with larger phenomena or debates like AI or machine learning when publishing.

Despite the recent trend, algorithms dominate the topical structure of the corpus, as seen in Figure 8. By design of the visualisation, abstracts that mention algorithms cannot be over-represented in any part of the map because their baseline representation is already so high. Conversely, they could easily be underrepresented in some parts of the corpus, for example if a field did not use the term algorithm. That is not what we observe; algorithms are pervasive (Figure 8 Left).

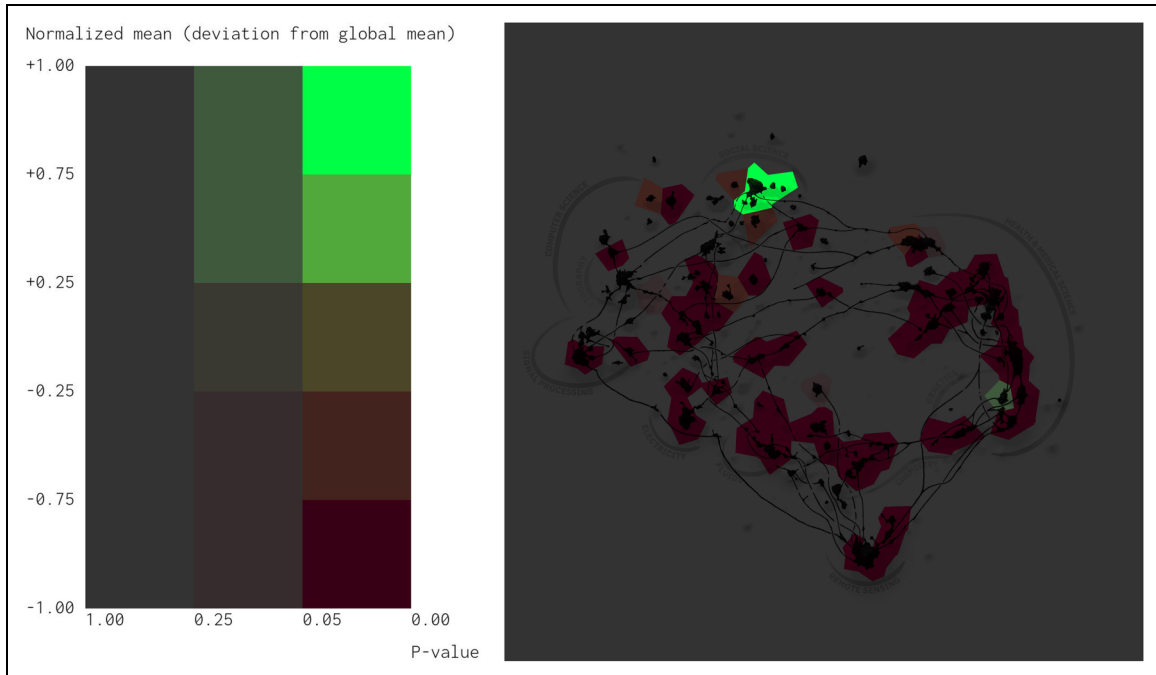


Figure 5. Left: colour scale for the datascape. If the statistical test is significant (p -value < 0.05) the cell is displayed in green (the documents retrieved by the query are overrepresented) or dark red (underrepresented). If the test is not significant but the p -value is below 0.25, a muted green or red shade is displayed as a visual cue. Otherwise, the cell is in grey. Right: The colour scale applied to the cells for a query about “fake news”. The code used to generate the Voronoi cells, perform the statistical analysis, and produce the images is available online as a Jupyter notebook.⁷

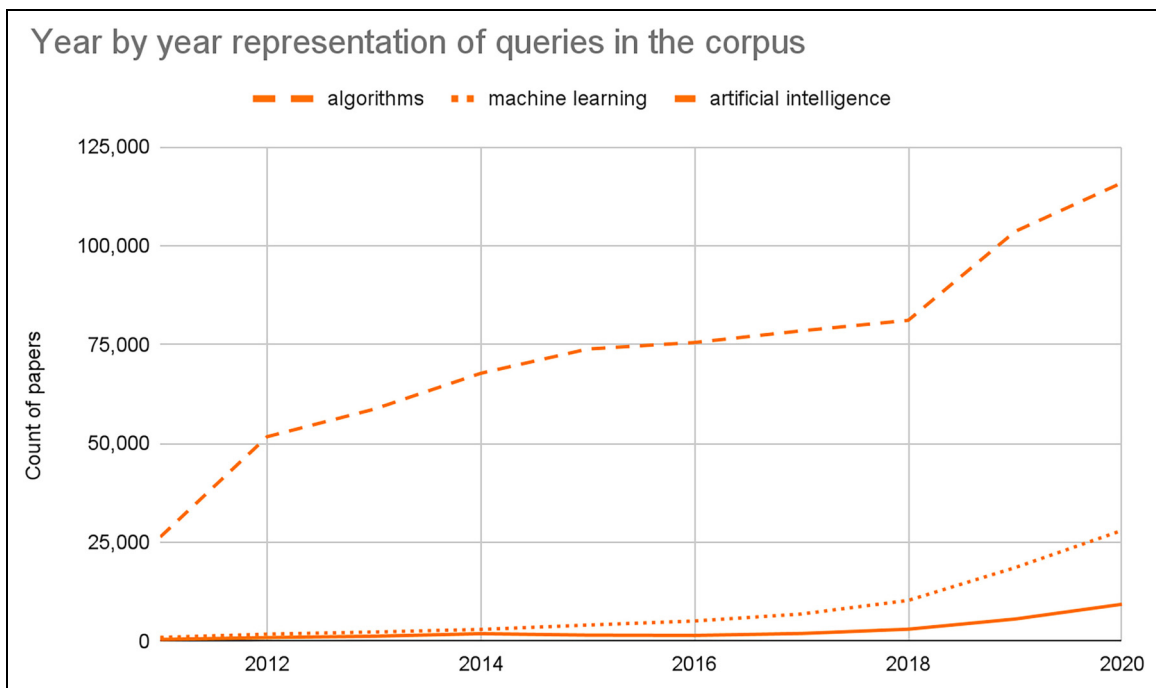


Figure 6. How are the queries about algorithms, machine learning and artificial intelligence represented in the corpus?.

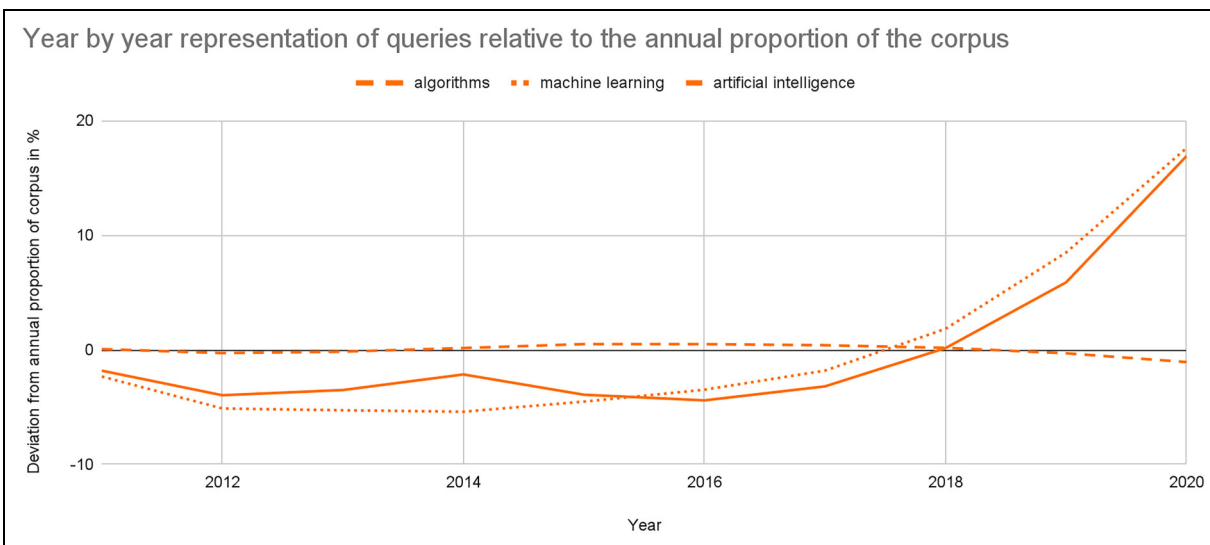


Figure 7. How are the queries about algorithms, machine learning and artificial intelligence represented relative to the annual growth of abstracts in the corpus? Measured percentual deviation in the annual proportion of the subset from the annual proportion of the total corpus.

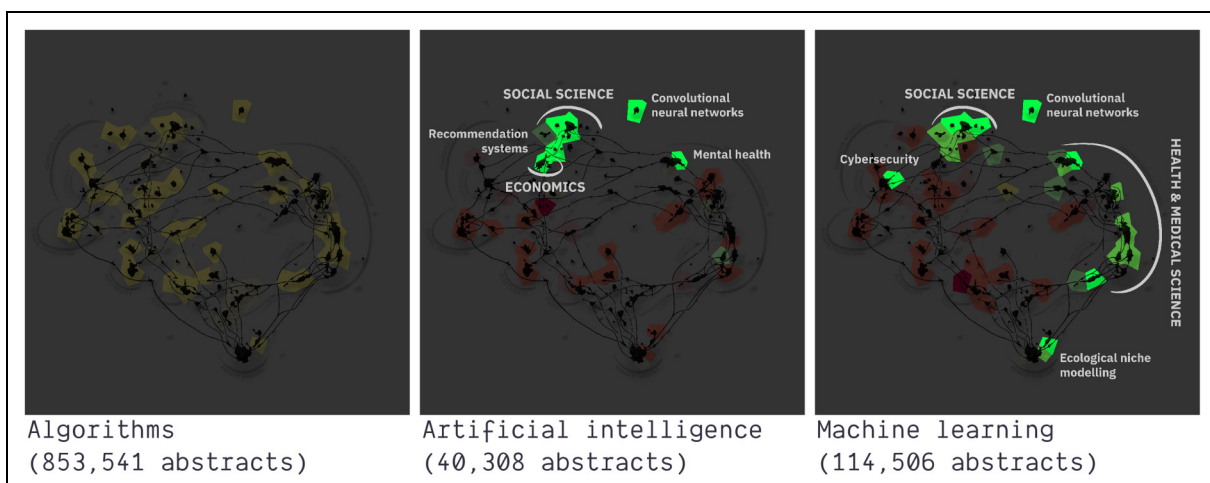


Figure 8. The three queries defining the corpus. The expressions overrepresented in the abstracts selected by the indicated query are highlighted in bright green. Left: the abstracts mentioning algorithms are neither over- nor underrepresented in any topic (dark yellow colour). Centre: abstracts mentioning AI are overrepresented in a few fields around social science and economics. Right: abstracts about machine learning are overrepresented in a croissant-shaped series of topics ranging from cybersecurity to health and medical science through social science.

The contrary is the case for AI, which is overrepresented in a few areas and around some isolated landmarks (Figure 8 Centre). First, AI is overrepresented in the social science and economics areas, where it is both used as a methodological term and studied as an object of critical analysis. Second, AI is mentioned around an isolated landmark about machine learning techniques such as convolutional neural networks and deep belief networks, which corresponds to the fact that such techniques have been particularly central in recent AI discourse. Third, the mental health landmark is also highlighted, which is mainly because its vocabulary on artificial intelligence overlaps.

Remarkably, AI is underrepresented in most of the other areas. Compared to mainstream media coverage, where it features prominently, AI is quite a niche term in the scientific literature. The same is true for machine learning, although not to the same extent (Figure 8 Right).

How do algorithms have agency in the scientific literature?

We can thus read the map as an answer to our question about what algorithms are doing in the scientific literature

and as highlighted in Figure 9, that answer comes with 12 broad thematic flavours signified by the annotated *areas* of the map. The overall finding from our qualitative annotation of the 235 underlying bridge and landmark queries, on which the thematic area annotations are based, is that algorithms are overwhelmingly applied in efforts to solve problems that have nothing to do with AI to begin with. In health and medical science they diagnose diseases and optimise treatments; in genetics they sequence genomes and biomes; in chemistry they model molecular dynamics; in remote sensing they detect earthquakes, predict weather patterns, or remove cloud cover on satellite imagery; in materials they improve tensile strength and model new composites; in flight they improve aerodynamics and fuel efficiency, or detect fake radar warnings; in fluids they solve differential equations; in electricity they distribute power from fluctuating sustainable energy sources like solar or wind to the grid; in signal processing they manage wireless networks or filter acoustic noise; in cryptography they detect phishing and generate watermarks; and in economics they schedule production flows and price securities. These are just examples, but they are indicative of the overall trend. The dominant way in which algorithms acquire agency through scientific discourse is as a solution to non-algorithmic problems across a wide range of application domains. The only real exceptions are computer science, where algorithms are developed, and social science, where we have annotated a landmark on socially responsible AI. In both cases, however, there are still plenty of examples where algorithms are applied to solve problems. In computer science they enable parallelisation and constitute key tools in cybersecurity, for example, and in social science they are used in contexts like bibliometrics, text classification, e-learning, spam detection, sentiment analysis and recommender systems.

If there is an overall topical division in the network it is perhaps between fundamental and applied science, although the different fields of application where AI solves different problems are, as already stated, the dominant structure. Figure 10 thus shows how the *making* of algorithms is overrepresented on the left side of the map in the computer science area (with landmarks such as “multi-objective evolutionary algorithms”, “nature-inspired meta-heuristic algorithms” or “indoors localisation systems”), while the *use* of algorithms is overrepresented on the right side in the health and medical science and remote sensing areas (notably around the “DNA sequencing for oncology” and “hydrological modeling” landmarks).

However, the fact that the semantic structure of the map mainly reflects different fields of application, with some division between more applied and more fundamental research on algorithms, does not mean that AI issues are completely absent. It just means that they are not prominent enough to assert themselves in the semantic structure. As we show in Figure 11, where we zoom in on the parts of

the corpus where abstracts mentioning ethics are overrepresented, there are indeed landmarks like “social network analysis”, “malware & spam detection”, “socially responsible AI” or “fake news detection & sentiment analysis” where algorithms are clearly an issue. Yet, ethics is underrepresented in the quasi-totality of the rest of the map and even when abstracts mention ethics, it does not necessarily mean that an algorithm is the issue. Some abstracts are about using algorithms to improve or assess the ethics of something else.

Where are the controversies?

So, although we do not retrieve known artificial intelligence controversies as issue language directly from the annotated semantic network, we can still query for them directly through the datascape. We therefore constructed 18 queries based on reported issues from the AIAAIC (see also Appendix A). As shown in Figure 12, the results of these queries vary from issues like appropriateness or dual-use, that have no statistically significant distribution in the corpus, over surveillance or gender and racial discrimination, that we find very selectively represented around landmarks in health science, to issues like safety or privacy, that we find overrepresented in multiple areas.

What we find across all 18 queries, however, independently of their different distributions in the corpus, is that while they do make it possible to retrieve abstracts where AI technology is the source of controversy, they make it equally possible, and often easier, to retrieve abstracts where AI technology is framed as a solution to the issue in question. A case in point is the issue of appropriateness. The incidents recorded by the AIAAIC on appropriateness typically begin with someone questioning whether there is a need for the use of AI in a concrete situation. This was for example the case with Nvidia’s AI-enabled eye tracker in early 2023.¹ While some abstracts in our corpus do engage in this way with the appropriateness of using AI for a given task, such as for iris recognition (Aginako et al., 2017) or prediction of cardiovascular disease from smart watch data (Kim, 2021), they tend to define appropriateness very differently, namely as high prediction accuracy, i.e., whether a method yields reliable results and not as compliance with principles for socially responsible AI. Others, for example in the context of mental health detection (Abd Rahman et al., 2020), define appropriateness as the choice of methods that are up to the job in the sense that they validly measure what they are supposed to measure. Again, nothing directly to do with social responsibility. There are some exceptions to this, such as discussions on how to use machine learning to restore gait functionality for patients with powered prostheses in a responsible way (Lemoyne and Mastroianni, 2019), but many abstracts deal with appropriateness as a complete non-AI issue and mobilise machine learning simply as a tool to assess whether some procedure

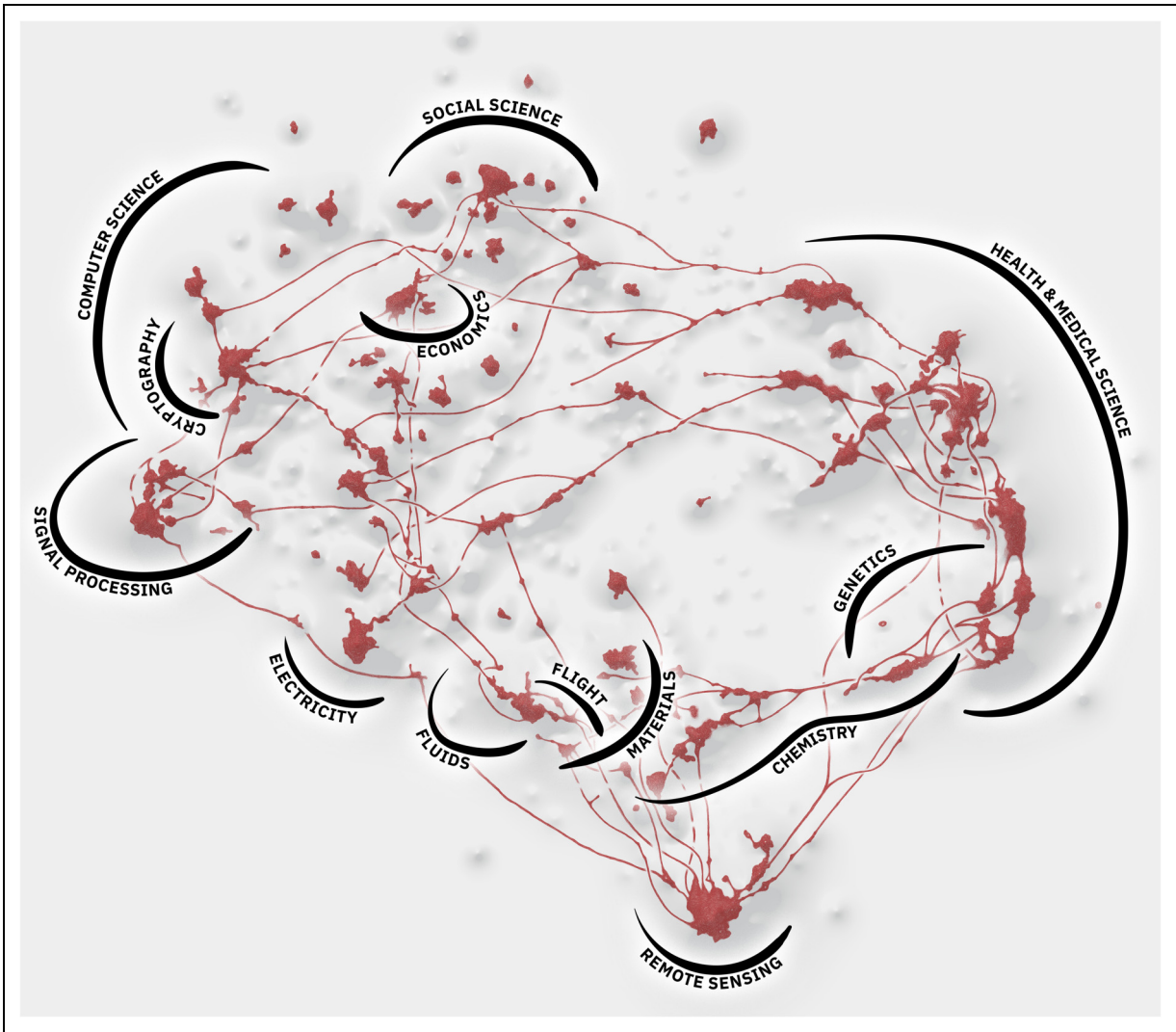


Figure 9. The main annotated areas of the map are highlighted in black. They broadly correspond to scientific fields. We have intentionally kept the borders of the areas vague, allowing them to overlap and transition into one another where the landmarks and bridges dictate it. It is also worth noting that the areas do not entirely cover the map, as some landmarks do not fit in this thematic breakdown. For instance, the “social science” cluster (on top, see also Figure 2) loosely encompasses clusters about sociological subfields (“social network analysis”), techniques used in sociology but not only (“principal component analysis”), sociology of algorithms (“socially responsible AI”), adjacent fields (“text search systems”) and some more distant fields (“fuzzy logic”).

has been appropriate or not. The issue could be the appropriate use of blood transfusions, in which case machine learning simply provides a better method for assessing whether doctors are indeed administering blood transfusions appropriately (Yao et al., 2019).

Some of the issues are hard to query precisely for. This is true for racial and gender bias and discrimination, where the queries return abstracts on statistical discrimination and bias. It is also true for employment, because it can refer to a demographic variable or, in some fields, simply the fact that a technique is used (employed). Yet, bias and discrimination is one of the issues where algorithms have clearest agency as a problem, for example in relation to

racial bias in predictive policing (e.g., Brantingham et al., 2018) or gender bias in machine translation (e.g., Prates et al., 2020). For most other queries, the capacity of algorithms to solve problems tend to overshadow their capacity to cause them. They cause problems with freedom of expression when algorithmic content moderation leads to shadow banning (e.g., Elkin-Koren, 2020), but results for that query mainly focus on algorithms being useful when detecting speech that is sexist or hateful (e.g., Mozafari et al., 2020; Rodríguez-Sánchez et al., 2020). Algorithms are also staged as a solution to misinformation, for example to detect bots (e.g., Ferrara, 2017) or by having recommender systems on social media prioritise trusted

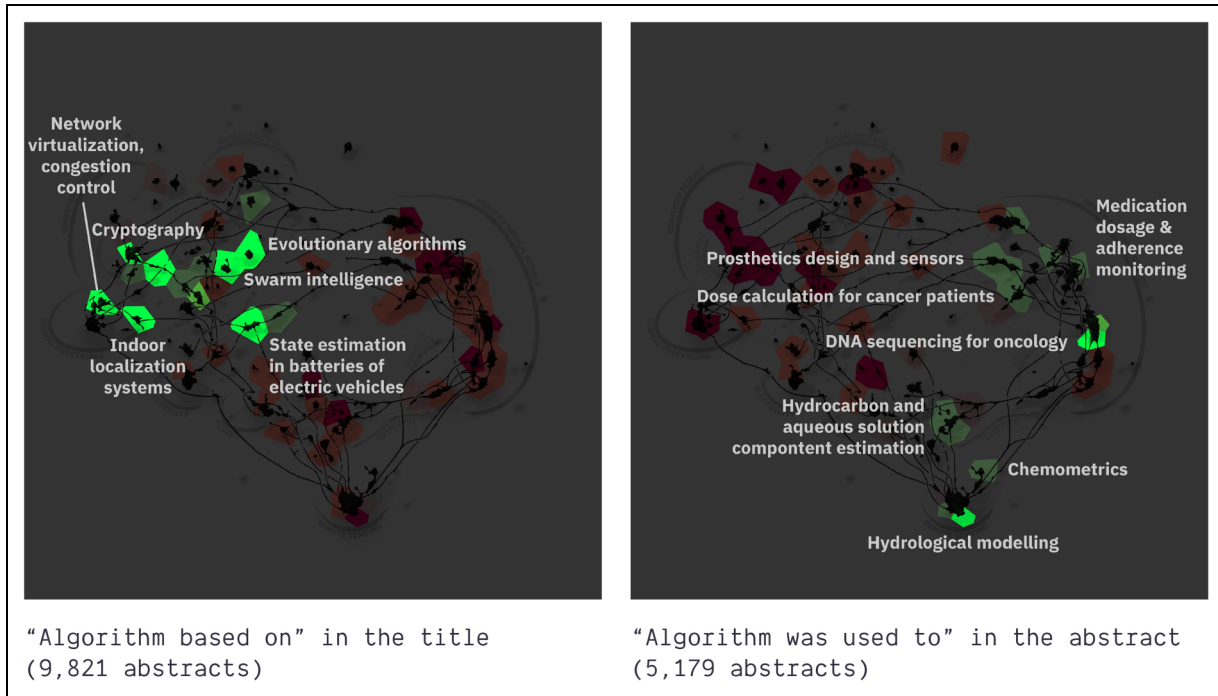


Figure 10. The expressions overrepresented in the abstracts selected by the indicated query are highlighted in bright green. The making of algorithms is more represented on the left side of the map, while the use of algorithms is more represented on the right side.

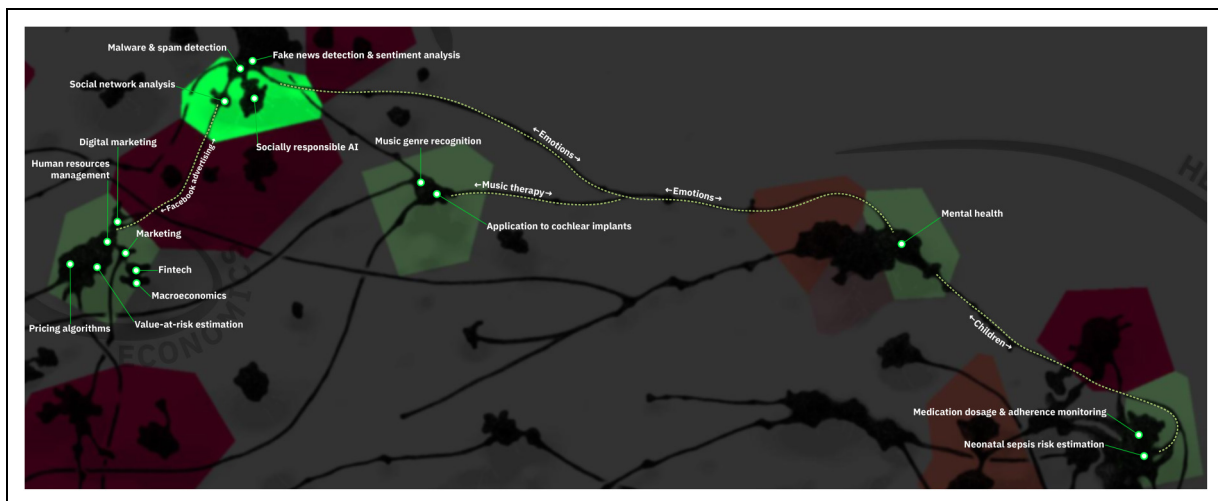


Figure 11. The landmarks where abstracts mentioning ethics are represented in a statistically significant (bright green) and non-significant (muted green) way. Relevant annotations for represented landmarks are highlighted in white.

sources (e.g., Pennycook and Rand, 2019). Fairness is a problem in many fields of science, for example in relation to signal distribution in electrical grids or wireless networks, and again algorithms are staged as a way to improve fairness. AI technology can of course be used for surveillance, but if it is surveilling things like cancer or hepatitis, the context is less problematic. Similarly for

privacy, where federated learning and differential privacy techniques are enabled by innovations in algorithm design, deep learning is used to improve cybersecurity, perform crop surveillance or solve safety issues in self-driving cars, and even when abstracts talk about dual-use, algorithms are not necessarily the problem. On the contrary, they could be instrumental in monitoring a potential

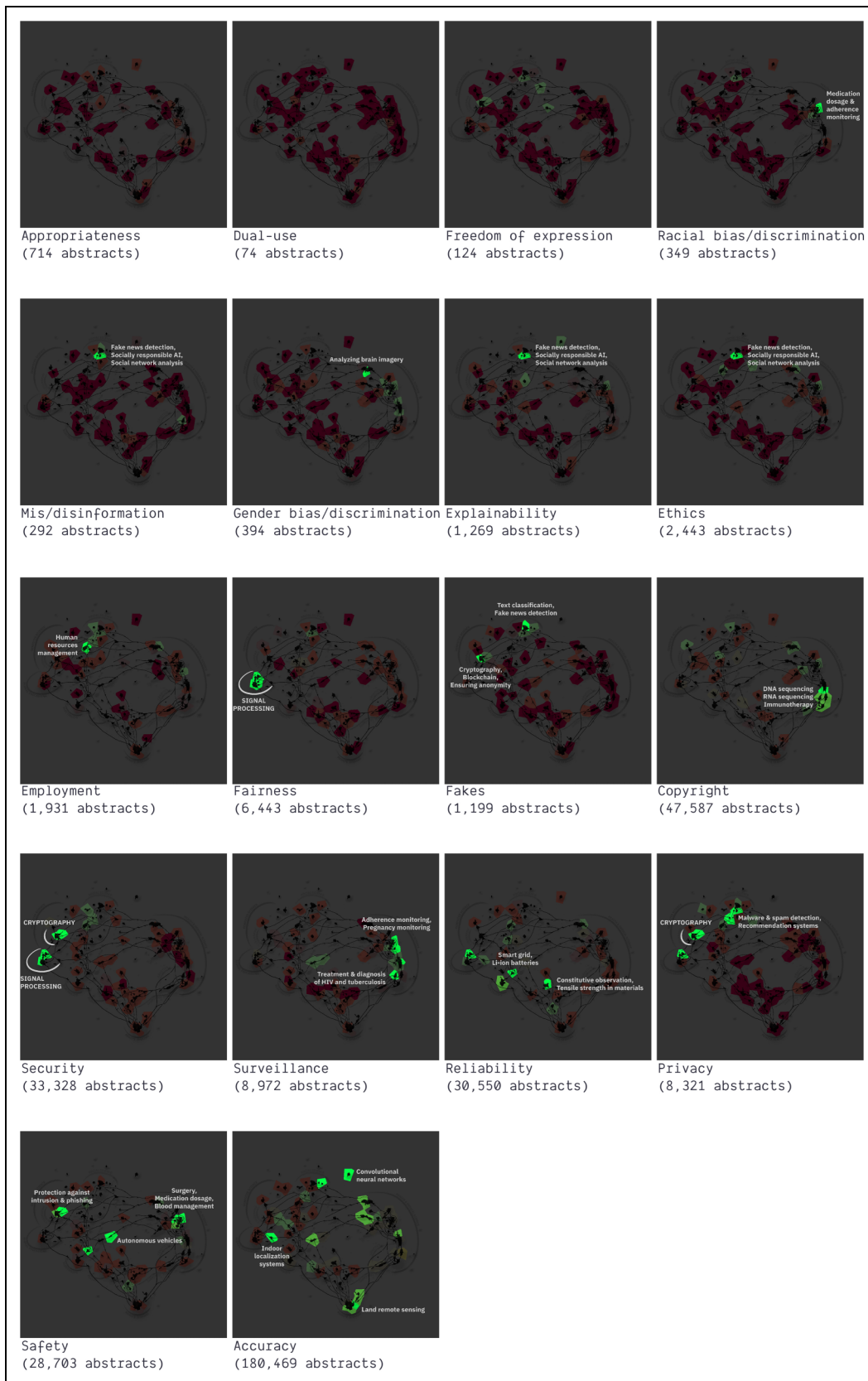


Figure 12. Eighteen subsets of the corpus containing controversial query terms and their distributions across the semantic network. Each query is derived from the incidents reported in the AIAIC database. Each visualization shows you where abstracts corresponding to the query in question are typically found in the basemap. Areas highlighted in bright green contain expressions that are statistically overrepresented in the abstracts corresponding to that query. All the visualisations are produced with the datascape from the last step in Figure 2. Full colour scale in Figure 5, full queries in Appendix 1.

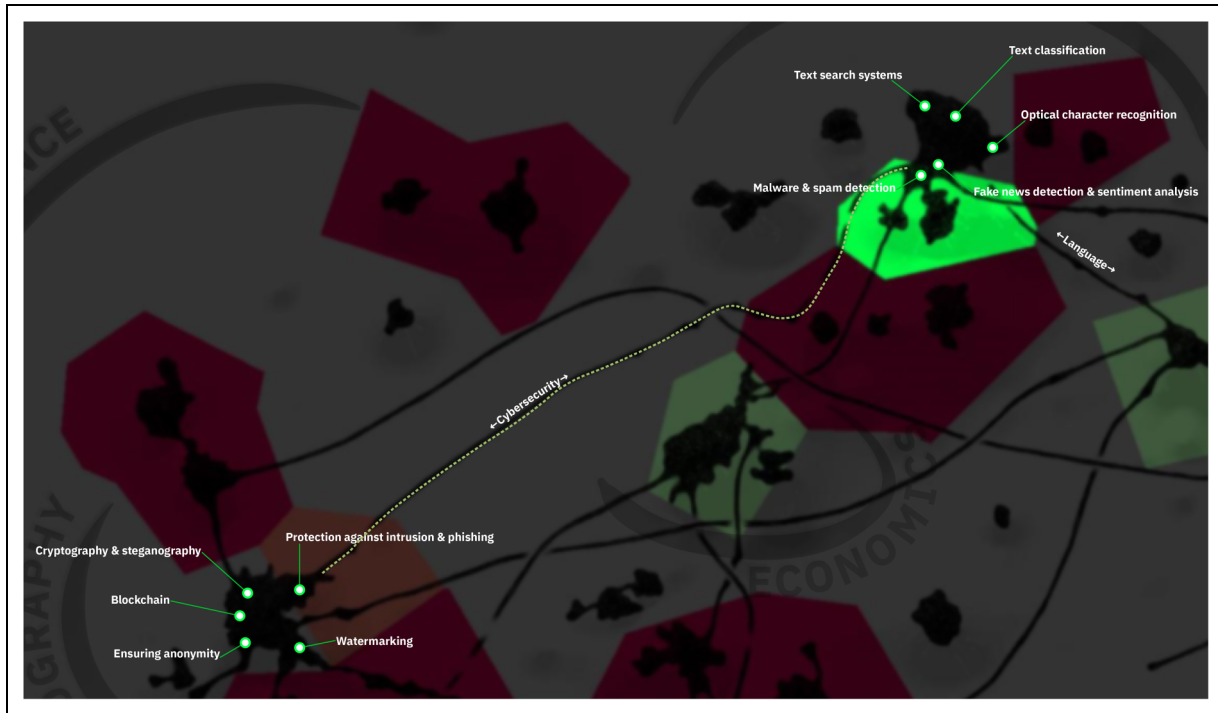


Figure 13. The landmarks strongly represented in abstracts about fakes (bright green) with relevant annotations. We find them in two distinct areas that we have originally annotated as cryptography (bottom-left) and social science (top-right). Seen in isolation, the extent to which landmarks like malware, spam and fake news detection, text search and optical character recognition are indicative of the social sciences is of course debatable. Note that the bridge “cybersecurity” connects the two areas.

Table 1. Breakdown of the number of abstracts about fakes by mention of fake news and by mention of detection.

Abstracts about fakes	About detection		Not about detection		Total	
About fake news	165	14%	91	8%	256	21%
Not about fake news	511	43%	432	36%	943	79%
Total	676	56%	523	44%	1199	100%

dual-use situation, such as transition from tobacco to cannabis use (e.g., Hu et al., 2021), or for managing risk in the development of products intended for dual-use (e.g., Zubova et al., 2021).

The fact that AI technologies are rarely staged as controversial in the scientific literature becomes particularly evident when we explore the issue of fakes. The term “fake” appears in the title of 72 (7%) of the 1045 incidents reported by the AIAAIC (accessed 2023-06-23). In some of these cases, AI stirs controversy when trending and recommendation algorithms on social media or search engines contribute to spreading conspiracies and disinformation or provide visibility for fake profiles. This was for example the case in the 2017 Las Vegas mass shooting incident where YouTube’s recommender algorithms kept promoting fake claims that the victims were actors in a staged plot (Levin, 2017). Another way in which AI has agency as a

problem is when generative models are used to produce fake content. There are controversial examples of deep fakes being used in political campaigns as early as 2018 in the AIAAIC database, such as when the socialist party in Belgium circulated a deep fake video of Donald Trump calling for the country to follow the U.S. out of the Paris climate accord² or when a female investigative journalist in India was harassed by Hindu nationalists who circulated a deep fake porn video of her.³ The entry *Jordan Peterson fake voice generator*⁴ reports that “the controversial psychologist” Jordan Peterson threatened legal action against “NotJordanPeterson.com, a website for generating AI audio clips of [him] saying whatever you want”, causing it to shut down. In other entries, such as *Audio deepfake fraudulently impersonates CEO*,⁵ the issue is the observed use of generative AI by fraudsters. In fact, in 58 of those 72 AIAAIC entries about fakes, the issue is AI-generated content.

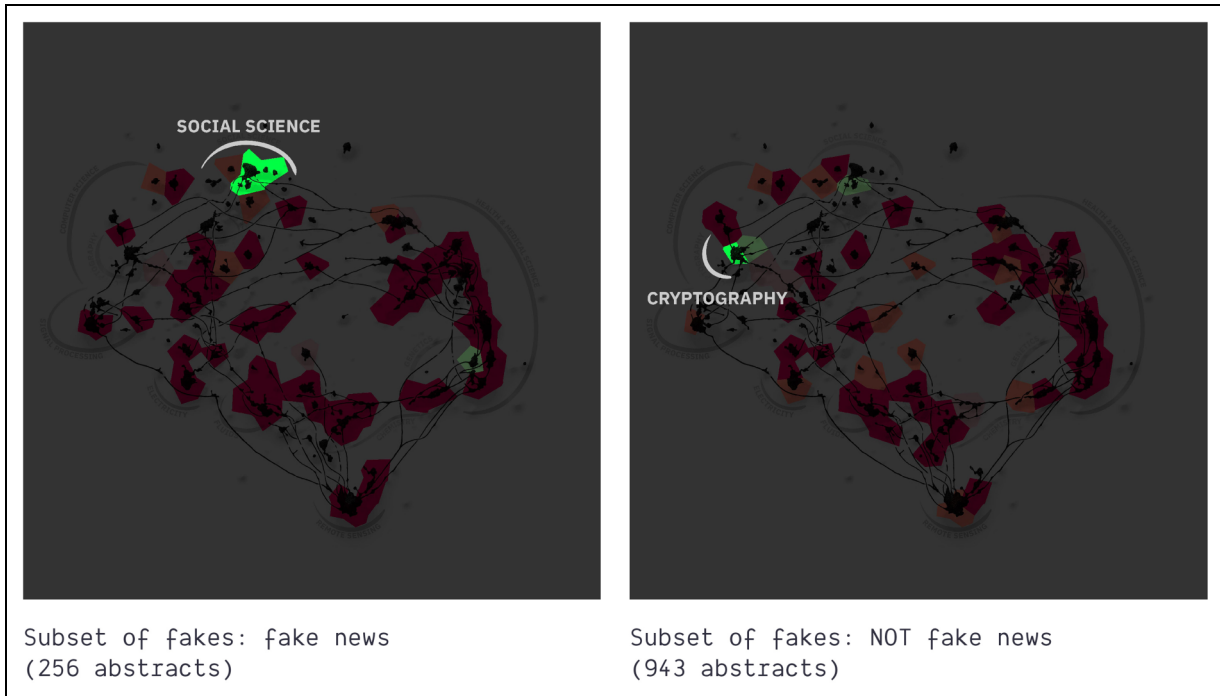


Figure 14. Two subsets of abstracts about fakes. About fake news (left) and not about it (right). The areas in green, representing topics overrepresented in the abstracts retrieved by each query, are disjoint. This suggests that the abstracts mentioning fake news are mostly in social science while the others are mostly in cryptography and cybersecurity. Full queries in Appendix 1.

It is to some extent possible to retrieve generative AI as the root cause of controversies about fakes in the scientific literature as well. We find abstracts that discuss the threat of deep fake evidence in courts (Maras and Alexandrou, 2019), the use of deep fake porn to silence political rivals (Maddocks, 2020) or deepfake generated disinformation and political deception campaigns (Vaccari and Chadwick, 2020). If we zoom in on the areas of the map where fakes are overrepresented (Figure 13), we find two distinct sets of landmarks: one in social science, with fake news detection and sentiment analysis, and one in cryptography. The former reflects the fact that AI has multiple forms of agency in relation to the issue of fake news. Notably, it far from always begins with AI-generated fakes and although the problem of fake news might be exacerbated by algorithms it might also be detected algorithmically. Besides fake news detection (Gravanis et al., 2019), machine learning is used to detect fake reviews (Yu et al., 2019; Mohawesh et al., 2021), fake profiles on Facebook (Gupta and Sahoo, 2020), fake Twitter followers (Cresci et al., 2015), etc., indicating once again that even within controversial topics where AI can certainly be a problem, it can also be the opposite.

Half of the abstracts about fakes are indeed about detection (56%, see Table 1) and 14% are about fake news detection specifically. This leaves 43% for the detection of other kinds of fakes. Fake news is an issue term in our corpus from around 2016, which resonates with the fact that it became a matter of public concern in the wake of the Brexit

vote and the 2016 American presidential election, and deep fakes only took off from 2019 onwards. The fact that fakes are mentioned in the corpus already from 2011 not only shows that fakes were a problem before AI was able to produce them but that algorithms were mobilised as a remedy to fakes before they became framed as culprits in specific versions of that issue as well. We thus see machine learning algorithms used to detect fakes in completely non-digital contexts like fake green tea (Li et al., 2015) fake ceramics (Praisler et al., 2013) or fake banknotes (Upadhyaya et al., 2018).

As shown in Figure 13, cybersecurity is the bridge that links abstracts about fakes in the social sciences and in cryptography. This partly reflects the fact that we see AI-based detection of fakes used to combat cyberthreats like manufactured biometric identifiers such as fake fingerprints or iris scans (Galbally et al., 2013). In some abstracts, fakeness refers to intentional deception that AI is tasked with unmasking as well as digital or fake agents in a cybersecurity context (Lavanya and Bhaskaran, 2015). Indeed, several abstracts show how these deepfakes can then be detected by the same technologies that produced them, such as generative adversarial networks (Guarnera et al., 2020). Those abstracts are always mainly about the AI technique used to perform the authenticity assessment. In that sense, authenticity is the main issue. However, in non-cybersecurity or misinformation contexts, fakes have nothing to do with deliberate deception but refers to some form of noise that AI can detect and filter out to improve a

process. For instance, “fake peaks from the acquired ... signal” to measure heart beats (Tran and Chung, 2015), or “fake defects” in display panels (Jeong et al., 2021). The paper is then focused on the improvement of that process (measuring heart beats or inspecting panels).

(Figure 14).

As shown in Table 1, 36% of the abstracts about fakes are neither about fake news nor detection. This is because AI-generated fakes can be used to solve problems in contexts like defensive cyber operations where troves of fake documents that are hard to comprehend can be artificially generated and used to inundate an adversary (Karuna et al., 2018); “fake-but-believable-versions of documents [that] can delay the attacker” (Han et al., 2021). It can also be used to generate fake identities to detect cheaters (Liu et al., 2019) or “fake contexts” as “a deception policy for privacy preservation” in smartphones (Zhang et al., 2016). In these ways, AI-generated fakes are staged as the solution to an ever-wider array of problems. Beyond cybersecurity, the most common scenario for that is the use of generative AI to expand datasets. The final goal may be to detect heart disease (Rath et al., 2021), fight malware (Mimura, 2020) or detect exoplanets (Flasseur et al., 2020) but the main issue is generally unrelated to AI, except that it demands data that is too scant, and that generative AI can make the data less scant.

From AI as the issue to AI as the political situation

We began this paper with two observations, namely that AI controversies have proliferated in the past decade and that these controversies are poorly reflected in the way mainstream media cover AI. As discussed, a mounting body of research points to the fact that mainstream media do not sufficiently stage AI as controversial and do not sufficiently situate the role of AI in relation to specific problems. Thus, to the extent that AI is indeed staged as controversial, it is typically in a sensationalist fashion where existential risks take precedence over more concrete problems. This has led to calls for more expert-informed media coverage.

While our analysis of what algorithms are doing in the scientific literature supports the notion that experts are more attentive to the specific ways in which AI technologies relate to different problems, it does not support the notion that experts are *per se* more focussed on the controversiality of such technologies. On the contrary, we find a clear preference for AI-driven solutions over AI-caused problems across the scientific literature. In that sense, it more than matches what others have found in mainstream media coverage. The difference is in the attention to situated problems, where the scientific literature is generally free of abstract and sensationalist narratives about existential threats or hyped technological utopias, and almost universally focussed on how algorithms matter in concrete circumstances. We

must therefore conclude that the problem of mainstream media coverage not staging AI as sufficiently controversial is not isolated to mainstream media but clearly present in the way scientific publications contribute to the issuefication of AI. Or rather, if it is a problem, then it is also a problem in the way scientists write about algorithms. Another possible conclusion is of course that AI technologies, broadly speaking, and in all their various applications, are considerably less controversial than the incidents reported by initiatives like the AIAAIC would suggest. To us, this calls for further research on the relationship between public AI debates, media coverage and the ways in which experts contribute to the issuefication of algorithms.

Drawing on Andrew Barry’s notions of political situations (Barry, 2012) and technological zones (Barry, 2006) we suggest that there might be a certain artificiality to current artificial intelligence controversies which has to do with the way they centre our attention around specific events and thus artificially limit our view of a much broader, interconnected landscape of non-AI problems and conflicts where algorithms still have agency. Barry specifically develops the notions of political situations and technological zones to correct a shortcoming in the analysis of knowledge controversies, namely the tendency to focus tightly on single case studies and thus miss the way in which transnational problems like climate change, in Barry’s example, bring a range of different controversies into relation both with one another and their various hinterlands. As he puts it: “the question of the relations between apparently distinct knowledge controversies has not been systematically addressed” and there is therefore a “lack of attention paid in existing studies of controversies to the historical antecedents of those controversies or to their association with a history of ongoing negotiations or conflicts of which they are taken to be a part” (Barry, 2012:229). Barry thus points to a methodological bias in controversy analysis where the benefits of following the connections laid bare by the contestation of specific knowledge claims outweighs the drawbacks of letting the actors engaged in this contestation unfold and thus delimit the problem. Perhaps this should also be a methodological note of caution when mapping AI controversies. A shift from understanding AI as the issue to understanding it as the political situation brings into potential view the situations where machine learning algorithms have agency in controversies where actors have assembled around issues that have nothing to do with AI in the first place. As a technological zone, then, AI weaves together controversies both where AI is the issue and where it is not. Barry saw the need for such a concept in relation to transnational issues like climate change where regulatory efforts cut across and thus come into tension with existing controversies that were not climate related in their origin.

Thinking about AI as a broader political situation, we suggest, can be seen as a contribution to Kitchin’s (2017)

strategies for unpacking the full socio-technical assemblage of algorithms and examining how they do work in the world. Many scholars, Kitchin included, have called for better algorithmic literacy to further public engagement with AI. On a basic level, this could mean acknowledging the presence of an algorithm in a system and building hypotheses about what its function might be (Rieder, 2017), whereas on a higher level it could mean the ability to assess and interact with algorithmic systems through informed decisions (Klug et al., 2023; Cotter, 2019). We believe that a reorientation from AI as an issue to AI as a political situation can help further this kind of literacy by bringing into view the perhaps not so obvious, in the sense of less controversial, ways in which algorithms have agency in relation to public problems. In the past decades, many authors have attempted to broaden and balance our views of AI by calling attention to its widespread controversiality (see for example Crawford, 2016). As a supplement to this wave of scholarship, we suggest that it is now time to *also* attend to the many instances and practices where AI is rather *uncontroversial*, yet important to society and our lives. As such, you might say that ours is a methodology for tracing the less controversial hinterlands behind the hotter public situations.

Limitations and further work

Throughout this paper we have used the terms algorithms, machine learning and artificial intelligence interchangeably although, as shown in Figure 8, there are significant differences in how they are represented in our corpus. It is important to acknowledge that the discussion on whether machine learning deserves to be called artificial intelligence and, if so, in what contexts is still a controversy in its own right. However, taking our cue from the way AI is increasingly used in popular discourse to denote any kind of incident involving machine learning algorithms or models, as witnessed, for example, in the way it is employed by the AIAAIC, we have taken the stance that talk of specific algorithms in the scientific literature is likely similar to what is popularly associated with AI. This involves the potential limitation that the search term *algorithm** does not in fact yield scientific abstracts about machine learning. Our analysis of the corpus, however, shows that, with very few exceptions, this is not a problem.

A more likely limitation, and one that we cannot discern the severity of directly from the corpus, is that we do not find abstracts that use or discuss specific machine learning algorithms without mentioning words like machine learning or algorithms. Including search terms like *convolutional neural networks* or *generative pre-trained transformers* would thus likely allow us to expand the corpus and could be done as a part of an iterative strategy where the corpus is used to source more search terms. As we have already discussed, we deemed it too risky to include

specialised search terms in our initial construction of the corpus, as it would likely be too heavily biased towards our personal knowledge of the field. Future work could explore the possibility of an iterative search strategy and thus draw on the findings from our analysis, potentially in combination with an exploration of other scientific databases like Dimensions that have better coverage of the social sciences and humanities.

Another limitation of our analysis is its obvious media bias. The fact that something is written about or not written about in the scientific literature can, of course, not be taken as an indication of its controversiality in society more generally nor as the final say on how experts issuefy AI. We have already discussed the significant differences between what we find in the scientific literature and what similar mappings of mainstream media discourse on AI have found. It is important to acknowledge that genre conventions for how scientific abstracts are written in most fields do not prescribe a discussion of all the possible controversies that a finding might come into tension with. This likely has consequences for the kinds of agency that algorithms acquire and do not acquire through scientific discourse. To take an example from our exploration of the issue of fakes, one paper used AI to detect fake smiles “using this information to improve recognition algorithms, and eventually to build computer systems which are more responsive to human emotions” (Hossain et al., 2018). This counts, in our analysis, as a way in which machine learning algorithms are used to solve a non-AI problem. We take it as an example of how an algorithm is charged with the capacity to solve a problem rather than cause one, which, strictly speaking, is how this piece of scientific literature issuefy AI. Clearly, though, in this case the resulting development of AI systems that are more responsive to human emotions may very well – highly likely, in fact – become an issue for other actors elsewhere. We have deliberately chosen to take an affirmative approach to this obvious media bias (Marres, 2015) and consider the authors in our corpus as actors with a specific capacity for bringing AI into our lives through the medium of peer-reviewed science. Future work could therefore fruitfully investigate how experts contribute to the issuefication of AI through other media, for example scientific blogs, policy briefs or in their contributions to panels and public debate.

Acknowledgements

We would like to thank the project team for helpful feedback and commentary on various stages of this paper. Special thanks to Snorre Ralund and Johan Irving Søltoft, who participated in the curation of datasets and methods development for this part of the project.




Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research presented in this article was supported by VELUX FOUNDATIONS, Algorithms, Data & Democracy Jubilee Grant (no. 342044).

ORCID iDs

Anders Kristian Munk  <https://orcid.org/0000-0002-5542-3065>
 Mathieu Jacomy  <https://orcid.org/0000-0002-6417-6895>
 Matilde Ficozzi  <https://orcid.org/0009-0001-8660-5734>

Supplemental material

Supplemental material for this article is available online.

Notes

1. Entry AIAAIC0929 -2023- <https://www.aiaaic.org/aiaaic-repository/ai-algorithmic-and-automation-incidents/nvidia-eye-contact> (last accessed January 2024).
2. Entry AIAAIC0391 -2018- <https://www.aiaaic.org/aiaaic-repository/ai-algorithmic-and-automation-incidents/deepfake-donald-trump-calls-for-climate-agreement-exit> (last accessed January 2024).
3. Entry AIAAIC0152 -2018- <https://www.aiaaic.org/aiaaic-repository/ai-algorithmic-and-automation-incidents/rana-ayyub-deepfake-porn-attack-doxing> (last accessed January 2024).
4. Entry AIAAIC0283 - 2019 - <https://www.aiaaic.org/aiaaic-repository/ai-algorithmic-and-automation-incidents/jordan-peterson-fake-voice-generator> (last accessed January 2024).
5. Entry AIAAIC0462 - 2020 - <https://www.aiaaic.org/aiaaic-repository/ai-algorithmic-and-automation-incidents/audio-deepfake-fraudulently-impersonates-ceo> (last accessed January 2024).
6. <https://github.com/jacomyma/bds-ai-issue-material/tree/main/images>.
7. <https://github.com/jacomyma/bds-ai-issue-material/tree/main/notebooks>.

References

- Abd Rahman R, Omar K, Mohd Noah SA, Mohd Danuri MSN and Al-Garadi M (2020) Application of Machine Learning Methods in Mental Health Detection: A Systematic Review. *IEEE Access* 8: 183952–183964. 10.1109/ACCESS.2020.3029154.
- Aginako N, Castrillón Santana M, Navarro JL, Martínez-Otaza J and Sierra B (2017) Periocular and iris local descriptors for identity verification in mobile applications. *Pattern Recognition Letters*. 91. 10.1016/j.patrec.2017.01.021.
- AIAAIC (n.d.) AIAAIC - AIAAIC Repository. Available at: www.aiaaic.org/aiaaic-repository (last accessed January 2024). Retrieve from Internet Archive: [<https://web.archive.org/web/20230925200329/https://www.aiaaic.org/aiaaic-repository>]
- Suchman L. (2023). The uncontroversial ‘thingness’ of AI. *Big Data & Society*, 10(2).
- Barry A (2006) Technological zones. *European journal of social theory* 9(2): 239–253.
- Barry A (2012) Political situations: Knowledge controversies in transnational governance. *Critical Policy Studies* 6(3): 324–336.
- Benbouzid B (2023) Fairness in machine learning from the perspective of sociology of statistics: How machine learning is becoming scientific by turning its back on metrological realism. In: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, Chicago, IL, USA, 12–15 June 2023, pp.35–43. New York: Association for Computing Machinery.
- Bender EM, Gebru T, McMillan-Major A and Shmitchell S (2021) On the dangers of stochastic parrots: Can language models be too big?. In: Proceedings of the 2021 ACM conference on fairness, accountability, and transparency, 3–10 March 2021, pp.610–623. New York: Association for Computing Machinery.
- Bouma G (2009) Normalized (pointwise) mutual information in collocation extraction. In: From Form to Meaning: Processing Texts Automatically: Proceedings of the Biennial GSCL Conference 2009, Potsdam, DE, 30 September 2009, pp. 31–40.
- Brantingham PJ, Valasik M and Mohler GO (2018) Does predictive policing lead to biased arrests? Results from a randomized controlled trial. *Statistics and public policy* 5(1): 1–6.
- Bunz M and Braghieri M (2022) The AI doctor will see you now: assessing the framing of AI in news coverage. *AI & Soc* 37: 9–22. <https://doi.org/10.1007/s00146-021-01145-9>.
- Callon M (1998) An essay on framing and overflowing: Economic externalities revisited by sociology. *The Sociological Review* 46(1): 244–269.
- Chuan CH, Tsai WH and Cho S (2019) Framing artificial intelligence in American newspapers. *AAAI workshop: AI, ethics, and society*. AAAI workshops. AAAI Press.
- Cotter K (2019) Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram. *New media & society* 21(4): 895–913.
- Crawford K (2016) Can an algorithm be agonistic? Ten scenes from life in calculated publics. *Science, Technology, & Human Values* 41(1): 77–92.
- Crépel M, Do S, Cointet JP, Cardon D and Bouachera Y (2021) Mapping AI issues in media through NLP methods. In: *CHR*, pp. 77–91.
- Cresci S, Di Pietro R, Petrocchi M, Spognardi A and Tesconi M (2015) Fame for sale: Efficient detection of fake Twitter followers. *Decision Support Systems* 80. 10.1016/j.dss.2015.09.003.
- Dandurand G, McKelvey F and Roberge J (2023) Freezing out: Legacy media’s shaping of AI as a cold controversy. *Big Data & Society* 10: 2. <https://doi.org/10.1177/20539517231219242>.
- Elgaard Jensen T, Kleberg Hansen AK, Ulijaszek S, Munk AK, Madsen AK, Hillersdal L and Jespersen AP (2019) Identifying notions of environment in obesity research using a mixed-methods approach. *Obesity Reviews* 20(4): 621–630.
- Elkin-Koren N (2020) Contesting algorithms: Restoring the public interest in content filtering by artificial intelligence. *Big Data & Society* 7(2): 2053951720932296.
- Ferrara E (2017) Disinformation and social bot operations in the run up to the 2017 French presidential election. arXiv preprint arXiv:1707.00086.
- Flasseur O, Denis L, Thiébaud É and Langlois M (2020) PACO ASDI: an algorithm for exoplanet detection and

- characterization in direct imaging with integral field spectrographs. *Astronomy & Astrophysics* 637: A9.
- Galbally J, Marcel S and Fierrez J (2013) Image Quality Assessment for Fake Biometric Detection: Application to Iris, Fingerprint, and Face Recognition, *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 710–724, doi: 10.1109/TIP.2013.2292332.
- Gravanis G, Vakali A, Diamantaras K and Karadais P (2019) Behind the Cues: A benchmarking study for Fake News Detection. *Expert Systems with Applications* 128. 10.1016/j.eswa.2019.03.036.
- Guarnera L, Giudice O and Battiato S (2020) DeepFake Detection by Analyzing Convolutional Traces. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2841–2850.
- Gupta BB and Sahoo SR (2020) Fake profile detection in multi-media big data on online social networks. *International Journal of Information and Computer Security* 12: 303. 10.1504/IJICS.2020.10026785.
- Han Q, Molinaro C, Picariello A, Sperli G, Subrahmanian VS and Xiong Y (2021) Generating fake documents using probabilistic logic graphs. *IEEE Transactions on Dependable and Secure Computing* 19(4): 2428–2441.
- Honnibal M and Montani I (2017) *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing*.
- Hossain MZ, Gedeon T and Sankaranarayanan R (2018) Using temporal features of observers' physiological measures to distinguish between genuine and fake smiles. *IEEE Transactions on Affective Computing* 11(1): 163–173.
- Hu M, Benson R, Chen AT, Zhu SH and Conway M (2021) Determining the prevalence of cannabis, tobacco, and vaping device mentions in online communities using natural language processing. *Drug and Alcohol Dependence* 228: 109016.
- Jacomy M (2021) *Situating Visual Network Analysis*. PhD Thesis, Aalborg University, DK.
- Jacomy M, Venturini T, Heymann S and Bastian M (2014) ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS one* 9(6): e98679.
- Jeong EY, Kim J, Jang WH, Lim HC, Noh H and Choi JM (2021) A more reliable defect detection and performance improvement method for panel inspection based on artificial intelligence. *Journal of Information Display* 22(3): 127–136.
- Karuna P, Purohit H, Ganesan R and Jajodia S (2018) Generating hard to comprehend fake documents for defensive cyber deception. *IEEE Intelligent Systems* 33(5): 16–25.
- Kim MJ (2021) Building a Cardiovascular Disease Prediction Model for Smartwatch Users Using Machine Learning: Based on the Korea National Health and Nutrition Examination Survey. *Biosensors* 2021(11): 228. <https://doi.org/10.3390/bios11070228>.
- Kitchin R (2017) Thinking critically about and researching algorithms. *Information, communication & society* 20(1): 14–29.
- Klug D, Steen E and Yurechko K (2023) How Algorithm Awareness Impacts Algospeak Use on TikTok. In: *Companion Proceedings of the ACM Web Conference 2023*, 30 April - 4 May 2023, pp. 234–237, New York: Association for Computing Machinery.
- Latour B, Jensen P, Venturini T, Grauwin S and Boullier D (2012) 'The whole is always smaller than its parts'—a digital test of Gabriel Tarde's monads. *The British journal of sociology* 63(4): 590–615.
- Lavanya S and Bhaskaran VM (2015) Effective data aggregation using EIGRP with adaptive time control scheme in WSN. *Advances in Natural and Applied Sciences* 9(12): 32–40.
- Lazer D, Kennedy R, King G and Vespignani A (2014) The parable of Google Flu: traps in big data analysis. *Science* 343(6176): 1203–1205.
- LeMoyné R and Mastroianni T (2019) Network Centric Therapy for Wearable and Wireless Systems. In: Dabov P (eds) *Smartphones: Recent Innovations and Applications*. New York: Nova Science Publishers, Hauppauge. Ch. 7.
- Levin S (2017) *YouTube Alters Search Algorithm over Fake Las Vegas Conspiracy Videos*. *The Guardian*, 6 October 2017. Available at: www.theguardian.com/us-news/2017/oct/06/youtube-alters-search-algorithm-over-fake-las-vegas-conspiracy-videos (last accessed January 2024). Retrieve from Internet Archive: [<https://web.archive.org/web/20231030205004/https://www.theguardian.com/us-news/2017/oct/06/youtube-alters-search-algorithm-over-fake-las-vegas-conspiracy-videos>].
- Li Y, Lei J and Liang D (2015) Identification of fake green tea by sensory assessment and electronic tongue. *Food Science and Technology Research* 21(2): 207–212.
- Liu Y, Chang CC and Huang PC (2019) Security protection using two different image shadows with authentication. *Math Biosci Eng* 16(4): 1914–1932.
- Maddocks S (2020) 'A Deepfake Porn Plot Intended to Silence Me': exploring continuities between pornographic and 'political' deep fakes. *Porn Studies* 7(4): 415–423. doi: 10.1080/23268743.2020.1757499.
- Maras MH and Alexandrou A (2019) Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos. *The International Journal of Evidence & Proof* 23(3): 255–262.
- Marres N (2012) On some uses and abuses of topology in the social analysis of technology (or the problem with smart meters). *Theory, Culture & Society*, 29(4–5): 288–310.
- Marres N (2015) Why map issues? On controversy analysis as a digital method. *Science, Technology, & Human Values*, 40(5): 655–686.
- Maslej N, Fattorini L, Brynjolfsson E, Etchemendy J, Ligett K, Lyons T, Manyika J, Ngo H, Niebles JC, Parli V, Shoham Y, Wald R, Clark J and Perrault R (2023) *Artificial Intelligence Index Report 2023*. Stanford University. Available at: <https://aiindex.stanford.edu/report/> (last accessed January 2024). Retrieve from Internet Archive: [<https://web.archive.org/web/20240117152446/https://aiindex.stanford.edu/report/>]
- Mimura M (2020) Using fake text vectors to improve the sensitivity of minority class for macro malware detection. *Journal of Information Security and Applications* 54: 102600.
- Mohawesh R, Tran S, Ollington R and Xu S (2021) Analysis of concept drift in fake reviews detection. *Expert Systems with Applications* 169: 114318.
- Mozafari M, Farahbakhsh R and Crespi N (2020) Hate speech detection and racial bias mitigation in social media based on BERT model. *PLoS one* 15(8): e0237861.
- Munk AK (2019) Four styles of quali-quantitative analysis: Making sense of the new Nordic food movement on the web. *Nordicom Review* 40(s1): 159–176.
- Olazaran M (1996) A sociological study of the official history of the perceptrons controversy. *Social Studies of Science* 26(3): 611–659.

- Palla G, Derényi I, Farkas I and Vicsek T (2005) Uncovering the overlapping community structure of complex networks in nature and society. *Nature* 435(7043): 814–818.
- Pennycook G and Rand DG (2019) Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences* 116(7): 2521–2526.
- Praisler M, Domnisoru D and Domnisoru L (2013) Chemometric method for the automated identification of cucuteni ceramics based on ATR-FTIR spectra. *European Journal of Science and Theology* 9: 249–256.
- Prates MO, Avelar PH and Lamb LC (2020) Assessing gender bias in machine translation: a case study with google translate. *Neural Computing and Applications* 32: 6363–6381.
- Rath A, Mishra D, Panda G and Satapathy SC (2021) Heart disease detection using deep learning methods from imbalanced ECG samples. *Biomedical Signal Processing and Control* 68: 102820.
- Rieder B (2017) Scrutinizing an algorithmic technique: The Bayes classifier as interested reading of reality. *Information, Communication & Society*, 20(1): 100–117.
- Rodríguez-Sánchez F, Carrillo de Albornoz J and Plaza L (2020) Automatic classification of sexism in social networks: An empirical study on twitter data. *IEEE Access* 8: 219563–219576.
- Roe J and Perkins M (2023) ‘What they’re not telling you about ChatGPT’: exploring the discourse of AI in UK news media headlines. *Humanit Soc Sci Commun* 10, 753 (2023). <https://doi.org/10.1057/s41599-023-02282-w>
- Singh VK, Singh P, Karmakar M, Leta J and Mayr P (2021) The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. *Scientometrics* 126: 5113–5142.
- Stilgoe J (2018) Machine learning, social learning and the governance of self-driving cars. *Social studies of science* 48(1): 25–56.
- Tancoigne E, Barbier M, Cointet JP and Richard G (2014) The place of agricultural sciences in the literature on ecosystem services. *Ecosystem Services* 10: 35–48.
- Thelwall M and Sud P (2022) Scopus 1900–2020: Growth in articles, abstracts, countries, fields, and journals. *Quantitative Science Studies* 3(1): 37–50. https://doi.org/10.1162/qss_a_00177.
- Tran TV and Chung WY (2015) A robust algorithm for real-time peak detection of photoplethysmograms using a personal computer mouse. *IEEE Sensors Journal* 15(8): 4651–4659.
- Upadhyaya A, Shokeen V and Srivastava G (2018) Decision tree model for classification of fake and genuine banknotes using SPSS. *World Review of Entrepreneurship, Management and Sustainable Development* 14(6): 683–693.
- Vaccari C and Chadwick A (2020) Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society* 6(1). <https://doi.org/10.1177/2056305120903408>.
- Venturini T, Baya Laffite N, Cointet JP, Gray I, Zabban V and De Pryck K (2014) Three maps and three misunderstandings: A digital mapping of climate diplomacy. *Big Data & Society* 1(2): 2053951714543804.
- Venturini T and Munk AK (2021) *Controversy mapping: A field guide*. Cambridge, UK: Polity Press.
- Whatmore SJ (2009) Mapping knowledge controversies: science, democracy and the redistribution of expertise. *Progress in Human Geography* 33(5): 587–598.
- Yao Y, Cifuentes J, Zheng B and Yan M (2019) Computer algorithm can match physicians’ decisions about blood transfusions. *Journal of Translational Medicine* 17: 1–5.
- Yu C, Zuo Y, Feng B, An L and Chen B (2019) An individual-group-merchant relation model for identifying fake online reviews: an empirical study on a Chinese e-commerce platform. *Information Technology and Management* 20: 123–138.
- Zhang L, Cai Z and Wang X (2016) Fakemask: A novel privacy preserving approach for smartphones. *IEEE Transactions on Network and Service Management* 13(2): 335–348.
- Zhang Y (2022) Algorithmic photography: A case study of the Huawei Moon Mode controversy. *Media, Culture & Society*, 44(4): 690–705.
- Zubova LV, Korovin EV, Smirnov AS, Kuzmin VN and Kurakov AV (2021) Development of Problem-Oriented Management and Decision-Making System and Optimization of Economic and Social Systems. *Webology* 18: 436–451.