Investigating the influence of AI research topics in the academic, public, and industry spheres

Glen Berman*, Kate Williams**, and Sandra Michalska***

* glen.berman@anu.edu.au
0000-0003-3249-0190
College of Engineering and Computer Science, Australian National University, Australia

** kate.williams@unimelb.edu.au
0000-0002-2882-1068
School of Social and Political Sciences, University of Melbourne, Australia

*** sandra.michalska@kcl.ac.uk
Policy Institute, King's College London, United Kingdom

The Artificial Intelligence research field sits at the intersection of several overlapping spheres (academia, industry, media), each with their own logics and commitments. The influence of research within these worlds is studied through a number of bibliometric methods, including citation metrics for measuring influence within academia, and counts of patents and news-media mentions for influence in industry and the media. Using a large-scale dataset of research outputs, we compare the topical content of outputs that are highly influential in each of these worlds.

We identify significant differences between the content of influential research in these worlds, indicating that the academic, industry and media worlds value different aspects of the Artificial Intelligence field. These differences provide new insights on the evaluation of research produced within the Artificial Intelligence field.

1. Introduction

Since 2010, propelled by advances in applying Machine Learning (ML) techniques to prediction tasks, the field of Artificial Intelligence (AI) has experienced a boom (Russell & Norvig, 2021). The production of AI research has rapidly increased (Liu et al., 2018). The field has received significant private and public investment (Abadi et al., 2020; Bughin et al., 2017), and widespread news media coverage (Ouchchy et al., 2020). Concurrently, the field has become the subject of public policy debate (Jobin et al., 2019) and social science research (Frank et al., 2019). Consequently, evaluation of AI research and AI research funding decisions are of widespread significance both within the AI research community, and in the broader academic and research sector.

Given the significance of AI research across multiple worlds—particularly those of academia, industry, and the media—the AI field can be conceptualised as sitting at the intersection of a range of spheres, each with their own logics and commitments. These different logics result in different aspects of AI research being valued differently across the spheres. The media sphere, for example, has closely attended to societal risks associated with AI developments (Chuan et al., 2019). Indeed, concerns have been raised about a potential disconnect between AI research which garners media coverage, and AI research which is currently the focus of commercialisation (Ouchchy et al., 2020). However, the intersection of these spheres is not uniform. Different institutions, and different individual researchers, are situated at different points of intersection, depending on a range factors, including sociopolitical context, institutional history, and funding model. A university in North America is likely to be subject to the logics of the academic, industry and media spheres in a way that is different to a university in Europe, or a commercial firm in Asia. Each of these institutions is thus likely to
value AI research differently. Reflecting this, previous research has demonstrated that the focus of AI researchers varies based on their affiliation; AI researchers affiliated with industry, in particular, have increasingly focused on Deep Learning (DL) techniques (Klinger et al., 2020). Yet, it is unclear whether the same topic areas within AI research are influential across the academic, industry, and media spheres. As such, this study addresses the following research question: what topics within the AI field are influential in the academic, industry, and media spheres?

In answering this research question, we adopt Elsevier’s definition of the AI field (Siebert et al., 2018), which includes research on high-level AI techniques (e.g. natural language processing, computing vision), the application of these techniques (e.g. in the health sciences), and social science research on their use and societal significance (e.g. in the economics field). We conceptualise influential scholarly outputs for each sphere as those that are ranked in the 99th percentile on a relevant metric (citations for the academic sphere, patent filings for the industry sphere, and media mentions for the media sphere). While these metrics are widely used, we note that they are incomplete proxies for influence. To identify topics within the AI field we use Latent Dirichlet Allocation (LDA) topic modelling, which is an unsupervised ML technique for extracting latent topics from a corpus of documents. In our use, the corpus consists of abstracts from a sample of 4,716 scholarly outputs. Our study demonstrates a novel use of altmetrics data alongside traditional citation data to identify the topic areas of highly influential publications across multiple spheres.

2. Related work

2.1. Research influence

One feature that increasingly shapes the scientific community is the evaluation of research (Shapin, 1995). Evaluation is laden with the logics and assumptions of the spheres in which it operates (Lamont, 2012; Williams, 2020). Within the distinct logics of the academic, industry, and media spheres, evaluation practices operationalise measurements of research ‘influence’ differently. Within specific institutions, meanwhile, evaluation practices are the product of that institution’s attempts to navigate these spheres. Little is known, however, about the specific research focus that appeals to the logics of the academic, industry, and media spheres.

Historically, attempts to understand research influence have primarily occurred within the logic of the academic sphere, relying on citation and authorship data (Lawani, 1981). This data is used in bibliometric studies designed to measure publication influence (Roemer & Borchardt, 2015, 2017). Such studies are paramount in assessing the impact of research in the academic sphere (Williams, 2020). Recently, in studies of the influence of scholarly outputs, researchers have also begun to incorporate alternative metrics (altmetrics). Altmetrics seek to record attention to, or engagement with, scholarly outputs in spheres beyond academia (Roemer & Borchardt, 2017), through monitoring of outputs on social and news media, and in patent filings (Haustein et al., 2014). Altmetrics can thus be used to contextualise or extend insights garnered from bibliometric analysis of traditional citation data (Klinger et al., 2020), enabling the influence of research in the industry and media spheres to be considered through measurement of the commercialisation of scholarly outputs (through patent filings) and media interest in outputs (through social and news media mentions). As academic institutions and national funders are increasingly subject to the logic of these spheres, there has been substantial interest in using altmetrics within these institutions to inform their own evaluations of research (Sugimoto et al., 2017). However, as standalone measures, altmetrics tend to lack credibility...
and thus are better suited to interpretation within suites of measures with an accompanying rationale (Cheung, 2013).

2.2. Identifying latent research topics

Where a sample of documents is very large and human reading and interpretation is not feasible, topic modelling offers a method for the unsupervised learning of latent structure across the sample (Anupriya & Karpagavalli, 2015; Grimmer & Stewart, 2013). Latent Dirichlet Allocation (LDA) is a widely used probabilistic approach to topic modelling (Grimmer & Stewart, 2013). LDA is premised on the notion that texts embed knowledge not only by conveying information explicitly through structured sentences, but also implicitly through how words co-occur with each other (Blei et al., 2003). In particular, the co-occurrence of words in documents can reveal information about shared context or themes in those documents (Gefen et al., 2017). LDA thus uses an inductive approach to identify latent topics. Topics are not externally defined by the researcher, but rather emerge bottom-up by measuring the co-occurrence of words within documents in a corpus (Blei et al., 2003).

In the context of studying research fields, LDA has been used to interpret the broad thematic focus of the field through modelling of the content of publication abstracts (Li & Lei, 2021). Studies in different domains, such as sociology (Giordan et al., 2018) and criminology (Vander Beken et al., 2021), and different application areas, such as recommendation systems (Amami et al., 2016) and classification systems (Kim & Gil, 2019) have applied LDA to the abstracts of scholarly outputs. Syed et al. (2017) considered whether topic modelling of abstracts produced comparable results to modelling of the full text of scholarly outputs, finding that for large document corpora, modelling of abstracts alone can produce quality topics. Louvigné et al. (2013) found that abstract content is sufficient for topic modelling of research due to the concise yet rich use of terms in abstracts.

3. Data and method

3.1. Sampling AI scholarly outputs

Our sample of AI scholarly outputs is derived from Elsevier’s AI dataset, which consists of 726,158 outputs indexed on Scopus (de Kleijn et al., 2017). We restricted our sample to the years 2014 to 2019, and removed outputs that did not include sector data or were duplicates, leaving a sample of 276,966 outputs. Scopus provided bibliometric data, including academic citations, which we use as a proxy for influence in the academic sphere. As proxies for the industry and media spheres we used patent citations and news media mentions. These metrics are not provided by Scopus, so instead are sourced from Altmetric1. We searched Altmetric for the DOIs of our sample of scholarly outputs, finding matches for 95,686 outputs (Altmetric’s coverage of scholarly outputs is incomplete). Of the matches we found, some were duplicate matches and 11 were erroneous, reducing the sample to 93,492 outputs. Finally, for each output we retrieved the abstract text from Scopus. Overall, 99.6% of abstracts were successfully retrieved, resulting in the final sample of 93,088 outputs used to identify highly influential scholarly outputs.

3.2. Identifying highly influential scholarly outputs

We conceptualised highly influential scholarly outputs in the academic, industry, and media spheres as those which were in the 99th percentile for academic citations, patent citations, and

1 https://www.altmetric.com/
news media mentions respectively. We note that high performance in these metrics is not indicative of inherent research quality, but rather of the reception of research outputs. For each sphere, scholarly outputs in the 99\textsuperscript{th} percentile of the relevant proxy metric were subset. Where a scholarly output was in the 99\textsuperscript{th} percentile of more than one metric, it was included in all relevant subsets. Scholarly outputs that were not in the 99\textsuperscript{th} percentile for any of the proxy metrics were placed in an ‘other’ subset. By design, this ‘other’ subset is larger than the three subsets of highly influential outputs. As such, in order to reduce the computational load of estimating topics across a very large corpus whilst ensuring a level of granularity in the topic model that would enable us to discern patterns between different subsets, for topic modelling we created a random sample of the ‘other’ subset of equal size to the number of unique scholarly outputs in the combined highly influential subsets (2,358 outputs). Table 1 shows the distribution of scholarly outputs across these subsets. The total number of unique scholarly outputs included in these subsets is 4,716.

Table 1. Distribution of scholarly outputs across the spheres of influence.

<table>
<thead>
<tr>
<th>Subset name</th>
<th>Definition of the subset</th>
<th>Number of outputs in subset</th>
<th>Percentage of outputs unique to that subset</th>
</tr>
</thead>
<tbody>
<tr>
<td>high-Academic</td>
<td>99\textsuperscript{th} percentile in academic citations (≥ 324 citations)</td>
<td>930</td>
<td>66.5%</td>
</tr>
<tr>
<td>high-Industry</td>
<td>99\textsuperscript{th} percentile in patent citations (≥ 3 citations)</td>
<td>908</td>
<td>70.0%</td>
</tr>
<tr>
<td>high-Media</td>
<td>99\textsuperscript{th} percentile in news–media mentions (≥ 5 mentions)</td>
<td>868</td>
<td>88.7%</td>
</tr>
<tr>
<td>other</td>
<td>Random sample of outputs not in any other subset</td>
<td>2,358</td>
<td>NA</td>
</tr>
</tbody>
</table>

3.2. Estimating a topic model

To prepare research abstracts for topic modelling English stop words were removed from the abstracts, and the abstract text was decomposed into a document-term matrix, which records the number of times each unique word (term) occurs in the abstract. In the LDA approach individual terms are represented as a multinomial probability distribution across topics. Abstracts (documents) are also represented as a multinomial distribution across topics, which is determined by summing the distribution across topics of all terms in the abstract. As such, the number of topics to identify in a corpus must be externally set by the researcher. Given this, an important aspect of the LDA approach is identifying an optimal number of topics for a given corpus, which is usually achieved by estimating topic models for a range of different numbers of topics, and comparing the resulting models. Topic models were estimated for between 5 and 30 topics, with the optimum number of topics calculated using Gibbs sampling (Phan et al., 2008) and two semantic coherence measures (Deveaud et al., 2014; Griffiths & Steyvers, 2004). Figure 1 reports the performance of these topic models. A topic model with 14 topics was found to perform best across these metrics, and was thus selected for interpretation.

Figure 1. Topic model performance using Gibbs sampling and semantic coherence measures
4. Results

4.1. Description of topic model

For each topic within an LDA topic model, a list of terms that are most likely to indicate that topic can be generated using per-word-per-topic probabilities. Similarly, the per-document-per-topic probabilities can be derived to indicate the most likely topics that represent a given abstract (Blei et al., 2003). Figure 2 combines these calculations to show the 5 terms most associated with each topic, and the distribution of topics across all 4,716 abstracts, where each abstract is assigned to its most dominant topic. The terms associated with many of the topics highlights the significance of ML, and particularly DL, techniques in the AI field. Several topics refer to categories of tasks in which these techniques are used: the key terms associated with topics 1 and 12 appear to refer to natural language processing; the terms associated with topics 5 and 8 appear to refer to image processing; and the terms associated with topics 2 and 9 refer to classification tasks (e.g. between categories). Additionally, topics refer to particular application domains for AI techniques: topics 4 and 10 appear to refer to applications in the health domain.

4.2. Topical focus in the academic, industry and media spheres

For each of the subsets of scholarly outputs described in Table 1 (high-Academic, high-Industry, high-Media, other), we can compare how the latent topics identified are distributed, as shown in Figure 3. Note that these subsets contain some overlap: while the majority of scholarly outputs in each of the high-influence subsets are unique, some may be included in more than one subset. Nonetheless, as can be seen in Figure 3, topic focus varies between the three subsets and the ‘other’ category. The high-Academic and high-Industry subsets share roughly similar distributions of the topics. In both of these subsets, topics focused on DL techniques (topics 8 and 14) are most dominant. The high-Media subset, meanwhile, has a different distribution. The topics associated with application of AI in the health domain (topics 4 and 10) share a far greater proportion of the topical focus in the high-Media subset than in the other subsets. Conversely, topics focused on DL are less represented in the high-Media subset than the other subsets. Notably, the topical distributions within each of the high-influence subsets diverge from the distribution within the ‘other’ subset. To confirm the
statistical significance of these findings we also ran a Multinomial Logistic Regression of the distribution of topics across different spheres, which is reported in Table 2.

Figure 2. Distribution of topics across the corpus, with top terms per topic shown

Figure 3. Distribution of topics across the spheres

Table 2. Multinomial Logistic Regression between spheres and latent topics (ref: other)
5. Discussion

5.1. Divergence in the academic, public, and industry spheres

This paper explored the topical foci of AI research that display patterns of influence across academic, public, and industry spheres. The findings demonstrate that within the academic, industry, and media spheres different AI research topics have been influential. We identified statistically significant differences in research focus among various spheres of influence, with particularly stark differences in the topical focus of publications that were endorsed in the news-media sphere compared to the academic and commercial spheres.

In the academic and industry spheres, DL techniques, particularly for image processing tasks, are most influential. Meanwhile, in the public sphere, applications of AI in the medical domain are most influential. These results highlight a distinction between fundamental AI research and the application of that research in applied fields. It appears that in the public sphere, applications of AI research are most influential, whilst in industry and academic spheres this is not the case.

More broadly, although the Elsevier AI dataset is designed to incorporate research on all AI techniques, the extent to which keywords associated with all topics reflect ML or DL techniques reflects the centrality of these approaches to AI research. These results can be mobilised by academic institutions and national research funders to inform AI research investment strategies, and by critical scholars in further analyses of the AI field.

5.2. Limitations

This is a preliminary study. Our sampling strategy may have produced a dataset that is unrepresentative of the broader AI field. This is both an intended consequence of our focus on
highly influential scholarly outputs, and may also be an indirect result of the random sample of ‘other’ outputs included in the dataset. Our identification of influential outputs is also dependent on the bibliometric and altmetric data provided by Scopus and Altmetric. Validating their metrics was outside the scope of this study. Additionally, our definition of highly-influential is binary (99th percentile on a certain metric), and flattens distinctions in influence. In future research more graduated definitions of influence may be helpful. Finally, the quality of topics identified in our model has not been validated by expert review, which may be useful in larger future studies (Wallach et al., 2009).

6. Conclusion

This study shed light on the use of multiple metrics in evaluation, while offering insights into the topical content of a significant field of research that increasingly shapes contemporary society. It also demonstrated the utility of combining topic modelling with bibliometric and altmetric analysis in the study of complex or hybrid research fields. Overall, the collection of terms within a topic proved to be semantically coherent, demonstrating the potential for content abstraction from the publications’ abstracts using a topic modelling approach. Our findings also contribute to the growing discourse on the significant role of DL (or more broadly neural networks) within the AI field, by highlighting the dominance of these approaches in high performing publications. The study suggests that greater attention to capturing how different areas of AI research are taken up in different spheres is required.

Open science practices

We are not aware of an open database of Artificial Intelligence documents. Thus, we are restricted to using a closed dataset, gained from Scopus.

Author contributions

Conceptualization: KW & GB
Data curation: GB & SM
Investigation: KW, GB, SM
Methodology: GB, KW, SM
Visualization: GB, SM
Writing – original draft: GB & KW
Writing – review & editing: GB, KW, & SM

Competing interests

The authors declare that they have no competing interests.

Funding information

The research was facilitated by KW’s ESRC research grant ES/V004123/1.

References


