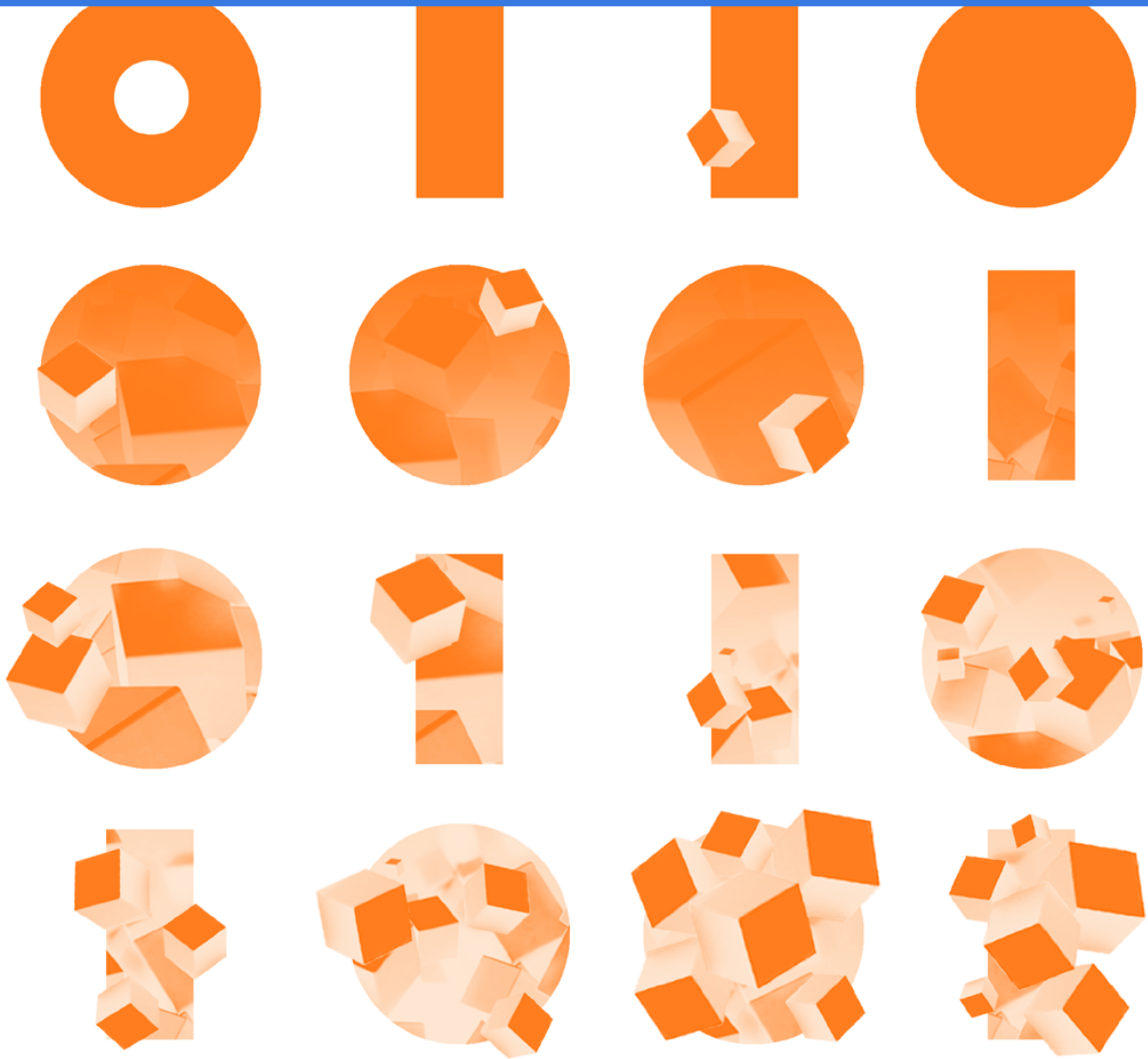
CONFIDENTIAL UNDER EMBARGO UNTIL TUESDAY **11 DEC 9:00 AM** CET.



Artificial Intelligence: How knowledge is created, transferred, and used

Trends in China, Europe, and the United States

ELSEVIER

Contents

Foreword

Artificial Intelligence: How knowledge is created, transferred, and used

Dan Olley,
Chief Technology Officer (CTO),
Elsevier, United States

"In recent years, artificial intelligence, or AI, has gained a surge in attention from policy makers, universities, researchers, corporations, media, and the public. Driven by advances in big data and computing power, breakthroughs in AI research and technology seem to happen almost daily. Expectations, but also fears, are mounting about the transformational power of AI to change society. In this whirlwind of attention and development, terms are getting confused. "artificial intelligence," "machine learning," and "data science" are often used interchangeably, yet they are not the same. AI is often intuitively understood as an umbrella term to describe the overall objective of making computers apply judgment as a human being would. Themes, such as deep learning, drop out of the AI umbrella to become their own research fields and technologies.

The confusion of terms, in a field with such potential to transform lives, needs to be addressed to ensure that policy objectives are correctly translated into research priorities, student education matches job market needs, and media can compare the knowledge being developed in various countries and regions across the globe. This is exactly the challenge we have set ourselves to tackle with this report. After all, we are an information analytics company focused on research and health, with data assets that can provide valuable insight into these important issues.

Powered by extensive datasets from our own and public sources—examined by our data scientists by applying machine learning on high-performance computing technology and validated in close collaboration with domain experts from research institutions and industry around the world—we have characterized the field of AI in a structured and comprehensive way. We then used this characterization to understand how AI knowledge is created, transferred, and used worldwide, with a focus on the "big 3" geographies: China, Europe, and the United States. We looked well beyond the traditional bibliometrics of published journal articles, examining also conferences, preprints, education, and competitions.

As I look at the resulting report, what most resonates with me is the section on approaches to AI, ethics, and responsible innovation. Traditional machine learning techniques rely on a human to decide what facets of the data are the most important to the model they are building. However, new techniques rely on the machine itself to decide what is important in the data to drive the required outputs. This is a fundamental shift as the focus moves from the design of the software program to the design of the training and testing data. This is important because as AI algorithms and models get more complex, there has understandably been a rise in the call for explainability. Why are we getting a certain result? How and what has the machine seen as important in the data? Is there any unconscious bias in the result?

Given the natural preconception that computers work with linear programs to give finite results, people often want to understand the "program flow" of the model. While there is some extremely valuable work going on to look inside the "black box" of modern machine learning techniques, this report clearly reveals a need to reset public preconceptions of how machines work with these new techniques, and the probabilistic results they give, to be able to properly discuss topics of ethics and bias. This change in mindset will shift the focus of the discussion to be as much about how we are designing our training of these machines to cover questions of ethics and bias as it is about peering into the models we have created to try and explain what has happened.

This is exactly what we now do at Elsevier with so-called "data squads": new algorithms are developed by a multi-skilled team that combines knowledge of the machine learning algorithms being used, the domain being worked on, and software engineering, testing, and ethics. In this way, we ensure that we design the machine's "training curricula" for the algorithm's intended purpose, while being able to mitigate any unintended consequences.

With this report, we aim to make a contribution to the responsible development, dissemination, and use of AI knowledge for the benefit of society. This report marks the start of a wider engagement of RELX, both on our online AI resource center where more in-depth insights are available, and through our collaborations in the research community and beyond. As CTO of Elsevier, I look forward to further engaging with you in the future."

Foreword

A source-based approach to measuring AI publication volume

Dr. Raymond Perrault,
Senior Technical Advisor,
Artificial Intelligence Center at
SRI International, United States

"Counting publications in AI is difficult, as the field is notoriously tricky to bound. Russell and Norvig[1] point out two main axes over which work is dispersed. The first goes from reasoning at one end to behavior at the other. The second restricts explanations to those that can be shown to closely reflect processes in humans (i.e., the cognitive science end) to those that are constrained by a broader appeal to rationality and optimization, and are more suitable to applications. Another obvious dimension is from research on new techniques to their applications in a wide range of domains. Since AI has absorbed basic techniques from so many fields (e.g., logic, probability and statistics, optimization, photogrammetry, neuroscience, and game theory, to name a few) and its methods are being applied in so many other fields (e.g., speech recognition, computer vision, robotics, cybersecurity, bioinformatics, and healthcare) it is not easy to draw a line between AI and fields both upstream and downstream from it.

What should or should not be considered AI also changes over time. Before the late 1980s, natural language processing (based on Chomskian linguistics, related parsing techniques, and first-order semantics) was definitely part of AI, and speech recognition (based on signal processing and Hidden Markov Models) was not. Both subareas are now largely driven by machine learning, and so are clearly within mainstream AI.

If there is a basis for drawing a line around AI, I believe it rests in the social fabric of the field, as expressed by the sources where new work appears, namely its journals and conferences, tied together by researchers who tend to work in one or two subareas at a time and mostly publish in a small set of related sources. As one of these sources, the AI spectrum of conferences ranges from the traditional venues for symbolic AI, e.g., IJCAI,[2] AAAI,[3] ICAPS,[4] and KR;[5] to major venues for machine learning and probabilistic reasoning, e.g., NIPS,[6] ICML,[7] and UAI;[8] to more independent application conferences such as KDD[9] and SIGIR.[10]

Basing counts of publications on sources provides a way to systematically and transparently describe what is included in an area (e.g., AI), or a group core of areas (e.g., the subareas of AI, or all of computer science), and to systematically vary the breadth and granularity of the specifications of the cores. All the information necessary is in indexes such as Scopus®. Alternatively, this could be done by training classifiers operating on publication content, always with ground truth given by the core-based tags.

This report follows this approach and applies multiple ways to shape and structure the field of AI. It is a very welcome contribution to understanding and monitoring the dynamics of an ever-emerging field. Systematizing and benchmarking the approaches over different sources and cluster algorithms would be interesting future research."

1 Russell, S., Norvig, P. *Artificial Intelligence*: A Modern Approach. 3rd ed. Essex, UK: Pearson Education Limited; 2014.
2 International Joint Conference on Artificial Intelligence.
3 Association for the Advancement of Artificial Intelligence.
4 International Conference on Automated Planning and Scheduling.
5 Principles of Knowledge Representation and Reasoning.
6 Neural Information Processing Systems.
7 International Conference on Machine Learning.
8 Uncertainty in Artificial Intelligence.
9 Knowledge Discovery in Databases.
10 Special Interest Group on Information Retrieval.

Foreword

Defining AI: new approaches help with AI ontologies

Prof. Enrico Motta,
Professor of Knowledge Technologies, The Open University, United Kingdom

"Disciplines do not exist per se. They emerge because of a collective construction process, whereby a community of researchers comes together, formulating and sharing common objectives, methods, and conceptualizations. Hence, disciplines are essentially about research communities. As these evolve, so do the associated disciplines. Thus, attempts at characterizing disciplines are in my view more successful if they follow a bottom-up approach, focusing less on top-down definitions than on identifying the relevant body of work.

Given this premise, I am very happy to endorse this report produced by the Elsevier team, which provides an operational characterization of the field of AI, in terms of 600,000 documents and over 700 field-specific keywords. This is an impressive piece of work that, to my knowledge, provides the most comprehensive characterization of AI outputs produced so far. Crucially, in contrast with manually developed taxonomies of research areas, which inevitably end up reflecting the specific viewpoints of the experts involved in the process, this characterization is data-driven, using machine learning and text mining techniques to classify documents and identify the relevant keywords. Thus, in my view, the report enjoys greater validity, providing a more objective reflection of the variety of existing contributions to the AI field.

In addition to its scientific value, there is also no doubt that this report will be a very valuable practical resource for people who wish to explore this space. For example, it will be very interesting to use this comprehensive characterization of the AI field to get a better understanding of key trends and topics, especially when the relevant body of work may be spread across different research communities, as it is the case, for instance, with work in the highly important area of AI ethics.

On a personal level, this work is also very exciting for me because it provides the basis for interesting new research. One of my main research areas concerns the use of AI technologies to develop innovative solutions that can help people to make sense of the dynamics of scientific research. Within this broad context, my team has developed an original approach to the automatic generation of taxonomies of research areas and, for example, it would be extremely interesting to investigate to what extent these different methods can cover the research space and to what extent they can be combined to improve accuracy. This is just one example of the many interesting possibilities for further research opened up by this work.

In sum, this is not just an excellent piece of work, but also the start of a very interesting line of research. I congratulate the Elsevier team for their tremendous work and I look forward to further developments in this space."

Executive summary

The growing importance and relevance of artificial intelligence (AI) to humanity is undisputed: AI assistants and recommendations, for instance, are increasingly embedded in our daily lives. However, AI does not seem to have a universally agreed definition. Our classification methodology contributes to the understanding of an evolving field with a shifting structure. AI clusters around the areas of Search and Optimization, Fuzzy Systems, Natural Language Processing and Knowledge Representation, Computer Vision, Machine Learning and Probabilistic Reasoning, Planning and Decision Making, and Neural Networks.

While the field spans several domains and can be viewed from different standpoints, such as teaching, research, industry, and media, there seems to be little overlap in vocabulary between these perspectives. Industry tends to emphasize algorithms, possibly for efficient gains in time and human labor. The increasing societal relevance of AI and potential ethical concerns raised by the growing use of algorithms reflect the visibility of applications and ethics themes in the media, which makes AI more imperative and intuitive to the public. Interestingly, ethics keywords are also more heavily represented in teaching, potentially as a result of public interest and some government mandates, like in The Netherlands. In AI research, ethics keywords are currently not explicitly visible, which poses the question of whether ethical analysis is forthcoming among AI researchers, whether such discussions are conducted outside of the AI field, or whether they take place outside of research altogether. This observation is noteworthy, as responsible innovation in AI is crucial to ensure safe and fair outcomes for all.

The apparent lack of a common language across perspectives calls into question the quality of understanding and communication across the AI field. With closer and instant collaboration across geographies and sectors, research dialogue shifts away from traditional sequential translation and towards parallel dialogues, online and through media and social media channels. New stakeholders, such as students, freelancers, and citizens, become involved in research, for example, on competition platforms like Kaggle. A common language and understanding would better connect actors in the AI ecosystem.

AI has also emerged as an area of importance for national competitiveness. Several national and international AI policies and strategies have been put forth in recent years, as both causes and consequences of growing AI research ecosystems. This has led to increased scientific output through a variety of dissemination modes, including publications, preprints, conferences, competitions, and software.

There are strong regional differences in AI activity. China aspires to lead globally in AI and is supported by ambitious national policies. A net brain gain of AI researchers in China also suggests an attractive research environment. China's AI focuses on computer vision and does not have a dedicated natural language processing and knowledge representation cluster, including speech recognition, possibly because this type of research in China is conducted by corporations that may not publish as many scientific articles. It shows robust growth of its research and education ecosystems, with a rapid rise in scholarly output and similar research usage as other regions. China's AI research has a rapidly increasing yet still comparatively low citation impact, which could be a symptom of regional, rather than global, reach. This is also apparent through its relatively low levels of international collaboration and mobility in research, which yield a comparatively small but highly cited corpus of AI research. As in many other research areas, collaboration is key to success, as demonstrated by increasing discussions on global social media and growing international AI competition numbers.

Europe is defined in this report as the 44 countries belonging to the European Union (EU) and associated countries eligible for Horizon 2020 funding. It is the largest region in AI scholarly output, with high and rising levels of international collaborations outside of Europe, but appears to be losing academic AI talent, especially in recent years. The broad spectrum of AI research in Europe reflects the diversity of European countries, each with their own agenda and specialties. Focus areas of European AI research include genetic programming for pattern recognition, fuzzy systems, and speech and face recognition. Deep learning research in Europe appears less connected to other subfields than it is in other regions, and AI robotics in Europe appear to be embedded in the machine learning cluster.

The United States corporate sector attracts talent and is strong in AI research, possibly due to their cross-sector joint labs tradition. The United States academic sector is also robust, both in terms of scholarly output and talent retention. The country appears to be leading the way in international AI competitions, and United States researchers increasingly collaborate internationally on AI research. AI in the United States has a strong focus on specific algorithms and separates speech and image recognition into distinct clusters. The corpus shows less diversity in AI research than Europe but more diversity than China.

Among other key contributors in AI, we note the rapid emergence of India, today the third largest country in terms of AI publications after China and the United States. Iran is ninth in publication output in 2017, on par with countries like France and Canada. Last year, Russia surpassed Singapore and The Netherlands in research output, yet remains behind Turkey. Germany and Japan remain fifth and sixt largest producers of AI research globally.

In this report, we provide insights for the benefit of research evaluators, research funders, policy makers, and researchers. We use a bottom-up approach to delineate the research fields of AI and invite further collaborative research on corpus definition. Our analysis also raises several questions of interest for potential future investigations:

- Is there a relationship between research performance in AI and research performance in more traditional fields that support AI (such as computer science, linguistics, mathematics, etc.)?
- How does AI research translate into real-life applications, societal impact, and economic growth?
- Where do internationally mobile AI researchers come from and go to?
- How sustainable is the recent growth in publications and how will countries and sectors continue to compete and collaborate?

Highlights

Artificial intelligence research focuses on Search and Optimization, Fuzzy Systems, Natural Language Processing and Knowledge Representation, Computer Vision, Machine Learning and Probabilistic Reasoning, Planning and Decision Making, and Neural Networks.

CHAPTER 2

There is increasing societal relevance of AI, particularly notable in small but growing application fields like health sciences, agriculture, or the social sciences; high public interest is reflected in social media and blog mentions. Despite this societal relevance, ethics is not yet strongly reflected in the research corpus, although recent conferences reveal a growing focus on ethics.

CHAPTERS 3 & 5

The field has grown annually by 5.3% in the last decade and 12.9% in the last 5 years. It has emerged as an area of importance for national competitiveness, yet also sees growing international collaboration. Europe is still the largest actor in AI research, despite rapid growth and ambition from China, while the United States supports a strong corporate sector alongside academia.

INTRODUCTION & CHAPTER 3

China aspires to lead globally in AI and is supported by ambitious policies and rapid growth, especially in computer vision and fuzzy systems. A recent brain gain of AI researchers also suggests an increasingly attractive research environment, and citation impact is also growing. However, compared to other regions, China's research appears to have a regional, rather than global, reach.

INTRODUCTION & CHAPTER 3

Europe is the largest and most diverse region in terms of AI scholarly output, with high and rising levels of international collaborations outside of Europe. However, Europe appears to be losing AI talent in recent years, especially in academia.

CHAPTER 3

The United States corporate sector attracts talent and is strong in research, possibly due to their cross-sector joint labs tradition. The academic sector is also robust, both in terms of scholarly output and talent retention.

CHAPTER 3

Among other key countries in AI research, we note the rapid emergence of India, today the third largest producer of AI publications after China and the United States.

CHAPTER 3

Introduction

The field of artificial intelligence (AI) is broad, dynamic, and rapidly evolving, and is producing technologies with enormous global societal implications.[11, 12, 13, 14] For example, advances in facial and speech recognition have produced virtual assistant technologies that are being integrated into daily life like Siri, Alexa, Google, iFLYTEK, and Baidu.[15] AI-based recommender systems have revolutionized online search optimization and digital ad targeting. In the realm of image interpretation, AI is improving medical image analysis for rapid and accurate diagnoses and treatment planning.[16] Research in AI is both theoretical and applied, and transcends traditional disciplinary boundaries, bringing together experts from diverse fields of study.[17]

Clarifying the scope and activity within this large field can help research leaders, policy makers, funders and investors, and the public navigate AI and understand how it has evolved over time. This effort may also provide clues as to where AI is headed and how policies might be shaped to continue making advances in a responsible way. For this report, Elsevier used High-Performance Computing Cluster (HPCC) developed at RELX and drew on their analytic expertise as well as insights from internal and external experts in AI research and application. This combined approach allowed us to ask, “How is knowledge in AI created, transferred, and used?”

Alessandro Annoni
Head of Digital Economy
Unit, Joint Research Centre,
European Commission

“We are only at the beginning of a rapid period of transformation of our economy and society due to the convergence of many digital technologies. Looking at the world of digital transformation, we live in an era that can be defined as the “Cambrian explosion of data”, and advanced data analytics are needed for us to navigate this world. AI is central to this change and offers major opportunities to improve our lives but ethical and secure-by-design algorithms are crucial to building trust in this disruptive technology. We also need a broader engagement of civil society on the values that need to be embedded in AI and the directions for future development.”

11 World Economic Forum. Artificial Intelligence and Robots. https://toplink.weforum.org/knowledge/insight/a1Gb0000000pTDREA2/explore/summary.
12 Abbany, Z. What good is AI for UN development goals? DW. May 16, 2018. https://p.dw.com/p/2xllV.
13 Schwab, K. *The 4th Industrial Revolution*. New York, NY: World Economic Forum. 2016.
14 Hager, G.D., et al. *Artificial Intelligence for Social Good*. Washington, DC: Computing Community Consortium; 2017. https://cra.org/ccc/wp-content/uploads/sites/2/2016/04/AI-for-Social-Good-Workshop-Report.pdf.
15 Adams, R.L. 10 Powerful Examples of Artificial Intelligence in Use Today. *Forbes*. 10 January 2017. https://www.forbes.com/sites/robertadams/2017/01/10/10-powerful-examples-of-artificial-intelligence-in-use-today/#5590a7c9420d.
16 Gray, A. 7 Amazing Ways Artificial Intelligence is Used in Healthcare. World Economic Forum. 20 September 2018. https://www.weforum.org/agenda/2018/09/7-amazing-ways-artificial-intelligence-is-used-in-healthcare/.
17 Cockburn, I.M., et al. The impact of artificial intelligence on innovation. National Bureau of Economic Research. Working paper No. 2449; 2018. http://www.nber.org/papers/w24449.pdf.

AI as a field brings together several domains—teaching, research, industry, and the media. AI discoveries and new technologies become core milestones in AI history, media reports influence public opinion, and the voices of various stakeholders influence policy. Breakthroughs bump up the hype (and research funding) surrounding AI, causing both excitement and concerns around adoption, including the potential for loss of jobs, privacy and control, misuse, and reaching "the singularity"—the point at which a machine can improve itself, independent of humans.[18, 19] With each advance, researchers, industry, and policy makers are asked to balance the transformational potential of AI with human safety and privacy.

While AI is a high priority on the agendas of policy makers and research and industry leaders and attracts daily news attention, it also lacks a universal definition. In the broadest terms, AI refers to the creation of machines (agents) that think and act like humans.[20, 21, 22, 23, 24] We can also differentiate between weak AI, i.e., machines that can simulate thinking within a narrow context to accomplish a specific task, and strong AI, i.e., intelligent machines that can reason. Yet, per Stanford's AI100 report,[25] "the lack of a precise, universally accepted definition of AI probably has helped the field to grow, blossom, and advance at an ever-accelerating pace."

The dynamic nature of AI is reflected in the so called "AI effect," which, according, to Hofstadter,[26] means that "AI is whatever hasn't been done yet." Today, emphasis is often on what AI can do: practitioners in AI focus on "the problems it will solve and the benefits the technology can have for society. It's no longer a primary objective for most to get to AI that operates just like a human brain, but to use its unique capabilities to enhance our world."[27] This focus on applications also means that many AI research outputs are found in non-AI journals or conferences. For these reasons, Elsevier took a "bottom-up" approach to characterize AI research, starting its analysis from the various domains in which AI is applied rather than relying on a single definition of AI.

The structure of the document is as follows. In the remainder of this chapter we give an overview of recent national policies in AI, reflecting the importance of AI to governments. Chapter one describes how we have, in the absence of a clear definition for AI, identified the relevant body of research published. In chapter two, we provide information on research areas that together make up AI. In chapter three, we use the research corpus from chapter one to identify global and regional trends as well as explore knowledge transfer. Chapter four takes a look at AI education, and chapter five reflects on ethics in AI. Finally, in the conclusion we suggest areas for further research.

18 Walsh, T. *Machines that Think*. Amherst, NY: Prometheus Books; 2018.
19 Tegmark M. *Life 3.0: Being Human in the Age of Artificial Intelligence.* New York, NY: Knopf; 2017.
20 McCarthy, J., et al. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. 1955. http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf.
21 Cellan-Jones, R. Artificial intelligence - hype, hope and fear. BBC News, p. 3. 16 October 2017. http://www.bbc.co.uk/news/technology-41634316.
22 Russel, S., Norvig, P. *Artificial Intelligence: A Modern Approach*. 3rd ed. Essex, UK: Pearson Education Limited; 2014.
23 Searle, J.R. Minds, brains, and programs. *Behav Brain Sci*. 1980;3(3):417-457. https://doi.org/10.1017/S0140525X00005756.
24 Encyclopedia Britannica. Artificial intelligence. Encyclopedia Britannica. 2018. https://www.britannica.com/technology/artificial-intelligence.
25 Stanford. *AI100 Report*. 2017. https://ai100.stanford.edu/2016-report/section-i-what-artificial-intelligence/defining-ai.
26 Hofstadter, D. *Gödel, Esher, Bach: An Eternal Golden Braid*. New York, NY: Basic Books, Inc. 1979.
27 Marr, B. The Key Definitions of Artificial Intelligence (AI) That Explain Its Importance. *Forbes*. 14 February 2018. https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/#6268887e4f5d

Introduction

National strategies and policies in AI

The capacity for AI research, technology, and application is seen as vital to national competitiveness, security, and economic strength. In the last two years alone, several countries and regions have developed and released AI strategic plans, essentially setting up a race to become the global leader in the field.[28] These strategies generally call for more investment to build the AI workforce and research and development capacity; anticipate how AI will change jobs and economies; and examine the social, economic, and ethical implications of AI.

AI policies developed as part of these national strategies vary widely from country to country, but focus on several elements: governance and regulation, ethics, security, and research, among others. Here we describe some of the specific research and innovation policies in AI, and the differences between countries and regions across the world. It is worth noting that where the United States and European governments seem to take a supportive role with AI policies that encourage research and industry, the Chinese government takes a more active role in determining the direction of AI in the country. Several countries, including Canada, the United States, China, Japan, and several in Europe allocate dedicated funding to achieving their strategies.

China issued its *New-Generation Artificial Intelligence Development Plan*[29] in July 2017, with key targets for the AI field through 2030 and the goal to become a world leader in AI theory, technology, and application. The three-year action plan focuses on strengthening its manufacturing capabilities and support systems and attracting and training a skilled AI workforce. The Chinese government budgeted over $2 billion for major R&D programs in 2018[30] and announced a $2.1 billion investment into an AI technology park in Beijing. In addition to these R&D investments, large datasets (consistent with the size of the Chinese population) and a relaxation of data regulations have created an advantage for China. Chinese corporate giants such as Baidu, Alibaba, and Tencent are also investing in AI research, alongside investment firms such as Sinovation Ventures,[31] which established an AI Institute in 2016.

In April 2018, the European Commission (EC) outlined a three-pronged approach to AI: increase public and private investment in AI, prepare for socio-economic changes, and ensure an appropriate ethical and legal framework. They also called for cooperation across member states as a "European AI Alliance." The EC announced that it would increase its AI research investment to €1.5 billion for the 2018-2020 period under the Horizon 2020 program. Per the commission, "this investment is expected to trigger an additional €2.5 billion of funding from existing public-private partnerships, for example, on big data and robotics."[32] The European Union (EU) member states also signed a Declaration of Cooperation on Artificial Intelligence[33] on issues such as research, socio-economic challenges, and legal and ethical frameworks. The importance of AI to the EC is visible through the Joint Research Centre's 2018 AI report, which investigates a broad range of industrial, business, and research activities (including patenting, frontier research, venture capital, start-ups, and public funded projects).[34]

28 Dutton, T. An Overview of National AI Strategies. *Medium Politics + AI*. 28 June 2018. https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd.
29 The State Council. The People's Republic of China. China issues guidelines on artificial intelligence development. 20 July 2017. http://english.gov.cn/policies/latest_releases/2017/07/20/content_281475742458322.htm.
30 China to spend over USD 2 billion in R&D this year. *The Economic Times*. 7 January 2018. https://economictimes.indiatimes.com/news/international/business/china-to-spend-over-usd-2-billion-in-rd-this-year/articleshow/62403032.cms.
31 Sinovation Ventures AI Engineering Institute. http://ai.chuangxin.com/.
32 European Commission. Artificial intelligence: Commission outlines a European approach to boost investment and set ethical guidelines. 25 April 2018. http://europa.eu/rapid/press-release_IP-18-3362_en.htm.
33 European Commission. EU Member States sign up to cooperate on artificial intelligence. 10 April 2018. https://ec.europa.eu/digital-single-market/en/news/eu-member-states-sign-cooperate-artificial-intelligence.
34 Craglia M. (Ed.), Annoni A., Benczur P. et al. 2018. *Artificial Intelligence: A European Perspective*, Luxembourg: Publications Office

Many AI strategies have also emerged at the national level in EU member states in recent years, resulting in a diversity of plans and approaches in the region. France recently declared AI a national priority[35] and announced a strategic plan For a *Meaningful Artificial Intelligence*.[36] In March 2018, Italy released *Artificial Intelligence at the Service of Citizens*[37] and the German government is due to release a national AI strategy in December 2018.[38] The United Kingdom (UK) published its Industrial Strategy[39] in November 2017 and its Artificial Intelligence Sector Deal in April 2018.[40] Other European countries that have recently released national strategies or reports on AI include Finland (*Finland's Age of Artificial Intelligence*[41]), Denmark (*New Strategy to Make Denmark the New Digital Frontrunner*[42]), and Sweden (*National Approach for Artificial Intelligence*[43]).

Interest in AI in the United States (US) was signaled by the release of a report from the National Science and Technology Council (*Preparing for the Future of Artificial Intelligence*[43]) in October 2016. The report noted that unclassified research on AI was being managed through the Networking and Information Technology Research and Development programme, supported by several federal funding agencies. At the time of the report, federal investment in unclassified AI research was estimated to be at US$1.2 billion and it was recommended that future investment should focus on basic research and long-term, high-risk initiatives, as the private sector investment in R&D would be limited. *The National Artificial Intelligence Research and Development Strategic Plan*[45] that accompanied the report set several objectives for federally funded AI research, such as ensuring effective human-AI collaboration, developing shared public datasets, and measuring and evaluating AI technologies through standards and benchmarks. In 2018, the White House hosted the "Artificial Intelligence for American Industry"[46] summit, which promoted a "free market approach to scientific discovery that harnesses the combined strengths of government, industry, and academia" and examined "new ways to form stronger public-private partnerships to accelerate AI R&D." AI was included as a priority area in FY19 budget, particularly funding for projects focused on transportation, healthcare, workforce training, and military applications.

35 AI for Humanity. French strategy for artificial intelligence. https://www.aiforhumanity.fr/en/.

36 Artificial intelligence: Making France a leader. 30 March 2018. https://www.gouvernement.fr/en/artificial-intelligence-making-france-a-leader.

37 The Agency for Digital Italy (AGID). Artificial Intelligence at the Service of Citizens. 2018. https://ia.italia.it/assets/whitepaper.pdf.

38 AI Hub Europe. Exclusive: German AI-Strategy Paper in English. 26 July 2018. http://ai-europe.eu/exclusive-german-ai-strategy-paper-in-english/.

39 HM Government. Industrial Strategy: Building a Britain Fit for the Future. 2017. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/664563/industrial-strategy-white-paper-web-ready-version.pdf.

40 Department for Business, Energy & Industrial Strategy, Department for Digital, Culture Media & Sport. Policy paper: AI Sector Deal. 26 April 2108. https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal#executive-summary.

41 Ministry of Economic Affairs and Employment. Finland's Age of Artificial Intelligence. 2017. http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf.

42 Ministry of Industry, Business and Financial Affairs. New Strategy to Make Denmark the New Digital Frontrunner. 30 January 2018. https://eng.em.dk/news/2018/januar/new-strategy-to-make-denmark-the-new-digital-frontrunner/.

43 Government Offices of Sweden. National Approach for Artificial Intelligence. 2018. https://www.regeringen.se/informationsmaterial/2018/05/nationell-inriktning-for-artificiell-intelligens/.

44 Executive Office of the President National Science and Technology Council Committee on Technology. Preparing for the Future of Artificial Intelligence. October 2016. https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

45 National Science and Technology Council, Networking and Information Technology Research and Development Subcommittee. The National Artificial Intelligence Research and Development Strategic Plan. October 2016. https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf.

46 The White House Office of Science and Technology Policy. Summary of the 2018 White House Summit on Artificial Intelligence for American Industry. 10 May 2018. https://www.whitehouse.gov/wp-content/uploads/2018/05/Summary-Report-of-White-House-AI-Summit.pdf .

Strategic planning in AI is underway in several other countries.[47] Canada became the first country to release a national AI strategy in 2017. The United Arab Emirates also launched an AI strategy in 2017, the first country to do so in the Middle East. In 2018, India released its national AI strategy,[48] while Japan unveiled its Artificial Intelligence Technology Strategy in March 2017[49], and the South Korean government announced a five-year plan to invest in and strengthen AI research and development. AI plans and programmes are in various stages of development in Malaysia, Singapore, and Taiwan. Mexico and Russia have released research priorities and strategic outlines while Tunisia and Kenya have formed task forces to examine the development of AI in Africa. Several Nordic and Baltic countries formed a regional collaboration in 2018 to develop AI capacity. Together, these efforts underscore the growing recognition by individual countries and regions of the potential impact of AI on society and human life and the need to develop knowledge and expertise in this field.

47 Dutton, T. An Overview of National AI Strategies. *Medium Politics + AI*. 28 June 2018. https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd.

48 NITI Ayog. Discussion Paper: National Strategy for Artificial Intelligence. June 2018. http://www.niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf.

49 Strategic Council for AI Technology, Artificial Intelligence Technology Strategy, March 2017, http://www.nedo.go.jp/content/100865202.pdf.

Dr. Zhiyun Zhao
National New-Generation Artificial Intelligence Development Research Center, Ministry of Science and Technology of the People's Republic of China, Institute of Scientific and Technical Information of China (ISTIC)

What is the China National New Generation Artificial Intelligence Development Research Center?
In July 2017, the State Council issued its "New-Generation Artificial Intelligence Development Plan," which proposed that a new Artificial Intelligence Planning and Promotion Office would be run through the Ministry of Science and Technology. The Ministry would be responsible for "promoting the construction of artificial intelligence think tanks and supporting various think tanks to carry out labor, [as] research on major issues of intelligence provides strong intellectual support for the development of artificial intelligence." To implement the plan, the Ministry of Science and Technology coordinated research strengths across relevant internal and external departments, and established the National New-Generation Artificial Intelligence Development Research Center. This Center is a high-end AI research platform established to accelerate AI development planning and strengthen the strategic research support of AI development on a national scale. By bringing together both domestic and foreign research forces, especially young AI talent, we will establish a stable and sustained strategic research team to further strengthen AI research and evaluation.

What is China's AI strategy? How can it be realized? How can it evolve to remain successful?
China's new AI strategy aims to establish the first mover advantage through top level and systematic AI deployment in three steps. By 2020, China's overall AI technology and application will be globally competitive. By 2025, we expect to achieve major breakthroughs in the basic theory of AI, and our AI technology and application will be among the world's best. By 2030, China will be the world's major innovation center in AI theory, technology, and application. The establishment of these objectives was based on the current strong foundation of AI development in China. China's AI development strategy puts forward an overall framework of "building a system, grasping dual attributes, adhering to the trinity, and strengthening the four major supports."[50] This strategy considers the current status of AI technology and the overall economic and social development of China.

What is China's AI policy? How is it determined? How can it adapt to the fast-changing AI landscape?
In order to follow through with our AI strategy and achieve our "three-step" goals, we have increased the resource allocation and special policies for AI. First, it is necessary to make full use of the existing funds and other stock resources, to increase the support of the central financial funds to guide multi-channel capital investment in the market, and to build several international leading innovation bases in the AI field. Second, we need to propose special safeguard measures through laws and regulations, ethical norms, key policies, intellectual property rights and standards, regulatory assessment, labor training, and popular science. We also need to integrate industrial policies, innovation policies, and social policies to achieve coordination of incentive development and rational regulation. Of course, we also fully realize that the continuous improvement and acceleration of AI development means that its impact on economic, social, legal, ethical, and other aspects cannot be clearly defined in the short term. Creating and implementing policy at such a fast pace is a global challenge.

50 State Council Issued Notice of the New Generation Artificial Intelligence Development Plan. 8 July 2017. https://flia.org/wp-content/uploads/2017/07/A-New-Generation-of-Artificial-Intelligence-Development-Plan-1.pdf

Chapter 1

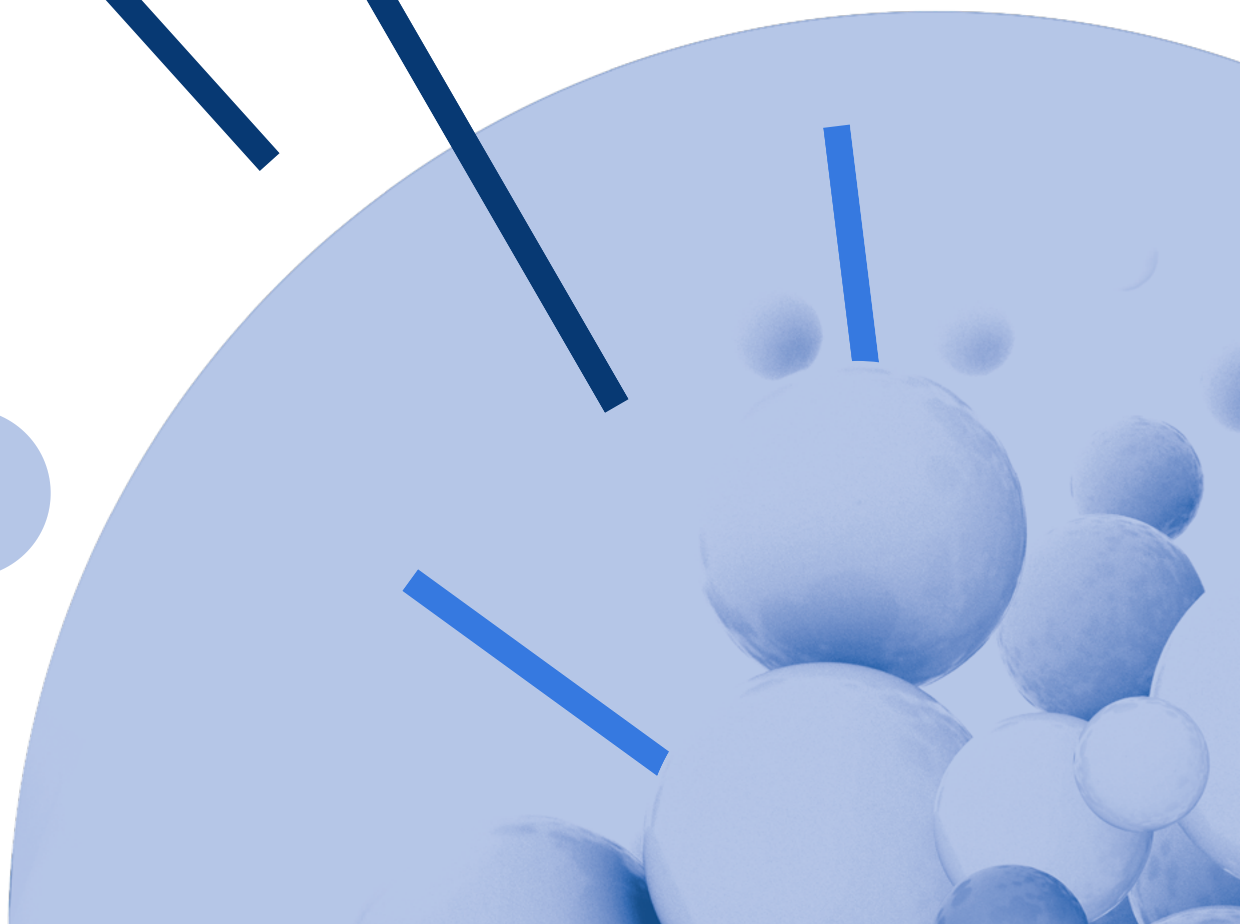
Identifying Artificial Intelligence research

The AI field has multiple definitions, but lacks a universally agreed understanding. AI means different things to different people: there are more differences than commonalities in how AI is spoken about in education, research, industry, and the media. This chapter describes our methodology for characterizing the field and determining what is in and what is out of scope.

Highlights

More than 600,000 AI scholarly publications extracted using AI technologies.

Core AI keywords with a large proportion of AI scholarly publications include Back-propagation Neural Network, Genetics-based Machine Learning, Cohen-Grossberg Neural Networks, Back-propagation Algorithm, Neural Networks Learning.



Using AI to define AI

Examining which words are used to talk about AI from the perspectives of teaching, research, industry, and the media, we see that there is not one common definition for AI: its meaning differs depending on the outlook with which it is approached.[51,52,53]

In many studies analyzing research dynamics, either a journal category or a keyword approach verified by experts is used to define a research area.[54] Extracting keywords from bodies of text from different perspectives (see Figure 1.1) allows us to reduce personal bias as well as take a view of the field that goes beyond research only. The width and breadth of the AI field, combined with its undefined and pervasive nature, however, makes manual approaches challenging and time-consuming. Therefore, following consultation with external AI experts, we chose to employ supervised AI techniques to further gain speed and efficiency. This methodology also allowed us to maintain the width and breadth of AI keywords, while sharpening the precision of the resulting research corpus of publications. The details of our methodology are explained in separate technical documentation on the Elsevier AI Resource Center.[55]

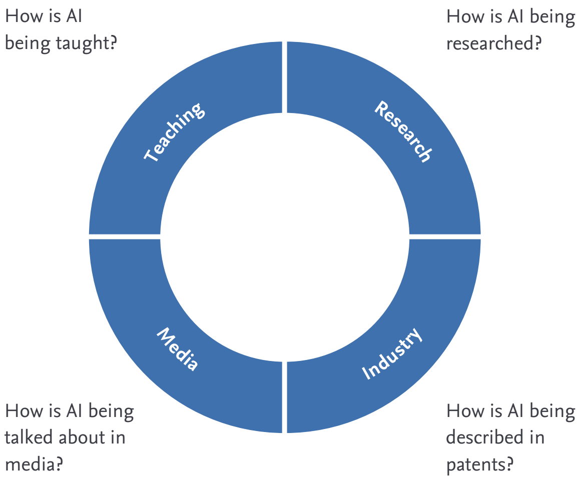


FIGURE 1.1
We extracted keywords from texts reflecting four perspectives on AI to define the field.

51 McKinsey Global Institute. Artificial Intelligence – The Next Digital Frontier? June 2017. https://www.mckinsey.com/~/media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx.
52 Clarivate Analytics. Artificial Intelligence – The Innovators and Disruptors for Next Generation Digital Transformation. 11 September 2017. https://clarivate.com/blog/ip-solutions/artificial-intelligence-innovators-disruptors-next-generation-digital-transformation/.
53 OECD. OECD Science, Technology and Industry Scoreboard 2017: The Digital Transformation. Paris, France: OECD Publishing; 2017. https://doi.org/10.1787/9789264268821-en.
54 See e.g., Elsevier. A Global Outlook on Disaster Science. 2015. https://www.elsevier.com/__data/assets/pdf_file/0008/538091/ElsevierDisasterScienceReport-PDF.pdf; Elsevier. Sustainability Science in a Global Landscape. 2017. https://www.elsevier.com/__data/assets/pdf_file/0018/119061/SustainabilityScienceReport-Web.pdf; or the above-mentioned reports from McKinsey Global Institute and Clarivate Analytics.
55 Elsevier. Artificial Intelligence Resource Center. https://www.elsevier.com/connect/ai-resource-center.

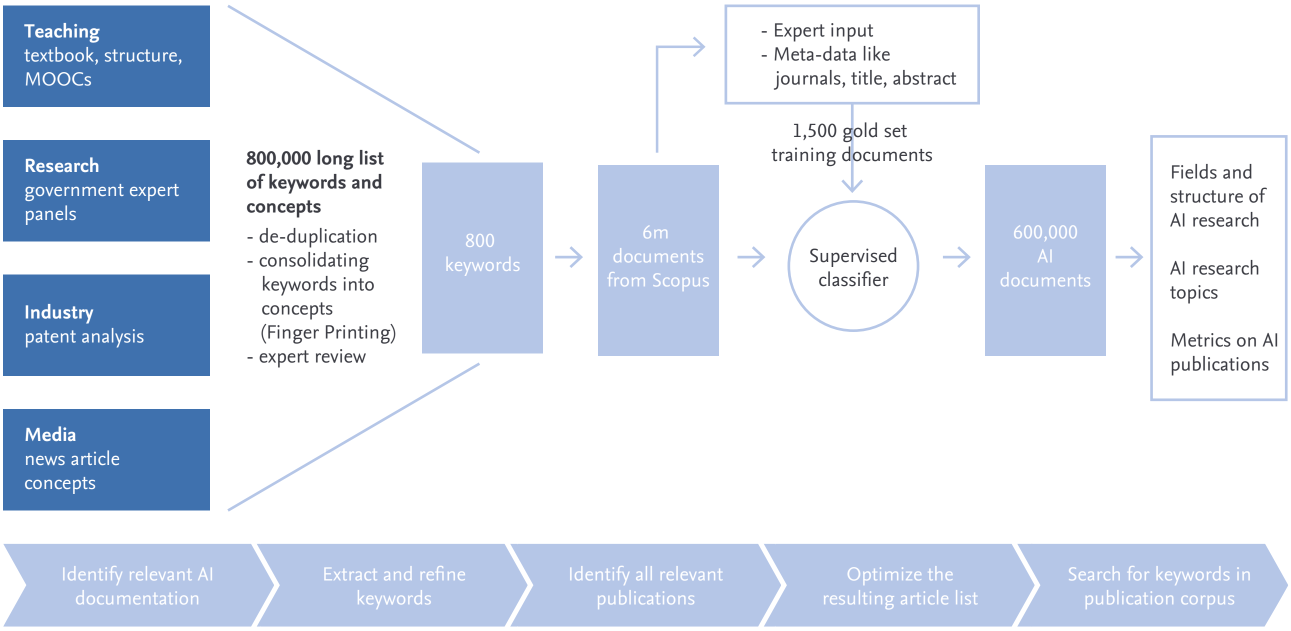


FIGURE 1.2
Process followed for selecting relevant AI publications for our analyses.

We mined the text and structure of representative books, the syllabi of massive open online courses (MOOCs), patents, news items and included keywords from research experts. To identify meaningful concepts, we used the Elsevier FingerPrint Engine™, which reduced this long list to 20,000 concepts. The list of concepts was shortened to 797 unique keywords following manual review (Figure 1.2).

We searched for each keyword in the titles, abstracts, and keywords of documents included in a Scopus May 2018 dataset, retrieving 5.7 million unique documents, including many false positives related to application terms (e.g., “finite elements”), broad terms (e.g., “ethical values”), or similar terms from other fields (e.g., “neural networks” in biology).

The Elsevier Fingerprint Engine™ identifies concepts and their importance in any given text by using a wide range of thesauri and data-driven controlled vocabularies covering all scientific disciplines, and by applying a variety of natural language processing (NLP) techniques. The advantage of using this technology is that the resulting terms are of high quality and more representative than standard sets of keywords, which often contain duplicates, synonyms, and irrelevant terms

We used supervised machine learning with further expert input on the training data set to eliminate false positives from the corpus while retaining relevant AI documents. The 797 keywords were ranked as high, medium, or low with regards to relevancy to the core field of AI and were assigned a respective weight. Figure 1.3 provides examples of the keywords and their ratings, alongside their share of AI and non-AI publications.

We employed a standard machine learning approach to train and evaluate our classifier model. In parallel, 1,500 documents were manually classified by internal experts as either "AI" or "non-AI" to use as reference and training input for the algorithm to determine the classification. This gold set of documents was randomly partitioned to keep a subset of known answers out of the example data used to train the model. These holdout examples were then fed into the trained classifier to obtain predictions for those documents. These predictions were then compared to the known class for each example, revealing that the model identified AI documents with 85% precision compared to the set of documents initially classified by AI experts. The complete set of 5.7 million documents was run through the model to generate predictions that were used to reduce the number identified as AI documents to approximately 600,000.

In summary, our method still requires a pre-defined, trusted input of AI documents as either a clustering starting point or a gold standard for classifier training. Those inputs could be sets of articles/conference papers or a group of authors. We used a keyword and a supervised classification approach to identify them. Further research might evaluate and optimize starting points and algorithms to shape and structure the field more sharply and broadly.

In line with recommendations of leading associations, like CRA[56] or Informatics Europe,[57] we stress that the results should not be used to assess individual researchers' productivity or performance. Rather, the metrics provide aggregate, descriptive trends and findings at the institutional or country level.

56 Computing Research Association. https://cra.org/.

57 Esposito, F., et al.; for the Informatics Europe Research Evaluation Working Group. Informatics Research Evaluation. 20 October 2017. http://www.informatics-europe.org/component/phocadownload/category/10-reports.html?download=63:research_evaluation_draft_20oct17.

More than 600,000 AI scholarly publications extracted using AI technologies.

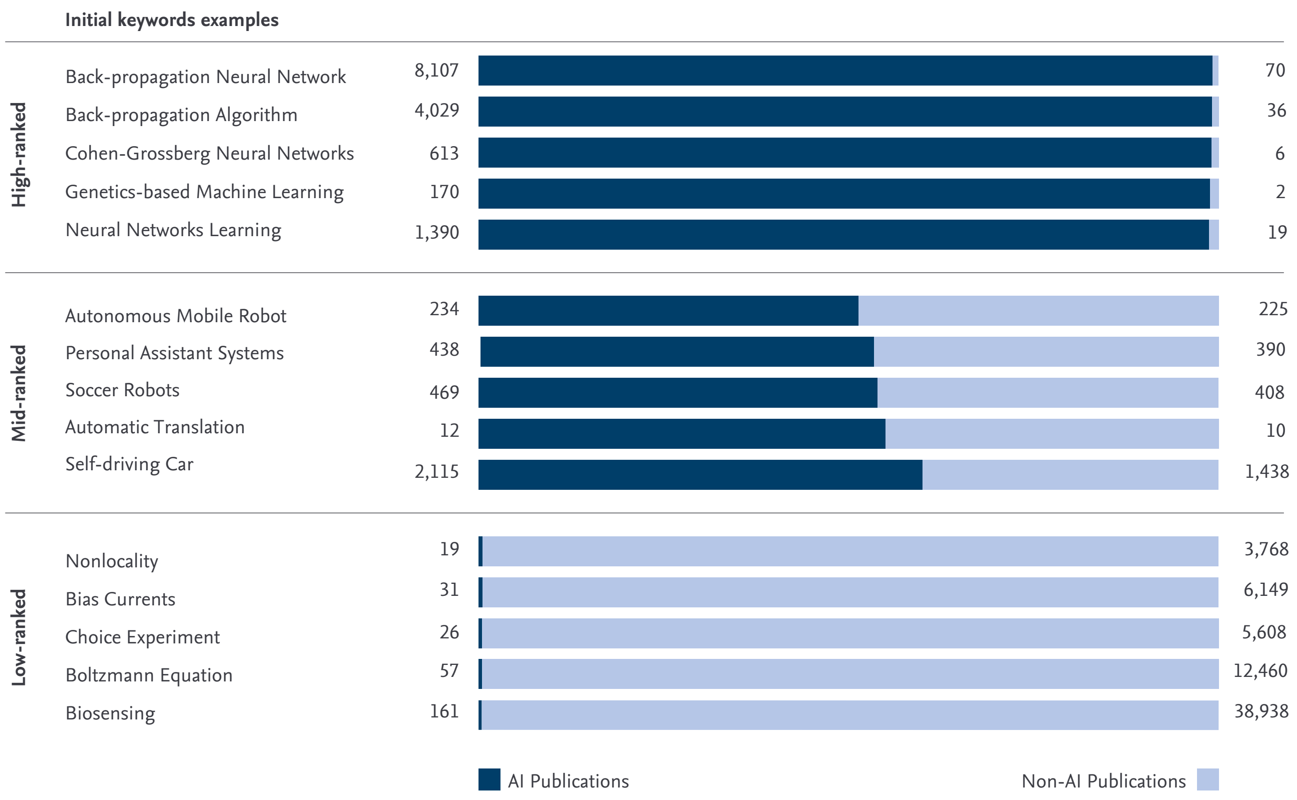


FIGURE 1.3
High-, mid-, and low-ranked keywords with number of AI and non-AI publications, 1998-2017; sources: Scopus and Elsevier Fingerprint Engine.

Chapter 2

Artificial Intelligence: a multifaceted field

The vocabulary used by actors from each perspective (teaching, research, industry, and media) reveals more divergence than commonality, while comparing keyword co-occurrences along the document set reveals the global structure of the field of AI in terms of subfields. This chapter presents an overview of our methodology and key findings on the composition of AI.

Highlights

Keywords shared across all 4 perspectives:

– Artificial Intelligence
– Deep Learning
– Machine Learning
– Neural Network
– Reinforcement Learning
– Speech Recognition

SECTION 2.1

Artificial Intelligence focuses on: Search and Optimization, Fuzzy Systems, Natural Language Processing and Knowledge Representation, Computer Vision, Machine Learning and Probabilistic Reasoning, Planning and Decision Making, and Neural Networks.

SECTION 2.2

2.1 Teaching, research, industry, and media perspectives

In the previous chapter, we explained how we selected keywords and concepts describing AI, and how we used these to find the relevant research corpus in Scopus. Comparing the keywords from the four different perspectives, we find little overlap in the way AI is spoken about in education, research, industry, and the media. The four perspectives only share six broad and general keywords, most of which relate to learning: "Artificial Intelligence," "Deep Learning," "Machine Learning," "Neural Network," "Reinforcement Learning," and "Speech Recognition." Figure 2.1 shows that each perspective has at least 30% "unique" keywords, with up to 69% in industry, suggesting that the understanding of AI varies by perspective. This raises a question about communication: how can it be effective in the absence of a common language?

Keywords describing societal issues or ethics appear only in the perspectives of teaching and media, possibly due to government mandates (course curricula) and a new emerging two-way dialogue between society and research (social media). Industry differentiates strongly between software and hardware and media focuses on "strong AI" with its "own personality." The physical embedding of AI and the idea of a personalized, "strong" AI is one driver for AI hype in the media. Teaching provides broad overviews of approaches, architectures, or tools. Many experts in research currently focus on neural networks.

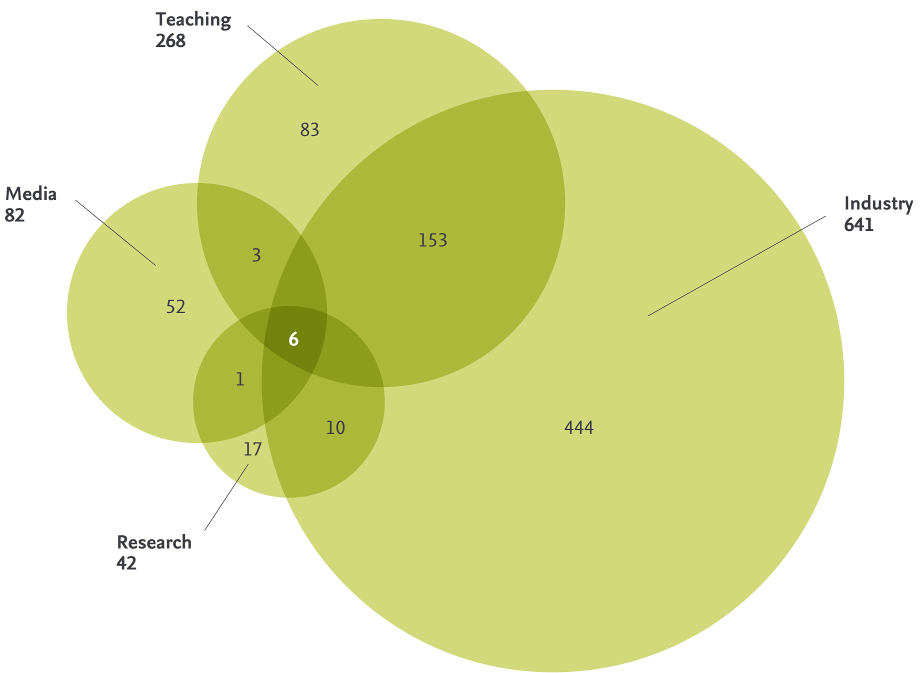


FIGURE 2.1
Keyword mapping (number of keywords) between AI perspectives.

2.2 Seven AI research clusters

We aimed to provide more depth to our subsequent analyses by structuring AI into research areas, using an unsupervised clustering technique.[58] This approach maps the keywords of all perspectives into clusters and illustrates their connections, based on co-occurrence within the documents. Co-occurrence indicates that those clusters do not stand alone, but strongly relate to each other, e.g., neural networks in a computer vision document. How do capabilities connect with each other and to application fields? The resulting graph illustrates the subfields of AI (Figure 2.2) and their connections through co-occurence in scholarly publications. On the Elsevier AI Resource Center,[59] the graph is interactive, allowing users to browse individual connections and clusters, by region and over time.

As shows in Figure 2.2, AI seems to cluster around the areas of Search and Optimization, Fuzzy Systems, Natural Language Processing and Knowledge Representation, Computer Vision, Machine Learning and Probabilistic Reasoning, Planning and Decision Making, and Neural Networks. Societal application fields, such as self-driving cars or robotics, are embedded into Planning and Decision Making as they have fewer underlying publications. The clusters seem to focus on statistics-based AI. Knowledge-based capabilities, such as "Ontologies or Semantics," do not form a cluster on their own, but are embedded in other clusters, predominantly in "Natural Language Processing and Knowledge Representation." Further research might investigate the sensitivity of this approach to the number of keywords and related publications in terms of normalized proportions over time. The strong growth of publications in recent years within the learning system field might outweigh knowledge-based approaches from more than 15 years ago.

Figure 2.2 illustrates the breadth of industry keywords (green), especially in the areas of "Fuzzy Systems" and "Computer Vision," whereas specific research keywords appear in "Neural Networks," teaching keywords in "Search and Optimization," and media keywords in fields such as "Planning and Decision Making" and "Natural Language Processing and Knowledge Representation."[60] The relatively low proportion of media-driven keywords could indicate that these are not key AI research fields, or that they are still in their research infancy, representing only a fraction of AI documents.

The online interactive graph[61] allows the exploration of connections and co-occurrences through time. For instance, it shows the intensification of the two clusters "Machine Learning and Probabilistic Reasonin" and "Neural Networks." It also reveals that the clusters "Deep Learning" in 2003 and "Swarm Intelligence" in 2000 have no co-occurring keywords but grow to become visible nodes on the graph in more recent years. Co-occurrences illustrate that certain learning system and neural network approaches are predominantly used in specific application fields, like "Recurrent Neural Networks" with "Natural Language Processing and Knowledge Representation." They also show that the keyword "Convolutional Neural Networks" is linked with "Computer Vision" and "Collaborative Filtering" with "Recommender Systems." Some connections indicate potential hierarchical relations, such as "Artificial Intelligence" co-occurring with the keyword "Neural Network," and further co-occurring with specific forms of neural networks.

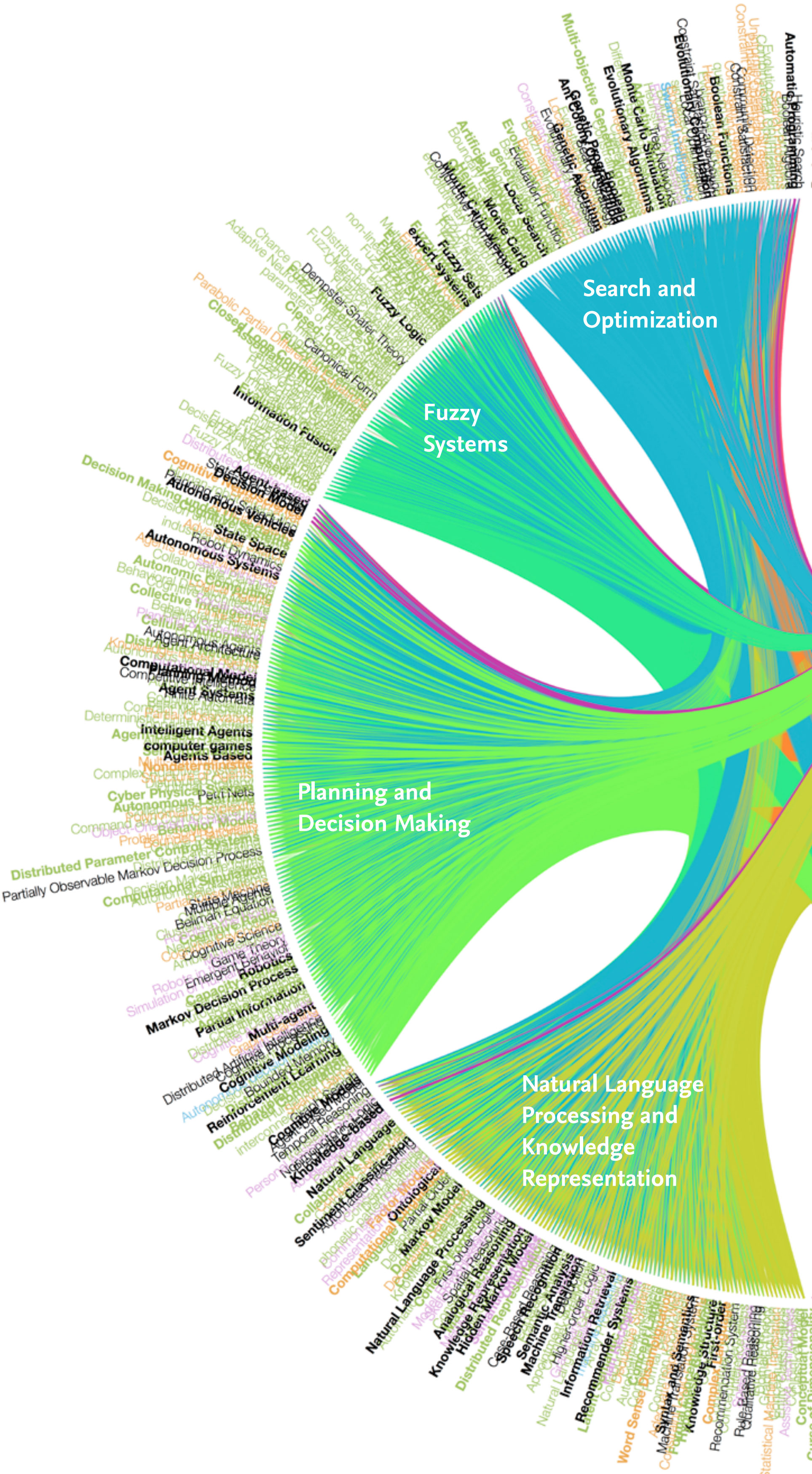
58 Louvain clustering:
https://perso.uclouvain.be/vincent.blondel/research/louvain.html
"The Louvain method is a simple, efficient and easy-to-implement method for identifying communities in large networks...The method is a greedy optimization method that attempts to optimize the 'modularity' of a partition of the network...The original idea for the method is due to Etienne Lefebvre who first developed it during his Master thesis at UCL (Louvain-la-Neuve) in March 2007...The method was first published in: 'Fast unfolding of communities in large networks,' Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre, Journal of Statistical Mechanics: Theory and Experiment 2008 (10), P10008 (12pp) doi: 10.1088/1742-5468/2008/10/P10008. arXiv: http://arxiv.org/abs/0803.0476."

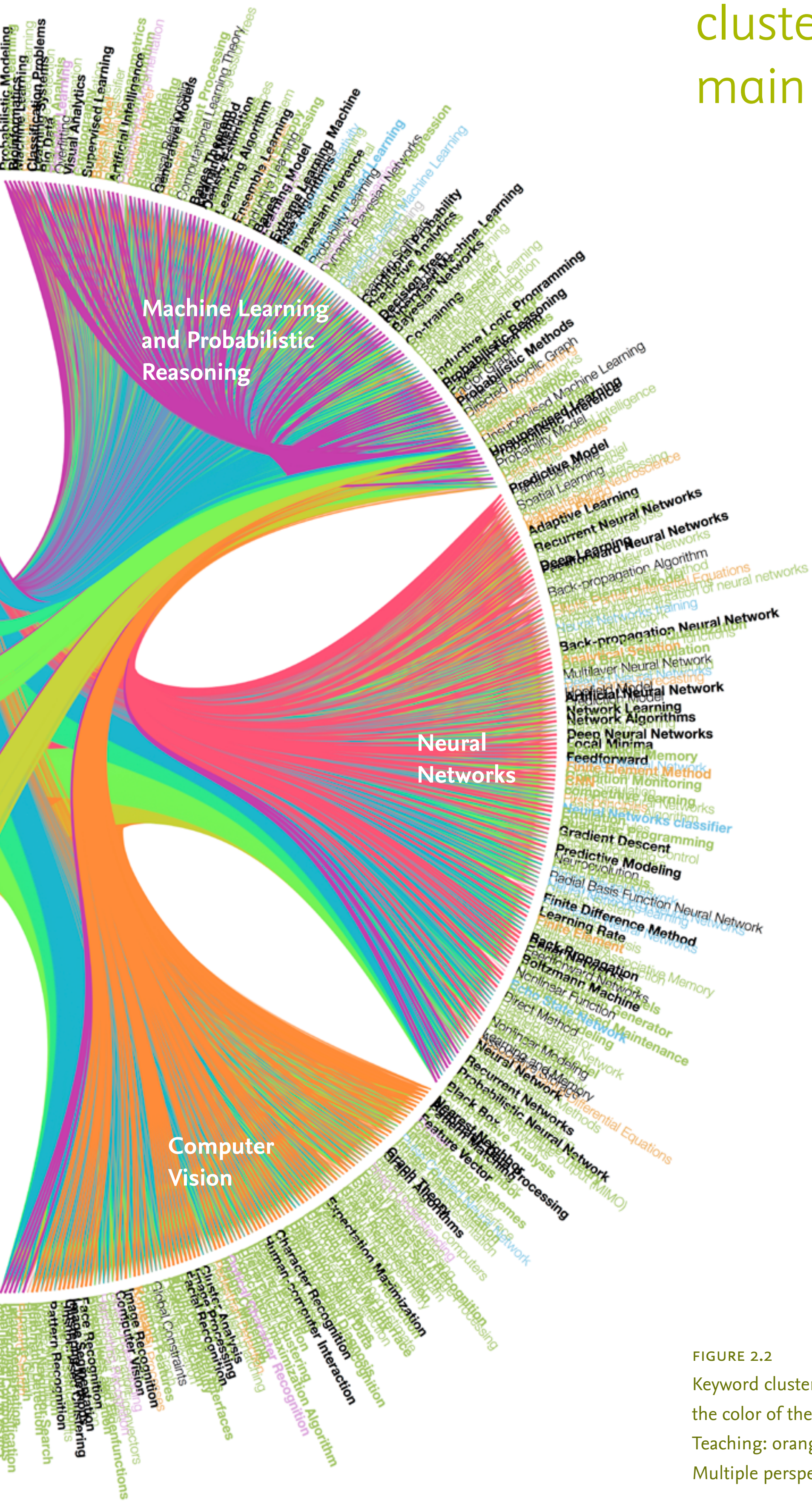
59 Elsevier. Artificial Intelligence Resource Center.
https://www.elsevier.com/connect/ai-resource-center.

60 Learn more about these in Science Direct Topic Pages:
https://www.sciencedirect.com/topics/index
Fuzzy Systems: https://www.sciencedirect.com/topics/chemical-engineering/fuzzy-systems; Speech Recognition:
https://www.sciencedirect.com/topics/neuroscience/speech-recognition;
Computer Vision: https://www.sciencedirect.com/topics/food-science/computer-vision-technology (specific application field) or Face recognition:
https://www.sciencedirect.com/topics/neuroscience/face-recognition;
Learning Systems: https://www.sciencedirect.com/topics/chemical-engineering/learning-systems;
Neural Networks: https://www.sciencedirect.com/topics/veterinary-science-and-veterinary-medicine/neural-network-software.

61 Elsevier. Artificial Intelligence Resource Center.
https://www.elsevier.com/connect/ai-resource-center.

In summary, our co-occurrence analysis reveals a sub-structure of the AI field, determined by its document corpus and keyword selection. Those might influence the weighting and share of subfields, such as knowledge-based fields. Further research might explore normalized approaches to compare subfields per year and investigate the sensitivity of keywords to the structure of the field. In chapter 3 we will use these clusters to present global and regional trends in AI.





The AI research field clusters around seven main research areas.

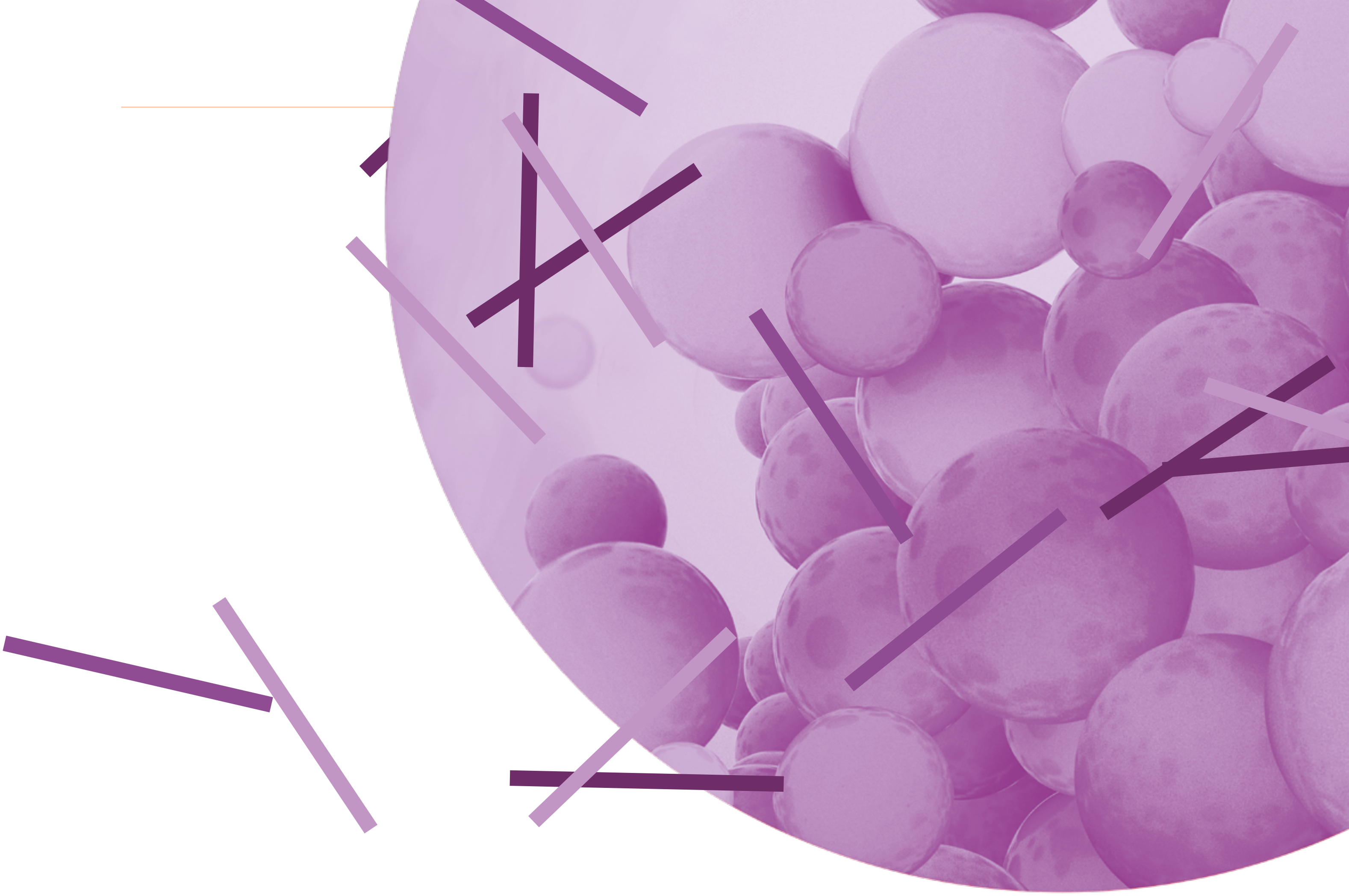
FIGURE 2.2

Keyword clusters and co-occurrences in the AI field, 2017; the color of the keyword represents its originating perspective: Teaching: orange, Research: blue, Industry: green, Media: pink, Multiple perspectives: black. source: Scopus.

Chapter 3

Artificial Intelligence research growth and regional trends

The purpose of this chapter is to identify and illustrate developments in AI research for three large geographies – China, Europe, and the United States. It investigates research outputs (including articles, conference papers, preprints, and competitions) and the resulting impact of scholarly publications, measured in the form of citations and downloads. Cross-sector research collaborations and researcher mobility analyses illustrate knowledge transfer. Analyses of subject fields, publication sectors, and top institutions help understand growth drivers and key players in the global research arena.



Highlights

AI research publications have grown by 12.9% annually over the last 5 years.

SECTION 3.1

arXiv preprints in core AI categories have grown 37.4% annually over the last 5 years, and especially fast in “Machine Learning” and “Computer Vision and Pattern Recognition”.

SECTION 3.1

China drives a lot of the global AI growth in publications, and also shows strong increases in citation impact.

SECTION 3.2

China has a strong focus on Computer Vision. Robotics belongs to Machine Learning and Probabilistic Reasoning in Europe and the United States.

SECTION 3.2

Over 70% of recent corporate AI research in the United States is published as conference papers. Academic-corporate collaborations in the United States account for 9% of AI publication, with high volume and citation impact from Microsoft and IBM.

SECTION 3.4

3.1 Global trends in AI research

Counting peer-reviewed publications is a common and easily understood measurement of research output. This section aims to give an overall view on all types of scholarly output, indexed by Scopus, namely journal articles (further referred to as articles), conference papers, and others, like review or survey papers. The following analysis is based on the refined corpus of more than 600,000 AI publications from 1998 to 2017, retrieved from Scopus (May 2018) following the method explained in chapter 2. In this chapter, we also examine preprints, conferences, and competitions.

Comparators selection and rationale

Many countries have recognized AI as an innovation driver. For a comprehensive global view of the dynamics of the AI field, we selected comparable regions over a period of 20 years (1998-2017) for analyses of various research dimensions (e.g., output, number of researchers, funding). As AI is now included as a key topic within innovation and research policies in many countries, it was important that the regions chosen for analysis connect to defined policy spheres. This consideration led us to choose Europe, including the 28 European Union Member States and affiliated countries under the EU's Horizon 2020 research funding program, such as Turkey and Israel. As the analyses illustrate, emerging countries like India, or smaller countries like Singapore, are not less relevant for a comprehensive view on AI but would require a different comparative structure.

The graph in Figure 3.1 illustrates the overall growth of the AI research field with now approximately 60,000 publications per year. Globally, the field of AI has shown strong growth of 12.9% in the last 5 years. Many AI historical timelines exist in literature, highlighting key events and discoveries along the 60-year journey of the field, including the "AI winters,"[62] understood as periods of disillusionment with the technology. From 2005 onwards for instance, research in neural networks starts winning vision and speech competitions, and by 2009 is dominant against some of the benchmark sets.[63] Around 2014-2015, several good review (survey) papers on deep learning start to appear.[64]

The development of the AI field can be seen as occurring in four phases of five years each, with the new economy and Internet emerging around 2000 alongside several of today's corporate players, like Amazon or Google. The Think Tank Eurasia Group and Sinovation Ventures[65] and Dr Kai-Fu Lee[66] identify four areas of AI: Internet AI (recommender systems), Business AI (fraud detection, financial forecasting), Perception AI (smart devices), and Autonomous AI (new hardware applications, like self-driving cars).

AI research publications have grown by 12.9% annually over the last 5 years

62 Wikipedia. AI Winter. https://en.wikipedia.org/wiki/AI_winter.
63 Computer Vision. http://people.idsia.ch/~juergen/vision.html.
- The NORB Object Recognition Benchmark. https://cs.nyu.edu/~ylclab/data/norb-v1.0/.
- The CIFAR Image Classification Benchmark. http://www.cs.toronto.edu/~kriz/cifar.html.
- The MNIST Handwritten Digits Benchmark. http://yann.lecun.com/exdb/mnist/.
- The Weizmann & KTH Human Action Recognition Benchmarks. http://www.nada.kth.se/cvap/actions/.
- Chinese characters from the ICDAR 2013 competition. http://www.nlpr.ia.ac.cn/events/CHRcompetition2013/competition/Home.html.

64 Historical overviews: Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Networks*. 2015;61:85-117. https://doi.org/10.1016/j.neunet.2014.09.003; Review (survey) article: LeCun, Y., et al. Deep learning. Nature. 2015;521:436-444. http://www.nature.com/articles/nature14539.
65 Eurasia Group, Sinovation Ventures. China embraces AI: a close look and a long view. December 2017. https://www.eurasiagroup.net/files/upload/China_Embraces_AI.pdf.
66 Lee, K-F. *AI Superpowers: China, Silicon Valley, and the New World Order*. New York, NY: Houghton Mifflin Harcourt; 2018. https://aisuperpowers.com/about/about-the-book.

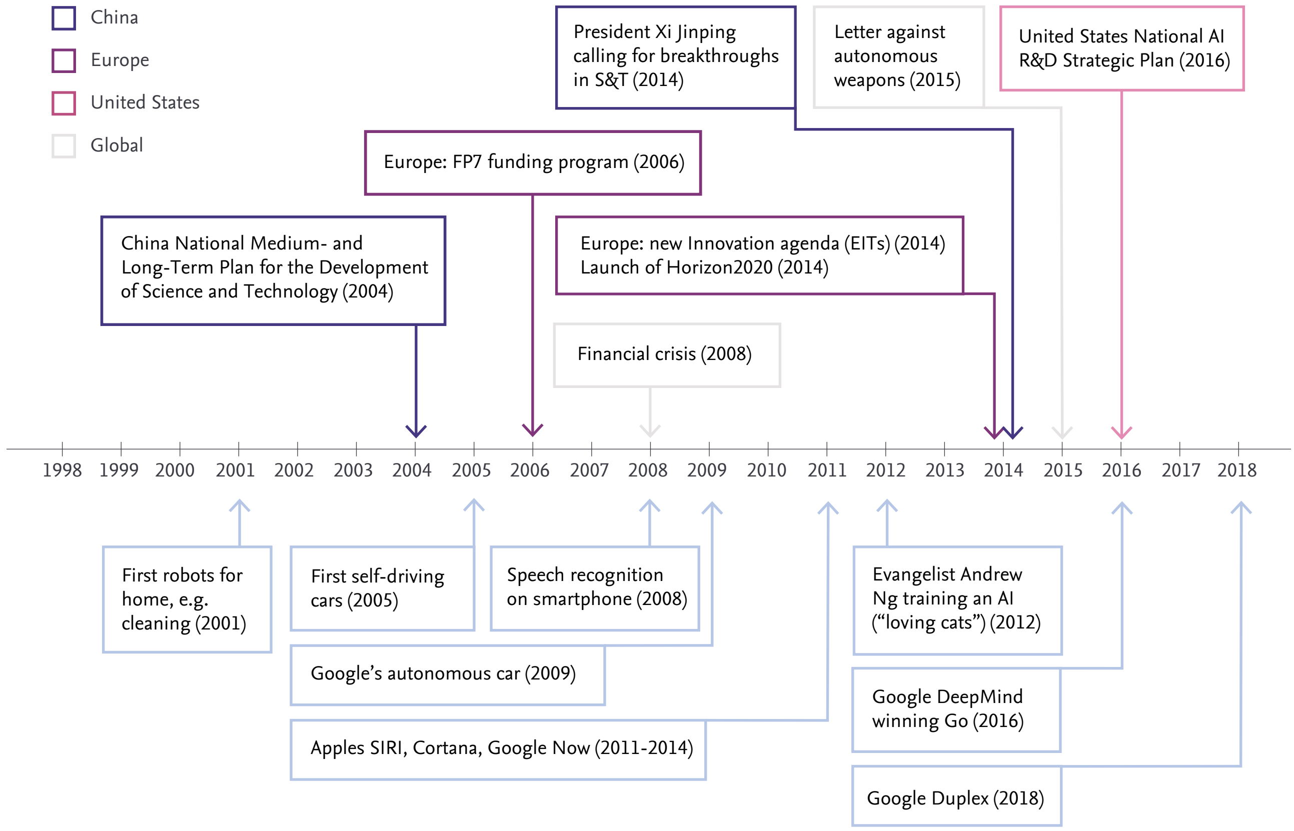


FIGURE 3.1
Selected AI-relevant policies and events (upper panel) and technology breakthroughs (lower panel), 1998-2018.

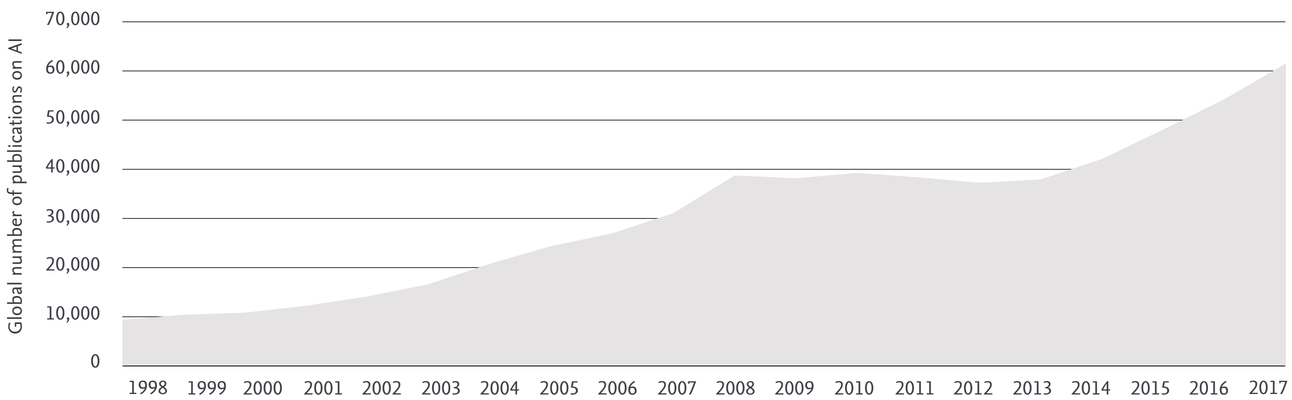


FIGURE 3.1
Annual number of AI publications (all document types), 1998-2017; source: Scopus.

The growth of research in the AI general capabilities of computer vision, neural networks, and machine learning systems is also apparent in the growth of publications (e.g., articles and conference papers) by co-occurrence cluster as illustrated in Figure 3.2. These research fields seem to explain the steep increase in publications after 2012. From the AI ecosystem, we see the rise of graphical processing units (GPUs) and the launch of ImageNet in 2012, a big open database with image training data that might have helped ignite this development.

Diachronic development in the number of publications by cluster do not show big differences between articles and conference papers. While the field of "Computer Vision" seems to benefit from developments in "Machine Learning and Probabilistic Reasoning" and "Neural Networks," "Natural Language Processing and Knowledge Representation" and other capabilities are less affected.

Preprints are another mechanism for disseminating AI research, and are typically used to circulate preliminary research outputs pending formal publication. arXiv is a popular academic preprint repository that has become an increasingly important channel for research dissemination in many fields of science and mathematics.[67] To examine trends in AI preprints over time, we first needed to determine which preprints should be considered AI research. Including arXiv categories obviously related to AI (e.g., cs.AI – Artificial Intelligence or stat. ML – Machine Learning) would miss important categories like computer vision and pattern recognition. Therefore, relying on titles and abstract text from arXiv, we used a refined list of 142 keywords and 12 arXiv subject areas designated by experts as having high relevance to the field of AI.

Within the context of the growth of the arXiv corpus, preprints in the 12 core AI subject areas have grown significantly as a percentage of the number of preprints in arXiv as a whole. In 1998, these 12 categories together account for only 149 preprints, or 0.62% of all preprints submitted to the arXiv repository. With gradual increases from 1998 to 2014, this percentage then rises sharply starting in 2015; in 2017, preprint submissions in these 12 categories account for more than 12% of all preprints submitted to arXiv.

Looking at the arXiv preprints submitted to the 12 core AI subject categories, we attempted to discern changes to submission patterns. Have AI researchers focused on different types of AI research over time, based on the number of preprints submitted to each subject category? Figure 3.3 depicts the proportion of preprints submitted to each category over time.

arXiv preprints in core AI categories have grown by 37.4% annually over the last 5 years.

67 Cornell University Library. arXiv monthly submission rates. https://arxiv.org/stats/monthly_submissions. Accessed 3 September 2018.

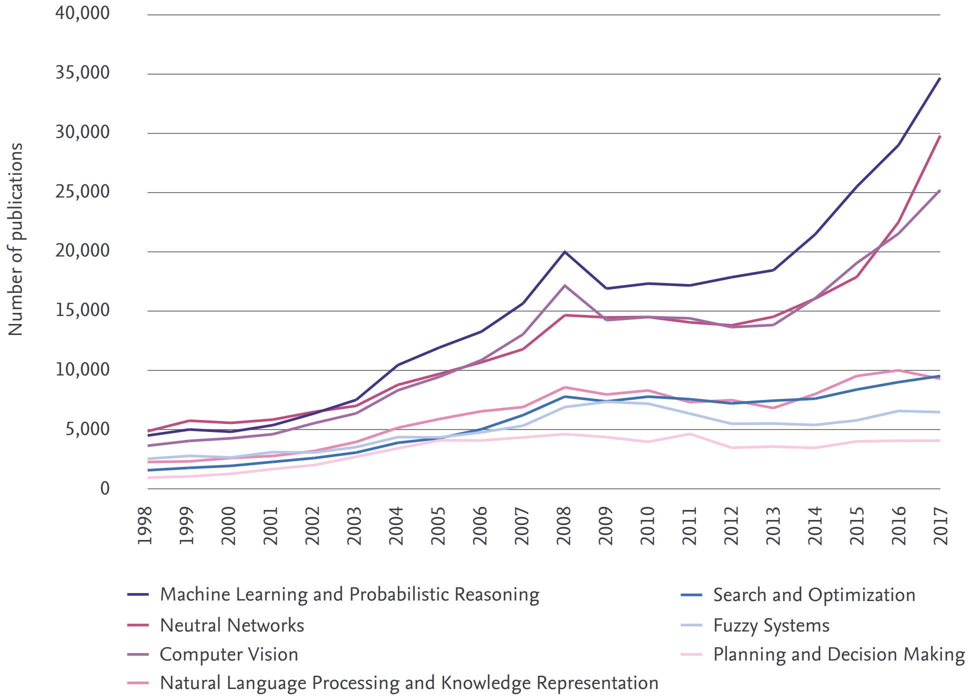


FIGURE 3.2
Annual number of AI publications by keyword co-occurrence cluster (all document types), 1998-2017; sources: Scopus and Elsevier clustering.

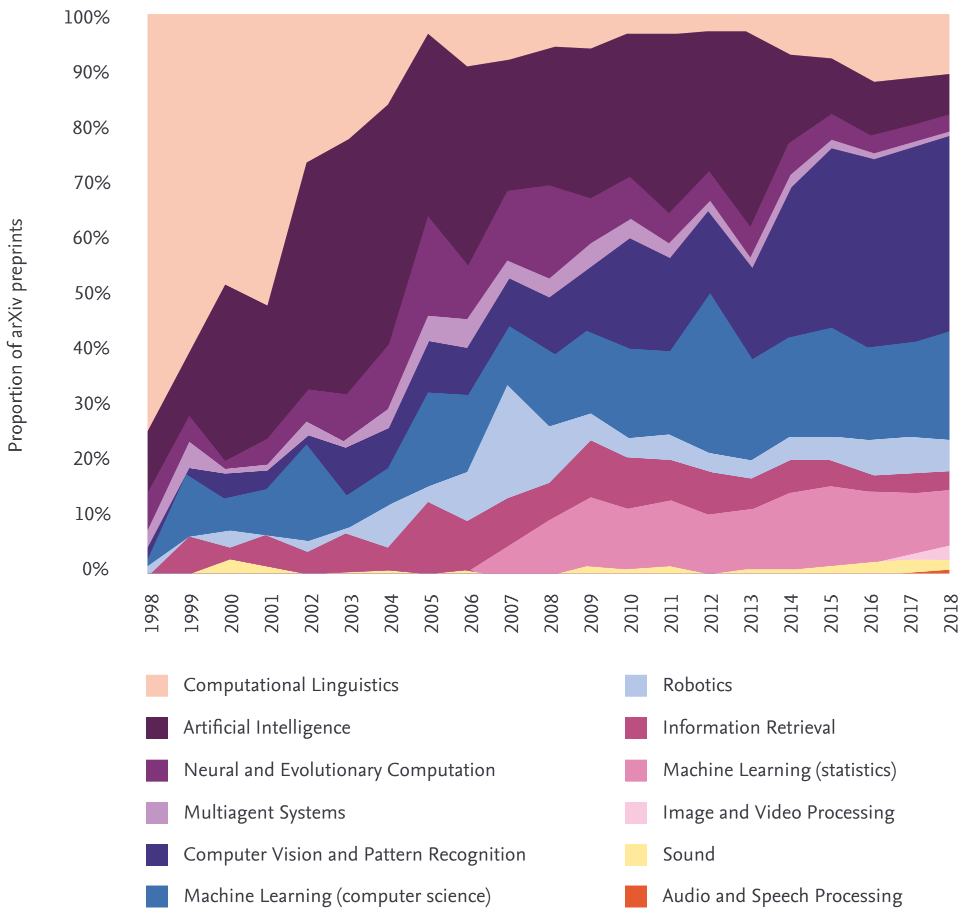


FIGURE 3.3
Proportion of arXiv preprints submitted in core AI categories, per category, 1998-2017; source: arXiv.

The volume of preprints in “Machine Learning” and “Computer Vision and Pattern Recognition” has grown rapidly in recent years.

The analysis of arXiv preprints in any of the 12 core AI subject areas shows dramatic growth in content relating to these topics, even relative to the growth of arXiv itself. Preprints in subject areas relating to core AI concepts account for 11.6% of all arXiv content in 2017, and 15.1% of submissions to date for 2018—a dramatic change from only a few years ago (2015: 5.61% of all arXiv content). This growth might be attributable to increased attention, funding, and research in the core AI areas, but it might also be indicative of the rise of arXiv as an important and trusted tool for research dissemination in these areas, as large AI research labs like Google DeepMind adopt the platform.

Research focus has likely shifted within the core AI fields over the past 20 years. More traditionally, computational linguistics and natural language processing research dominates arXiv submissions within these subject areas in 1998 (112 of the 149 papers submitted in all 12 categories, or 75.2%). While that area is still a factor in the AI research landscape, the arXiv data also points to a dramatic rise in the fields of computer vision and pattern recognition (from 1.3% of core AI submissions in 1998 to 32.7% in 2018) and machine learning (1.3% in 1998 to 17.8% in 2018)—both of these areas focus on the application of deep learning technologies.

Additionally, platforms like arXiv seem to be increasing the specificity allowed to researchers by adding new and more precise subject area designations, for example, distinguishing between statistics and computer science research in machine learning (started in 2007, 10.8% in 2018) or adding subject categories (“Computer Science - Sound” was added in 2004, and both “Audio and Speech Processing” and “Image and Video Processing” were both added in 2017).

Both arXiv preprint and Scopus publication analyses illustrate the evolution of the AI field, based on areas the platforms’ researchers are focusing on. While more generic terms like “Artificial Intelligence” see their submission rates erode on arXiv over time, they are actually emerging as umbrella terms.

Dr. Roberto M. Cesar Jr.
Adjunct Coordinator, São Paulo Research Foundation (FAPESP), Brazil

AI research output: beyond traditionally published papers

"AI and machine learning have attracted increasing attention in recent years, building into a kind of unforeseen revolution that has re-organized the scientific community, private sector, government, and society. Many intellectual tasks are currently being automated by AI processes, reflecting a culmination of the efforts and advances made across many different scientific communities (including computer scientists, engineers, and neuroscientists, among many others) working in research institutions and companies all over the world.

AI open-source libraries and training data sets are being produced, shared, and used interchangeably by researchers, programmers, and students from various disciplines. To better understand this phenomenon, it is important to recognize that AI and machine learning methods typically involve four fundamental elements:

1. Learning and classification algorithms.
2. Data to train and to evaluate the algorithms.
3. Data scientists to code, set up the software, and prepare the data.
4. Computer hardware to store and run the code.

The unique characteristics of the AI field make it challenging to evaluate its development. Traditionally, research advances in computer science and related fields are disseminated as papers published in journals and conference proceedings. Thus, commonly used research performance indicators have included the number of published papers and citations. However, it is now clear that these indicators only cover a fraction of the advances being made within each of the four AI research elements described above. In fact, the AI R&D community has adapted and expanded over time to include multiple disciplines devoted to increasing research efforts that integrate all four elements.

Many groups in academia and industry plan, as part of their research activities, the production and release of datasets and machine learning-specific libraries, making them available at certain times through papers submitted to peer-reviewed journals. This "roll out" is like the planned advertising campaign in the production of a new movie, which often involves "watch the movie, read the book, listen to the soundtrack, and buy the t-shirt." Therefore, it is essential to develop new indicators that track the interim release of AI open-source libraries and public datasets and can better describe the AI research landscape than published papers alone.

Initiatives that help us understand the development of the AI field are important, not only so that we can remain up-to-date on the research advances being made, but also so we can analyze the possible outcomes of this ongoing revolution and its impact on society."

3.2 Regional research trends in AI

The rise of China

As shown in Figure 3.4, Europe is still the largest contributor to AI research but continues to lose publication share. The United States is regaining ground lost in the last five years. China is bound to overtake Europe in publication output in AI in the near future, having already overtaken the United States in 2004.

Figure 3.5 illustrates that other individual countries are showing strong development in AI. For instance, India emerges as the third largest country in AI research in the last five years. Other emerging countries, like Iran, appear among the top 10 countries in AI research. Established research nations like Japan are also growing in terms of AI publication output, but with less vigour than the United States or China. Full country-level data is available through the Elsevier AI Resource Center.[68]

68 Elsevier. Artificial Intelligence Resource Center. https://www.elsevier.com/connect/ai-resource-center

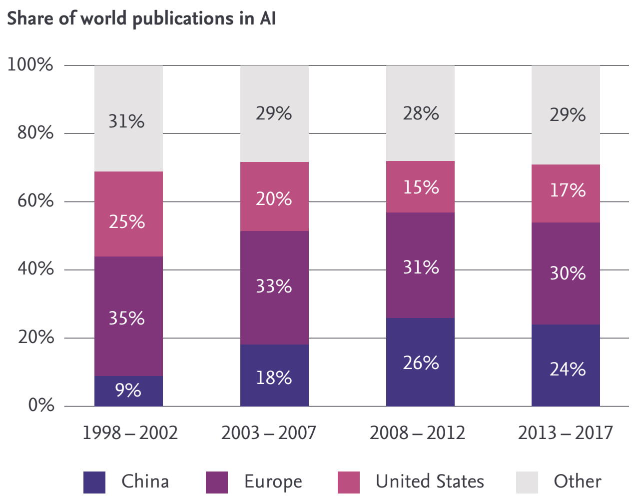


FIGURE 3.4
Share of global publication output in AI (all document types) for periods 1998-2002, 2003-2007, 2008-2012, and 2013-2017, per region; source: Scopus.

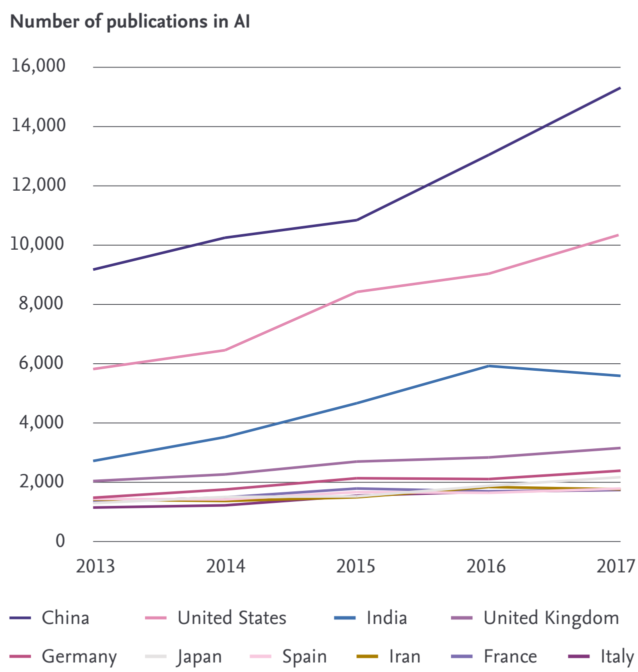


FIGURE 3.5
Publication output per country/territory (all document types), 2013-2017; source: Scopus.

In Europe and the United States, AI research has a stronger focus on health while in China the emphasis is on agriculture.

Success in AI in application fields, like the health sciences, mobility, or agriculture, fuels interest and growth in AI research. This section investigates the specialization of regions in AI research fields and clusters and reveals the focus on AI applications in medicine in Europe and the United States

The purpose of the *OECD Fields of Research and Development (FORD)* categories are to break down R&D expenditure and personnel by fields of research and development. FORD categories are used to classify R&D by fields of inquiry, namely, broad knowledge domains based primarily on the content of the R&D subject matter.

The *Relative Activity Index (RAI)* approximates the specialization of a region by comparing it to the global research activity in the AI field. RAI is defined as the share of a country's publication output in AI relative to the global share of publications in AI. A value of 1.0 indicates that a country's research activity in AI corresponds exactly with the global activity in AI; higher than 1.0 implies a greater emphasis, while lower than 1.0 suggests a lesser focus.

Nearly 60% of AI research publications fall within the natural sciences, which is also seeing the fastest growth rate. Other fields, like the agricultural sciences, also show strong growth but on a smaller base (~2%). Figure 3.6 reveals China's strong specialization in AI in the agricultural sciences, and the United States' focus on the medical and health sciences. Europe and the United States' apparent emphasis on the humanities refers to a very low number of publications and may be influenced by language.

Relative research focus per region

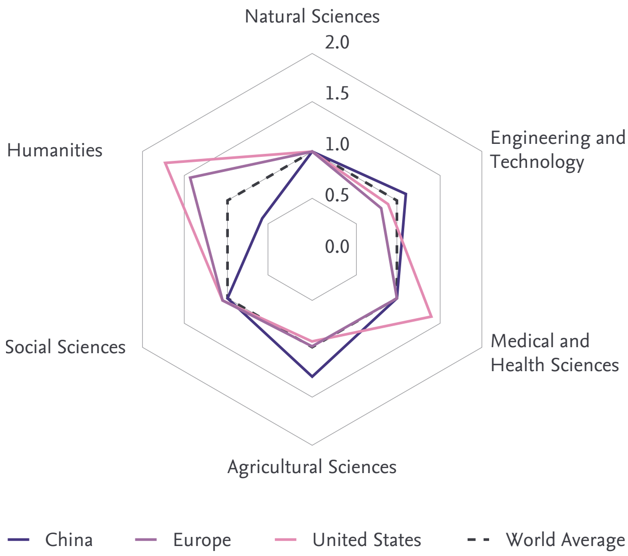
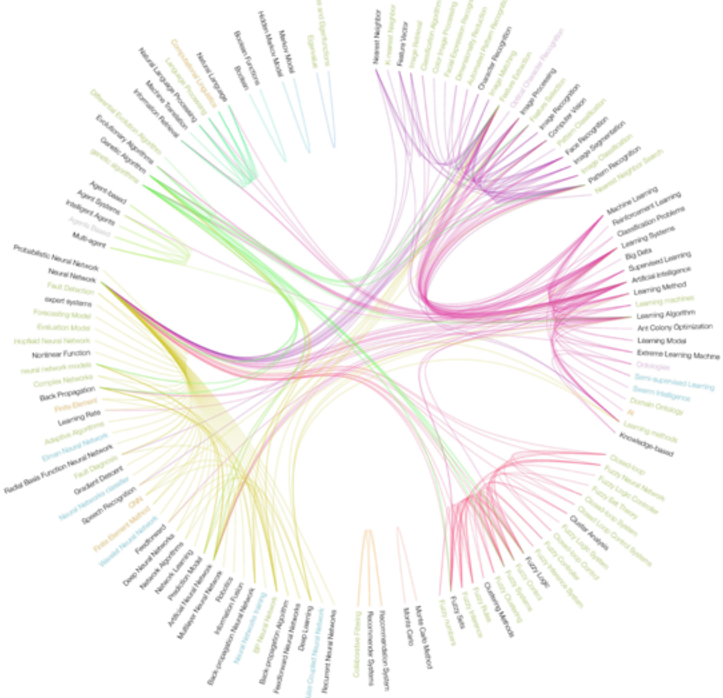
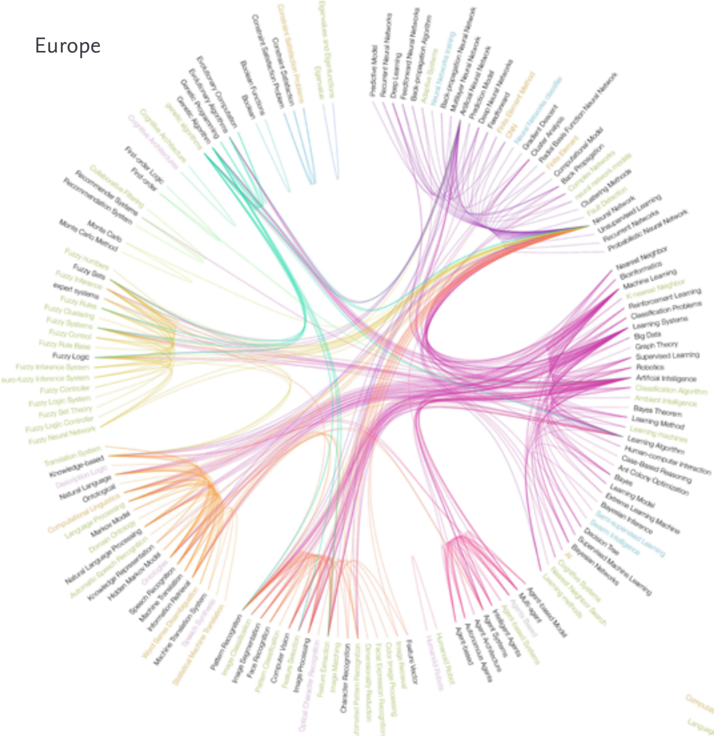


FIGURE 3.6
Relative Activity Index (RAI) of publications (all document types) per FORD category per region, 2017; dashed line indicates world average; source: Scopus.





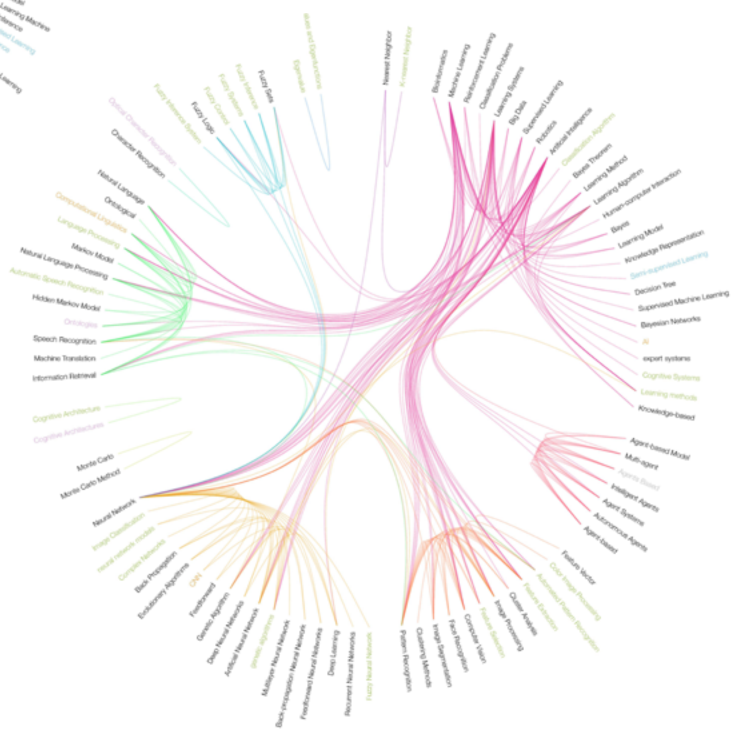


FIGURE 3.7
Keyword co-occurrences with 500+ shared publications (all document types) for China, Europe, and the United States, 2017; sources: Scopus and Elsevier clustering.

A comparison of keyword co-occurrences (Figures 3.7) illustrates how each region's AI research specializes, helping identify common interests and differentiation, such as shared "Fuzzy Systems" clusters but distinct clusters for several types of research under the term "Neural Network."

The United States has a thinner cluster structure, due to its overall lower volume of publications. This includes a less differentiated field compared to the strongly industry-influenced clusters "Fuzzy Systems" and "Computer Vision" in China and Europe. China's most apparent difference from Europe and the United States is the lack of a "Natural Language Processing and Knowledge Representation" cluster. This might be due to low publication volumes for China in this area, as research on this topic may be driven by corporations (which publish fewer papers than universities) in that country. Expert interviews confirm the strong Chinese focus in the area of "Face Recognition." Among Chinese publications, the "Neural Network" cluster appears very differentiated, including in prediction models and backpropagation, as well as robotics. In Europe and the United States, robotics is part of the "Machine Learning and Probabilistic Reasoning" cluster. In China and Europe, we identify additional clusters on "Genetic Programming" and "Evolutionary Algorithms" for topics like "Pattern Recognition." Further details on regional specialization is obtained through analysis of publications per year and co-occurrence clusters for each region (Figures 3.8-3.10).

Robotics belongs to Machine Learning and Probabilistic Reasoning in Europe and the United States.

Strong Chinese focus on Computer Vision.

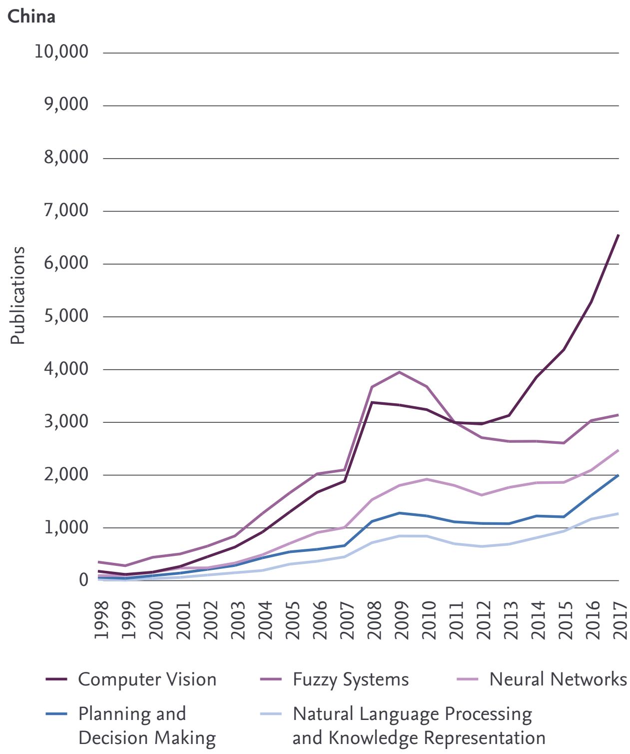


FIGURE 3.8
Annual publications per cluster for China (all document types), 1998-2017; sources: Scopus and Elsevier clustering.

China has a clear focus on the area of "Computer Vision," with very rapid recent growth, and sees a flattening of its research in "Fuzzy Systems," which drove China's publication growth in the first decade. "Machine Learning and Probabilistic Reasoning" and "Search and Optimization" impact all subfields, yet "Computer Vision" particularly benefits from developments in those areas. The spikes in 2009 are due to strong conference expansion in the field of engineering around that time. With the rise of neural networks, China seems to shift away from research topics in engineering, like "Fuzzy Systems."

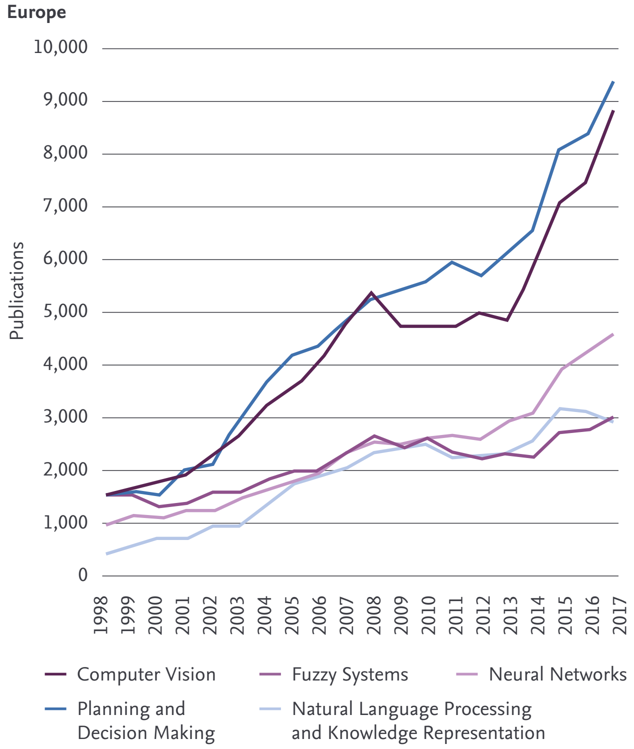


FIGURE 3.9
Annual publications per cluster for Europe (all document types), 1998-2017; sources: Scopus and Elsevier clustering.

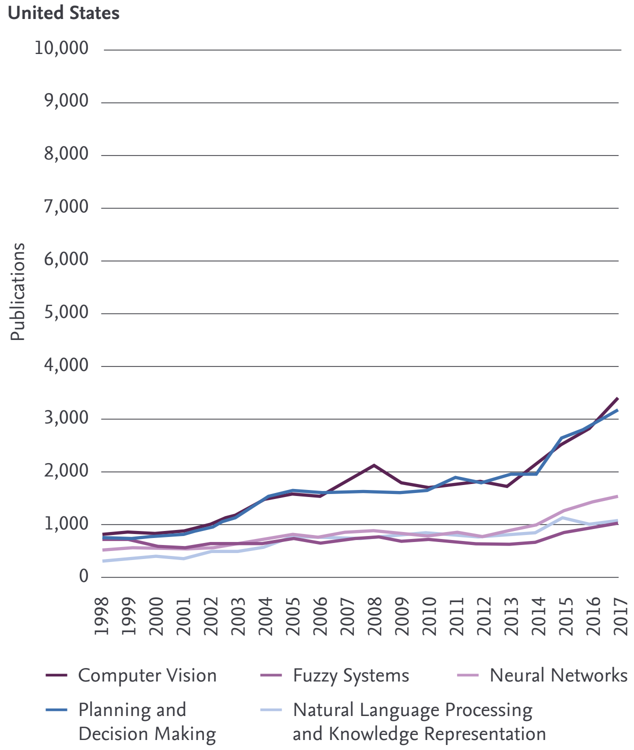


FIGURE 3.10
Annual publications per cluster for the United States (all document types), 1998-2017; sources: Scopus and Elsevier clustering.

Europe and the United States show similar cluster patterns, with the areas of "Planning and Decision Making" and "Computer Vision" strongly driving the AI field. Publications from Europe focus more on "Planning and Decision Making" than on "Computer Vision." "Neural Networks" research is rapidly growing in terms of journal articles but less so in conference papers across all regions, whereas "Natural Language Processing and Knowledge Representation" research shows stronger growth in conference papers across the regions.

In addition to the influence of language, the differences in AI research specialization between China and the United States might also result from different priorities; in China, we see a focus of AI research on agriculture and in the United States on health. "Planning and Decision Making" is applied to automated driving systems, reinforcement learning, robotics, human-computer interface, computer games and films, logistics, and mobile networks. A possible explanation may be found in the long industrial tradition in Europe and United States.

Prof. Fredrick Heintz
Associate Professor of Computer Science
at Linköping University Linköping,
President Swedish AI, Sweden

The importance of conferences to the AI field

"In computer science, especially in fast moving areas such as artificial intelligence, there is a long tradition of high-impact and high-prestige conferences. This has led to most results first being published at international conferences with thorough peer review and low acceptance rates. For the top conferences in the field, it is common to have an acceptance rate of 15-20%. This makes the conferences both highly competitive and timely. The main advantages of conferences over journals are that they have a fast turnaround time, they reoccur every year, and they have a clear submission deadline. Many researchers also publish in journals, often merging and extending conference papers because journals provide more space to share details. As the importance of bibliometrics continues to increase, more researchers are also publishing in journals. For the AI/machine learning part of the Wallenberg AI, Autonomous Systems and Software Program (WASP), Sweden's largest individual research program, we have explicitly set a goal to increase the number of publications from Swedish researchers at the top general AI and machine learning conferences, namely AAAI, ICML, IJCAI, and NIPS.[69] Achieving this goal will increase the presence of Swedish researchers at these venues and set the standard for the researchers in WASP—to aim for the top general conferences in the field."

The AI conference landscape

Key AI conferences, and specifically their calls for papers, give an early indication of the current trends in AI research. Figure 3.11 is comprised of over 300 keywords manually extracted from the call for papers from the top 10 AI conferences in 2018, suggested by the Stanford AI Index.[70] The focus on "Learning" and "Machine Learning Systems" continues, but we also see a strong interest in multi-agent topics.

As illustrated in Figure 3.12, the AI conference landscape is complex: conferences overlap across subfields, with strong connections between core AI and the field of data mining. AI conferences also touch upon associated fields such as mathematics, statistics, brain science, robotics, computer graphics, linguistics, cognitive science, social science, bioinformatics, computer systems, or high-performance computing.[71] Similarly, as noted by Raymond Perrault in this report's foreword, traditional conferences in symbolic AI use the term "Artificial Intelligence" while newer AI conferences use machine learning and probabilistic reasoning terms and/or connect with more independent application conferences.

69 Association for the Advancement of Artificial Intelligence, International Conference on Machine Learning, International Joint Conferences on Artificial Intelligence, Neural Information Processing Systems
70 Artificial Intelligence Index. 2017 Annual Report. http://cdn.aiindex.org/2017-report.pdf.
71 ML, DM, and AI Conference Map. 2015. Updated 25 November 2017. http://www.kamishima.net/archive/MLDMAImap.pdf.

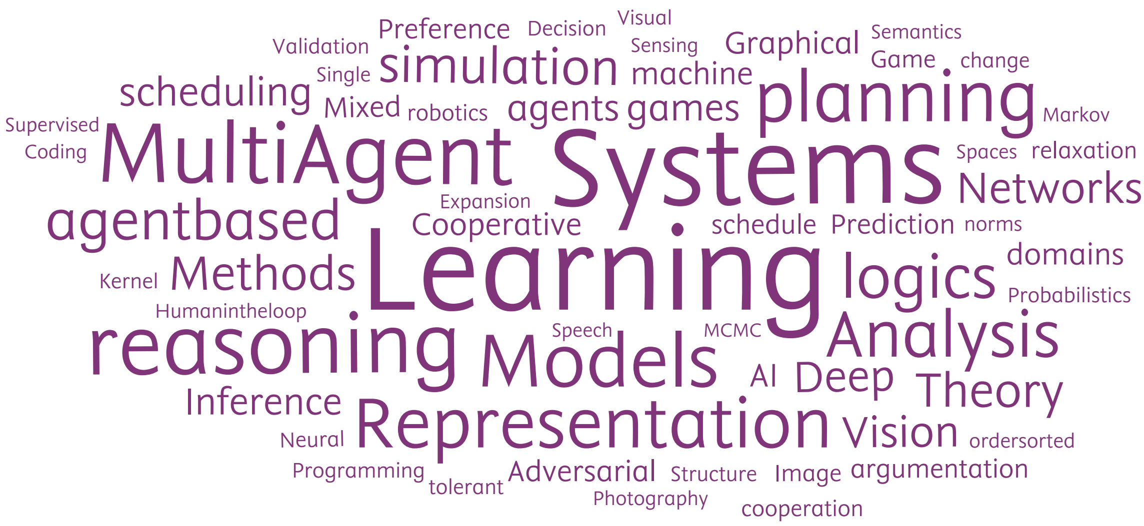


FIGURE 3.11
Keyword cloud from calls for papers by the 10 key AI conferences in 2018 suggested by the AI Index.

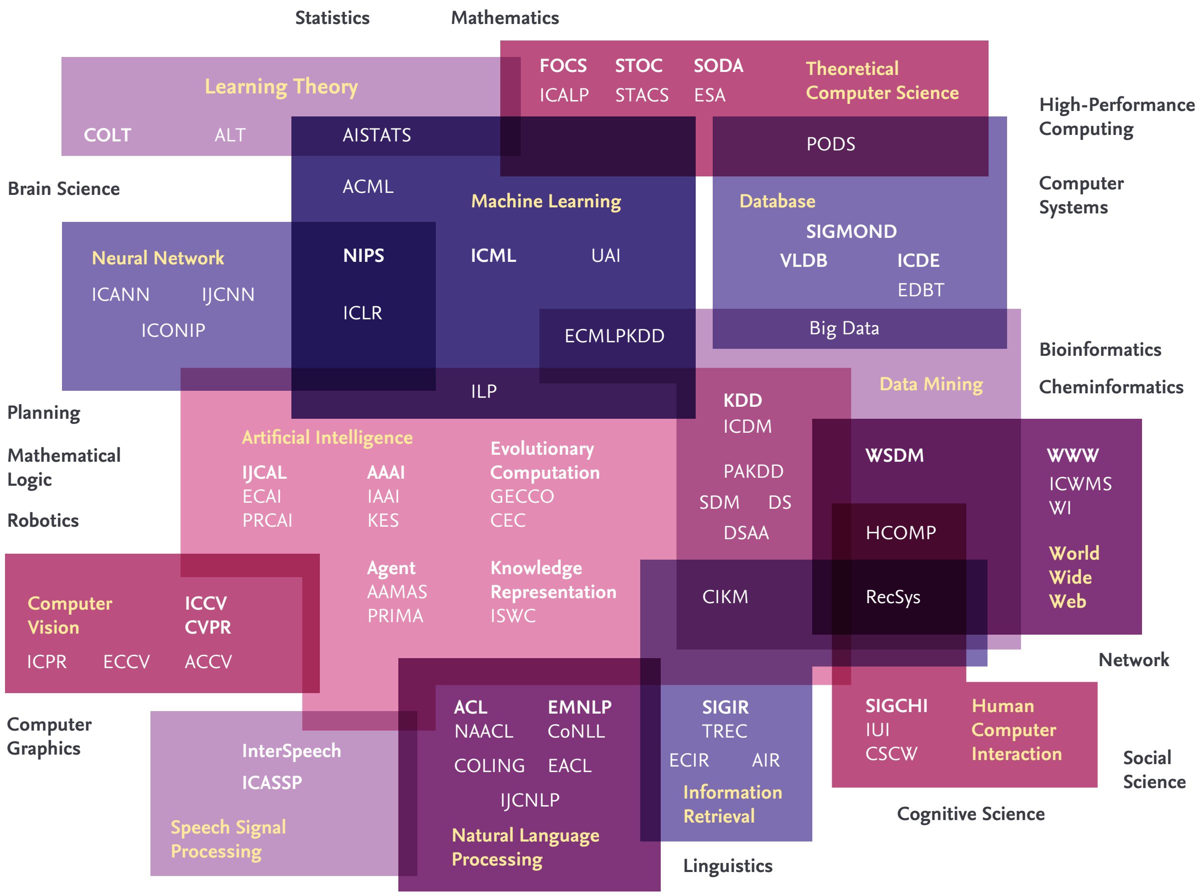


FIGURE 3.12
Landscape of AI conferences, courtesy of Prof. Toshihiro Kamishima; National Institute of Advanced Industrial Science and Technology (AIST), Japan; source: kamishima.net.

Nearly half of DBLP AI conference papers include an author from Europe.

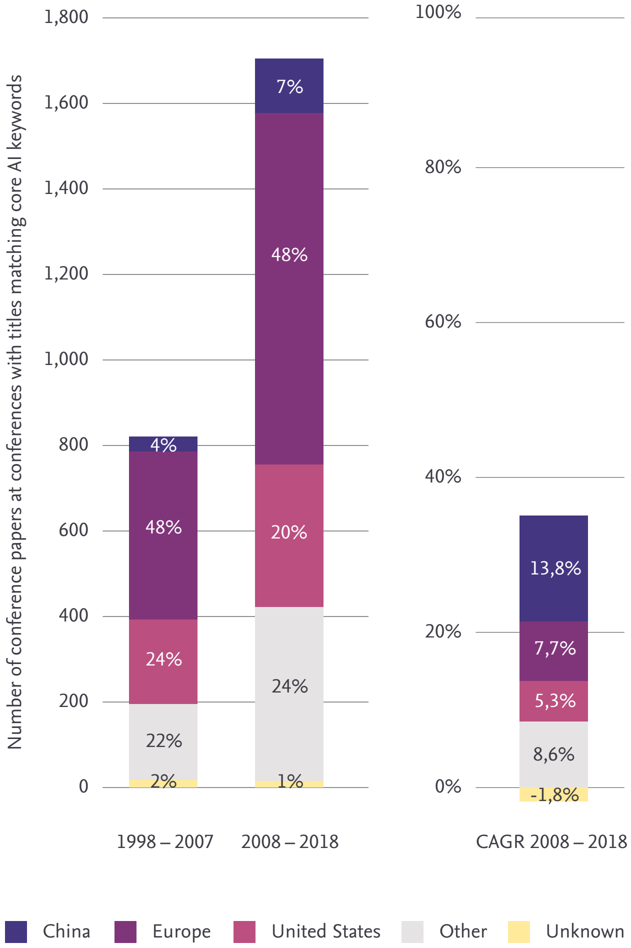


FIGURE 3.13
DBLP-tracked conference papers with core AI terms in their titles, by region, 1998-2017.

We dive deeper in the diachronic and regional trends of AI conferences using data from the Digital Bibliography & Library Project (DBLP) Computer Science Bibliography website.[72,73,74]

Looking at a very narrow subset of AI-related conferences that contain the 142 core AI keywords in their titles, we see that once again China has seen the most dramatic increase in conference papers over the two decades. However, this increase is not significantly different than the overall increase in conference papers in the region over the same years. In fact, for each of the regions except the United States, the growth of conference papers with core AI terms in their titles is less than the growth in DBLP-tracked conference papers overall, and the difference for the United States is not significant (see Figure 3.13).

If there is an increasing amount of AI-related academic activity, as measured by number of conference papers with core AI keywords in their titles, it is not apparent in the data examined from the DBLP. However, multiple issues with the DBLP data, including incomplete coverage of some computer science topics, make it impossible to draw definitive conclusions. Further research to understand how well the DBLP corpus reflects real-world conference research activity, and which areas of computer science have better coverage in the database, is needed to better measure research activity in this area.

72 dblp Computer Science Bibliography. https://dblp.uni-trier.de/.
73 Ley, M. DBLP: some lessons learned. *Proceedings of the VLDB Endowment.* 2009;2(2):1493-1500. doi: 10.14778/1687553.1687577.
74 dblp Computer Science Bibliography. Statistics – Records in DBLP. https://dblp.uni-trier.de/statistics/recordsindblp.html.

We gain further insight on various research output types by examining conference papers across sectors. In all regions, academia is by far the biggest contributor, matching its share of conference papers within Scopus almost directly with the overall share.

In China, the corporate sector has a higher share of conference papers among all publications and the government sector has the lowest (see Figure 3.14). In Europe, the corporate sector has only a slightly higher share of conference papers and the government sector a slightly lower one, but the differences are less pronounced than for China (see Figure 3.15). In the United States, the corporate sector has a consistently higher share of conference papers. The government sector starts with a comparatively high share of conference papers, which declines in recent years (see Figure 3.16).

Over 70% of recent corporate AI research in the United States is published as conference papers.

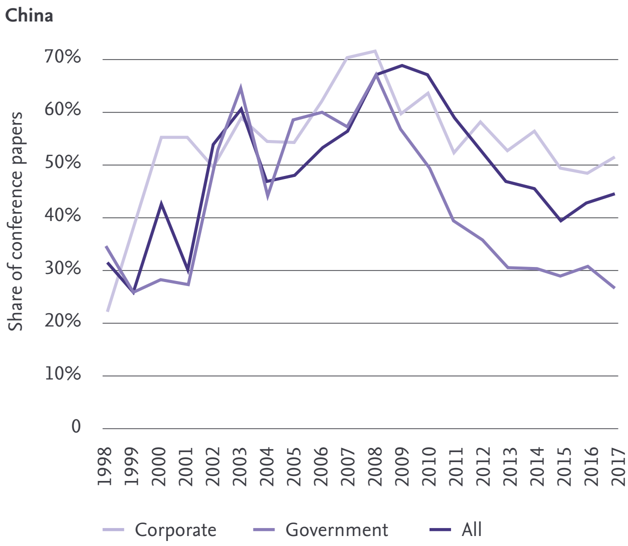


FIGURE 3.14
Share of conference papers in AI per sector for China, 1998-2017; source: Scopus.

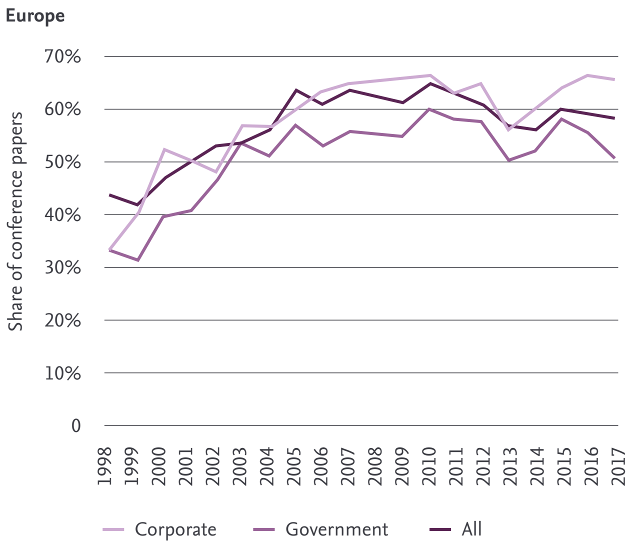


FIGURE 3.15
Share of conference papers in AI per sector for Europe, 1998-2017; source: Scopus.

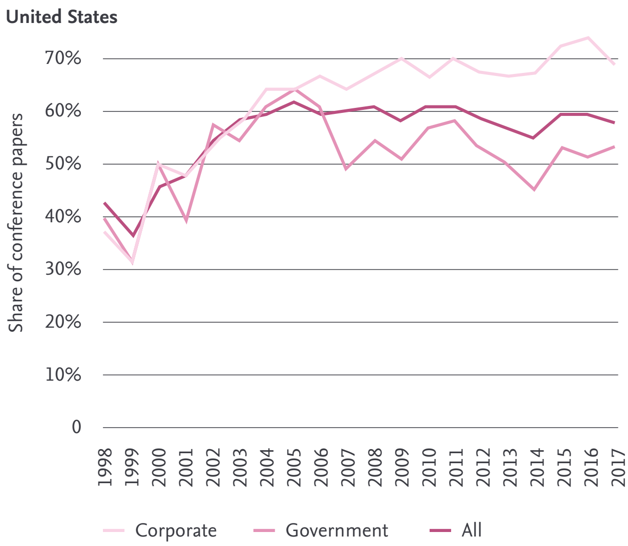


FIGURE 3.16
Share of conference papers in AI per sector for United States, 1998-2017; source: Scopus.

Key Contributors (academic and corporate institutions)	Number of publications (all)	Field-Weighted Citation Impact
China		
Chinese Academy of Sciences	9,064	1.64
Tsinghua University	5,287	1.77
Harbin Institute of Technology	4,306	1.08
Shanghai Jiao Tong University	4,084	1.19
Zhejiang University	3,739	1.11
Europe		
Universite Paris Saclay (France)	3,214	1.67
INRIA Institut National de Recherche en Informatique et en Automatique (France)	2,312	2.77
Imperial College London (United Kingdom)	1,790	2.39
Universite Pierre et Marie Curie (France)	1,738	1.38
University of Granada (Spain)	1,736	1.22
United States		
Carnegie Mellon University	4,038	3.04
Massachusetts Institute of Technology	2,312	3.49
Microsoft USA	2,648	5.84
IBM	2,351	2.45
Stanford University	2,235	5.54

FIGURE 3.17

Top 5 institutional contributors per region by number of AI publications (all document types), 2013-2017; source: SciVal.

Field-Weighted Citation Impact (FWCI) is an indicator of the citation impact of a publication. It is calculated by comparing the number of citations actually received by a publication with the number of citations expected for a publication of the same document type, publication year, and subject. FWCI is always defined with reference to a global baseline of 1.0 and intrinsically accounts for differences in citation accrual over time, differences in citation rates for different document ages (e.g., older documents are expected to have accrued more citations than more recently published documents), document types (e.g., reviews typically attract more citations than research articles), and subjects (e.g., publications in medicine accrue citations more quickly than publications in mathematics).

Within the regions, we identify key institutions based on number of publications and FWCI. This information should be seen in the context of the overall regional output and citation impact to gain insight into the institutional structure of a region, i.e., a region with several mid-sized contributors in AI might appear lower in such a list compared to regions with big, centralized research organizations. The major 100 contributors to AI publication output represent 41% (99k of 241k) of the global AI corpus and hold 32% (109k of 338k) of the global conference papers. China stands out in the top 100 with over one-third of major contributors (37), while the United States (19) and Europe (21) together hold another one-third and the remaining countries hold the final one-third. The three key regions hold 75% of the world's contributors to AI publications. Figure 3.17 illustrates a few top contributors per region. The United States not only has two major corporate contributors, but Microsoft USA is also an outstanding contributor to citation impact. All top five contributors have citation impacts three to five times higher than the world average. Europe is dominated by French institutions, followed by British and Spanish institutions. France and Italy have strong national governmental research organisations with CNRS [75] and CNR. [76]

Other top contributors in China (in order of the number of AI publications) are the universities in Huazhong, Beihang, Northeastern, Southeast, Wuhan, Xi'an Jiaotong, Dalian, South China, and Xidian. In the United States, the following universities also make a sizeable contribution to the global AI corpus of research: Southern California, Georgia Institute of Technology, Illinois at Urbana-Champaign, Berkeley, Harvard, Maryland, Washington, Texas at Austin, Michigan, and Columbia. In Europe, we note the following universities as top contributors to AI publications: Edinburgh (United Kingdom), Leuven (Belgium), Politecnica de Catalunya (Spain), Oxford (United Kingdom), University College London (United Kingdom), Politecnica de Madrid (Spain), Manchester (United Kingdom), Technical University of Munich (Germany), Lisboa (Portugal), and Delft (the Netherlands). Other countries stand out in the top 100, such as Singapore, Iran, Canada, Taiwan, Hong Kong, Japan, and Australia, each with two major contributing institutions. Although they do not cluster as a key region, they might be important players to consider. Other countries like Germany might be underrepresented due to their federal research structure compared to peers like France, the United Kingdom, or Spain.

75 Centre National de la Recherche Scientifique.
76 Consiglio Nazionale delle Ricerche.

Interview

Prof. Chuan Tang
Chengdu Library and Information Center, Chinese Academy of Sciences (CAS), China

How do AI researchers recognize excellent research? What indicators do they look for?
The indicators that AI scholars value include research that has been presented at top academic conferences, is being done at a large scale, has won international first-class competitions, and has undergone rigorous peer review.

How does Chinese science contribute to the advancement of AI in China and globally?
The level of AI research in China has rapidly increased in recent years, and China is now considered to be the most active country in this competitive field. At the top academic conferences on AI held in 2015,[77] the number of research papers published by Chinese institutions ranked second, exceeding 20% of all papers published. In its National Artificial Intelligence Research and Development Strategic Plan (October 2016),[78] the White House pointed out that the number of scientific research papers on deep learning in China has surpassed that of the United States. At the 2017 conference from the Association for the Advancement of Artificial Intelligence (AAAI), the number of papers submitted by Chinese researchers was the highest in the world. AI research in China has risen to win international attention for Chinese researchers, so much so that the AAAI adjusted its 2017 schedule to accommodate the Chinese New Year. However, in the areas of AI basic theory, key technologies, global influence, and leading figures in the field, China lags behind the world's leading research institutions. The key players driving the development of deep learning are not Chinese scholars. Companies such as Baidu are hiring foreign experts to take charge of their AI-related business ventures.

What are the main problems, obstacles, and difficulties involved in the development of AI in China?
Among the key elements needed to develop the field, China has abundant policies in place and capital support, as well as advantages in terms of data volume and an application market that is unmatched in any other country. The problems, obstacles, and difficulties are concentrated in core technologies and talent, as well as employment and competition.

- The AI chip is at the core of the industry, with the highest technical requirements, added value, and strategic positioning. However, China's chip industry has been weak for some time, and there is a serious lack of chip design and chip foundry capabilities.
- China lacks long-term efforts in basic AI research. Most researchers tend to follow trends in Western countries and work on improving existing technologies. There is a lack of long-term research in promising areas of basic science, particularly in areas without obvious benefits in the short term. Although the number of published AI papers from China has surpassed that of other countries, their global influence is limited, resulting in lack of impact of the underlying AI technology beyond China.
- The development and application of AI technology requires a high-quality talent team. Per recent statistics, the total number of people involved in AI in China is only over 50,000, ranking 7th in the world. Only 38.7% of those involved in China's AI sector have more than 10 years of industry experience, and there are few domestic educational institutes with related majors such as machine learning. There is also uneven distribution of AI talent technology systems in China. Talent is concentrated in the application phase, while the infrastructure and technology layers remain weak. In general, China still lacks leading talent with international influence, innovative and entrepreneurial talent to promote industrial development, and an industry capable of applying AI skills and tools.

77 US, China most active in AI research, report finds. Nikkei Asian Review. 9 December 2016. https://asia.nikkei.com/Business/Science/US-China-most-active-in-AI-research-report-finds.
78 National Science and Technology Council, Networking and Information Technology Research and Development Subcommittee. National Artificial Intelligence Research and Development Strategic Plan. October 2016. https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf.

The special case of AI competitions

In parallel with or complementary to academia, competitions are another important arena for dissemination of AI research, and are also used as a vehicle for recruitment, training, and collaboration. For this purpose, we examined Kaggle,[79] a leading platform for hosting public data and machine learning competitions, and a home to a dynamic community of data scientists and machine learning experts.

Competition rewards range from knowledge to prestige to financial incentives and vary by competition category. Featured competitions usually have a financial reward, recruitment competitions offer jobs, and research competitions address complex problems and can contribute to breakthroughs within the community. For example, a competition was hosted for an algorithm that could identify the Higgs boson within particle collisions at CERN.[80,81] Looking at incentives, 36% of competitions indicate knowledge as the reward; these competitions are primarily used as educational tools and reflect the collaborative nature of the field. Competitions with financial incentives do represent a sizeable percentage of competitions, however. Jobs represent 1% of competition rewards, and are usually hosted by Silicon Valley companies and corporations. The financial rewards offered on Kaggle vary greatly, and the number of entries to competitions like those on Kaggle does not necessarily correlate with the amount of prize money offered. High financial rewards seem to lead to increases in membership, but many competitions offering non-financial rewards have had more submissions than those offering a job or financial reward.

79 Kaggle. https://www.kaggle.com/.

80 CERN: Conseil Européen pour la Recherche Nucléaire, the European Organization for Nuclear Research.

81 Jepsen, K. The machine learning community takes on the Higgs. *Symmetry*. 15 July 2014. https://www.symmetrymagazine.org/article/july-2014/the-machine-learning-community-takes-on-the-higgs.

75% of Kaggle datasets are created in the United States.

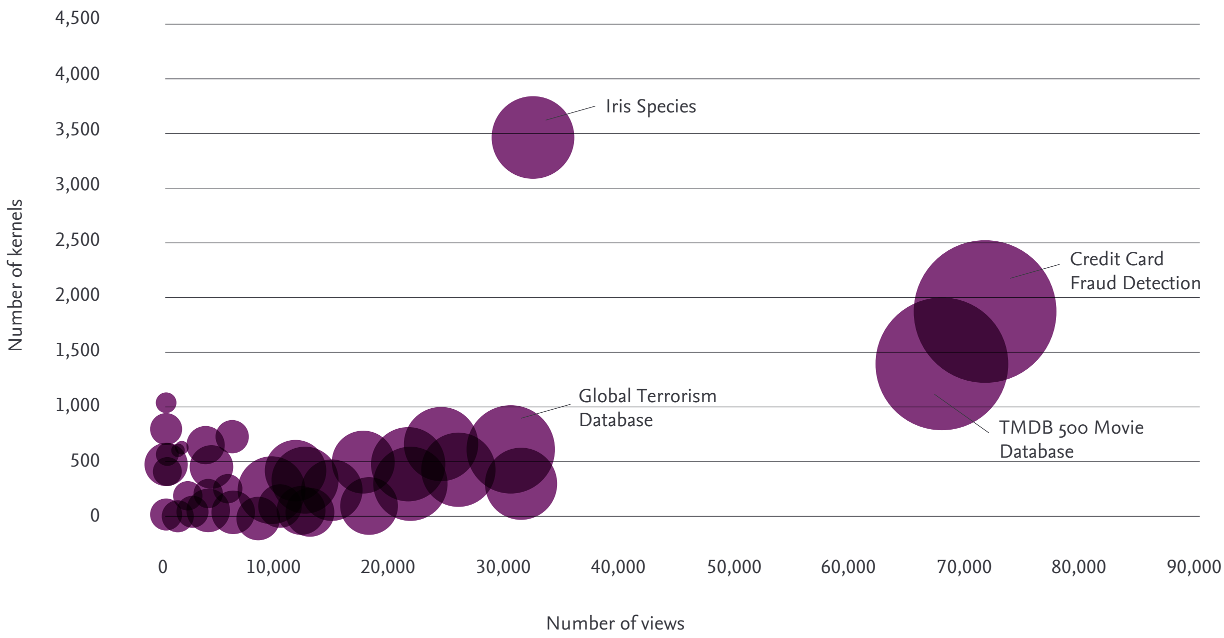


FIGURE 3.18
Kaggle dataset views, downloads (size of nodes), and kernels, 2010-2018; source: MetaKaggle, published under CC BY-NC-SA 4.0.[82]

25% of Kaggle user survey respondents are located in the United States.

Looking at the organizations that have uploaded datasets by region, it is apparent that Kaggle is heavily dominated by the United States, with 1,074 of the 1,441 datasets provided by organizations based within the United States. These numbers do not reflect the amount of sets uploaded by individuals, as out of the 9,572 datasets uploaded, only 1,441 were uploaded by organizations. Figure 3.18 analyzes Kaggle datasets and provides further insights into the community, in particular showing that some of the most downloaded and viewed datasets are not associated with competitions but are simply robust enough to allow users to continually contribute to their analysis.

- **Views** on Kaggle indicate the number of times users view a dataset online, and as such are an indicator of potential interest.
- **Downloads** track the number of times users download a dataset and are therefore an indicator of further interest.
- **Kernels** are online notebooks in which code can be edited or run. The number of kernels is therefore an indicator of usage.

In 2017, Kaggle conducted a survey of its users to gather information about the community, and received responses from 1.6% of Kaggle users. Most Kaggle users are in Asia, the United States, and Europe. These three regions account for nearly 70% of users, with over one-quarter in the United States. Within Europe, there are at least 400 respondents each from the United Kingdom, France, and Germany. The dominance of these nations within the AI sector is reflected in the distribution of the 50 top ranked users, with 10 users from France, 9 from the United Kingdom, and 6 from Germany. Asia, excluding China, also accounts for a large share of survey respondents. India accounts for 16.8% of total respondents and 9.3% of top 150 ranked users. In the Kaggle survey, only 3% of the respondents are from China, and there are only 10 Chinese nationals in the top 100 users, with 40% of them residing overseas. The lack of Chinese users may be related to the relative obscurity of the website within that country, with the sheer volume of rival local websites such as Alibaba's TianChi (天池) website, the Latest Activities & TianChi Competition page, and DataCastle, with 75,208 registered users. Other popular competitions are held by the Data Foundation, Kesci, China Computer Federation, Biendata, Big Data Research Center, Hack Data, Soda, and many more.[83] This could indicate a preference in China for national over international competitions.

82 https://creativecommons.org/licenses/by-nc-sa/4.0/.

83 Zhihu. What are the data competition and related competition websites at home and abroad? https://www.zhihu.com/question/36374964.

3.3 Regional research impact and usage comparison

Computer science research is disseminated in a variety of publication types (e.g., journals, conferences, etc.) and forms (e.g., software, code, etc.). Thus, while article citations may not fully capture research impact in the AI field, they nevertheless play a relevant role, especially for comparative benchmarking of entities on scholarly impact. Article downloads also offer an interesting perspective on scholarly usage, revealing a different dimension of engagement from those that read an article but may not systematically publish or cite an article (e.g., students, practitioners, corporate researchers, the public, etc.). While citation impact is a lagging indicator, as the accumulation of citations takes time, measuring downloads allows for insights on impact immediately after the publication of an article.

Regional inequalities in citation impact do not apply to download impact.

For ease of comparison, the Field-Weighted Citation Impact (FWCI) and Field-Weighted Download Impact (FWDI) of articles published by researchers in the comparator regions were rebased to the global annual FWCI and FWDI values within AI, such that the FWCI and FWDI of all AI research articles equals 1.0 for all years. Figure 3.19 reveals that regional inequalities in citation impact are not reflected in download impact, suggesting comparable usage of each region's research. While China's FWCI is still below that of Europe and the United States, it shows tremendous growth over the past two decades, from half the world average to reaching the world average in recent years. Europe's FWCI remains stable over the period, comfortably higher than the global average. The United States' FWCI is the highest among regions, remaining between one and a half to two times as high as the global average over the period. The 2016-2017 dip in FWCI for the United States may be due to incomplete citation data, although there seems to be a slight decreasing trend following a 2014 peak.

FIGURE 3.19
Rebased AI Field-Weighted Citation Impact (FWCI, bold lines) and Field-Weighted Download Impact (FWDI, dotted lines) (all document types) per region, 1998-2017; source: Scopus.

Beyond citation and downloads, it is now possible to track the online attention received by research. The increasing diversity of scholarly communication outlets and connectivity of the research community now extends to the media or blog mentions as well as discussions on social media channels. The PlumX dashboard[84] provides deeper and broader insights into mentions and captures for a variety of research outputs, such as publication usage (e.g., downloads), mentions (e.g., news references), social media (e.g., Facebook Likes, Twitter re-tweets), or captures (e.g., Mendeley, GitHub).

PlumX Metrics provide insights into the ways people interact with research output (articles, conference papers, book chapters, etc.).

PlumX Metrics use 50 sources, including Scopus, SSRN, arXiv, SciELO, Airiti, PubMed, YouTube, Vimeo, GitHub, Patents, and more. Plum analyzes and covers more than 40 media sources, such as bepress, ORCID, VIVO, RSS Feeds, DOIs, PubMed, Books, SSRN, arXiv, SlideShare, SoundCloud, YouTube, Vimeo, Patents, Clinical Trials, GitHub, SourceForge, Dryad, figshare, and web pages.

As we see from AI competitions, the research community expands into non-institutional researchers and open-source platforms. The lines between AI research and application development blur. Mentions and captions are new channels to understand, earlier and in different ways, the interest in research publications (e.g., journal articles or conference papers, which comprise the majority of the AI corpus, in addition to a smaller number of book chapters and review papers, and a negligible number of other publication types).

As expected given the substantial proportion of AI research in the natural sciences, this field contributes the largest number of online mentions (235k), yet when this is normalized for corpus size, has comparatively few mentions per publication. This probably stems from the specificity of the AI research topics, compared to the general interest in more societally relevant fields. Indeed, AI application fields, like agriculture and health sciences, are relatively strongly discussed and mentioned in social media, blogs, and news,

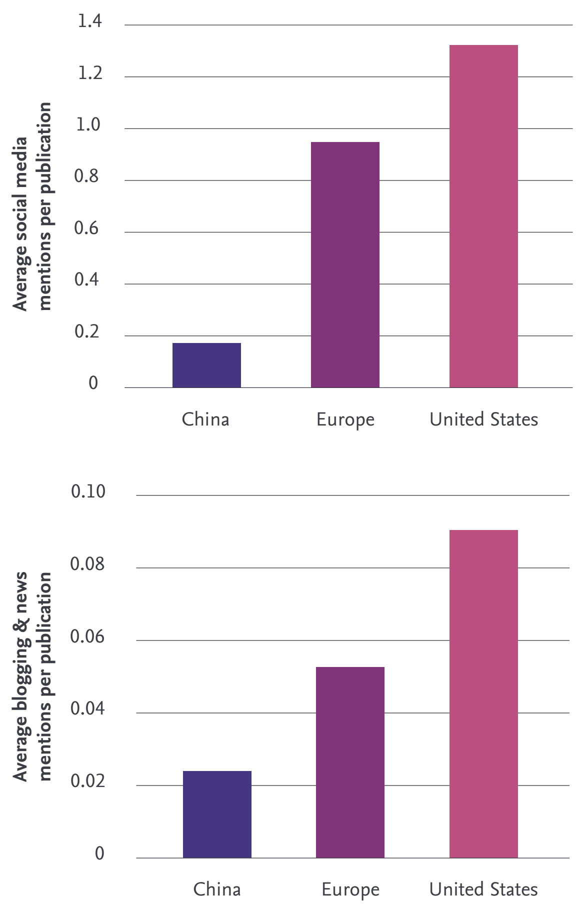


FIGURE 3.20
Average online mentions per publication per region, 2008-2018; source: PlumX dashboard.

but with a smaller base of research publications compared to the overall field of AI.

In the same way as for AI subject categories, we look at the online mentions for research coming from the different regions. Figure 3.20 reveals that China's AI research is comparatively less discussed on social media (potentially due to access restrictions in that country) and in blogs/news, although language and coverage may influence the latter. Europe has more mentions than China, in particular on social media. Media/blog coverage of European research may also, to some extent, be negatively influenced by the variety of languages spoken in the region relative to the predominantly English sources covered. This can also account for the United States' highest media outreach. From a global point of view, these dynamics underline the preponderance of English as the current lingua franca of AI research, as well as the comparatively lower international visibility and outreach of Chinese AI research.

Further exploration by region or subject is possible on the PlumX Dashboard accessible via the Elsevier AI Resource Center.[85]

84 Plum Analytics. PlumX Dashboards. https://plumanalytics.com.
85 Elsevier. Artificial Intelligence Resource Center. https://www.elsevier.com/connect/ai-resource-center.

3.4 AI knowledge transfer

Prof. Tieniu Tan
Institute of Automation, Chinese Academy of Sciences (CAS), China

"Governments can enforce policies and regulations, provide sufficient funding, and develop and maintain adequate infrastructure to support the artificial intelligence field. Different countries may have different initiatives and strategies and compete for AI talent, but they should also collaborate. In our era, international collaboration is essential—no country can thrive in the AI field in isolation."

Next to research publications, research collaboration is a core element of scholarly communication and knowledge transfer between regions, disciplines, and sectors. Collaborations have become the cornerstone of innovation and excellence, crossing borders, disciplines, and communities. Developments are propelled by low-cost travel, high-speed internet connectivity, mobile technology, social media, public engagement, and funding programs that encourage scholars, communities, and policy makers to expand their networks beyond their immediate working environments and traditional spheres of influence.

AI talent migrating from one sector to the other is another way of knowledge transfer, especially in emerging fields. A recent study by LinkedIn Talent solutions[86] showed the outflow from academia into the corporate sector. The study hinted at other aspects of talent recruiting through up-skilling and sourcing talent from adjunct fields. On the other side, academia is facing higher competition on research talent[87] as noted by *The Guardian* and the *Financial Times*. This section analyzes patterns of the three key regions—China, Europe, and the United States—within academia across regions, between the corporate sector and academia, and along the different formats of transfer, such as researcher migration and collaboration.

AI grew from cross-discipline inspirations and cross-sector work, e.g., between academia and corporations. Collaborative research tends to be more impactful in terms of citation rates. Collaboration can usually be detected from the patterns of co-authorship of published articles or the acknowledgements within them. While co-authorship is not the only form of collaboration, particularly in fields such as the social sciences and arts and humanities, it can be quantified with reasonable robustness and is the basis for the indicators discussed in this section. Research collaboration is analyzed through the proxy of publications resulting from the efforts of two or more authors. Collaboration can be further subdivided into the following types: international collaboration, national collaboration, or institutional collaboration.

Single authorship is declining across all regions and AI research is becoming more collaborative. Europe and the United States are increasingly collaborating internationally. For the United States, this brings not only an expansion in publication share, but also higher citation impact. International collaboration in Europe in contrast drives mainly publication share. China is reducing its institutional collaboration and shifting to national and international collaboration. Its international collaboration brings more citation impact than for the United States and Europe. In the direct comparison of international collaboration (Figure 3.21) across the three regions, we see Europe's strong increase in publication volume and China's success in increasing volume and citation impact through international collaborations.

86 Henriques, P. 3 Unconventional Strategies for Recruiting Machine Learning Talent. LinkedIn Talent Blog. 15 August 2018. https://business.linkedin.com/talent-solutions/blog/trends-and-research/2018/recruiting-machine-learning-talent.

87 Sample, I. 'We can't compete': why universities are losing their best AI scientists. The Guardian. 1 November 2017. https://www.theguardian.com/science/2017/nov/01/cant-compete-universities-losing-best-ai-scientists.; Ram, A., UK universities alarmed by poaching of top computer science brains. Financial Times. 9 May 2018. https://www.ft.com/content/895caede-4fad-11e8-a7a9-37318e776bab

Strong growth of international collaboration in AI research over the last two decades.

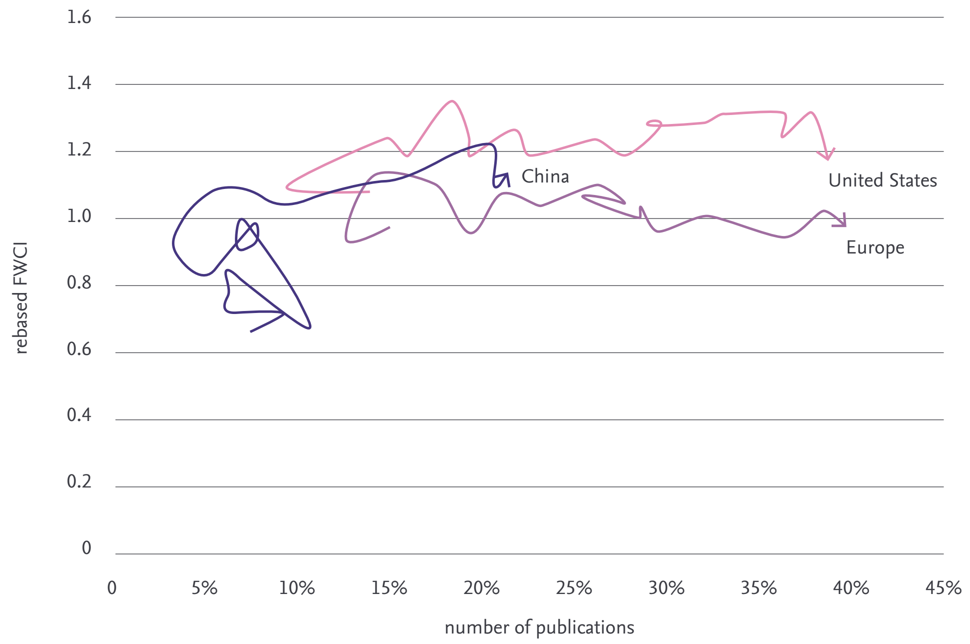


FIGURE 3.21
Number of publications from international collaborations (all document types) and their rebased Field-Weighted Citation Impact (FWCI), 1998-2017; source: Scopus.

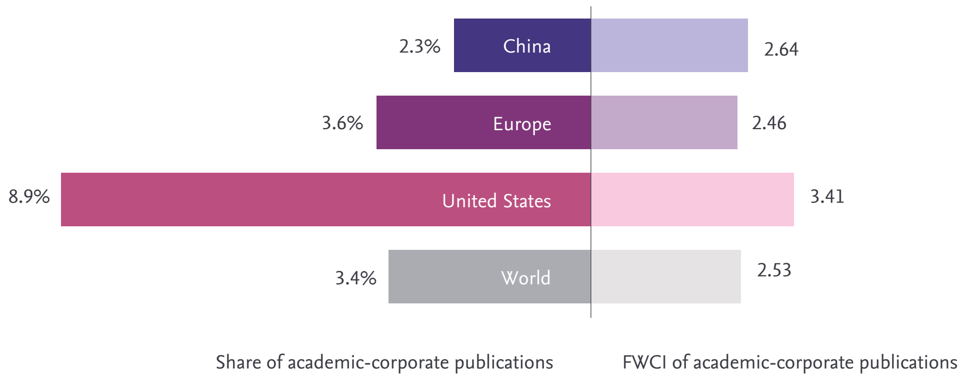


FIGURE 3.22
Academic-corporate share of publications (left-hand side, dark color) and their Field-Weighted Citation Impact, FWCI (right-hand side, light color) (all document types), 1998-2017; source: Scopus.

More than 90% of AI research is produced by the academic sector. Yet academic-corporate collaboration, analyzed here through the proxy of publications with authors across both sectors, plays a key role in terms of knowledge transfer and innovation. Quick research transfer into applications is a key goal of governments and innovation programs, stimulating economic development and job creation.

Globally in all sectors, academic-corporate collaboration receives higher citation rates, and this is also the case for AI research conducted in each region (Figure 3.22). Such cross-sector collaborations are particularly prominent in the United States, accounting for nearly 9% of their output in the field with a citation impact of more than thrice the world average. This can be explained by the strong United States AI corporate sector, with companies such as Microsoft and IBM contributing significantly to AI scholarly output and impact. China is below the 3% global average share of academic-corporate publications, and Europe slightly above it, with both regions reaping similar citation impact benefits these collaborations.

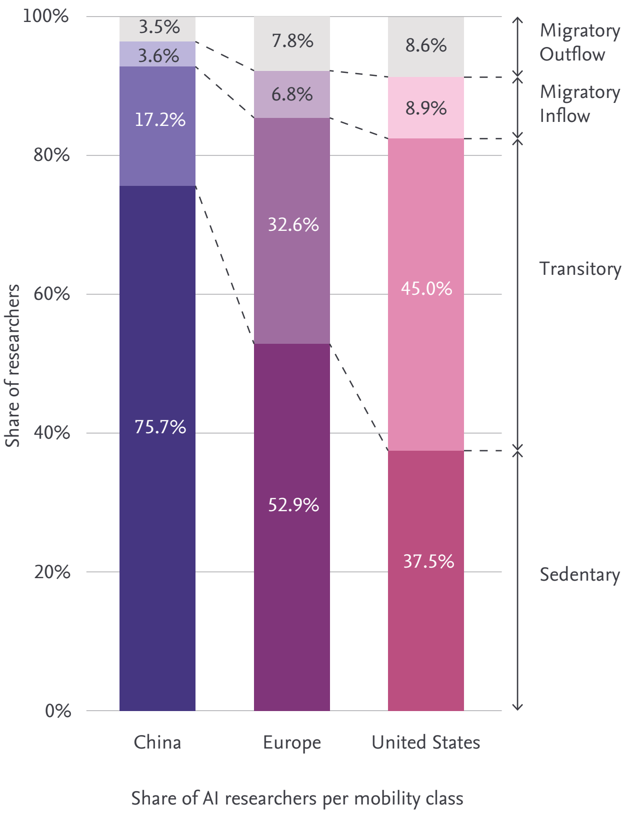


FIGURE 3.23
Share of AI researchers (%) per mobility class, 1998-2017; source: Scopus.

9% of scholarly output by academic-corporate collaborations in the United States

Beyond research collaborations, researcher mobility indicates knowledge exchange—in person and as researchers physically relocate to other regions. Figure 3.23 illustrates the shares of each mobility class per region.

The approach presented here uses Scopus author profile data to derive a history of active authors. Based on the affiliations recorded in each author's publications over time, authors are assigned to a mobility class defined by the type and duration of observed moves:

- Migratory — researchers who stay abroad or in the region for two years or more.
- Transitory — researchers who stay abroad or in the region for less than two years.
- Sedentary — researchers with only a regional affiliation in Scopus during the period 1998–2017.

Relative productivity is calculated by dividing the average number of publications per researcher for the specific mobility class in the region by that of all researchers from the region. Relative impact is calculated by dividing the average FWCI for the specific mobility class in the region by that of all researchers from the region.

Europe sees a net outflow over the 20-year period, with 7.8% migratory outflow over 6.8% migratory inflow. China has a small inflow surplus (0.1 percentage point), with 3.5% outflow and 3.6% inflow, while the United States has a net gain of 0.3 percentage point. Recent articles from the United Kingdom[88] and The Netherlands[89] highlight this effect. To better understand the impact of these effects, Figures 25-27 explore relative productivity and impact.

Figure 3.24 shows that China has a very high level of sedentary researchers with a rather low relative citation impact and relative productivity compared to migratory or transitory researchers. On top of the ~25% that come to the country, 17% are staying more than 2 years, but still bring productivity and impact benefits. Through researcher mobility, China is gaining in relative productivity and relative impact.

Figure 3.25 shows that migrating researchers increase research productivity in Europe. Europe is gaining impact and productivity from its migratory balance, even if 1 percentage point more people are flowing out of Europe than into it. Similar to China, sedentary researchers in the region show lower citation impact levels compared to migratory and transitory researchers.

As demonstrated by Figure 3.26 the United States attracts impactful researchers and holds the lowest share of sedentary researchers. Nevertheless, the high citation impact of sedentary researchers might indicate a reason for international inflow into the country. While outflowing and transitory researchers have lower relative citation impact, they also have higher relative productivity.

88 Sample, I. Big tech firms' AI hiring frenzy leads to brain drain at UK universities. *The Guardian*. 2 Nov 2017. https://www.theguardian.com/science/2017/nov/02/big-tech-firms-google-ai-hiring-frenzy-brain-drain-uk-universities.

89 Universiteit Leiden. Holger Hoos in NRC about AI brain drain. 28 August 2018. https://www.universiteitleiden.nl/en/news/2018/08/holger-hoos-about-ai-braindrain-in-nl.

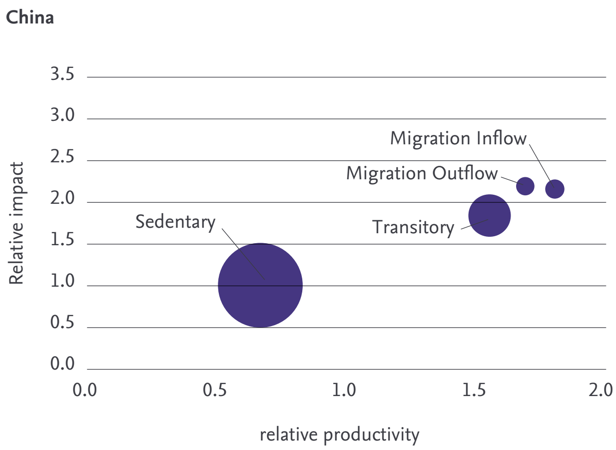


FIGURE 3.24
Relative productivity and relative impact per mobility class for China, 1998-2017; bubble size represents the percentage of researchers in each mobility class; source: Scopus.

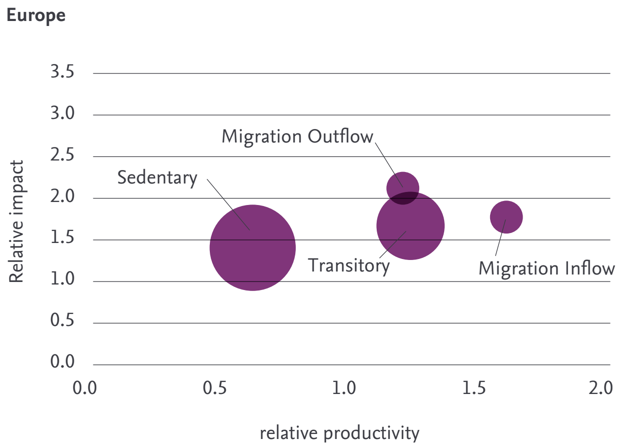


FIGURE 3.25
Relative productivity and relative impact per mobility class for Europe, 1998-2017; bubble size represents the percentage of researchers in each mobility class; source: Scopus.

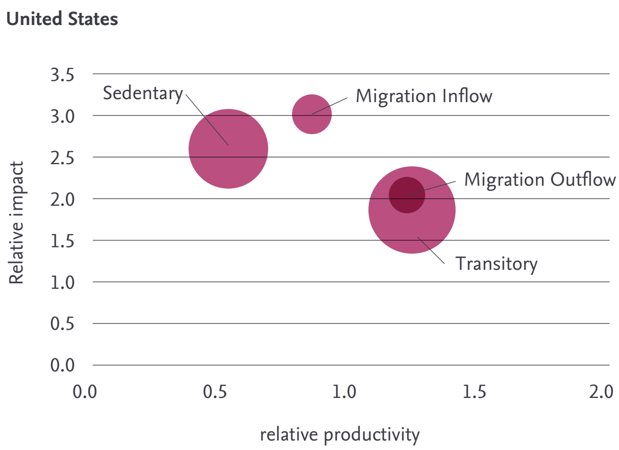


FIGURE 3.26
Relative productivity and relative impact per mobility class for the United States, 1998-2017; bubble size represents the percentage of researchers in each mobility class; source: Scopus.

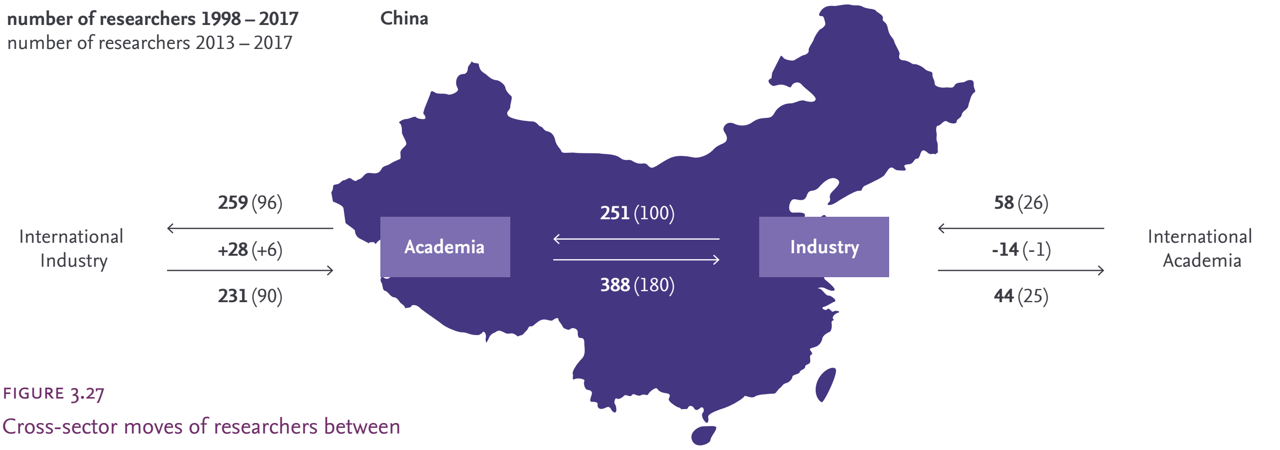


FIGURE 3.27
Cross-sector moves of researchers between academia and industry, either domestically or internationally, for China, 1998-2017; source: Scopus.

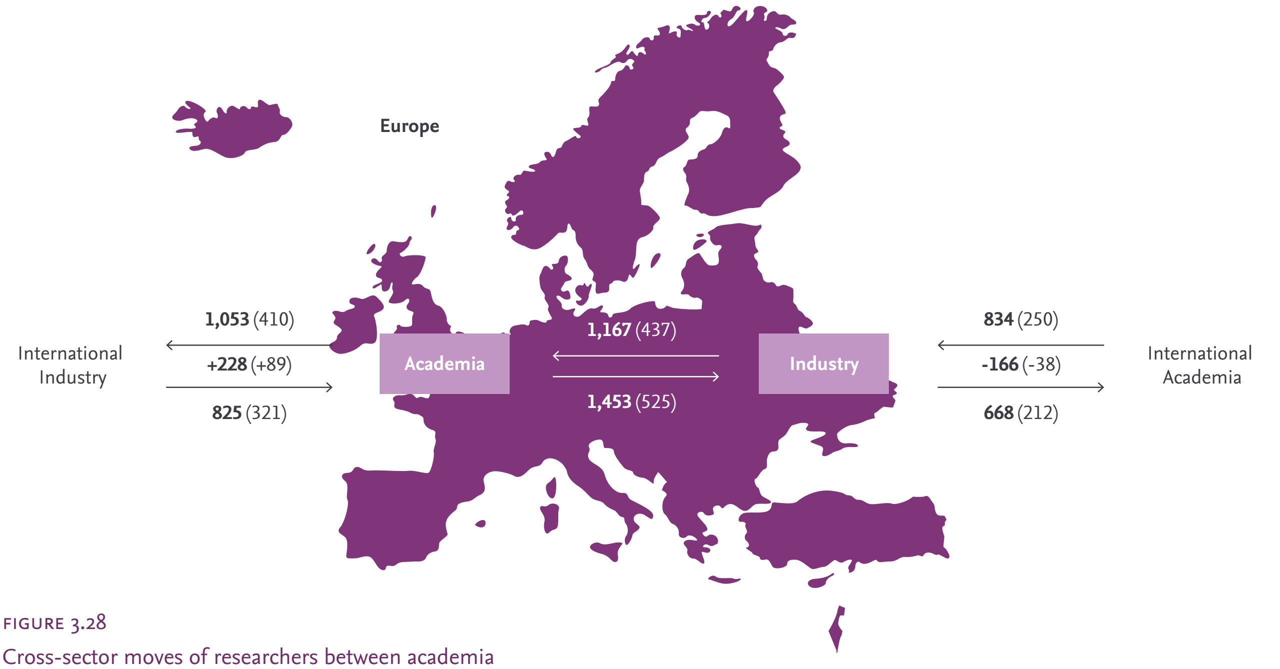


FIGURE 3.28
Cross-sector moves of researchers between academia and industry, either domestically or internationally, for Europe, 1998-2017; source: Scopus.

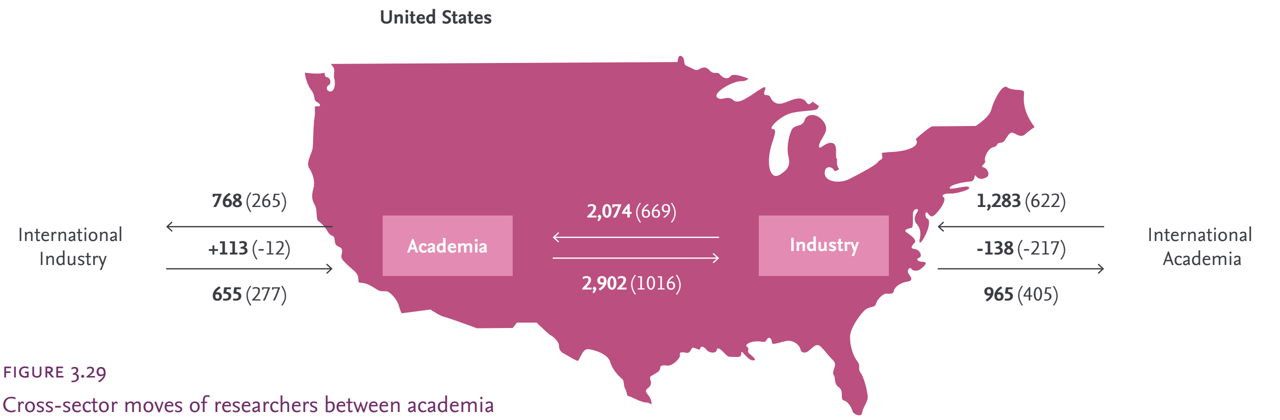


FIGURE 3.29
Cross-sector moves of researchers between academia and industry, either domestically or internationally, for the United States, 1998-2017; source: Scopus.

Industry in the United States attracts by far the most AI talent from international academia.

Researcher mobility is not only constrained to geographical movements—researchers can also move between sectors. The analyses in Figures 3.27-3.29 provide insights into the cross-sector mobility of researchers between academia and industry, both within a country and internationally. Overall, there were more moves by researchers from academia to industry than vice versa across all regions in most of the 20 years. This might speak to the academic mandate to educate for societal impact. For the last five years (2013-2017), along the accelerated growth of research publications, the situation changes and speaks to the brain drain discussion, for instance in Europe, but the time frame is too short to confirm a full trend shift. Europe seems to be losing academic talent, while attracting AI talent for local industry. While Europe sees 38 more researcher movements from international academia to Europe than vice versa, it faces, in regional comparison, strong net outflow of academic talent to international industries.

The United States achieves a net inflow of AI talent, both in academia and industry. Industry in the United States has attracted by far the most AI talent in the last five years. Whether they are coming from the outflow of academics from Europe and China requires further investigation.

In summary, AI research is a global competition among established and emerging research regions. Research boundaries and formats are blurring around conferences, corporate contributions, competitions, and social media dialogue. Research collaboration and researcher mobility, both across geographies and sectors, contributes to knowledge transfer and yields citation impact benefits. Analyzing those data more systematically should provide further transparency into AI research dynamics and might also help unveil the impact of AI as a general-purpose technology.

Chapter 4

Artificial Intelligence education

Educating enough AI talent, fast enough to satisfy corporate and research demand, is a key challenge. Digital education formats provide important support. They not only resonate with AI-interested audiences but also offer lower entry barriers for students across the globe. Next to Open Machine Learning platforms like AI competitions, there are many popular Massively Open Online Courses (MOOCs) offering self-learning facilities. This chapter presents a brief overview of the online education space as well as a case study on AI education at the Institute of Automation, Chinese Academy of Sciences.

Highlights

As of December 2017, there were over 9,000 MOOCs offered by over 800 universities worldwide.

SECTION 4.1

Most graduates at the Institute of Automation, Chinese Academy of Science select applied programs such as Pattern Recognition and Intelligent Systems, Computer Application Technology, and Control Theory.

SECTION 4.2

4.1 A brief overview of online AI education

As of December 2017, there were over 9,000 MOOCs offered by more than 800 universities worldwide.[90] Additionally, private sector companies are increasingly partnering with MOOC platform providers to provide courses. A few rise to the top in terms of traction gained with number of learners, in particular, Coursera, edX, XuetangX, Udacity, and FutureLearn.[91] Unfortunately, other than per-course anecdotes about participation and graduation rates, none of these platforms appear to provide detailed metrics about the number of AI courses provided or the number of learners over time in their AI courses.

Coursera released the top 10 most popular courses on their platform for 2017; 3 of the top 10 relate to AI or machine learning.[92] Interestingly, all three courses are not produced by a university, but by a private company, whose founder has significant industry credentials.[93] One of the most popular MOOC platforms, Udacity, delivers its courses with very little university involvement.[94]

Google, Microsoft, and Nvidia have all launched their own online learning platforms relating to AI and machine learning to help democratize AI and ensure that their recruitment pipelines for engineers with these skills remains full, as well as to promote adoption of their hardware and cloud platforms.[95]

90 Shah, D. By the numbers: MOOCs in 2017. Class Central. 18 January 2018. https://www.class-central.com/report/mooc-stats-2017/.
91 Ibid.
92 Sinha, N. Year in review: 10 most popular courses in 2017. Coursera Blog. 14 December 2017. https://blog.coursera.org/year-review-10-popular-courses-2017.
93 Young, J.R. Andrew Ng is probably teaching more students than anyone else on the planet (without a university involved). EdSurge. 7 June 2018. https://www.edsurge.com/news/2018-06-07-andrew-ng-is-probably-teaching-more-students-than-anyone-else-on-the-planet-without-a-university-involved.
94 Paterson, J. Despite overall setbacks, one MOOC on AI gains ground. Education Dive. 15 June 2018. https://www.educationdive.com/news/despite-overall-setbacks-one-moocon-ai-gains-ground/525812/.
95 Bhatia, R. What does Google, Microsoft stand to gain from launching free MOOCs in AI. Analytics India. 11 April 2018. https://www.analyticsindiamag.com/what-does-google-microsoft-stand-to-gain-from-launching-free-moocs-in-ai/.

Interviews

Lynda Hardman
Director, Amsterdam Data Science,
Past President, Informatics Europe, and
Manager Research & Strategy, Centrum
Wiskunde & Informatica (CWI)

How can talent shortage in AI be addressed?
Addressing the AI talent shortage is key to the success of AI, particularly in Europe where coordinated large-scale initiatives are currently lacking. We require more capacity for both teaching students and funding researchers. Additionally, coordinated funding should build on existing structures and networks rather than create new ones to avoid fragmentation of scarce expert resources. The principles of computing science and AI—in addition to the societal implications of automated decision-making based on data—must be taught in schools. We need to educate not only an elite group of AI experts, but also politicians and policy makers, so that they fully understand the implications of decisions made by data-driven algorithms.

A report like this is an invitation to explore facets of AI. What discussions and future research do you find most thought-provoking and would you like to see?
Turning decisions over to automated algorithms requires an understanding of the implications beyond the technical, to also include the ethical, societal, and political. From a computer science perspective, we need to ensure that results are explainable, unbiased, and transparent. I strongly believe that “responsible AI” is an asset, particularly for Europe. Following the Informatics Europe recommendations on automated decision-making, this report offers a number of indicators to help us navigate the field. I am pleased to see Elsevier raising awareness of the AI field through this report, which reveals the richness of a discipline that goes beyond data-driven solutions.

Frank van Harmelen
Professor, Knowledge Representation
& Reasoning in the Computer Science
department (Faculty of Science) at the Vrije
Universiteit, Amsterdam, The Netherlands

Which perspectives particularly draws your attention?
The education perspective. It is interesting to realize that we include ethical aspects in today’s curriculum, not only because of government mandates, but also because it is seen as increasingly important by researchers themselves. I realized that I’m not aware of any specific AI ethics journals or sections. I also recognized that students’ expectations (towards an AI education) are strongly influenced by social media, whereas our curricula are based on proven scientific insights and practical case studies. This poses an interesting challenge that we are trying to address in our new cross-institute Amsterdam School of Data Science, which offers a broad and innovative set of data science courses. We also see AI massive online open courses (or MOOCs) as a great opportunity for education as well as their own research area within the field of education. AI MOOCs have evolved from a way to transfer knowledge to a method of research co-creation. In general, the question remains, “How can AI education keep pace with its fast evolution?”

A report like this is an invitation to explore facets of AI. What discussions and future research do you find most thought-provoking and would you like to see?
AI has a great potential in advancing science itself. Once we make AI concepts and processes transparent and explainable, it will not only accelerate the development of new scientific hypothesis, but also justify the testing of those hypotheses with the help of AI. I see a future where computers move from being simple tools used in science to becoming our “colleagues” in science.

4.2 Case study on AI graduates in China

To illustrate trends in AI education in China, we analyze data from the Institute of Automation, Chinese Academy of Sciences (IA, CAS). From 2013 to 2017, there were 140-160 students annually enrolled at IA, CAS to pursue postgraduate study. Most graduates at IA, CAS selected applied programs, such as Pattern Recognition and Intelligent Systems, Computer Application Technology, and Control Theory. School recommendation and dispatch plays a large role in the destination of graduates of AI higher education at IA, CAS.

At IA, CAS, there are five common subject areas related to AI: control engineering, control theory and control engineering, pattern recognition and intelligent systems, computer application technology, and computer technology. Figure 4.1 shows the subject area distribution of the graduate students at IA, CAS from 2013 to 2017. Control engineering is the only subject area in which masters students study. Pattern recognition and intelligent systems has the largest number of graduate students followed by computer application technology and control theory and control engineering. From 2013 to 2017, control engineering is the only masters students' subject area.

In China, university and research institutes are also responsible for contacting employers and/or creating job hunting plans for undergraduates and graduates. The options are: dispatch (by the university or research institute), postgraduate (which usually means pursuing a PhD degree), secondary dispatch, or going abroad (Figure 4.2). Most IA, CAS graduates benefit from primary or secondary dispatch, being sent by IA, CAS directly to work. It seems easy for graduates in AI to find a job in China. Only a few of them pursue a PhD degree, and few go abroad.

In summary, in AI, education can't wait years until the knowledge consolidates. The demand for AI talent has now moved beyond existing capacities. Policy governance succeeds in some regions, while others build new online formats of education, exploration, and integrated research. While ethics are part of some AI curricula, it might need to receive much higher attention to educate responsible graduates and drive responsible innovation.

Prof. Zhenan Sun
Institute of Automation,
Chinese Academy of Sciences
(CAS), China

"China's artificial intelligence research has developed very fast in recent years, increasing its global significance within the field. China has unique advantages in applied technology research and development, for example, in the area of face recognition. AI education has been receiving more and more attention in recent years, not only in universities, but also in vocational colleges, and even in secondary and primary schools. This growing AI talent base will result in even greater future development of the AI field in China."

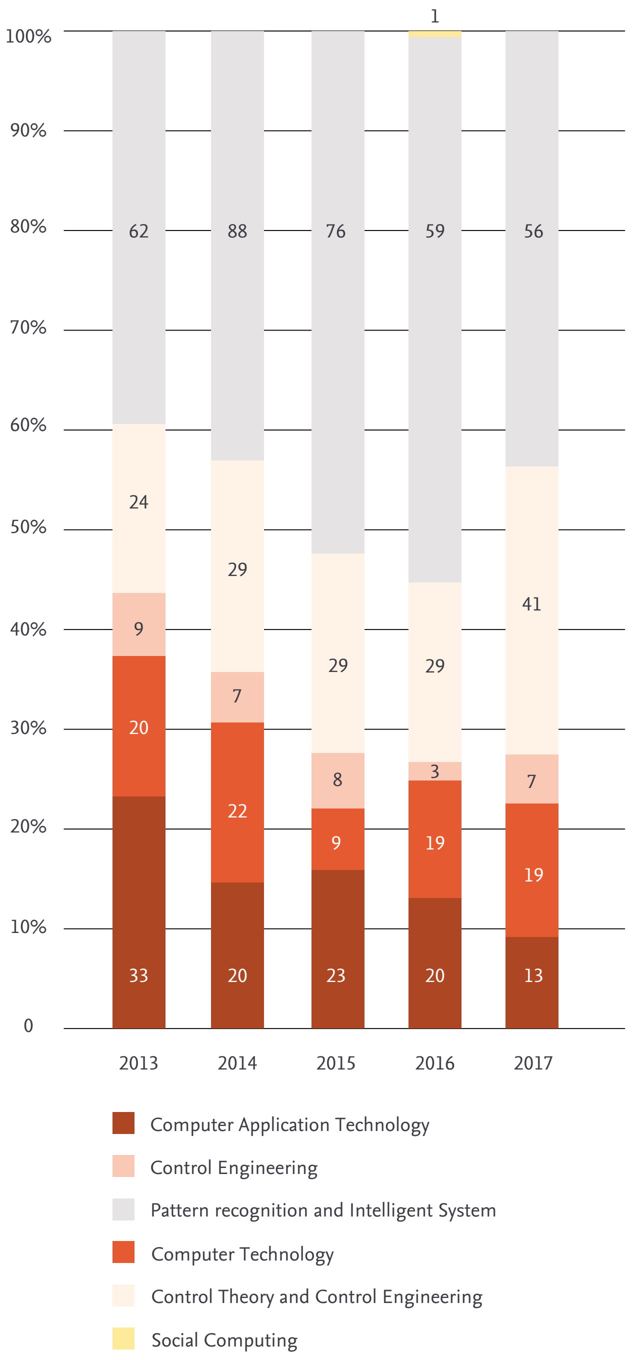


FIGURE 4.1
Subject areas of IA, CAS graduate students, 2013-2017; source: IA, CAS.

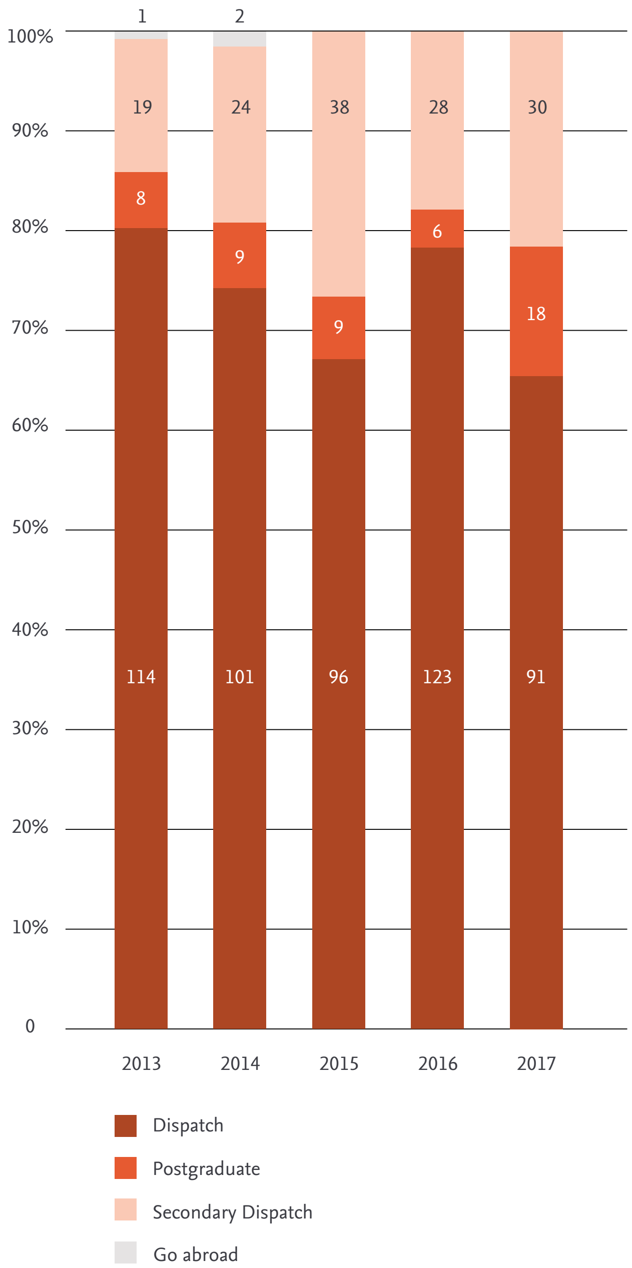


FIGURE 4.2
Types of graduate outcomes at IA, CAS, 2013-2017; source: IA, CAS.

Interview

Yuichiro Anzai
Senior Advisor, Director, Center for Science Information Analysis, Japan Society for the Promotion of Science (JSPS), Chairman, Strategic Council for AI Technology

What is your background in and relation to the AI field?
My PhD thesis was on systems theory and control theory and my first work in AI was on game-playing programs—writing computer programs and analyzing human behaviors so that the programs could eventually work as human players. I did a postdoc at Carnegie Mellon University starting in 1976, where I worked on human learning processes and machine learning. In the 1990s, during the so called "AI winter," not many people talked about AI, but I continued to work in the area. Interest was really rekindled when we learned about Prof. Geoffrey Hinton at the University of Toronto and researchers at Google making breakthroughs in the use of artificial neural networks for image recognition. I have always been "chasing two rabbits at once": AI and cognitive science. These two fields were separate in the late 1980s, but we are now at a time when we can integrate them again. Today, the Japanese government is planning to launch a project that could make breakthroughs for AI to be truly useful to humans—it's a good time to integrate. But there is a lack of researchers who know both sides.

As chair of the Japanese government's AI strategy council, can you tell us about topics you address?
The Japanese AI strategy council was established in April 2016 by the Prime Minister. One of the first things we did was to publish an industrialization roadmap for what we here in Japan refer to as "Society 5.0." This roadmap has four pillars: productivity, healthcare, mobility, and infrastructure. We are very concerned about how to grow AI talent. We need to think about education—from primary school to university to lifelong adult education.

Many countries form national AI strategies and see AI as part of competitiveness. Does there need to be a national strategy to bring together various players?
We see the increasing strength of the United States, China, and Europe. In Japan, we need to emphasize our own strengths, including manufacturing, the Internet of Things (IoT), robotics, and cyber-physical systems. We also need to put more emphasis on applied AI within the service industry, healthcare and other areas, and prioritize the use of AI in finance and cybersecurity. Agriculture is also an important field where supply chains can be further optimized.

In terms of AI research areas, where do you see underinvestment in the field?
The most underdeveloped area is the connection of deep learning with more contextual or symbolic processing. Human communication ability is far ahead of anything deep learning can do right now. From a cognitive science perspective, BDI (belief, desire, intention) is very difficult to infer externally. AI needs to integrate all of that information. Another crucial aspect is behavior justification, such as the ability of AI to explain its inference processes.

Can you tell us about the talent need and how AI may change society from a job perspective?
Narrow-minded AI talent will not take us very far—we need people with diverse skills. Also, in terms of nurturing innovation more broadly, we need to change the structure of industries to make it easier for start-ups to grow. We also need to change the employment system, especially in Japan, as it traditionally has been quite rigid. It's quite natural that job descriptions will change through technological innovation. It's a lesson from history. Jobs will change, and some jobs will disappear, but new ones will appear. I am optimistic, but we need to change the educational system in response.

Chapter 5

The imperative role of ethics in Artificial Intelligence

Experts suggest that ethics and AI appear in the public debate in three ways: the purpose of AI, the ways to incorporate beliefs, desire, and intentions into AI, and the abuse of AI. This chapter strives to explore initial insights into existing data on ethics, and underscores the need for a deeper dialogue and investigation into ethical aspects of AI for a comprehensive view and understanding of the field.

Highlights

Ethics are not an explicit consideration in AI research.

SECTION 5.1

Addressing the ethics of AI requires collaboration between technologists, policy makers, civil society, and other stakeholders.

SECTIONS 5.2

5.1 Ethics and AI

Within our keyword analysis and categorization in chapter 2, we found three rather general ethics-related keywords in the set of 797 (0.4%): “Ethical Values,” “Social Issues,” and “Social and Ethical Issues”—these only appeared in the perspectives of teaching and media. Given their importance in the comprehensive understanding of the field, we wondered and explored how to understand this. Are these actually non-AI terms, yielding non-AI publications, or are they relevant for the discussion and worth a separate thought? We pursued the latter question.

Stahl[96] and colleagues manually reviewed the ethics literature, also using “Ethics” as a keyword term, and identified further ethics categories, like “Privacy.” This corresponds to observations in calls for papers from leading 2018 AI conferences (following Stanford’s AI Index).[97]

From those conference papers, we manually extracted 326 additional keywords (>200 overlapped with the existing 797 keywords list), of which 22 were ethics-related (~5%): 10 referred directly to ethics and AI keywords (e.g., “Trustable AI,” “Explainable AI,” or “Values in Multiagent Systems”); 7 referred indirectly to it (e.g., “Belief-Desire-Intention Models,” “Logics for Norms”); and 5 used ethics-related keywords (e.g., “Human-aware Planning,” “Agents Competing Against Humans”). These terms were mostly spotted in multi-agent systems conference papers. Another indication for an increasing trend of ethics-related keywords and publications are the increasing publication numbers in the Scopus results for 2017/2018 on the identified keywords, such as “Explainable AI.”

Most of the terms focus on ethics/value approaches from within (multi-agent) systems, in contrast to potential regulatory, external approaches to enforce ethics, like policies. Some terms describe research fields rather than specific solutions and might require further semantic clarification. It is interesting to note that the accepted papers from AAMAS[987] (a multi-agent conference) refer to similar general terms, like those identified by Stahl et al in their review paper, such as “Value,” “Security,” “Privacy,” etc. This might suggest a base for further research.

96 Stahl B.C., et al. The ethics of computing: a survey of the computing-oriented literature. *ACM Computing Surveys*. 2016;48(4). Article No. 55. doi: http://dx.doi.org/10.1145/2871196.
97 Artificial Intelligence Index. 2017 Annual Report. http://cdn.aiindex.org/2017-report.pdf.
98 AAMAS 2018 Stockholm. Accepted Papers: Main Track. http://celweb.vuse.vanderbilt.edu/aamas18/acceptedPapers/.

Prof. Dame Wendy Hall
Professor of Computer Science in Electronics and Computer Science, University of Southampton, Director of the Web Science Institute, United Kingdom

Ethics are crucial to AI

Since its conception in the 1950's AI has experienced a seasonal cycle, with big spurts of funding when new AI technologies show potential for major breakthroughs, followed by the so-called AI winters due to the funding droughts that follow when the impact does not come through fast enough to sustain the funding. We are currently in a period of AI hype, driven by recent developments in computer processing power, the availability of big data, and the evolution of Deep Learning. These recent developments, however, build on research conducted a long time ago: expert systems were initially developed in the 80s, and work on multidimensional Neural Networks has been ongoing for a while. These peaks and troughs in AI funding are caused by the length of the AI development cycle: it overpromises and underdelivers within the short terms of funding cycles, as it depends on other advances to come to fruition. For these reasons, it is essential to continue to support long-term, blue-sky research in AI between the hype cycles.

The advent of the World Wide Web and subsequently the Semantic Web has helped advance research: anyone with access to the Internet can now use collaborative documents to work across geographies and time-zones,we have powerful search engines to help us find the information that we want when we want it, and social media or other messaging applications to facilitiate communication between scientists. At the same time however, we have seen the growth of criminal and anti-social behaviour on the Internet such as the dissemination of fake news, the promulgation of propaganda, and even the proliferation of terrorism. These are issues that technical solutions alone cannot solve. Human intervention is necessary to tackle the ethical and social issues that access to an open and free internet brings. Special attention is needed to guide AI to avoid perpetuating prejudice. For instance, training sets must be carefully selected to minimise bias in algorithms, whose future uses and applications cannot be predicted. A successful AI is an inclusive AI, accounting for diversity of gender, age, or origin. The good news is that people can come into AI from a variety of backgrounds - interdisciplinarity is necessary for developing responsible and ethical AI.

Ethical concerns raise the issues of responsible innovation and regulation, and the challenge of finding the right balance between the two. Too little regulation may lead to unforeseen consequences, with potentially devastating impact given the growing presence of AI in our daily lives. Yet too much regulation can stifle innovation.

A successful AI strategy depends on data, computing capacity, education, talent, and diversity. Ethical considerations must also be embedded at the onset of any AI developments, to enable and ensure responsible innovations in the long term.

Machine Ethics are both prominent and relevant for AI.

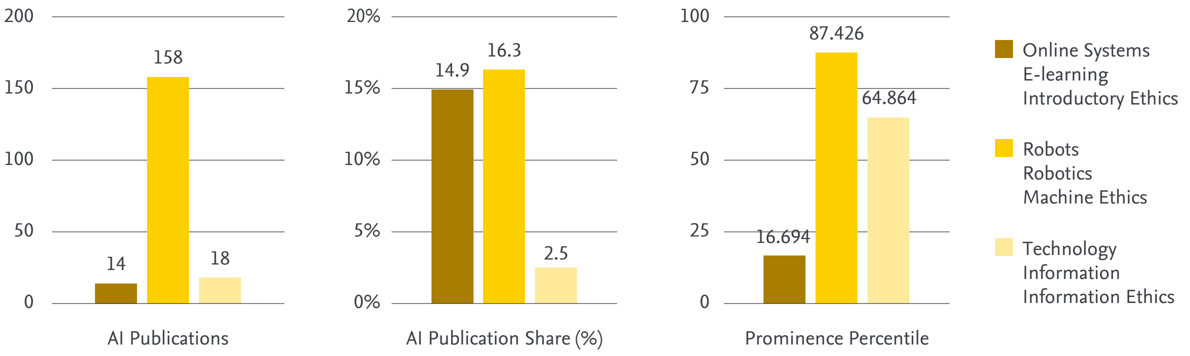


FIGURE 5.1
Ethics topics with relevant AI publication share, 2013-2017; source: SciVal.

Since ethics is an established research field, we explore the full range of Scopus data in form of SciVal Topics of Prominence for topics regarding ethics, and ethics in AI. The results include "Bioethics," "Business Ethics," "Ethics in Special Diseases," "Machine Ethics," "Ethics in Research," and "Ethics in Social Science" aspects, such as cultures. Overall, we see that ethics-related topics do not show a strong momentum (high prominence percentile). Only 22 topics include any of our AI papers and among those, 3 seem to be large enough and relevant for the AI discussion. Only one topic seems to be relevant and with sufficient momentum: "Machine Ethics."

Topic Prominence in Science
SciVal Topics are a collection of documents with a common intellectual interest—a "research problem." Topics can be large or small, new or old, and growing or declining, and can evolve. New topics can be born, and many topics are inherently multidisciplinary. Old topics may be dormant, but still exist. Across all of Scopus, SciVal clustered ~97,000 global research topics based on direct citation patterns. Overall, we find 33,671 topics with at least one AI publication (35.2%) and 4,212 topics with an AI publication share >10% (4.4%). Prominence is a new indicator that shows the current "momentum" of a topic by looking at very recent citations, views, and CiteScore values. It predicts and helps researchers and research managers identify topics that are likely to be well funded, given the correlation between prominence level and funding.

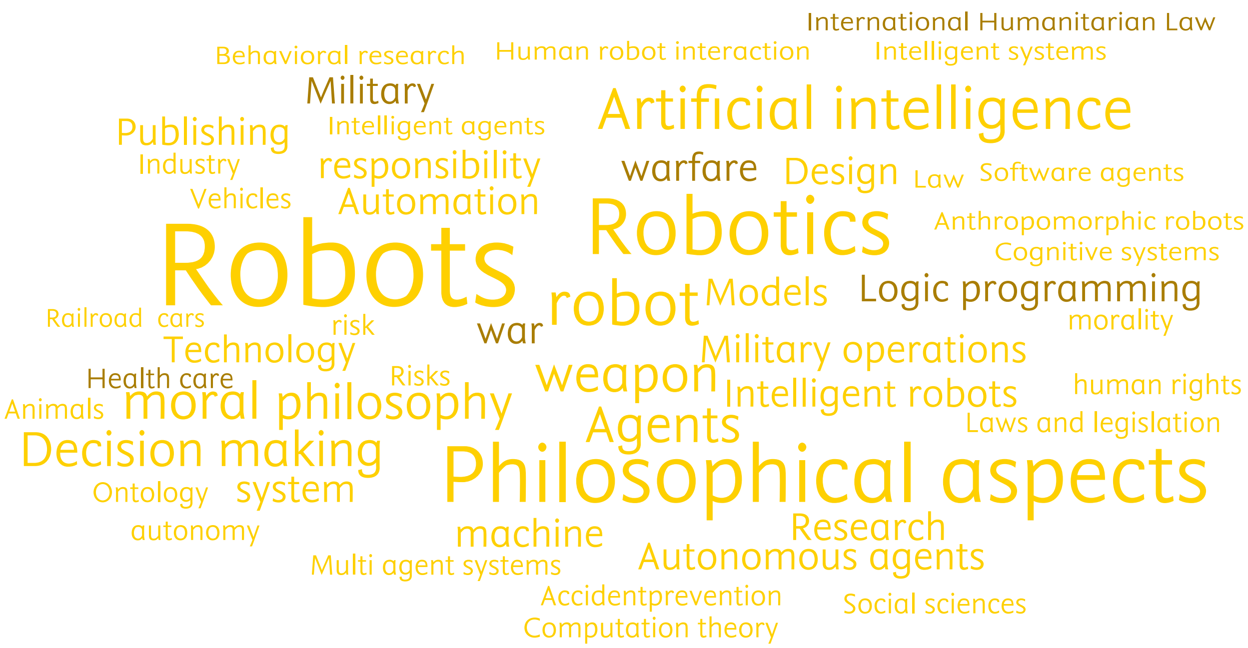


FIGURE 5.2

Most relevant key phrases among the topics "Robots," "Robotics," and "Machine Ethics" 2013-2017, with font size indicating relevance and color indicating growth (positive in yellow and negative in brown); source: SciVal.

Exploring this topic through its key phrases (Figure 5.2) reveals that it is predominantly concerned, and increasingly so, with "Robotics" and "Philosophy." This seems to suggest a focus on human-machine interactions and potential ethical, societal, and legal consequences of AI applications in robotics.

This short analysis supports initial findings on the disconnect between AI and ethics in joint research publications. The insights from conference papers indicate stronger interest in the question of how to integrate normative systems into AI. While the purpose and misuse of AI might be more of a political discussion, initial expert exchange invites exploration of reasons such as the acceptance guidelines of AI journals on ethics-related papers and corporate behavior in patenting, given the litigation risks with ethical topics.

5.2 AI for the good and AI doing good: questions on ethics and AI

The steadily growing stream of policy-related publications on artificial intelligence (AI)[99, 100, 101, 102] and continuing media interest highlight the importance of considering ethical and social issues in relation to AI. As the preceding section shows, the attention to ethics in the public discourse is not reflected in the research literature. There is growing attention to some specific concerns, notably privacy, trust, fairness, transparency, and accountability, but it is not clear whether addressing these individually can ensure that the socio-economic benefits of AI warrant the ethical costs.

To address ethical issues of AI, it is important to understand the underlying concepts, both in terms of AI and in terms of ethics. An autonomous vehicle, a facial recognition system, or an insurance claim checking algorithm are all examples of AI, but they have very different technical capabilities and the ethical questions they raise differ greatly. In addition to a better understanding of what is meant by AI, which this analysis set out to do, it also must be clear what is meant by ethics.

The current debate on the ethics of AI does not always provide the necessary depth and sometimes neglects the fact that there are millennia of ethical philosophical debate. The much-discussed question of the ethics of autonomous cars, frequently linked to the so-called trolley problem[103, 104] where a vehicle must decide between two different options, all of which injure humans and raise ethical concerns, is an example of this problem. The ethical quality of a human driver is determined by the consequences of what they do, but also by their understanding of the situation and their disposition to act in certain ways. From a philosophical perspective, this constitutes a complex mix of different ethical positions that needs to be understood to come to an ethical evaluation of the action. It is unclear how this complexity, namely the difference between action, its justification, and internal states, can be reflected in AI. This points to the larger question of how the ethics of AI can be looked at in a way that is philosophically sound as well as practically relevant.

ORBIT,[105] the Observatory for Responsible Research and Innovation in Information and Communication Technologies (ICT) is a project funded by the UK Engineering and Physical Sciences Research Council (EPSRC). Members of the ORBIT team who have contributed to this text are listed.

Observatory for Responsible Research and Innovation in ICT (ORBIT):
Margherita Nulli (left), ORBIT Project Officer;
Bernd Stahl, Investigator, Director of the Centre for Computing and Social Responsibility at De Montfort University;
Martin De Heaver, Managing Director;
Marina Jirotka, Investigator, Professor of Human Centred Computing;
Carolyn Ten Holter (right), Marketing Officer;
Paul Keene (not in photo), Online Director

Another key question is the level at which the ethics of AI needs to be addressed. Much of the technical work that aims to address the ethics of AI looks at an ethical quality at the technical level (e.g., the algorithm behind an application) and the individual and organizational responsibilities in developing and deploying these.[106] While this is no doubt important work, it may well be blind to key ethical issues related to ownership, intellectual property, economic distribution, and political power that AI also raises. The former tries to ensure that the consequences of AI use are ethically acceptable, i.e., that AI does good, while the latter focus on the broader socio-economic and political consequences. When addressing the ethics of AI, both aspects are important, but it is not clear whether and how they can be governed simultaneously.

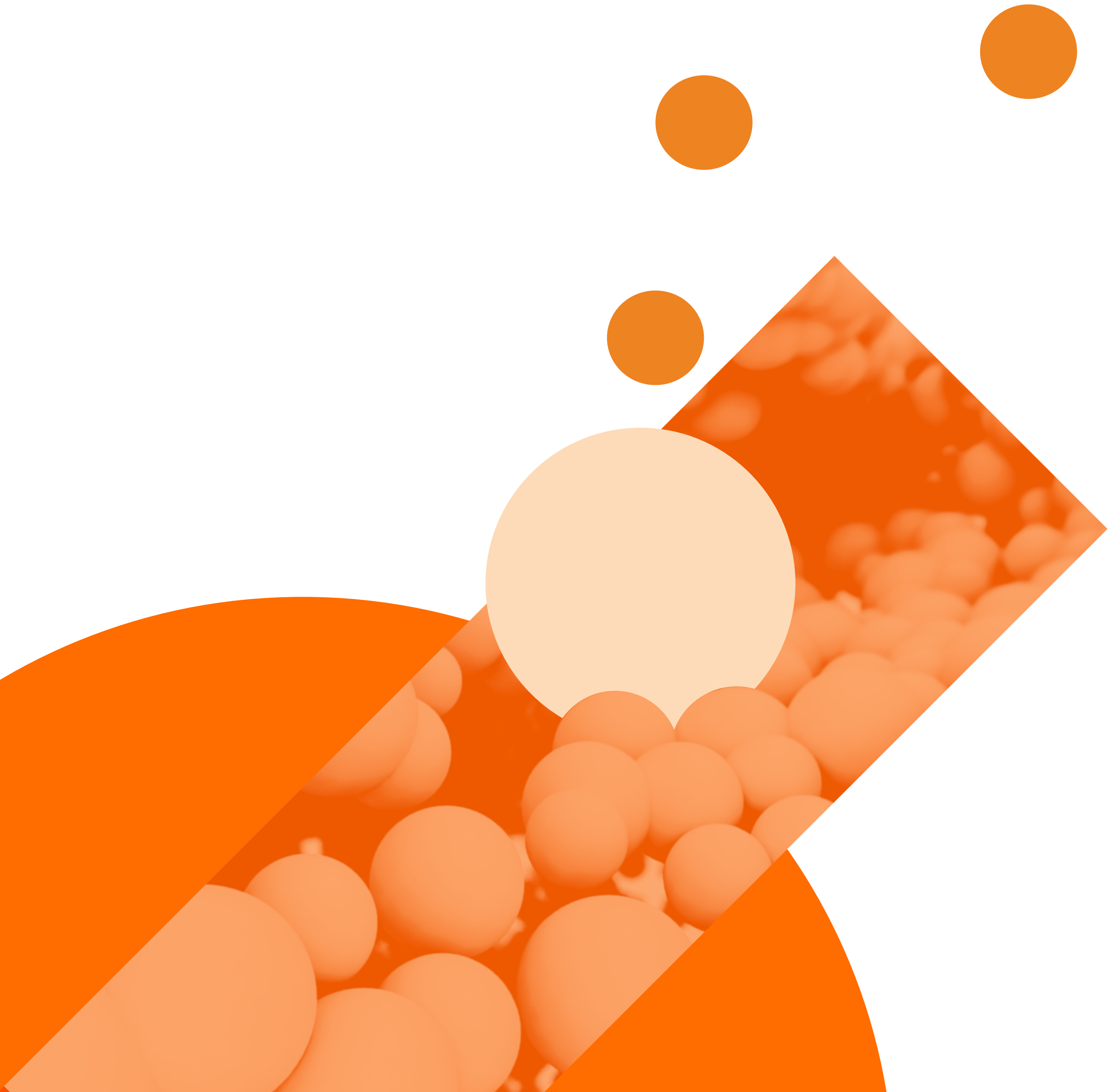
A final question we would like to raise is that of the novelty of the ethics of AI. The debate of ethical issues of computing goes back to the very beginnings of digital technology.[107,108] The same is true about AI.[109] This raises the question whether in the current discussion of AI one can draw on existing insights, tools, and methods, or whether AI poses ethical problems of a fundamentally novel nature that require radically new thought.

While we have tried to highlight some open questions with regards to ethics and AI relevant to this report, we believe that the following key points require further attention:

- The ethics of AI needs to be aware of and build on existing ethical thought. However, it also needs to work in partnership with technical and business communities to shape the ethical outcomes. This should involve existing processes, such as those developed in responsible research and innovation in ICT.[110]
- Ethics is not a fixed set of rules that determine good and bad but is thoroughly embedded in social context. To ensure the benefits of AI while addressing the pitfalls, appropriate governance mechanisms need to be developed.[111]
- AI is not just a technical tool. Due to its potentially enormous impact, addressing the ethics of AI requires collaboration between technologists, policy makers, civil society, and other stakeholders.

Overall, identifying and addressing the ethics of AI will be a large task and societally driven. To do this, the clarification of concepts, as done in this study, is important. Further research could look at the temporal development of different types of AI research as well as ethical questions. The technical capabilities of AI are likely to develop rapidly, calling for constant reflection as well as empirical insights into social and ethical consequences. Research such as the present report should therefore be dynamic, open to all, and ethical questions should form a natural component of all AI research and education.

99 European Group on Ethics in Science and New Technologies. Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems. Luxembourg: Publications Office of the European Union; 2018. https://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf.
100 Executive Office of the President National Science and Technology Council Committee on Technology. Preparing for the Future of Artificial Intelligence. October 2016. https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.
101 House of Lords Select Committee on Artificial Intelligence. Report of Session 2017–19. AI in the UK: ready, willing and able? HL Paper 100; 2018. https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf.
102 House of Commons Science and Technology Committee. Robotics and artificial intelligence. Fifth Report of Session 2016-17. HC 145; 2016. https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/145/145.pdf.
103 Foot, P. *Virtues and Vices and Other Essays in Moral Philosophy*. Oxford, UK: Oxford University Press; 2002.
104 Wolkenstein, A. What has the Trolley Dilemma ever done for us (and what will it do in the future)? On some recent debates about the ethics of self-driving cars. *Ethics Inf Technol*. 2018;20(3):1–11. https://doi.org/10.1007/s10676-018-9456-6.
105 www.orbit-rri.org.
106 FAT/ML. Principles for Accountable Algorithms and a Social Impact Statement for Algorithms. https://www.fatml.org/resources/principles-for-accountable-algorithms.
107 Weizenbaum, *J. Computer Power and Human Reason: From Judgement to Calculation*. New York, NY: W. H. Freeman & Co Ltd; 1976.
108 Wiener, N. *The Human Use of Human Beings*. New York, NY: Da Capo Press; 1954.
109 Dreyfus, H.L. *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge, MA: MIT Press; 1972.
110 Jirotka, M., et al. Responsible research and innovation in the digital age. *Comm ACM*. 2017;60(5):62-68. https://doi.org/10.1145/3064940.
111 Winfield, A.F., Jirotka, M. Ethical governance is essential to building trust in robotics and AI systems. *Philos Transact A: Math Phys Eng Sci*. 2018. http://eprints.uwe.ac.uk/37556.



Concluding remarks and future research

Exploring a dynamic, emerging, complex, and changing field like AI is a fascinating endeavor, and we hope our report provides useful insights into the field as well as inspires further research and exploration of the field and its applications and implications.

The exchange around this report has made it clear that innovation, driven by AI as a field of technological capabilities and applications, is not only a technological challenge, but is largely driven by data, computing infrastructure, and societal acceptance. In this, AI is probably not different to previous general-purpose technologies and might benefit from that experience (e.g., definition, evolution cycles, success factors, societal impact, ethics, etc.).

We hope that this report provides a glimpse into the multifaceted nature of AI to help knowledge exchange and dialogue among stakeholder groups. We also anticipate that these insights may inform research and funding strategies.

We understand that, given the evolving nature of the field, we need ways to stay up-to-date. The Elsevier AI Resource Center[112] offers a platform to provide further insights, connect to others' work, and foster further research and discussion. We particularly look forward to engaging in efforts in the following directions.

1 Scoping and structuring of the AI field

Other perspectives, data sources, and algorithms could help advance the scoping and structuring of the field and contribute to a broader approach to identify and shape emerging research fields.

→ For this, we will continue our semantic research and innovation around AI ontologies.

2 Monitoring the emergence and dynamics of AI research

A basket of relevant metrics will help to build trust in systematic analysis. Aligning and agreeing on appropriate ways to monitor the evolution and impact of the field will stay a core focus.

→ We will continue our analytical efforts and help partners establish and run AI monitors.

3 Knowledge transfer and impact in other societal sectors

Different application sectors accelerate in different regions and require differentiated AI capabilities (e.g., "Computer Vision" versus "Search and Optimization").

→ We will provide examples of AI applications and illustrate their impact on societal challenges.

4 Facilitating the dialogue for responsible innovation

We gained awareness about the disconnect of ethical topics and AI. This includes the challenges of data bias and the need for more systematic dialogue.

→ We will explore ways to support this dialogue, such as roundtables or in our journals.

112 Elsevier. Artificial Intelligence Resource Center. https://www.elsevier.com/connect/ai-resource-center.

Interview

Prof. Ingrid Ott, Karlsruhe Institute of Technology (KIT), Chair in Economic Policy, Member of the Commission of Experts for Research and Innovation (EFI) in 2014-2018, Germany

AI seems to lack a universally accepted definition. What is the best way to navigate such a field?
Despite the abundance of AI research activity, the notion of AI is still fuzzy. To avoid misunderstandings in communication, a comprehensive conceptualization of AI is therefore indispensable. A good approach is to understand the field from the bottom up by integrating the perspectives of various disciplines and actors and being aware of the dynamics associated with the evolution of the technology field.

AI connects sectors. What does this look like in practice?
I see AI as a typical general-purpose technology (GPT), like the steam engine, electricity, or information and communication technologies (ICT). As such, it is characterized by pervasiveness, i.e., it will diffuse into almost any part of our economies and lives. Pervasiveness allows for linking to far more or less isolated fields such as nanotechnology, the most recent GPT since ICT. Nanotechnology is especially successful in designing new materials, nowadays used in completely heterogenous contexts, e.g., for coating prostheses, rotor blades of windmills, or outer walls of ocean giants. My point with this example is that GPTs like AI are the binding element between such diverse industries as the health and life sciences, green energy, logistics, or arts. Throughout the early stage of technology development and the associated strong potential for further improvement, efficiency gains can quickly be realized if the needs of the application sectors are coordinated. Platforms that bundle the needs of the different actors—and the platform design—are of special importance.

What role does innovation play in the broader AI ecosystem, with strong industry influence on the one hand and huge societal impact on the other?
I see the largest economic potential of AI in the complementarity of innovation processes between AI and downstream industries that perpetually and mutually fuel themselves (so-called innovational complementarity). The associated implications of future AI go far beyond mere technological or economic considerations. To get a grasp of this feeling, I find it helpful to look back at the implications of today's well-established GPTs. Electricity has made value creation independent from access to daylight; the use of ICT allows for remote work. But the potential of GPTs may only be exploited if at the same time family life is re-organized accordingly. This also causes friction within the existing social security system. Both examples also highlight the necessity of secure and stable access to complementary infrastructure as essential conditions. Frictions on the level of complementary technologies thus affect productivity of the GPT.

A report like this is less of a conclusion, and more of an invitation to explore the facets of AI. What discussions and future research are most thought-provoking, and would you like to see?
Like any key technology, AI also has the potential of being "Janus-faced." Its further development and diffusion come with challenges and opportunities. The abovementioned complementarities make AI-enhanced production processes not only more complex but also more vulnerable to abuse. We thus must continuously develop the institutional settings under which the technology is developed and used, without being naïve or anxious. I strictly plead for extensive basic understanding of the functioning logic of AI not only for AI developers but also for those who apply AI. What I have in mind might be called "AI literacy," which I see as an essential capacity even at the level of private users. The direction of technological change is shaped by us!

Appendices
Artificial Intelligence experts

Alessandro Annoni
Head of Digital Economy Unit
Joint Research Centre
European Commission

Dr. Roberto M. Cesar Jr.
Adjunct Coordinator
São Paulo Research Foundation
(FAPESP)
Brazil

Prof. Frank van Harmelen
Professor, Knowledge
Representation & Reasoning
Vrije Universiteit Amsterdam (VU)
The Netherlands

Dr. Yuichiro Anzai
Senior Advisor, Director, Center for
Science Information Analysis, Japan
Society for the Promotion of Science
(JSPS)
Chairman, Strategic Council for AI
Technology
Japan

Prof. Dame Wendy Hall
Professor of Computer Science in
Electronics and Computer Science
University of Southampton
Director of the Web Science
Institute
United Kingdom

Fredrick Heintz
Associate Professor of Computer
Science
Linköping University
President Swedish AI Society
Expert Member High Level Expert
Group on Artificial Intelligence
European Commission
Sweden

Prof. Lynda Hardman
Director, Amsterdam Data Science,
Past President, Informatics Europe,
Manager Research & Strategy,
Centrum Wiskunde & Informatica
(CWI)

Martin De Heaver
Managing Director
Observatory for Responsible Research and Innovation in ICT (ORBIT)
United Kingdom

Marina Jirotka
Investigator Observatory for Responsible Research and Innovation in ICT (ORBIT),
Professor of Human Centred Computing
United Kingdom

Margherita Nulli
ORBIT Project Officer
Observatory for Responsible Research and Innovation in ICT (ORBIT)
United Kingdom

Carolyn Ten Holter
Marketing Officer Observatory for Responsible Research and Innovation in ICT (ORBIT)
United Kingdom

Paul Keene
Online Director
Observatory for Responsible Research and Innovation in ICT (ORBIT)
United Kingdom

Prof. Ingrid Ott
Chair in Economic Policy
Karlsruhe Institute of Technology (KIT)
Member of the German Expert Commission on Research and Innovation (EFI) 2014-2018
Germany

Prof. Enrico Motta
Professor of Knowledge Technologies
The Open University
United Kingdom

Dr. Raymond Perrault
Senior Technical Advisor
Artificial Intelligence Center at
SRI International
United states

Bernd Stahl
Investigator Observatory for
Responsible Research and
Innovation in ICT (ORBIT),
Director of the Centre for
Computing and Social Responsibility
at De Montfort University
United Kingdom

Prof. Chuan Tang
Associate Researcher
Chengdu Library and
Information Center
Chinese Academy of Sciences
(CAS)
China

Giuditta De Prato
Team leader / Scientific Officer
Digital Economy Unit
Joint Research Centre
European Commission

Prof. Zhenan Sun
Institute of Automation (IAS)
Chinese Academy of Sciences
(CAS)
China

Dr. Zhiyun Zhao
Director New Generation
Artificial Intelligence
Development Research Center
Party Committee Secretary
Institute of Scientific and
Technical Information of China
(ISTIC)
China

Prof. Tieniu Tan
Institute of Automation (IAS)
Chinese Academy of Sciences
(CAS)
China

Elsevier and other contributors

Program Directors
Maria de Kleijn
Nick Fowler

Program Manager
Clive Bastin

Content & Analytics
Sarah Huggett
Mark Siebert
Jorg Hellwig
Polly Allen
Amrita Purkayastha
Basak Candemir
Jeroen Baas
Jeroen Geertzen
Kyle Defrancesco
Milan Splinter
Stephanie Faulkner
Thomas Gurney
Wei Wang
Yu Hsuan Chou

Subject matter experts
Dan Olley
Ron Daniel
Paul Groth
Anthony Scerry
Jessica Coxs
Curt Kohler
Sweitze Roffel
Rinke Hoekstra
George Tsatsaronis

Engagement
Dante Cid
Anders Karlsson
Stephane Berghmans
Ann Gabriel
Xiaoling Kang
Karina Lott
Federica Rosetta
Lesley Thompson
Max Voegler

Communications
Marianne Parkhill
Sacha Boucherie
Taylor Stang

Writer
Stacey Tobin

Design
Edenspiekermann

Appendices

Methodology

Methodology and rationale

Our methodology is based on the theoretical principles and best practices developed in the field of quantitative science and technology studies, particularly in science and technology indicators research. The Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems (Moed, Glänzel and Schmoch, 2004)[113] gives a good overview of this field and is based on the pioneering work of Derek de Solla Price (1978),[114] Eugene Garfield (1979),[115] and Francis Narin (1976)[116] in the United States, and Christopher Freeman, Ben Martin, and John Irvine in the UK (1981, 1987),[117] and in several European institutions including the Centre for Science and Technology Studies at Leiden University, The Netherlands, and the Library of the Academy of Sciences in Budapest, Hungary.

The analyses of bibliometric data in this report are based upon recognized advanced indicators (e.g., the concept of relative citation impact rates). Our base assumption is that such indicators are useful and valid, though imperfect and partial measures, in the sense that their numerical values are determined by research performance and related concepts, but also by other, influencing factors that may cause systematic biases. In the past decade, the field of indicators research has developed best practices that state how indicator results should be interpreted and which influencing factors should be taken into account. Our methodology builds on these practices.

A body of literature is available on the limitations and caveats in the use of such bibliometric data, such as the accumulation of citations over time, the skewed distribution of citations across articles, and differences in publication and citation practices between fields of research, different languages, and applicability to social sciences and humanities research.[118]

Document types

We use all document types to provide a comprehensive view of the field, including articles and conference paper breakdowns when needed:

- Research Article
- Book Chapter
- Newspaper Article
- Report
- Review
- Conference Paper

Comparators

The report focuses on China, Europe, and the United States to provide regional insights from large entities with comparable research output. Recognizing that research performance is often tied to funding levels, we define Europe as the 28 countries in the European Union (EU: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom) and an additional 16 that are eligible for Horizon 2020 funding (Albania, Armenia, Bosnia and Herzegovina, Faroe Islands, Georgia, Iceland, Israel, Moldova, Montenegro, Norway, Serbia, Switzerland, the former Yugoslav Republic of Macedonia, Tunisia, Turkey, and Ukraine).

Counting

All analyses make use of whole counting rather than fractional counting. For example, if a publication has been co-authored by one author from China and one author from the United States, then that publication counts towards both the publication count of China, as well as the publication count of the United States. Total counts for each country are the unique count of publications.

113 Moed H., et al. *Handbook of Quantitative Science and Technology Research*. Dordrecht, Germany: Kluwer; 2004.

114 de Solla Price, D.J. (1977–1978). “Foreword,” *Essays of an Information Scientist*, Vol. 3, v–ix.

115 Garfield, E. Is citation analysis a legitimate evaluation tool? *Scientometrics*. 1979;1(4):359-375.

116 Pinski, G., Narin, F. Citation influence for journal aggregates of scientific publications: theory with application to literature of physics. *Information Processing & Management*. 1976;12(5):297-312.

117 Irvine, J., et al. Assessing basic research: Reappraisal and update of an evaluation of four radio astronomy observatories. *Research Policy*. 1987;16(2-4):213-227.

118 Elsevier. Research Metrics Guidebook. 2018. https://www.elsevier.com/research-intelligence/resource-library/scival-metrics-guidebook.

Fingerprinting

We use the Elsevier Fingerprint Engine®[119] based on Natural Language Processing (NLP) techniques to identify the main topics and concepts from unstructured text. This includes scientific articles, abstracts, funding announcements and awards, project summaries, patents, proposals, applications, and other sources. The unstructured text is mapped to a ranked set of standardized, domain-specific concepts that define the text, known as a Fingerprint. By aggregating and comparing fingerprints, the engine looks beyond metadata.

Identifying preprints in artificial intelligence (AI)

The arXiv preprint metadata corpus is available via their public API using metadata queries, or by OAI-PMH for bulk download. For this analysis, it was downloaded via OAI-PHM on August 20, 2018 and included 1,424,193 documents. Of those records, 1,129 were found to be invalid (missing data, including unassigned primary keyword or year). This is less than 0.08% of records. For our keyword search, we used case-sensitive search that was mindful of word boundaries for abbreviations like "AI," but a case-insensitive search that ignores word boundaries for full terms. While not a full solution for word stemming, this allows us to find pluralization of these terms.

The first arXiv pre-prints that match the "core AI" keyword list were added in 1992 in the field of high-energy physics (subject codes hep-ph and hep-th). However, few documents match these terms prior to 1998: the 36,708 matching documents from 1998 forward represent more than 99% of all matching documents in arXiv. Our analysis focuses on submissions to arXiv 1998-2018, which includes 1,354,190 documents.

We ranked all arXiv subject areas based on the percentage of documents with the primary subject area that matched at least one keyword. Separately, we asked subject matter experts to indicate which arXiv categories they would consider to be highly related to core AI fields. The experts returned a list of 15 arXiv subject areas. Of the top 12 subject areas ranked, 11 were included in the list provided by the AI subject matter experts. The one subject area that was not included, cs.SD or "Computer Science – Sound," has a cryptically short name. In 2017, an Engineering subject area for "Audio and Speech Processing" was added; prior to that time, all audio processing computer science articles would have been included in cs.SD. We believe this was an easy category for our team of experts to miss without this information.

The 13th arXiv subject area on our ranked list was from Biology, "Neurons and Cognition." This research area is known to have many false positive results because of non-AI discussions of neural networks. Beyond that result, broader fields like "Human Computer Interaction" and "Emerging Technologies" appeared on the list, and while our 12th ranked subject area, "Robotics," had 19.9% matching documents, all other categories had fewer than 15.1%.

The three remaining categories that experts determined to be aligned with AI, but that had less than 15.1% matching documents, included "Social and Information Networks," "Computer Science and Game Theory," and "Condensed Matter - Disordered Systems and Neural Networks." These categories are possibly broader than others, which dilutes any focus on core AI research. Our team of experts might have interpreted these subject names differently than they are being used by the arXiv research community. Alternatively, it is possible that the list of 142 keywords is skewed away from research in these fields. Future research plans include establishing more robust methods for identifying AI research from titles and abstracts.

Inclusion of hypercollaborative articles

While hypercollaborative articles may represent extreme outliers in co-authorship data, they are included in all the analyses since they remain proportionally few and because they are counted only as a single internationally co-authored article for each country contributing to the article, and for each country pairing.

Measuring cross-sector researcher mobility

The approach presented here uses Scopus author profile data to derive a history of cross-sector mobility of active author affiliations recorded in their publications and to assign them to mobility classes defined by the type and duration of observed moves.

119 https://www.elsevier.com/solutions/elsevier-fingerprint-engine.

How are individual researchers unambiguously identified in Scopus?
Scopus uses a sophisticated author-matching algorithm to precisely identify articles by the same author. The Scopus Author Identifier gives each author a unique ID and groups together all the documents published by that author, matching alternate spellings and variations of the author's last name and distinguishing between authors with the same surname by differentiating on data elements associated with the article (such as affiliation, subject area, co-authors, and so on). This is enriched with manual, author-supplied feedback, both directly through Scopus and via Scopus' direct links with ORCID.

How are mobility classes defined?
For any given researcher, the publications of that researcher during the period are categorized as either Arrivals, Departures, or Domestic based on the author's affiliation during the period. Separately, publications are also categorized as either Academic or Industry depending on the type/sector of their institutional affiliation. We track researcher movement across sectors by analysing changes in the researchers' affiliations over time.

For comprehensiveness, although we do not start "counting" researcher movements prior to the period, if a researcher's portfolio predates the period of analysis, his or her initial category (e.g., Domestic Academic) is determined by the latest publication prior to the period. For example, if Researcher A publishes under an academic affiliation in 2016 and then publishes under a corporate affiliation in 2017, we count that as an academic-to-industry move for 2017. Moreover, if a researcher moves multiple times between academia and industry during the period, each move is counted separately toward that year's total cross-sector movement, with the limitation that a researcher can move a maximum of once in each direction per year. For instance, returning to our previous example, suppose Researcher A ping-pongs between the sectors frequently, publishing under an academic affiliation in 1999, a corporate affiliation in 2005, an academic affiliation in 2007, and then another corporate affiliation later in 2017. For this series of publications, we would count 1 move of academic to corporate in 2005 and 2017 each and 1 move of corporate to academic in 2007.

One complicating factor for such analyses is that some authors publish with two or more different affiliations, revealing their attachment to both academia and industry. These individuals could possibly publish with either or both affiliations, depending on the specific studies on which they are working. Therefore, the most important aspect to analyse is the net movement of researchers between the sectors in one direction after subtracting those that move in the other direction. This minimizes the influence of the fluctuations due to co-affiliation.

Measuring International Researcher Mobility
The approach presented here uses Scopus author profile data to derive a history of active regional authors. Based on the affiliations recorded in each author's publications over time, authors are assigned to a mobility class defined by the type and duration of observed moves.

How are mobility classes defined and measured?
The measurement of international researcher mobility by co-authorship in the published literature is complicated by the difficulties involved in teasing out long-term mobility (resulting from attainment of faculty positions, for example) from short-term mobility (such as doctoral research visits, sabbaticals, secondments, etc.), which might be deemed instead to reflect a form of collaboration. In this study, active researchers are broadly divided into three groups:

- Sedentary: active researchers whose Scopus author data for the period indicates that they have not published outside the region.
- Transitory: active researchers whose Scopus author data for the period indicates that they have remained abroad or in the region for less than two years.
- Migratory: active researchers whose Scopus author data for the period indicates that they have published outside the region.
- Inflow: researchers whose publication history indicates that they first published outside of the region and then published inside of the region.
- Outflow: researchers whose publication history indicates that they first published inside the region and then published outside of the region.

How do we characterize the mobility groups?
To better understand each mobility group, three aggregate indicators are calculated for each to provide insight into the scholarly productivity, impact, and seniority of the researchers within each group: the average field-weighted citation impact of the publications by authors in the group, the average relative productivity of authors in the group, and the average relative seniority of authors in the group. Field-weighted citation impact (FWCI) is a measure of publication impact based on citations and normalized against the average for publications of a similar age, type, and subject. Relative productivity is a measurement of the number of publications per year since the first appearance of each researcher as an author during the period, relative to all regional researchers in the same period. Relative seniority represents years since the first appearance of each researcher as an author during the period, relative to all regional researchers in the same period. All three indicators are calculated for each author's entire output in the period (i.e., not just those articles listing a regional address for that author).

Topic Prominence in Science
Through topics analyses, it is possible to identify emerging topics with high momentum and how these topics are related to a selected entity or group's research portfolio. Topics can be large or small, new or old, and growing or declining. The granularity of topics allows us to define the problem-level structure of science. Due to the way it is structured, topics do not need field weighting to be coherent collections and topics in social science and humanities are just as valid as in STEM areas, although they may be smaller and less prominent.

Appendices

Glossary of terms

Adaptive algorithm
An adaptive algorithm is an algorithm that changes its behavior at the time it is run, based on information available and an a priori defined reward mechanism.

Agent-based system technology
Agents are sophisticated computer programs that act autonomously on behalf of their users, across open and distributed environments, to solve a growing number of complex problems.

Compound Annual Growth Rate
CAGR is defined as the year-over-year constant growth rate over a specified period of time. Starting with the first value in any series and applying this rate for each of the time intervals yields the amount in the final value of the series.

$$CAGR(t_o, t_n) = (V(t_n)/V(t_o))^{\frac{1}{t_n - t_o}} - 1$$

$V(t_o)$: start value, $V(t_n)$: finish value, $t_n - t_o$: number of years.

Fingerprint
A ranked set of standardized, domain-specific concepts that define the text.

FWCI
Field-weighted citation impact (FWCI) is an indicator of mean citation impact and compares the actual number of citations received by a publication with the expected number of citations for publications of the same document type (article, review, or conference proceeding paper), publication year, and subject field. When a publication is classified in two or more subject fields, the harmonic mean of the actual and expected citation rates is used. The indicator is therefore always defined with reference to a global baseline of 1.0 and intrinsically accounts for differences in citation accrual over time, differences in citation rates for different document types (reviews typically attract more citations than research articles, for example) as well as subject-specific differences in citation frequencies overall and over time and document types.

FWDI
Field-weighted download impact (FWDI) is a replication of the FWCI calculation for downloads.

Machine learning
The process by which an AI uses algorithms to perform functions. It is the result of applying rules to create outcomes through an AI.

Natural language processing
Natural language processing (NLP) is a field of computer science, AI, and computational linguistics concerned with the interactions between computers and human (natural) languages, and, in particular, concerned with programming computers to process large natural language corpora.

Neural networks
A computational approach based on a large collection of neural units, loosely modeling the way a biological brain solves problems with large clusters of neurons connected by axons.

Optimization algorithms
A group of mathematical algorithms used in machine learning to find the best available alternative under the given constraints.

Pattern recognition
A branch of machine learning that focuses on the recognition of patterns and regularities in data, although it is in some cases considered to be nearly synonymous with machine learning.

Relative Activity Index
The relative activity index (RAI) approximates the specialization of a region in comparison to the global research activity in the AI field. RAI is defined as the share of a country's publication output in AI relative to the global share of publications in AI. A value of 1.0 indicates that a country's research activity in AI corresponds exactly with the global research activity in AI; higher than 1.0 implies a greater emphasis while lower than 1.0 suggests a lower emphasis compared to global activity.

Supervised learning
The machine learning task of inferring a function from labelled training data.

Text mining
The process of examining large collections of written resources to generate new information, and to transform the unstructured text into structured data for use in further analysis.

Appendices

Sources

arXiv[120] is an e-print service in the fields of physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, electrical engineering and systems science, and economics that is owned and operated by Cornell University, a private not-for-profit educational institution. arXiv is funded by Cornell University Library,[121] the Simons Foundation,[122] and member institutions.[123]

dblp computer science bibliography[124] is an online reference for bibliographic information on major computer science publications. It has evolved from an early small experimental web server to a popular open-data service for the computer science community. DBLP's mission is to support computer science researchers in their daily efforts by providing free access to high-quality bibliographic metadata and links to the electronic editions of publications.

As of May 2016, DBLP indexes over 3.3 million publications, published by more than 1.7 million authors. To this end, DBLP indexes more than 32,000 journal volumes, more than 31,000 conference or workshop proceedings, and more than 23,000 monographs.

Kaggle[125] is a crowd-sourced platform to attract and train data scientists. It is the world's largest community of data scientists and machine learners. Kaggle got its start by offering machine learning competitions and now also offers a public data platform, a cloud-based workbench for data science, and short-form AI education. On 8 March 2017, Google announced that they were acquiring Kaggle.

Plum Analytics[126] is dedicated to measuring the influence of scientific research with the vision of bringing modern ways of measuring research impact to individuals and organizations that use and analyze research.

ScienceDirect®[127] is Elsevier's full-text scientific journal platform. With an invaluable and incomparable customer base, the use of scientific research on ScienceDirect.com provides a different look at performance measurement. ScienceDirect.com has more than 14 million active users, with over 900 million full-text article downloads in 2018.[128]

SciVal[129] offers quick and easy access to the research performance of over 10,000 research institutions and 230 regions and countries. Using advanced data analytics technology, SciVal processes enormous amounts of data to generate powerful visualizations in seconds. The 170 trillion metrics in SciVal are calculated from 46 million publication records published in the 21,915 journals of 5,000 publishers worldwide.

Scopus®[130] is Elsevier's abstract and citation database of peer-reviewed literature, covering 71 million documents from more than 23,700 active journals, book series, and conference proceeding papers by 5,000 publishers.

Scopus coverage is multilingual and global: approximately 46% of titles in Scopus are published in languages other than English (or published in both English and another language). In addition, more than half of Scopus content originates from outside North America, representing many countries in Europe, Latin America, Africa, and the Asia-Pacific region.

For this report, a static version of the Scopus database covering the period 1996-2017 inclusive was aggregated by country, region, and subject defined by FORD subject areas.[131]

TotalPatent[132] offers the most patent content available from a single source and the tools to search, compare, and analyze results.

120 https://arxiv.org/.
121 https://www.library.cornell.edu/about.
122 https://simonsfoundation.org/.
123 https://confluence.cornell.edu/x/ALlRF.
124 https://dblp.uni-trier.de/.
125 https://www.kaggle.com/.
126 https://plumanalytics.com.
127 https://www.elsevier.com/solutions/sciencedirect.
128 https://www.elsevier.com/about/this-is-elsevier.
129 https://www.elsevier.com/solutions/scival.
130 https://www.elsevier.com/solutions/scopus.
131 Frascati Manual 2015. OECD Library. https://read.oecd-ilibrary.org/science-and-technology/frascati-manual-2015_9789264239012-en#page60.
132 https://www.lexisnexis.com/totalpatent/.

Amsterdam
Data Science

中国科学院
自动化研究所
CASIA
INSTITUTE OF AUTOMATION
CHINESE ACADEMY OF SCIENCES

FAPESP

ISTIC

KIT
Karlsruher Institut für Technologie

ORBIT

SRI International®

SWEDISH ARTIFICIAL
INTELLIGENCE SOCIETY

Elsevier is a registered trademark of Elsevier B.V. | RELX Group and the RE symbol are trademarks of RELX Intellectual Properties SA, used under license. © 2018 Elsevier B.V.

ELSEVIER