Better Intelligence about Artificial Intelligence

Eleven insights from a programme of research using data science and machine learning to track the evolution of AI

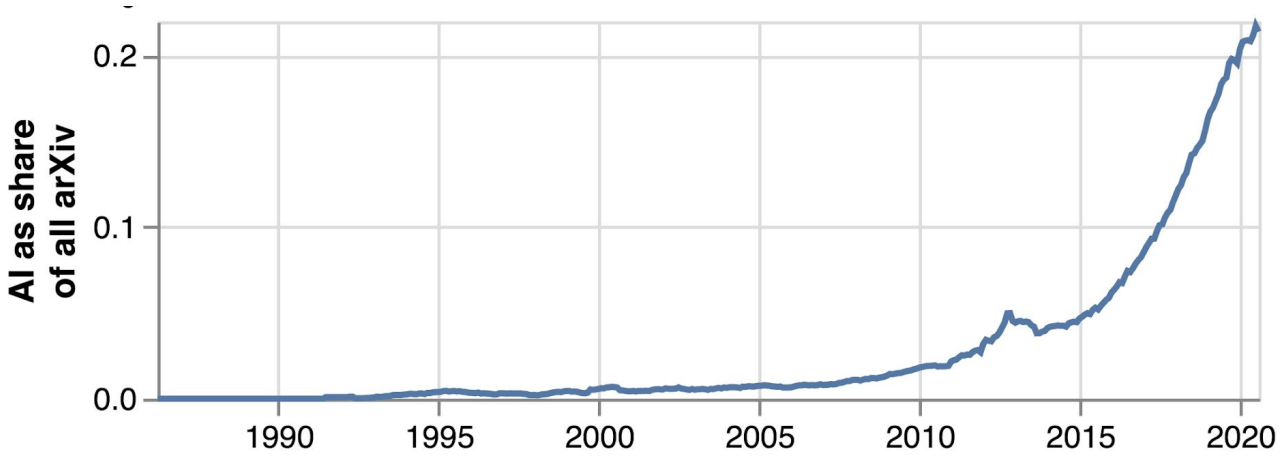
Juan Mateos-Garcia

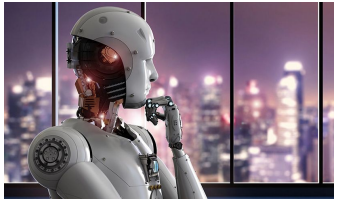
Collective Learning Group
ANITI, University of Toulouse
14 April 2021

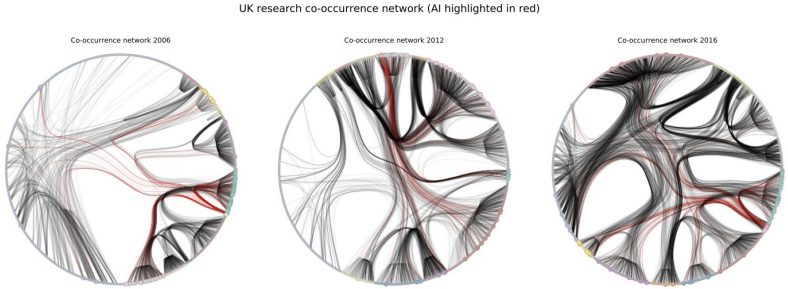
nesta

Summary

Dominant representations of AI progress are deterministic and disempowering







We need to open the black box of AI R&D: who is doing it, where is it going, what is its impact?

This requires moving beyond simple economics and indicators of AI. I present 11 findings from Nesta's research about this.

nesta

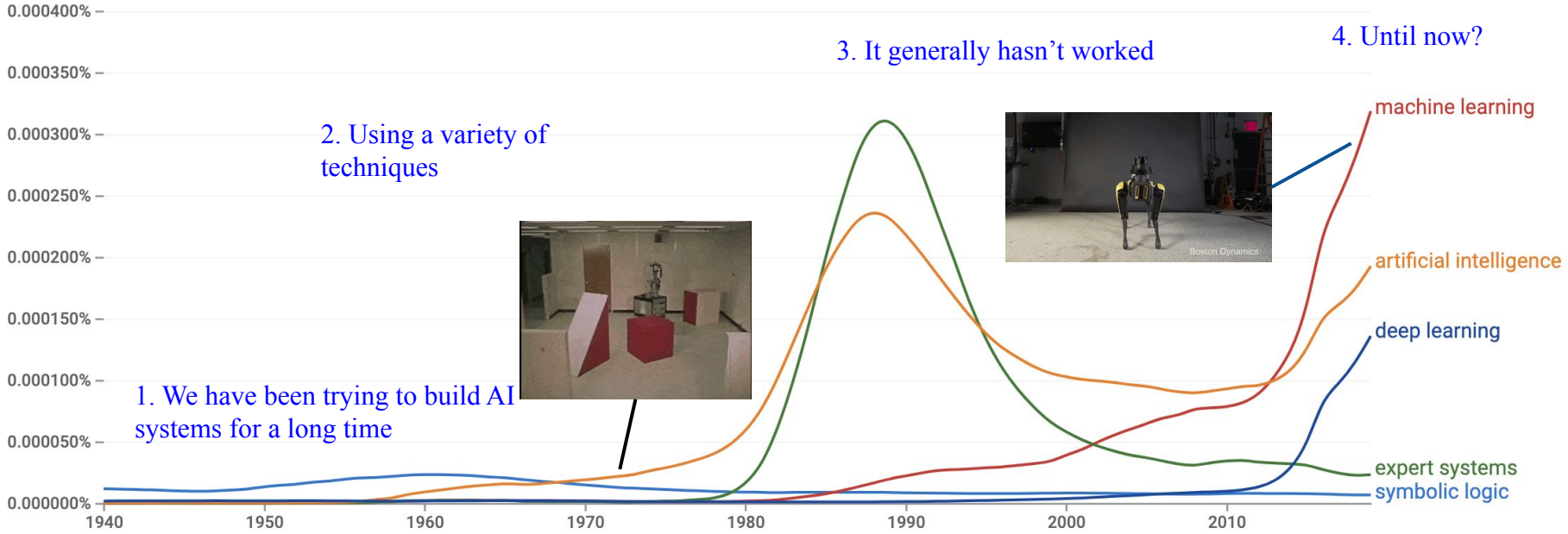
Click icons for interactive versions of the figures

Introduction

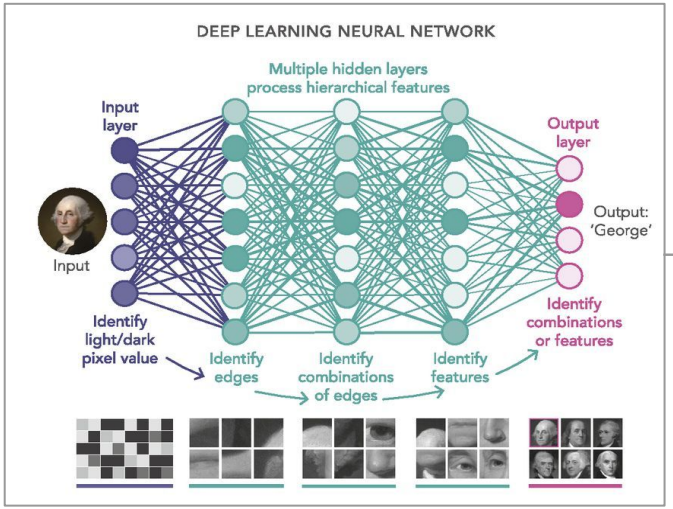
We are experiencing a new boom of interest in AI linked to the arrival of deep learning… but there are growing concerns about where AI is going.

Artificial intelligence

A working definition *Machines able to behave reasonably in a wide range of circumstances* (Brian Arthur)

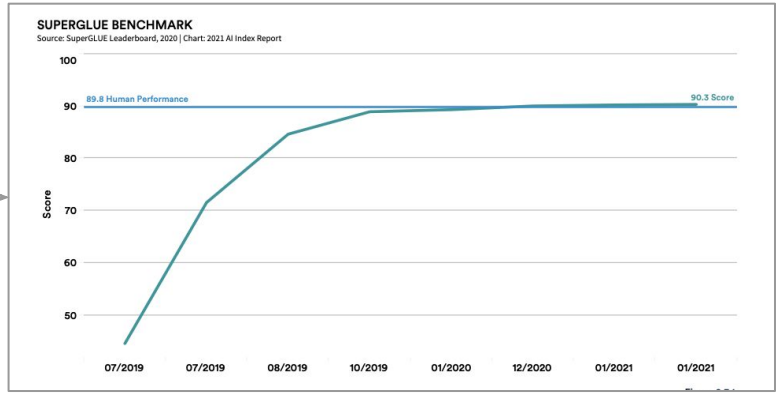


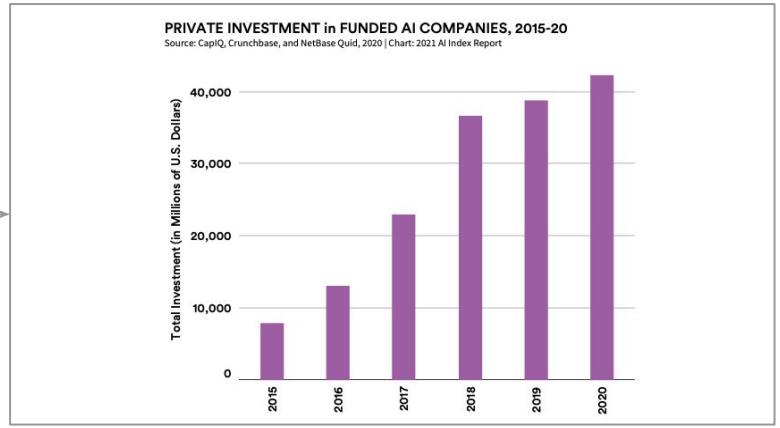
A modern approach



Data and compute intensive deep learning methods

Substantial improvements in performance vs human benchmarks





Increasing levels of commercial interest, investment and deployment

A variety of concerns

There are concerns about the trajectory and impacts of AI technologies

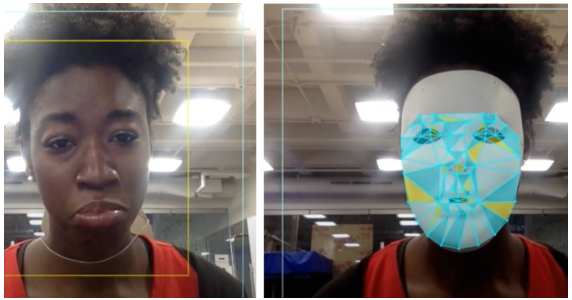
Economic

Acemoglu and Restrepo, 2019, Zuboff, 2018, Eubanks, 2018.

Technological / environmental

Marcus, 2019, D'Amour et al 2019, Strubell et al, 2019

Ethical



Buolamwini et al 2018, Bender et al, 2021, Paullada et al 2020

Two (idealised) frames

	General perspective	Take on AI	Measurement
1. Mainstream economics	■ Technological progress is a scalar ■ It requires complementary investments ■ It produces externalities	■ General Purpose Technology ■ Requires new skills and processes ■ Might displace labour and increase inequality	■ Levels of investment in AI and impact on productivity and employment
2. Complexity economics (with Science and Technology Studies)	■ Technological "progress" is a vector ■ The best technology does not always win ■ History and context matter	■ Hardware and data lotteries ■ Commercial and technological races ■ Lack of inclusion in the workforce can result in discriminatory systems	■ Composition of AI ■ Maps of the AI innovation system: who, where and why

We need data science methods to measure these dimensions of AI

Our programme of research

Co-authors

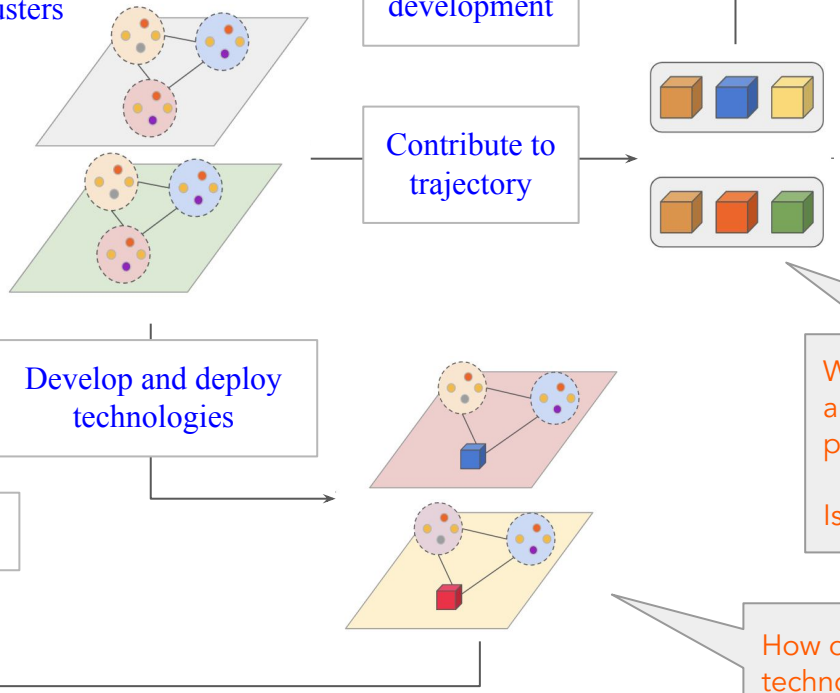
Individuals work in organisations that are part of ecosystems /clusters

Foundation for future development

Contribute to trajectory

Develop and deploy technologies

Rewards



Who participates (people, places) and who is excluded? What factors drive participation?

Where is the technology going and how is this driven by participants? (individuals, firms)

Is the field becoming narrower?

How does the nature of the technologies being developed drive / constrain impacts?

Methods

Is it about AI?

Semantic analysis

Computer Science > Computers and Society

[Submitted on 29 Oct 2020]

Machine Ethics and Automated Vehicles

Noah J. Goodall

Road vehicle travel at a reasonable speed involves some risk, even when using computer-controlled driving with failure-free hardware and perfect sensing. A fully-automated vehicle must continuously decide how to allocate this risk without a human driver's oversight. These are ethical decisions, particularly in instances where an automated vehicle cannot avoid crashing. In this chapter, I introduce the concept of moral behavior for an automated vehicle, argue the need for research in this area through responses to anticipated critiques, and discuss relevant applications from machine ethics and moral modeling research.

Comments:	12 pages
Subjects:	**Computers and Society (cs.CY)**
ACM classes:	K.4.1
Journal reference:	In: Meyer G., Beiker S. (eds) Road Vehicle Automation. Lecture Notes in Mobility. Springer, Cham (2014)
DOI:	10.1007/978-3-319-05990-7_9
Cite as:	arXiv:2010.15665 **[cs.CY]**
	(or arXiv:2010.15665v1 **[cs.CY]** for this version)

Submission history

From: Noah Goodall [view email]

[v1] Thu, 29 Oct 2020 15:14:47 UTC (471 KB)

What is the researcher affiliation / country / type of organisation?

What is the paper about?

Topic modelling

```
safety__vehicle__vehicles__driving__drivers: 0.157

applications__solutions__control__solution__r: 0.059

values__societal__moral__norms__artificial: 0.059
```

Fuzzy matching

arXiv.org

Microsoft Azure Cognitive Services

GRID

nesta

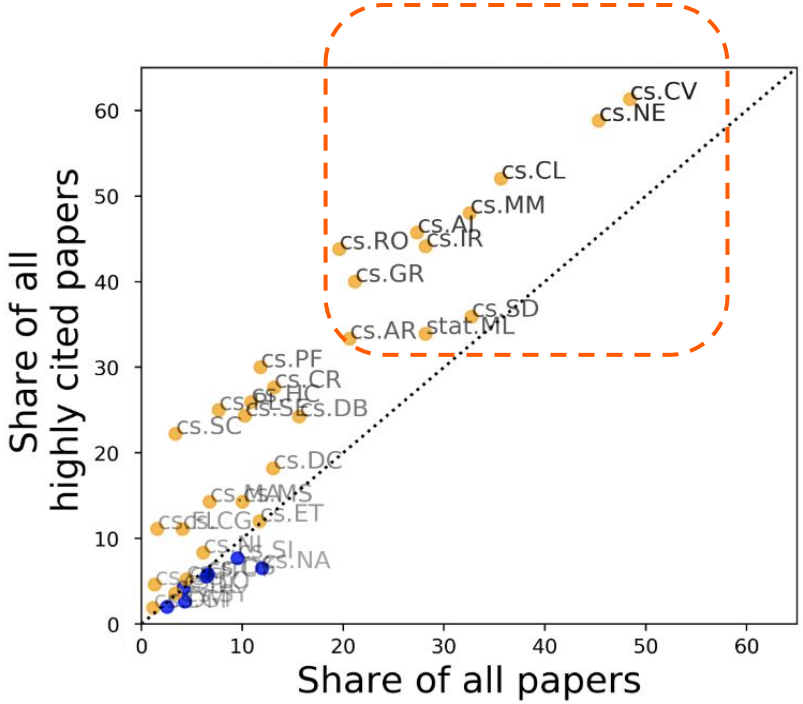
State of play

We are experiencing an AI revolution in volume and composition of AI R&D

1. AI displays the features of a General Purpose Technology

We use topic modelling and keyword extraction to identify AI and deep learning papers in the arXiv preprint corpus

- **Activity**
  - 80% of AI papers in arXiv published after 2012
- **Diffusion**
  - Deep learning being applied in all topics
- **Influence**
  - Deep learning overrepresented among highly cited papers in ~ all computer science subfields

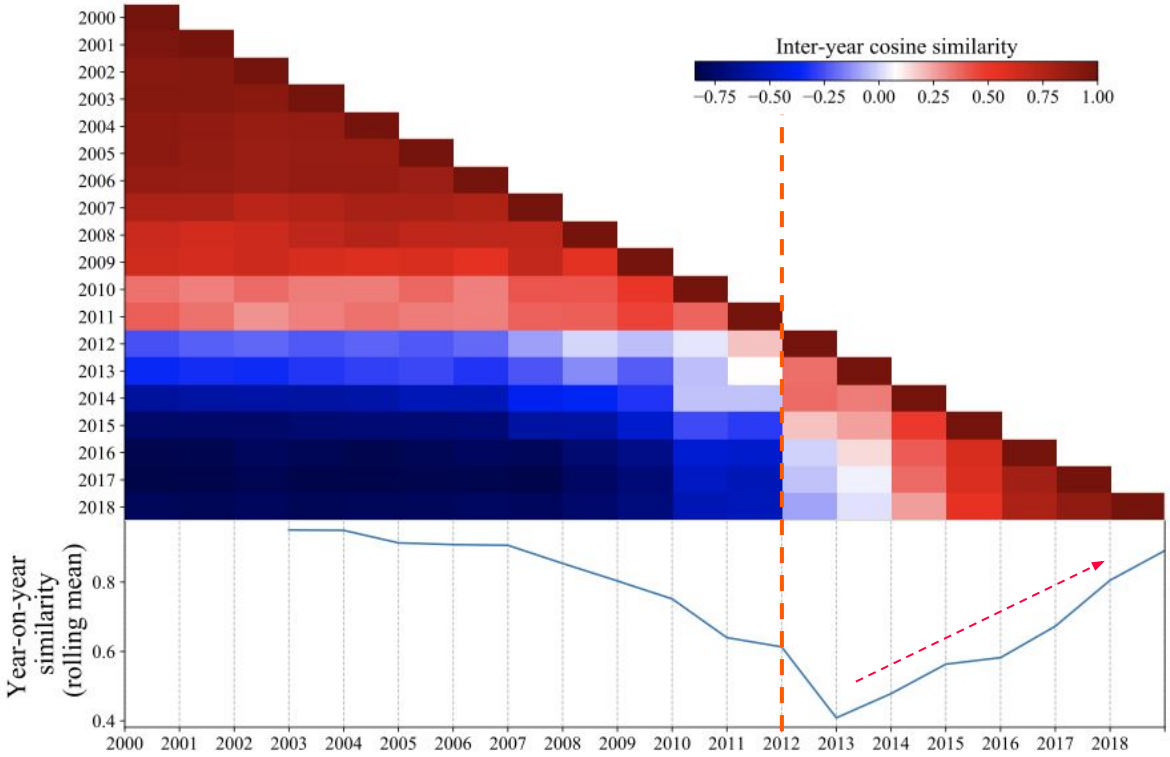


Deep Learning representation among highly cited papers (Y) vs presence in the data (x) [Reference]

2. Also an important change in its composition

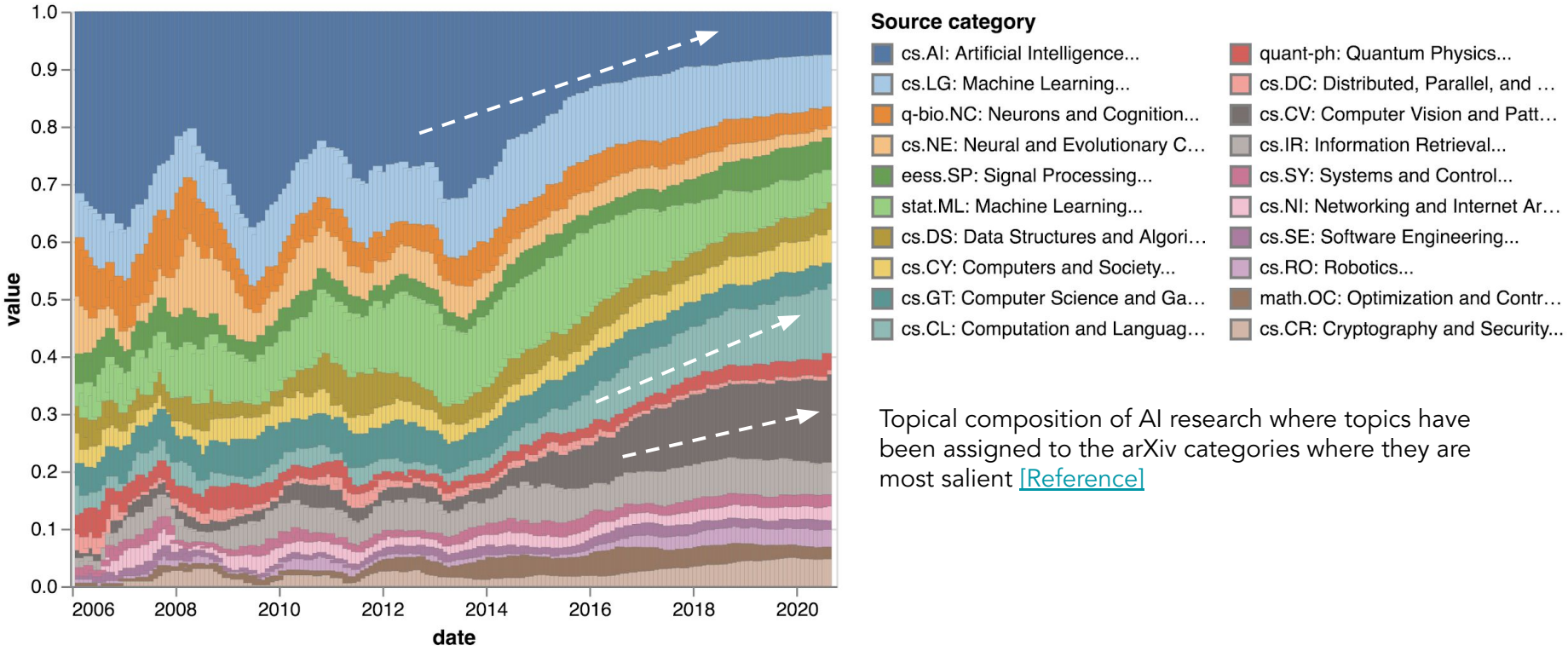
We analyse the evolution of the topical composition of AI research

- We see an abrupt break in the early 2010s.
  - Ca. publication of krizhevsky et al. 2012
- Evidence of consolidation afterwards



Intra-year similarities in thematic composition of AI research based on topic modelling [Reference]

3. The change in composition reflects the arrival of deep learning methods applied in areas such as computer vision and computer language, and a decline of symbolic methods.



Topical composition of AI research where topics have been assigned to the arXiv categories where they are most salient [Reference]

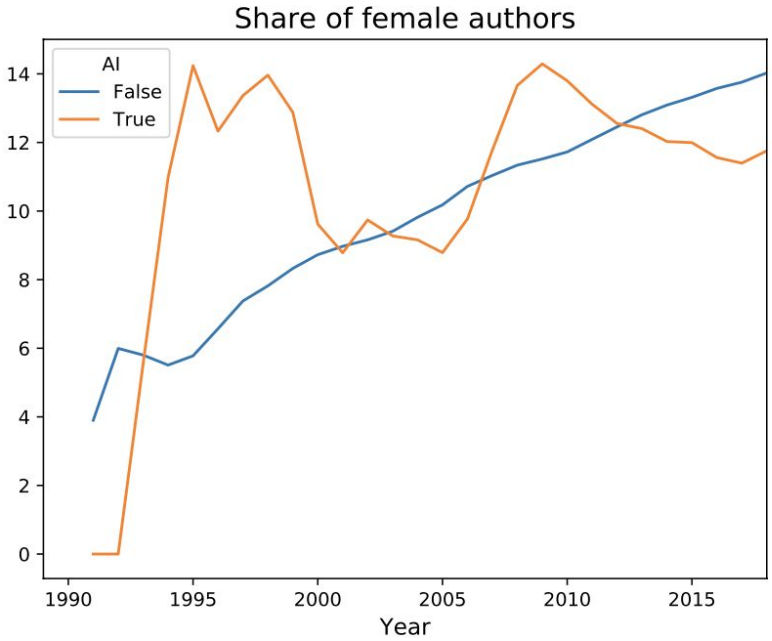
Who is doing it?

AI is increasingly dominated by a small number of organisations and clusters, and the diversity of its workforce has stagnated

4. In relative terms, female participation in AI research has stagnated since the mid-1990s

We use a gender inference system to estimate gender diversity in AI research workforce and explore differences across disciplines and countries

- Computer science and machine learning fields tend to be less diverse than applied fields like health, and companies less than universities
- Women have higher levels of participation in applied fields related to health and computing and society
- We find some evidence that papers involving female authors tend to involve more diverse mixes of disciplines and consider social aspects of AI deployment

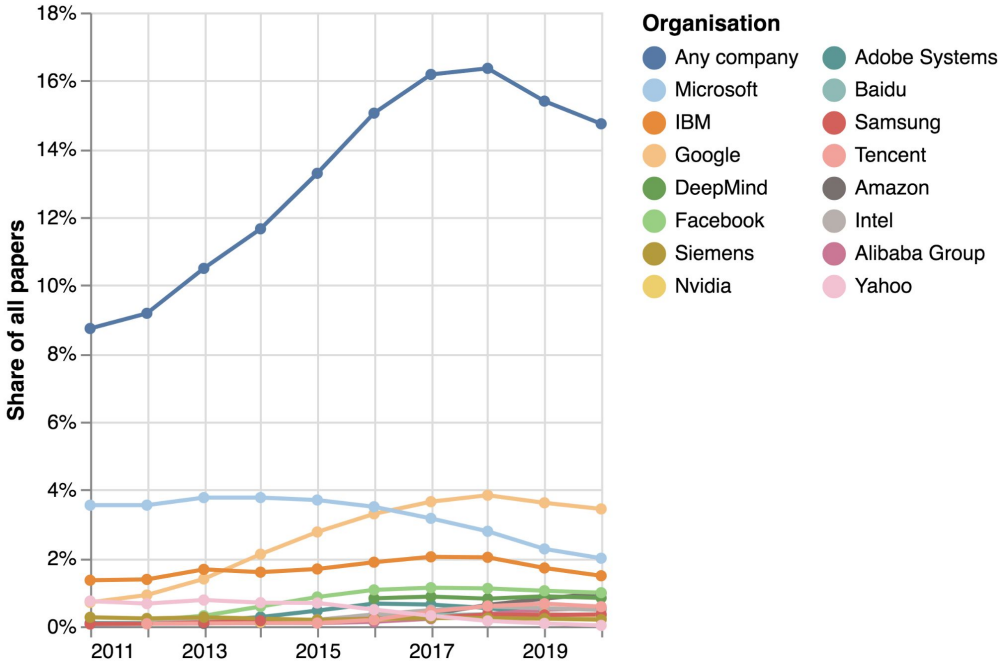


Share of individual female authors in the corpus of AI research [Reference]

5. Private sector companies are strongly involved and influential in AI research

We match arXiv with MAG and GRID to and scrape websites from key research labs

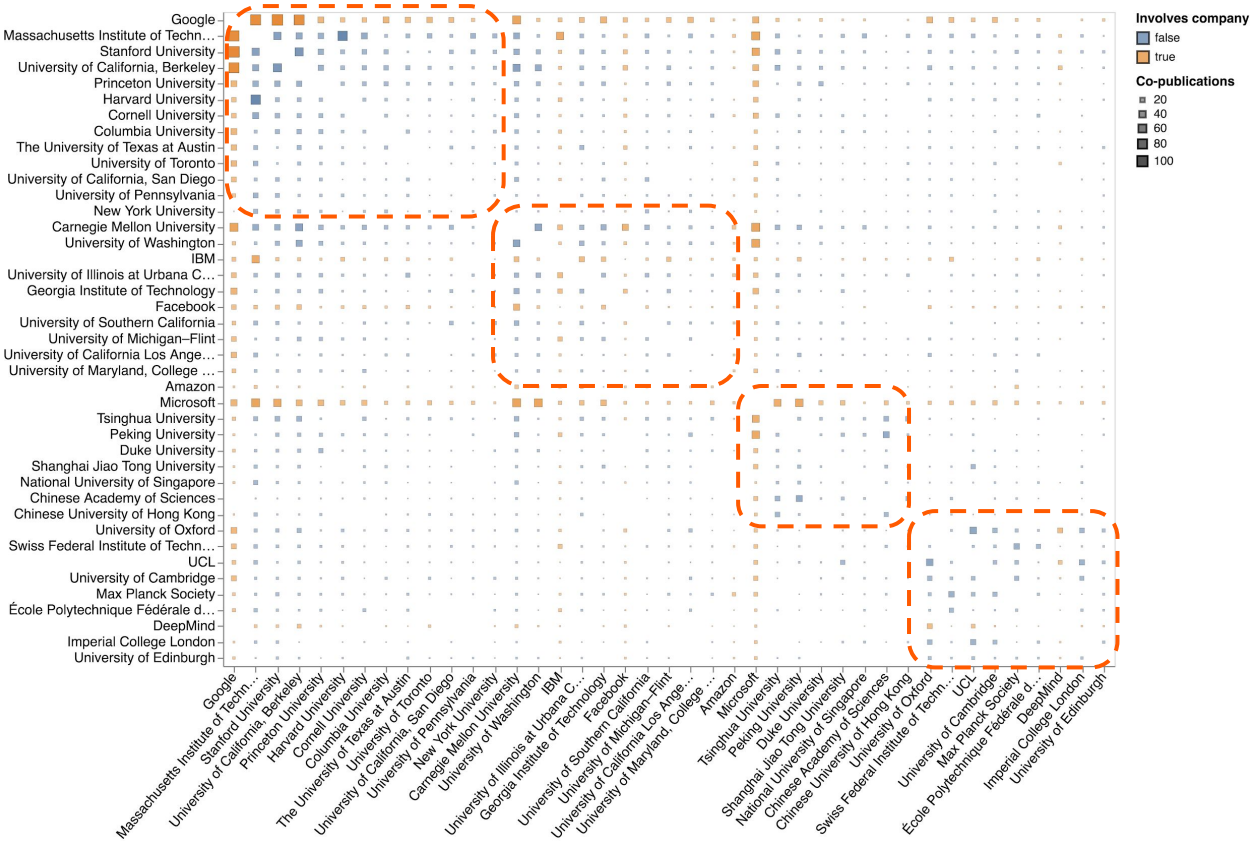
- 15% of AI papers in 2020 involved at least one company
  - Mostly technology companies in the USA and China
- Papers involving companies tend to be more influential
- 60% of company papers involve at least one academic institution (often elite ones)



Share of papers involving at least one company and top 15 companies by overall level of AI activity [Forthcoming]

Organisations [2]

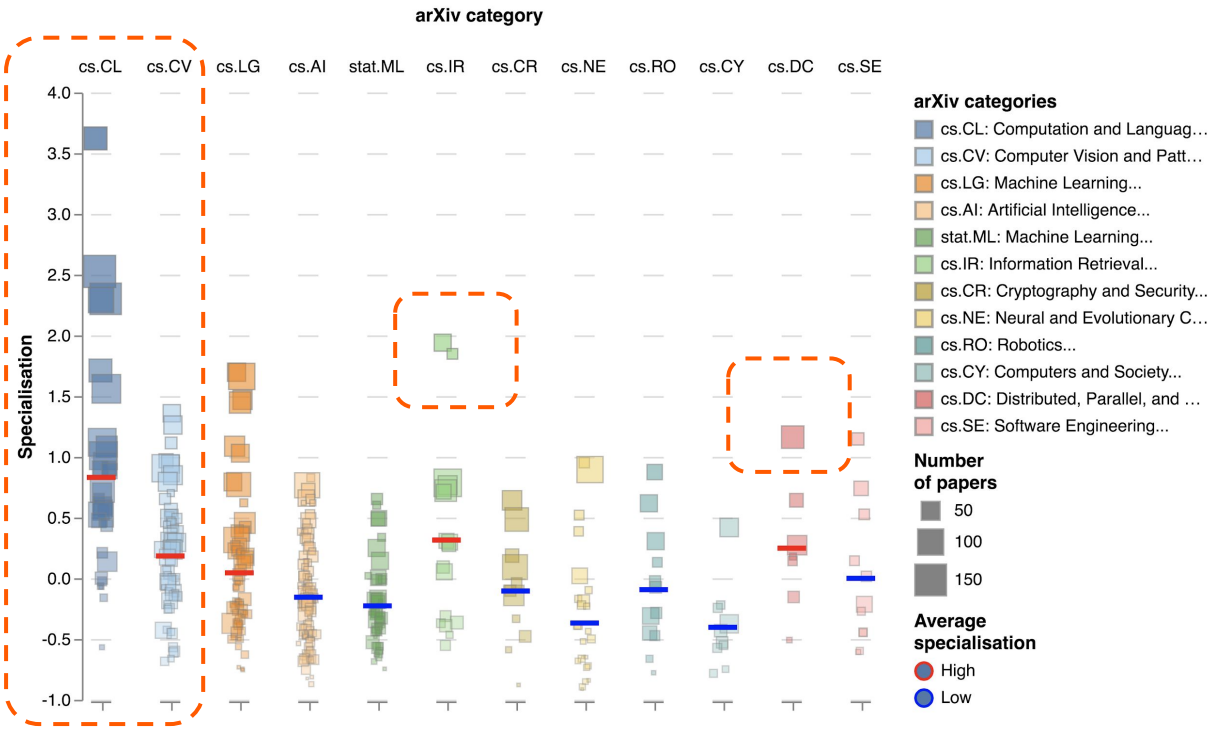
- 60% of company papers involve at least one academic institution (often elite ones)
- Communities of research collaboration organised around geographical lines



Research collaborations between top institutions sorted by the communities that different institutions are aligned with [Forthcoming]

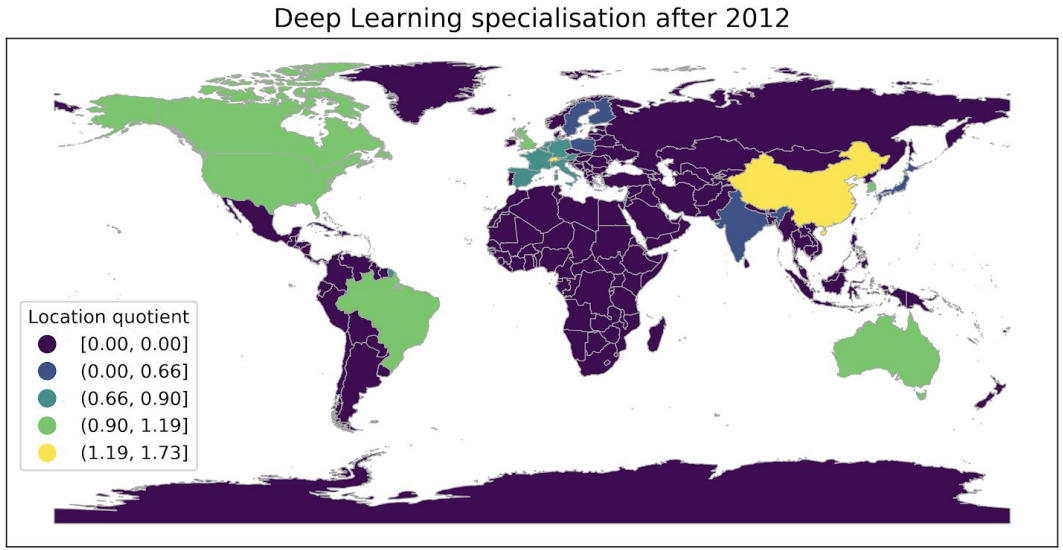
6. Private sector companies specialise in video and data intensive techniques and advertising and search use-cases

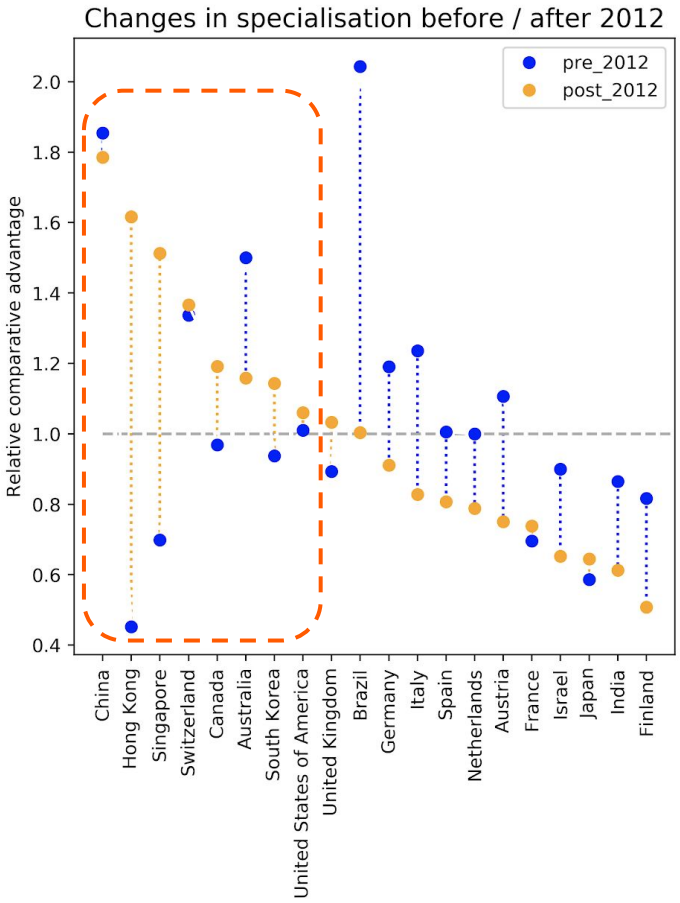
We compare the AI thematic specialisation (based on a topic model) between private and educational institutions



Relative specialisation in AI research topics in private companies [Reference]

Geography [1]





7.There has also been a shift in the geography of AI research towards the East (China, Singapore) and the West (Canada). Some evidence that EU countries are falling behind.

Relative specialisation in Deep learning research by country [Reference]

8. AI research (and specially state of the arts methods) tend to be strongly concentrated in regions with related capabilities.

We use machine learning to identify research and industrial capabilities related to deep learning and regress deep learning specialisation on them.

- In 2018, the top 30 regions globally accounted for ca. 60% of deep learning research activity
- This concentration has strengthened over time.
- Regions that co-locate related research and industrial capabilities tend to develop stronger Deep learning clusters after 2012.

	Model 1	Model 2	Model 3	Model 4
y	RCA_{DL,t_1}	RCA_{DL,t_1}	RCA_{DL,t_1}	RCA_{DL,t_1}
RCA_{DL,t_0}	0.12**	0.12**	0.122**	0.124**
	(0.06)	(0.06)	(0.06)	(0.06)
$arXiv_{sp}$	0.159**	0.16**	-0.026	-0.01
	(0.064)	(0.065)	(0.076)	(0.078)
$CrunchBase_{sp}$		0.01	-0.173**	-0.154**
		(0.043)	(0.069)	(0.072)
$arXiv_{sp} \times CrunchBase_{sp}$			0.274***	0.245***
			(0.09)	(0.094)
$arXiv_{sp} \times CrunchBase_{tot}$				0.205***
				(0.068)
$arXiv_{tot}$	0.091**	0.089**	0.094**	-0.078
	(0.041)	(0.042)	(0.04)	(0.08)
is_China	1.543***	1.546***	1.549***	1.584***
	(0.224)	(0.224)	(0.219)	(0.223)
R^2	0.149	0.147	0.156	0.166
n	451	451	451	451

Table 2 Dependent variable is RCA_{DL,t_1}. Standard errors in brackets are clustered by country. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

[Reference]

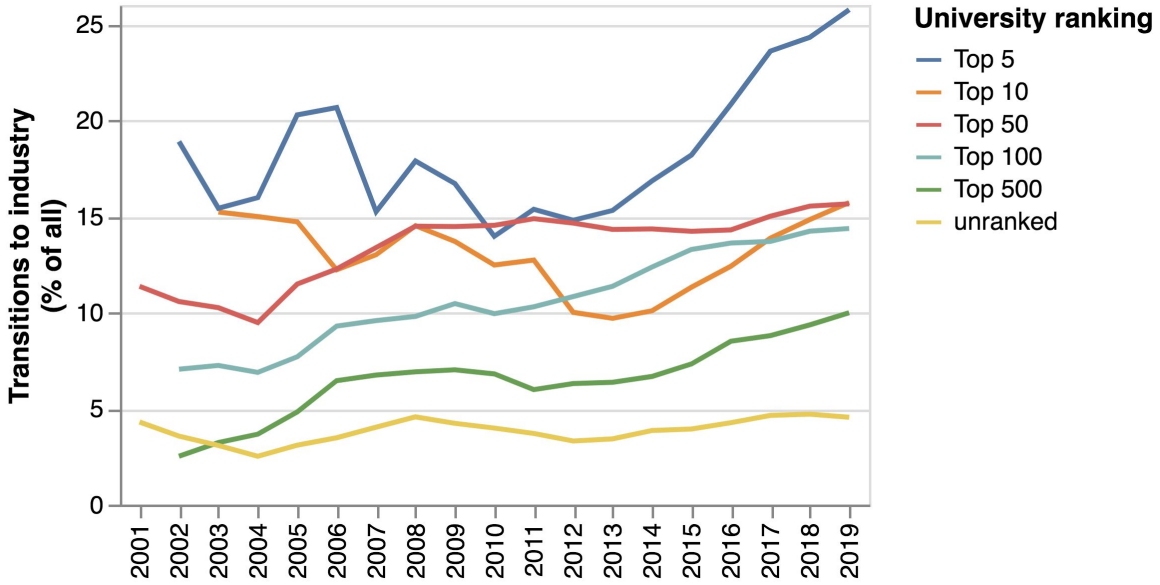
Where is it going?

AI research seems to be becoming thematically narrower and there is evidence of a brain drain of talent from academia to industry. A case study of AI applications against Covid-19 suggests limited impact

9. There is a growing flow of researchers from academia into industry

We measure researcher transitions between academia and industry, and model its drivers and impacts

- Influential (and male) researchers from top institutions are more likely to transition intotechnology companies
- Although these researchers experience a ‘citation bump’ after moving into industry this is offset over time when we compare them with similar researchers who stayed in academia



Share of job transitions to industry by university position in the Nature Rankings [Reference]

10. AI research seems to be getting thematically narrower

We calculate thematic diversity in AI research according to a variety of metrics and parameter sets

- We find evidence of stagnation and even decline in AI's thematic diversity in recent years
- Companies have narrower thematic profiles than educational institutions after we control for other factors.
- Elite academic institutions that collaborate more with companies tend to be less thematically diverse

Rao-Stirling	0	0	1	1	2	2
Company index	-0.09*	-0.26*	-0.02*	-0.12*	-0.31***	-0.72**
	(-1.28)	(-0.69)	(-0.32)	(-0.28)	(-3.66)	(-2.22)
Number of papers (log)	0.85***	1.12***	0.72***	0.99***	0.99***	1.37***
	(12.75)	(8.89)	(9.56)	(6.71)	(13.26)	(9.3)
Year	0.02*	0.0*	0.01*	-0.0*	0.04*	0.02*
	(0.39)	(0.08)	(0.19)	(-0.1)	(1.02)	(0.62)
R^2	0.44	0.72	0.32	0.64	0.58	0.82
N	564	564	564	564	564	564
Fixed Effects	No	Yes	No	Yes	No	Yes

Table 9: Regression results for various diversity metrics (Balance in top table, Weitzman in middle table, Rao-Stirling in bottom table) and parametre sets (see columns). t-values in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

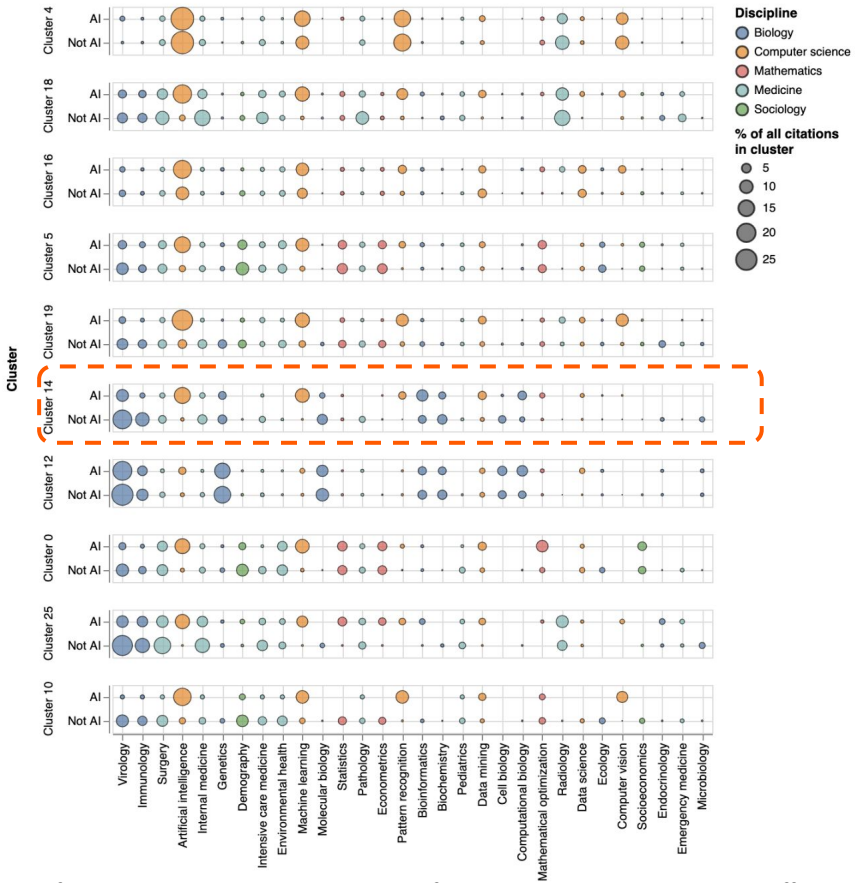
[Reference]

Covid-19 application

11. AI's failure to contribute to the fight against Covid-19 might be linked to its reliance on data-intensive techniques and insularity from other disciplines

We identify all open research related to Covi, cluster it into groups based on its focus and analyse AI-related activity inside it

- Most AI research is using computer vision to detect the virus in medical scans (a relatively low value application)
- AI research tends to neglect existing knowledge outside of computer science
- AI research receives less citations than comparable research not using AI



Share of citations to various disciplines from AI / not AI research in different topical clusters [Reference]

Conclusions

State of play

We are experiencing a revolution in AI research

Who is doing it?`

AI's trajectory is being shaped by a small number of (mostly private sector) organisations and tech clusters. Workforce diversity is stagnant

Policy implications

This evidence suggests we may need policies to preserve spaces for public interest AI research, make the research workforce more inclusive, maintain diversity in the technologies that are explored, increase crossover with other disciplines & support AI deployment beyond tech

Where is it going?

The set of techniques being explored is becoming narrower and industry is attracting the most influential researchers. AI has played a limited role in the fight against Covid-19

Research

- More data (patents, open source software, skills)
- Better models (including ABMs and models enabling causal analysis)
- Evidence of technical and economic impacts

Development & deployment

- Real-time data collection and analysis
- Build decision-support tools and interactive visualisations [arXlive]
- Explore opportunities and limitations of AI/ML methods in new domains and in collaboration with other disciplines [Mateos-Garcia, 2021]

Go to GitHub to access our code and data

Thank you!

Questions?

juan.mateos-garcia@nesta.org.uk
@Jmateosgarcia

nesta