The AI Spring of 2018

Sofia Olhede and Patrick Wolfe discuss the implications as nations race for AI dominance.

The past months have seen a whirlwind of reports on Artificial Intelligence (AI)—with experts Professor Dame Wendy Hall and Jérôme Pesenti having foreshadowed this activity by releasing their UK Government-commissioned report on AI and commercial growth[1] last October. March then saw French president Emmanuel Macron's announcement of an 'AI for humanity' strategy[2], following French mathematician and Fields Medalist Cedric Villani's report entitled 'For a meaningful artificial intelligence'[3]. Also mixed into this flurry of activity is the recent UK House of Lord's report 'AI in the UK: ready, willing and able?' released in April, which in turn was followed by the announcement of public investment geared towards the UK AI sector. Feeling dizzy yet? Throw in a February report on malicious use of AI[4] for good measure, and it's clear that the international stakes have never been higher.

Not to be left out of the international arms race in AI, the European Union is also weighing in. Late in April, 25 European countries signed an agreement to collaborate on AI, and to turn this goodwill into reality is the promise of a €1.5 billion investment into research for 2018-20[5], with an aim to invest €20 billion by the end of 2020.

As statisticians, do we care, and how do these developments relate to us? AI has in common parlance taken on a meaning much beyond its research connotations; it now seems to mean data science, machine learning, automation and anything remotely related to 'deep tech'. This means that actionable information from data, and automated decisions from data analysis, both fit squarely into what is now popularly viewed as AI. Thusly the discipline of Statistics will doubtless find itself caught up in these international developments.

The UK House of Lord's report is convinced that the UK can still lead in this field, building on a historically strong research programme. To make these developments societally friendly it proposes five principles: 1) AI should be developed for public good, 2) AI should be operated with intelligibility and fairness, 3) AI should not be used to diminish rights to data or privacy, 4) All citizens have a right to education to flourish with AI and finally 5) no AI should be given the right to automatically destroy or hurt human beings.

Bringing back shades of Isaac Asimov's three laws of robotics, these principles might seem at first glance to be straightforward, but the report draws out the

---

[1] https://www.gov.uk/government/publications/growing-the-artificial-intelligence-industry-in-the-uk

[2] https://www.aiforhumanity.fr/en/

[3] https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf

[4] Brundage, M. The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation, 2018, arXiv:1802.07228.

[5] http://europa.eu/rapid/press-release_IP-18-3362_en.htm

many shades of grey in these problems, discussing in detail intelligibility of AI, and the issues that follow from a lack of transparency. The meaning of transparency itself is also dissected in the report, highlighting the many issues of black-box algorithms, some of which we have discussed previously in this column[6]. The report goes on to recommend how academic research can more readily generate commercial value if IP policies are adjusted accordingly, and proposes national retraining schemes for the UK workforce.

How does the UK government's response to AI compare to its neighbors both near and far? Villani's report stresses that market forces alone cannot guarantee French AI developments. Like the Lords' report, it stresses the risks of black-box algorithms, and suggests that mechanisms must be put in place to use data for common good. It warns of the US brain drain from France, and suggests forming a national network of AI institutes, setting up a public laboratory for studying the future of work, and an AI Ethics committee, echoing the Nuffield Ada Lovelace Institute as well as the UK's new Centre for Data Ethics and Innovation. EU countries are looking to `Ethical AI' as a possible area of strength, since as the Villani report notes, AI investment volume is led by the United States and China, followed by Canada, the United Kingdom and Israel.

Finally, the report on Malicious AI by Brundage *et al* is interesting precisely because it has a very different focus. It highlights all our interest in trying to predict and preempt malicious use of AI, and stressing the need to grow a culture of responsibility in AI researchers. Threats are perceived to digital security, political security and physical security. As most technologies are dual-use, inevitably they impact both defense and offense for cyber attacks. As AI technologies develop, existing threats will grow more severe, new threats will be introduced and the typical nature of threats will change.

What does this all mean for the rest of society?

First, the strong push for ethical AI will no doubt lead to changes in the undergraduate curriculum at leading universities. Many first-tier Computer Science departments in the US are introducing ethics courses: Harvard University and MIT have together introduced a joint course on ethics for computer scientists; the University of Texas at Austin in turn now has 'Ethical Foundations of Computer Science'; and Stanford University is developing a related new course for 2019[7]. Statistics should follow suit: we must teach our students how to do ethics assessments, and the correct way of handling an ethics review board process. Areas such as medical statistics and social statistics already have a firm grasp of these topics, and in that sense are ahead of computer science.

---

[6] Olhede and Rodrigues, Fairness and transparency in the age of the algorithm, April 2017, Olhede and Wolfe, When algorithms go wrong, who is liable, 2017

[7] https://www.nytimes.com/2018/02/12/business/computer-science-ethics-courses.html

Second, opportunities will be coming Statistics' way. The current focus on interpretable AI is exactly in the bailiwick of statistics, where our wish for explanatory models, and rigorous proof can help us make sense of many currently opaque methods. As focus is placed on 'intelligible AI' (borrowing a phrase from the Lords' report), and this is precisely an opportunity for Statistics to contribute to make AI better.

Finally, given the pace of developments, much in these many reports will very quickly become redundant. The rest of society is not standing idle at this time. For example, legal experts are quickly aiming to learn how AI can both automate the profession, and also present new legal challenges[8]. As a profession we have a unique opportunity to use our understanding of data, algorithms, and the way these interact through models to help make sure the AI Spring of 2018 leads to a bountiful harvest.

[8] http://www.lawsociety.org.uk/support-services/research-trends/capturing-technological-innovation-report/