[dstl]

The Science Inside

The Alan Turing Institute

Lancaster University

[ ] Work with us

Please email us if you wish to be added to the mailing list.

[ ] Contribute

Have something to share, please let us know and we can feature it in a future issue.

[ ] Spread the Word

Please feel free to pass this onto your colleagues.

[ ] Contact Details

caiss@ lancaster.ac.uk
caiss@dstl.gov.uk

To email us at Dstl, scan QR code

Issue: 8 **Date: January 2024**

CAISS
Computation & AI for Social Science Hub

Newsletter in collaboration with The Alan Turing Institute and Lancaster University

CAISS Talk Series Reports

The November talk was a great event with Dr Lewys Brace from Exeter University discussing "Biases when exploring online extremist sub-cultures and the "inceolosphere": examples from the ConCel project". Incels, short for "involuntary celibate", is an online sub-culture where individuals define themselves by their inability to form sexual relationships with women. Recent years have seen an increase in the amount of work using large-scale, data-driven, analysis methods to understand such ecosystems, and this work typically utilises text data acquired from online spaces, which is then analysed using Natural Language Processing (NLP) techniques. However, there are several points along the road from data collection, through to interpretation of results where biases can emerge when using such methods on these online sub-cultures.

Lewys talked to us about how bias can appear by:

- Not selecting the "right" online spaces for gathering data
- The data collected may not be representative of the extremist ecosystem
- The initial "seed" list may be biased – (manual checking attempts to reduce this)
- Data cleaning - as these sites have a specific sub cultural language in use – interrogation of the data in depth helps with this
- Edgy humour can be a euphemism for racist, misogynistic and homophobic views, is it irony or genuine?
- Deciding on the measure to use can be problematic – using multiple measures helps with a "sanity check" and can offer additional insights.

The team used the Fisher Jenks algorithm (https://pbpython.com/natural-breaks.html) which uses an iterative approach to find the best groupings of numbers based on how close they are together; (i.e. based on variance from the group's mean) while also trying to ensure the different groupings are as distinct as possible (by maximizing the group's variance between groups). Analyses were also carried out at the micro-level to adopt a context-based approach i.e. integration of ideology with personal life experiences and the macro-level which can cause issues in this case with the use of hateful language. This was mitigated by using violent language and out group terms in the analysis.

A very engaging question and answer session followed covering many aspects of Lewys work such as: group isolation, whether Incels use the dark web (they tend not to), whether the groups can be infiltrated (no, people doing this are spotted, ridiculed and driven out), cross culture (groups are emerging in Japan and Russia), Incel demographics (young, white males in general) and how to track individuals over time. Further work in this area is ongoing using topic modelling and the idea of potential hybrid ideologies.

Report: https://crestresearch.ac.uk/resources/incel-and-the-incelosphere-an-overview-of-current-research-and-understanding/

UK Defence Science and Technology Laboratory

Continued from previous page:

CAISS were privileged to have Assistant Professor Xiao Hui Tao from the University of California Davies deliver our December talk.

Xiao Hui talked to us about how mobile phone data is used to monitor internal displacement within a country, in this case Afghanistan. This is especially relevant at present due to current world events. The forced displacement of people is a key cost of violence and Internal Displaced Persons (IDP) are hard to keep track off. This matters in terms of targeting aid more effectively and understanding likely locations of future instability in order to allocate forces or target specific programmes.

The "vast untapped resource" of mobile phone data was utilised to estimate violence induced placement in a granular manner. This was both methodological and substantive i.e. what was the overall effect of violence on displacement in Afghanistan, what factors affected the choice of destination and could the team confirm and test hypotheses from qualitative work gathered from surveys? A large amount of mobile phone data was used: 20 billion transactions, from the anonymised records of 10 million subscribers from Afghanistan's largest mobile phone operator from April 2013 to March 2017. 398 districts were identified and 5,984 violent events, 13,000 cell towers grouped by proximity into 1,439 tower groups. The results showed that for those in district on a violent day, there was an immediate and statistically significant increase in likelihood of leaving the district.

Results also showed that when looking at Islamic State violence versus Taliban violence, there was a larger impact for IS related violence than for the Taliban, this could be credited to the fact that IS have been known to target civilians when for example executions were filmed. There was also a larger impact with recently experienced violence and a smaller impact in provincial capitals.

When being displaced people were not just seeking economic opportunity. Half of those moving from a capital moved to other capitals or major cities. For those moving from non-capitals, more than half went to capitals or major cities with 30% moving to a provincial capital in the same province. The main driver was seeking safety rather than economic opportunities and this is consistent with the narrative. In non-capitals, violence resulted in people seeking safety close to home.

Xiao Hui talked specifically about some of the limitations and mitigating biases:

- There could be bias in the data sources
- Check and check again if the results contain bias
- People could be sharing mobile phones
- Are phones only being used by the wealthy
- Are women using phones in a patriarchal society?
- Is the displacement intra district rather than inter district?
- Are cell phone towers being destroyed resulting in data of false displacement?

The analysis of this data provided insight into the nature of violence-induced displacement in Afghanistan and helped to quantify some of the human costs of violence that would be difficult to measure using traditional methods such as surveys. While there are definite limitations to what can be observed through mobile phone data, conflict-prone regions are often also the places where traditional survey-based data are the least reliable and most difficult to obtain. This approach could complement traditional perspectives on displacement and eventually contribute to the design of effective policies for prevention and mitigation.

Link to paper: https://www.nature.com/articles/s41562-022-01336-4

Article Review

Hypotheses devised by AI could find "blind spots" in research

Could "Artificial Intelligence (AI) have a creative role in the scientific process" was a question posed in 2023 by a group of researchers in Stockholm. AI is already being used in literature searches, to automate data collection, run statistical analyses and even for drafting some parts of industry and academic papers. Sendhil Mullainathan, an economist at the University of Chicago Booth School of Business in Illinois has suggested using AI to generate hypotheses and stated "it's probably been the single most exhilarating kind of research I've ever done in my life".

AI could help with creativity as using large language models (LLM's) to create new text, even if it is inaccurate, it could lead to a statement such as: "here's a kind of thing that looks true"; when you think about it, this is exactly what a hypothesis is! These "hallucinations" are sometimes likely to be something that a human would not make and could aid thinking outside of the box.

Hypotheses are on a spectrum from concrete and specific to the abstract and general, using AI in areas where fundamentals remain hidden could generate insights. For example we know there is this behaviour happening, but we do not know why, could the AI identify some rules that could possibly be applied to this situation? James Evans, a sociologist at the University of Chicago says AI systems that generate hypotheses based purely on machine learning require a lot of data. Should we be looking to build AI that goes beyond "matching patterns" but can also be guided by known laws? Rose Yu, a computer scientist at the University of California, San Diego states that it would be a "powerful way to include scientific knowledge into AI systems".

Ross King a computer scientist at Chalmers University of Technology in Gothenburg is building robotic systems that perform experiments. Factors are being adjusted subtly in his "Genesis" system allowing these "robot scientists to be more consistent, unbiased, cheap, efficient and transparent than humans".

Hypothesis generation by AI is not new, in the 1980's Don Swanson pioneered "literature based discovery" with some software he created called "Arrowsmith" that searched for indirect connections and proposed for example that fish oil might help treat Raynaud's syndrome, where human circulation is limited in the hands. This hypothesis when taken forward was proved to be correct in that it decreased the bloods viscosity leading to improved circulation.

Data gathering is becoming more automated and automating hypothesis generation could become an important factor as there is more data being generated than humans can handle. Scaling up "intelligent, adaptive questions" will ensure that this capacity is not wasted.

So What? This approach could lead to valid hypotheses being developed which are clear and broad in areas where the underlying principals are poorly understood. A panacea perhaps to "researchers block" to unlock blind spots? For Defence this could mean helping to avoid group think, encourage more innovation outside of the chain of command and enabling things to be done differently in an often slow to change organisation. AI could prove to be a lot more useful than performing Literature Reviews.

Full article: Link here to Nature magazine

CAISS Bytes

Coming soon – March 2024, The CAISSathon

Where: The Alan Turing Institute, The British Library, London, UK.

When: 21st – 22nd March 2024

Theme of the event: Explainability

The social responsibility of Artificial Intelligence (AI) has been under increased scrutiny as it creeps into every facet of society. Understanding the potential ramifications and harm caused by AI is of key importance, particularly as AI technology used for facial recognition, loans and mortgages and job applications (amongst others), have already been shown to be biased against ethnic minority populations and, for example, those with disabilities. Within a Defence context, understanding how AI enabled technologies can facilitate decision making is of key interest. The need for explainable and transparent AI systems is one argument to uncover bias and prevent harm. However, does explainability solve these issues? Does understanding how AI makes it decisions provide enough evidence to negate the harm or at least provide indications of potential harm? Additionally, how understandable do such explanations need to be for expert and lay users?

The Computation and AI for Social Science Hub would like to invite you to participate in our first 'CAISSathon' to explore how explainability can be conceptualised and implemented. This event will propose a series of challenges that will be collaboratively addressed throughout the two days. It will bring together individuals from government, academia and industry to brainstorm and engineer potential solutions. Researchers will have an opportunity to put knowledge into practice and solve problems which have real life implications within Defence. At the end of the two days, a portfolio of potential solutions, research questions and collaboration will be established which will inspire future investigation and lead to insightful developments in this fast moving field. Additionally, there will be a prize awarded to the team who prepare the most innovative solutions.

For more information or to request an invite please email us, at any of the details on the first page

Why Algorithms pick up on our biases

Why do algorithms pick up on our biases? It could be argued that this is due to a 95 year old economic model that assumes people's preferences can be revealed by looking at their behaviour. However, the choices we make are not always what would be best for us. We might have a great wish list on our Netflix account which reflects our true interests, but watch the "trashy" shows that are easier to click on that Netflix sends us. All algorithms are built on what the user is doing, making predictions rather than realistic assumptions as revealed preferences can be incomplete and even misleading. Should algorithms be built with a move away from revealed preferences and encompass more behavioural science? Would this lead to an improvement in our welfare? Or do we just need to watch something "trashy" to de-stress at the end of the day?   Link: Nature Human Behaviour