Issue: 4 Date: May 2023

CAISS

Computation & AI for Social Science Hub

Newsletter in collaboration with The Alan Turing Institute and Lancaster University

AI UK, London March 2023 – CAISS visited the event

Around 3,000 delegates attended the QE2 Centre for AI UK. One of the most popular sessions dealt with the much hyped ChatGPT and was delivered by Gary Marcus, Emeritus Professor of Psychology and Neural Science at New York University. He began by stating that although we have a lot of individual AI solutions (for example, GPS) so far there is not a general purpose system that will do everything for us. ChatGPT is the one most advanced and reliable system to date, taking in massive amounts of data and has good guardrails, so it will not for example write an article on the benefits of eating glass! But is it the universal panacea?

Problems:

- It will make things up and it can even give references for fake information, there is an illusion that adding more information will mitigate the incorrect outputs.
- After completing eight million chess games, it still does not understand the rules.
- Driverless cars involves deep learning, this is not AI. This technology is just memorising situations and is unable to cope with unusual events. The system cannot reason in the same way that a human being does.
- If the circumstance is not in the training set it won't know what to do, in Chat GPT4 (which is the latest version) we do not know yet what that training data set is?

Positives:

- It can help with de-bugging, it can write pieces of code that are 30% correct and then humans can fix them, this is easier than starting from scratch, the "best use case".
- It can write letters, stories, songs and prose, it is fun, fluent and good with grammar.
- Large Language Models (LLMs) can be used to write articles – looks good but they have errors. If someone does not know the facts though it could be believed. But if it is a story and fiction, does this matter?

Worries and considerations:

Chat GPT is being used at scale, leading to misinformation and a possible polluting of democracy, there is an opportunity for fake information, potential discriminatory, stereotypical or even offensive responses. The 2024 US Presidential Election could be a concern, as the technology could be used by State Actors or as an advertising tool – leading to a spread of misinformation that appears plausible. It can write fictitious news reports, describe data etc. e.g. Covid 19 versus vaccines, the results will look authoritative. This could result in millions of fake tweets/posts in a day output via "troll farms". Large Language Models (LLM) without guardrails are already being used on the dark web. ChatGPT has been used in a programme to solve CAPTURES – when challenged the bot said it was a person with a visual disability! Already it is being used in credit card scams and phishing attacks.

Classical AI is about facts, LLM's do not know how to fact check e.g. Elon Musk has died in a car crash – we can check this as humans. With LLM's, as this is such a wide and fast moving area, should we be looking at them in the same way that we would look at a new drug? Possible controlled releases with a pause in place for a "safety check"?

Continued…

[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

DSTL/PUB149693

Continued from previous page:

AI literacy is important for future generations – understanding the limits is crucial, people still need to think in a critical way. Is a coordinated campaign needed to fully understand and warn about the limits of such technology?

Other presentations included Professor Lynn Gladden on Integrating AI for Science and Government, Public Perceptions of AI, how we can "do better in data science and AI", the on-line safety bill, creating economic and societal impact, what can data science do for policy makers and individual skills for global impact. Overall it was a fascinating two days with many opinions and high profile speakers under the overarching banner of open research, collaboration and inclusion.
Link: https://ai-uk.turing.ac.uk/

We told you last time about our forthcoming work on whether people can tell the difference between real and synthetic faces and what effect human or AI guidance would have on the responses. We asked what you thought. The results were as follows:

Question 1: Will participants be able to tell if a face is real or AI generated? Yes = 45%, No = 36%, Don't know = 19%
Question 2: Will participants rely more on human or AI guidance? Human = 64%, AI = 8.5%, Equally = 19% and Don't know = 8.5%
Thank you to everyone who took the time to answer the questions. Look out for our report later in the year when the study concludes.

CAISS Bytes

Did you know that according to Debora Nozza, Assistant Professor in Computing Studies from the University of Milan that the average person reads around 9,000 words per day, that equates to 200 million words in a lifetime. However, that would take up just 0.4 GB of data. You could carry your lifetime of words around on a small USB stick. Consider that around 44 billion GB of data are created each day on the web, I am sure you will agree that this is mind blowing! This is why we may need to turn to automated methods to help us make sense of all this data. Pass my reading specs I've a book to finish …..
https://deboranozza.com/

Italy banned ChatGPT for a while as the Italian watchdog had concerns over privacy. The watchdog said that not only would it block OpenAI's chatbot but it would also investigate whether it complied with General Data Protection Regulations. These concerns have now been addressed and it is available again.
Link: https://www.bbc.co.uk/news/technology-65431914

CAISS Paper Review 1.

Continued from previous page:

Review of Text as Data – Grimmer & Stuart

The paper "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts" by Justin Grimmer and Brandon M. Stewart, published in the Political Analysis journal in 2013, addresses the increasing use of automatic content analysis methods in political science research. The authors argue that these methods have the potential to offer significant advantages over traditional manual content analysis, but also pose important challenges that must be addressed.

The authors begin by outlining the benefits of automatic content analysis methods, including the ability to analyze large amounts of text quickly and accurately, the potential to detect patterns and relationships that would be difficult or impossible for human analysts to discern, and the ability to replicate findings across multiple studies. They also acknowledge, however, that automatic methods are not without limitations, such as difficulties in capturing the nuances of language, the potential for errors in coding, and the need for careful attention to issues of measurement and validity.

To address these challenges, the authors propose a framework for evaluating the quality of automatic content analysis methods, based on three key criteria: validity, reliability, and generalizability. They argue that these criteria should be used to assess the quality of automated methods in political science research, and provide a detailed discussion of how each criterion can be operationalized.

The authors also provide examples of how automated content analysis methods can be used in political science research, including the analysis of presidential speeches and legislative texts, the identification of ideological or partisan biases in news coverage, and the detection of patterns in social media data. They demonstrate how automated methods can be used to generate insights that would be difficult or impossible to obtain using manual methods, such as identifying the specific rhetorical strategies used by politicians to appeal to different audiences.

Finally, the authors acknowledge that the use of automated content analysis methods in political science research is still in its infancy, and that there is much work to be done to refine and improve these methods. They conclude by calling for continued research in this area, with a focus on developing more sophisticated and accurate methods for analyzing political texts, as well as exploring the potential for integrating automated content analysis with other data sources, such as survey data or experimental data.

In summary, Grimmer and Stewart's paper argues that automated content analysis methods offer great promise for political science research, but also pose important challenges that must be addressed. The authors provide a framework for evaluating the quality of automated methods, as well as examples of how these methods can be used to generate insights in political science research. They call for continued research in this area, with a focus on refining and improving these methods, and exploring their potential for integration with other data sources.

Link to paper: https://web.stanford.edu/~jgrimmer/tad2.pdf

Continued from previous page:

CAISS Paper Review 2.

Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. Justin Grimmer and Brandon M.Stewart

This paper although published nearly ten years ago still has some valid points in today's world as it discusses that "language is the medium for politics" and policy, whether spoken or written. In our quest to understand politics, from terrorist manifestos to peace treaties, we need to know what political actors are actually saying. The authors caution around using automated methods as the premise of applying careful human thought and robust validation are needed to ensure rigour. But with today's ever evolving technology is this still the case?

To understand politics we need to ascertain what is actually being said and by whom, in whatever medium it is delivered. However, the volume of material is massive and hiring people to read and code is expensive and scholars cannot do it all themselves. Automated content analysis methods can make this type of analysis possible. The authors do state that automated methods "amplify and augment" careful reading and thoughtful analysis, and their paper takes the reader though all the steps needed for this content analysis. Firstly acquiring the documents, pre-processing them and seeing if they meet the research objective, followed by classification, categorisation and then unpacking the content further. Automated content analysis methods can make the previously impossible possible. Despite the authors initial reservations they offer guidelines on this "exciting area of research" minimising misconceptions and errors and describe "best practice validations across diverse research objectives and models".

Four principals of automated text analysis are identified and the authors encourage revisiting these often during research, these are as follows:

1. All quantitative models of language are wrong – but some are useful. i.e. a complicated dependency structure in a sentence could change the meaning.
2. Quantitative methods for text amplify resources and augment humans.
3. There is no globally best method for text analysis. i.e. there are a lot of different packages available, one of which may suit a particular dataset better than another.
4. Validate, validate, validate. i.e. avoid the blind use of any one method without validation.

The authors point out that automated content analysis methods provide many tools that can be used to measure what is of interest, there is no one size fits all. Whichever tool is chosen needs to be content specific. New texts probably need new methods and ten years ago they identified that commonalities would allow "scholars to share creative solutions to common problems". Important questions could be answered by the analysis of large collections of texts, but if the methods are applied without rigour then few relevant answers will be forthcoming. When undertaking text analysis it is important to realise the limits of statistical models and the field of political science will be revolutionised by the application of automated models.

The overwhelming message of this paper is that textural measurement, the discovery of new methods and inference points allow us to build upon scientific interpretation and theory, and the journey does indeed continue at pace. Machine learning techniques have revolutionised our ability to analyse vast quantities of text, data and images rapidly and cheaply.

Link to paper: https://web.stanford.edu/~jgrimmer/tad2.pdf

More CAISS Bytes

Continued from previous page:

AI Fest - keynote report April 2023

A fantastic few days were had at the University of Exeter by delegates attending AI Fest. The theme of the event was on how to apply AI responsibility and ethically to Defence and Security with a focus on Human Centric AI. The challenge identified is how to link this to the Geopolitical arena. Keynote speaker Charlie Forte, Chief Information Officer for the MOD identified that collaboration and joint working is the way forward. He mentioned that the pace is relentless and we need to organically link everything together as we now have the opportunity to add huge value to the human experience. Air Vice Marshall David Arthurton agreed and reiterated that it is key that we adapt, we may need to embrace risks to ensure we do not get left behind. Paul Kealey, head of CIS Division at Dstl also joined the panel debate and agreed; he stated communication is key especially to win the publics opinion. Over the two days there were many talks, an exhibition and an opportunity to discuss issues at the cutting edge of AI development. This is a very exciting space to be involved with and will touch every aspect of our lives at some point; love it or loath it AI is here to stay and we need to be ready.

Did you see what we did?

The first review of the paper by Grimmer and Stuart was by the one and only ChatGPT, it did a passable job of it. We gave it the following prompt "*using 500 words summarise the paper text as data by Grimmer and Stewart written in 2013".* The second review was by Georgina Mason, CAISS newsletter editor. Please don't tell her if you preferred ChatGPT's version ☺

Have you used ChatGPT successfully to help with a piece of work, we would love to hear from you, tell us what it was and if you thought it was worth it? Please get in touch with the details on page one.

CAISS Out and About…. Conferences coming up…..

CREST international conference on Behavioural and Social Sciences in Security (BASS23), University of Bath, UK, from 11-13th July 2023.

The conference themes are Risk Assessment and Management; Gathering Human Intelligence; and Deterrence and Disruption. BASS23 registration is now open

Link here: https://crestresearch.ac.uk/opportunities/bass23-conference-call/