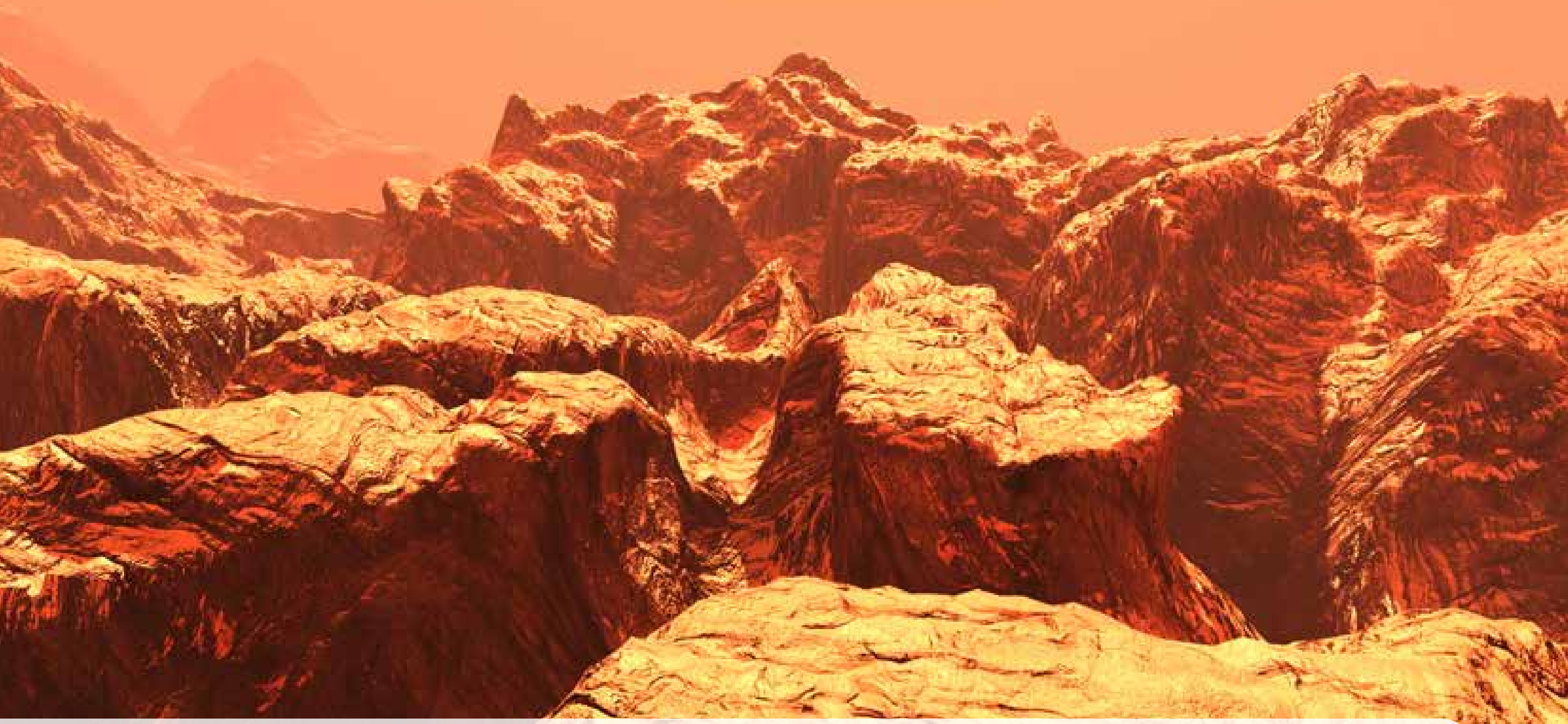
InspiredResearch

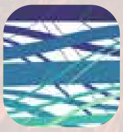
Summer 2023 Issue 23

Research News from the Department of Computer Science, University of Oxford

New study shows artificial intelligence could help locate life on Mars

Read more – p14





ChatGPT: a mathematical assistant in the making?

– p16

Empowering children in the age of datafication

– p19

Cohere: at the cutting edge of emerging AI

– p22

University of Oxford

Department of Computer Science

Inspired Research

is a twice-yearly newsletter published by the Department of Computer Science at the University of Oxford.

If you would like to learn more about anything you read in these pages, please get in touch: editorial@cs.ox.ac.uk

To subscribe to future issues, e-mail: editorial@cs.ox.ac.uk

To download previous issues, visit www.cs.ox.ac.uk/inspiredresearch

Editorial board

Kiri Walden (Managing Editor)
Suzanna Marsh (Editor)
Suvarna Designs (Designer)

Emma Dunlop
Huw Edmunds
Andreas Galanis
Stefano Gogioso
Suzanna Marsh
Jordan Summers
Carolyn Ten Holter
Kiri Walden
Dingmin Wang

Contributors

Simon Birnbach
Tristram Fenton-May
Leslie Ann Goldberg
Sebastian Köhler
Konrad Kollnig
Ulrik Lyngs
Ivan Martinovic
Pericle Salvini
Frieder Simon
Jack Sturgess
Carolyn Ten Holter
Ge Wang
Caroline Wood
Jun Zhao

CONTENTS

Page 3	Letter from the Head of Department
Page 4-9	News
Page 10	Alumni Profile
Page 11	GDPR: How a lack of enforcement came to undermine its ambition and reputation
Page 12-13	The self-regulating challenges of computing technology
Page 14-15	Artificial intelligence could help locate life on Mars
Page 16-17	ChatGPT: a mathematical assistant in the making?
Page 18	Entangled challenges: policy, publics and UK plc
Page 19	Empowering children in the age of datafication
Page 20-21	Using Inertial Sensors to Authenticate Users
Page 22-23	Cohere: at the cutting edge of emerging AI
Page 24	The role of data recorders in the reconstruction of a psychological accident

Letter from the Head of Department

Welcome to the Summer 2023 Issue of *Inspired Research*

To introduce this issue, I'd like to tell you about some research that has led to recent recognition for three of our faculty.

First, I'd like to tell you about research by **Professor Christian Coester** and his co-authors Sebastien Bubeck of Microsoft Research and Yuval Rabani of the Hebrew University. This research won a 'best paper' prize at the 55th Annual ACM Symposium on Theory of Computing which will took place in June in Florida. The paper refutes the 'randomized k-server conjecture', which had been a central open question in the field of online algorithms for several decades. In the k-server problem, there are k mobile servers located in some space. One by one, points of the space are requested, and each time an algorithm has to select a server to move to that point – without knowledge of future requests. It was conjectured that a randomized algorithm exists that is efficient, in the sense that the distance travelled by its servers is at most a log(k) factor larger than the optimum in hindsight (when all requests are revealed). The paper shows that no such algorithm can exist (no matter what, provably it doesn't exist). At a high level, this means that the lack of information is costlier than previously believed.

Second, I'd like to tell you about the research of **Professor Edith Elkind**, who has won the 2023 SIGAI Autonomous Agents Research Award. In connection with this award, Edith presented a plenary talk at the 22nd International Conference on Autonomous Agent and MultiAgent Systems in June. As the award citation explains, 'Her work provides fundamental understanding of economic paradigms in multiagent systems, with a particular focus on computational social choice and game theory.' The primary focus of Edith's current work is algorithmic challenges in multi-winner voting, where the goal is to select a fixed-size subset of alternatives from a given set. This captures settings such as identifying a set of candidates to shortlist for a job or selecting food items for a banquet or speakers for a conference. It is challenging to formulate what constitutes a fair solution in such settings, and Edith's work both proposes formal fairness criteria and develops efficient algorithms (voting rules) to identify solutions that satisfy these criteria. Achieving perfectly fair solutions on all instances tends to be (provably) computationally infeasible, so her work focuses on identifying structural constraints on voters' preferences that make these problems tractable.

Finally, I'd like to tell you about some research by **Professor Michael Bronstein** and his colleagues that was published in *Nature* in April. This research has impact on biology, so I will start with the biological context. Interactions between proteins are essential for most biological processes, so the ability to control protein-protein interactions is of great interest. While textbook depictions of protein binding often look extremely simple, the reality is more complex, making it hard to predict how and where binding events will occur. Professor Bronstein, together with collaborators from the EPFL protein engineering lab led by Professor Bruno Correia, developed a novel geometric deep learning framework called 'MaSif' which makes it possible to analyse protein structure and functional properties in order to computationally design protein interactions. In the *Nature* paper, they report brand-new protein binders that are designed to interact with four therapeutically relevant protein targets, including the SARS-CoV-2 spike protein.

I hope that you'll enjoy reading about more of our research in this issue.

Professor Leslie Ann Goldberg
Head of the Department of Computer Science
June 2023

NEWS

Andrzej Murawski to receive a Humboldt Prize

Associate Professor Andrzej Murawski has been elected to receive one of this year's Humboldt Prizes. The Humboldt Prize, the Humboldt-Forschungspreis in German, also known as the Humboldt Research Award, is an award given by the Alexander von Humboldt Foundation of Germany to internationally renowned scientists and scholars who work outside of Germany, in recognition of their lifetime's research achievements. Recipients are 'academics whose fundamental discoveries, new theories or insights have had a significant impact on their own discipline and who are expected to continue producing cutting-edge academic achievements in the future'. The prize is currently valued at €60,000 (about £50,000) with the possibility of further support during the prize winner's life, and can be used to support research projects carried out in cooperation with specialist colleagues in Germany. Andrzej will use part of his award to initiate a collaboration with Professor Roland Meyer from Technische Universität Braunschweig.

Professor Marta Kwiatkowska elected to American Academy of Arts

We are pleased to announce that Professor Marta Kwiatkowska has been elected as a member of the American Academy of Arts and Sciences.

Founded in 1780, the Academy honours excellence and convenes leaders from every field of human endeavour to examine new ideas, address issues of importance to the United States and the world, and work together, as expressed in their charter, 'to cultivate every art and science which may tend to advance the interest, honour, dignity, and happiness of a free, independent, and virtuous people.'

Marta joins the company of notable members – from the earliest members John Adams, Benjamin Franklin, Alexander Hamilton, and George Washington to Ralph Waldo Emerson, Maria Mitchell, and Alexander Graham Bell. International Honorary Members have included Charles Darwin, Albert Einstein, Winston Churchill, Wislawa Szymborska, Laurence Olivier, Mary Leakey, Gabriel Garcia Márquez, Akira Kurosawa, and Nelson Mandela.

Current members represent today's innovative thinkers in every field and profession, including more than two hundred and fifty Nobel and Pulitzer Prize winners. Marta has been invited to a formal induction in September, to be held in Cambridge, MA, where the Academy's headquarters are located.

New industry-academia research collaboration announced in Responsible Quantum Computing

The University of Oxford's Responsible Technology Institute (RTI) is delighted to announce a new research collaboration with the Quantum Computing and Simulation Hub (QCS), and EY.

Although significant progress has been made in quantum computing in recent years, there are many engineering challenges that remain before the commercialisation of the technology is a reality. Despite these challenges, it is the right time for the quantum computing community to engage with industry and with society on the technology's implications. This is necessary in order to ensure that society experiences not only the benefits of advanced quantum computing research, but also that principles of good governance, transparency and other aspects of responsible development are translated from the research environment into the commercial sector.

The ResQCCom project (Responsible Quantum Computing Communications) will focus on engaging with industry, with policymakers, and with the general public to discuss how quantum computing may affect society and how to prepare for this.

The project is being led by Professor Marina Jirotka, Professor of Human-Centred Computing and Director of the RTI, who commented, 'The rate of progress in quantum computing has accelerated enormously in recent years, and the industrial sector is rapidly becoming more developed. As potential use-cases become clearer, the importance of preparing for and anticipating its effects becomes more urgent.'

More information can be found at https://bit.ly/44FdxSj

University of Oxford researchers work together to protect COVID-19 orphans

Associate Professor Seth Flaxman (Department of Computer Science) and Professor Lucie Cluver (Department of Social Policy and Intervention) worked together alongside global organisation including the World Bank, WHO and NGOs such as Save the Children to change the lives of children who lost a parent or guardian due to COVID-19.

A new film, which is part of the Oxford Policy Engagement Network (OPEN) Research Stories series, shows how Seth and Lucie combined their experience of AI and machine learning and child and family social care to work together during the COVID-19 pandemic to improve global understanding of the number of children 'left behind' by losing a parent or guardian during the pandemic.

The film called *Researcher Stories: Covid-related Orphanhood* highlights the work of the Policy partnership called The Global Reference Group on Children Affected by COVID-19 and Professor Lucie Cluver, Department of Social Policy and Intervention said: 'Policy partnership is fundamental to everything we do.'

Having worked on AIDS orphanhood for 15 years, Lucie knew that research would be needed into the children orphaned by COVID 19. Combining epidemiology and demographics methods, they worked out that over 10.5M children have lost a parent or primary carer due to COVID 19. That figure could then be used by the World Bank, WHO and NGOs such as Save the Children to inform policy changes.

Associate Professor Seth Flaxman said: 'It's amazing to see how the estimates that come out of the studies we do can actually change policy and make a difference in the lives of children globally.'

Laura B Rawlings, Lead Economist at the Human Capital Project, World Bank said of the project: 'The work that we have done together has brought the importance of this crisis to the forefront of the policy dialogue. The support in terms of research and data, so that countries can respond accordingly is invaluable.'

Oxford's researchers and academics have a wealth of experience in engaging with policymakers and contributing to policy impact. This is another example of researchers from across faculties and subject areas who are working to share knowledge and expertise with the policymaking community so that, together, we can contribute to better policies that protect what is valuable and change the world for the better.

OPEN, who fund the Researcher Stories, is a growing network of researchers, doctoral students and professional services staff at the University of Oxford who share a vision of public policy powered by the world's best available research evidence and expertise.

Watch the film and find out more about the Researcher Stories series here: https://bit.ly/44oROhO

News in brief

Steve Hill and Lucy Sajdler were members of an interdepartmental Mathematical Physical and Life Sciences (MPLS) division team which enabled Oxford University to pick up one of three awards for UK universities at the Higher Education Business Continuity Network (HEBCoN) Awards 2023. The collaborative work last term involved staff across MPLS, who worked together to prepare for potential energy supply disruption this winter.

Second-year DPhil student Ruiwen Dong's paper 'The Identity Problem in $Z \wr Z$ is decidable' has been awarded Best Student Paper at the EATCS International Colloquium on Automata, Languages and Programming, which is one of the flagship conferences in Theoretical Computer Science. He will also receive the prestigious Kleene award for the best student paper at the Thirty-Eighth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS).

Christian Coester has co-authored a paper 'The Randomized k-Server Conjecture is False!', which has won a Best Paper award at STOC 2023: 55th Annual ACM Symposium on Theory of Computing. The paper refutes the 'randomized k-server conjecture', which has been a central open question in the field of online algorithms for several decades. Christian gave a talk about this paper for the online seminar series TCS+, which you can watch here: https://bit.ly/3pKY0BD

News in brief

A paper titled 'All's Well That Ends Well: Avoiding Side Effects with Distance-Impact Penalties' has received a Best Paper award at the NeurIPS Workshop on Machine Learning Safety. The paper investigates how the use of bespoke distance-impact metrics in the context of reinforcement learning, supports the prevention of side effects, whilst still permitting task completion. The authors of this work are Charlie Griffin, Joar Max Viktor Skalse, Lewis Hammond and Professor Alessandro Abate, all members of the Department of Computer Science, and of the Oxford Control and Verification group (OXCAV).

A paper, co-authored by two members of the Department has received the Outstanding Paper award at the 37th AAAI Conference on Artificial Intelligence (AAAI-23). AAAI is the flagship conference in AI, and is organised by the Association for the Advancement of Artificial Intelligence. The awarded paper, titled 'Misspecification in Inverse Reinforcement Learning', was written by Joar Skalse and Professor Alessandro Abate of the Oxford Control and Verification group (OXCAV). The publication is openly accessible on the arXiv at: https://bit.ly/3NPWu9d

Niki Trigoni, a Professor at the Department of Computer Science and CTO (Chief Technology Officer) at Navenio Ltd, has won the CI/TO of the Year title for the second year running. The award was made for 'leading a game-changing tech solution in solving real world problems built on award-winning science from University of Oxford'.

Oxford Professors awarded prestigious UKRI Turing AI World-Leading Fellowships

The University of Oxford's Professor Michael Bronstein (Computer Science) and Professor Alison Noble (Engineering) have been awarded prestigious UKRI Turing AI World-Leading Fellowships to conduct ground-breaking work on some of AI's biggest challenges. The two new Fellowships represent an £8 million investment as part of a suite of investments totalling £50 million announced by UKRI today, which will develop trustworthy and secure AI to help solve major challenges.

Professor Chas Bountra, Pro-Vice Chancellor for Innovation at the University of Oxford, said: 'I congratulate Professor Bronstein and Professor Noble on being awarded these prestigious Fellowships, and have no doubt they will achieve great things in such a rapidly developing field. AI has the potential to make a hugely positive impact on societies globally, and Oxford is at the forefront of this revolution. Our talented researchers, from a range of departments, are making tremendous advances in AI, from scientific breakthroughs through to security and ethics. This holistic approach can only bring benefits.'

Professor Michael Bronstein is an expert in theoretical and computational geometric methods for machine learning and data science, with his research encompassing a broad spectrum of applications ranging from computer vision and pattern recognition to biochemistry, drug design, and animal communication. Using the Fellowship, he will develop a novel mathematical framework for geometric and graph machine learning.

Michael said: 'I am greatly honoured to be chosen as one of the Turing AI World-Leading fellows. This grant will allow us to develop a novel mathematical framework for geometric and graph machine learning inspired by physical principles. Together with academic and industrial partners, we will apply these methods to some of the most challenging problems in the domains of drug and food design. In the longer perspective, I hope that our work will help develop new therapeutic protein molecules for diseases that are difficult to target with existing drugs, and assist in mapping the "dark matter" of food-based bioactive ingredients.'

Technikos Professor of Biomedical Engineering Alison Noble's research interests sit at the interface of AI with computer vision and clinical medicine. In recent years, her group has been at the forefront of international thinking in how to bring machine learning to ultrasound, working closely with clinical research groups in Oxford and overseas.

With the Fellowship, Professor Noble will work towards new AI for shared human-machine decision-making in healthcare imaging including studying ethics of the AI and trustworthiness (explaining decisions).

Kedar Pandya, Executive Director, Cross-Council Programmes at Engineering and Physical Sciences Research Council, said: 'The UK's expertise in the field of AI is a major asset to the country and will help develop the science and technology that will shape the fabric of many areas of our lives. That is why UKRI is continuing to invest in the people and organisations that will have wide-ranging benefit.'

The two new Fellowships build on the University of Oxford's previous success when the first five Turing AI fellows were announced in 2021. These included Professor Philip Torr, from the Department of Engineering Science, and Professor Michael Wooldridge, based at the Department of Computer Science.

Professor Sadie Creese speaks at Davos, WEF23

In January Professor Sadie Creese was invited to speak at the World Economic Forum annual meeting in Davos. Cybersecurity was a major topic of discussion and Sadie was one of several experts to warn that cyberattacks are increasing in sophistication and frequency.

In an interview during the event Sadie commented, ‘There's a gathering cyber storm. This storm is brewing, and it's really hard to anticipate just how bad that will be.’

The Annual Meeting 2023 coincided with the release of the Forum's 2023 Global Cybersecurity Outlook. The report found that business leaders are far more aware of the cyber threat than the year prior. In fact, 91% of respondents said they believe a far-reaching and catastrophic cyber event is at least somewhat likely in the next two years. However, the report concludes that organizations continue to face significant challenges when it comes to effectively addressing cyber concerns.

The Forum's report also notes that the potential targets for cyberattacks are increasing. Today, targets include not only government agencies or major corporations, but largely any organization that handles consumer data—no matter how small. Sadie said, ‘We need to accept that this is really about cyber resilience. There is no such thing as a hundred percent security. It's about resilience in the face of insecurity.’

The video of Sadie speaking as part of a panel of experts can be viewed here: https://bit.ly/46Oi5HR

Academics in the news

Largely thanks to the unprecedented boom of AI-related news following the launch of ChatGPT our academics have been almost constantly in the news in the last 6 months. Researcher Carolyn Ten Holter was interviewed on BBC Newshour, Professor Nigel Shadbolt co-authored an article on AI regulation for the Sunday Times, and Associate Professor Reuben Binns was quoted in an article about AI, but it turned out the article was written by AI and he'd never said any of it (a development Reuben himself found very interesting).

However, it's Professor Michael Wooldridge who has been kept the busiest. He's been interviewed for international newspapers and radio stations. In the UK he's appeared on BBC Politics Live, BBC Radio 4's Today programme, ITV's Tonight show, BBC World, BigThink, Adrian Chiles on Radio 5, Naked Scientists, BBC's The Briefing Room and others. In print he's turned up in *The Washington Post, The New Statesman, The Guardian* and many, many more.

While media appearances might be time consuming, they demonstrate our department's position as the ‘go to’ University Computer Science department in the UK for expert comment.

News in brief

Professor Edith Elkind has won the 2023 ACM SIGAI Autonomous Agents Research Award, for her work on computational social choice and algorithms for cooperative games, and extraordinary service to the community. The annual ACM SIGAI Autonomous Agents Research Award recognises researchers in autonomous agents whose current work is an important influence on the field. https://bit.ly/46K2wkw

Professor Marta Kwiatkowska has been awarded Poland's highest recognition of her work, a titular professorship. Marta attended a ceremony at the Presidential Palace in Poland, where she was given her award in person by the President of Poland, Andrzej Duda.

Professor Sir Nigel Shadbolt is one of 9 new experts appointed to advise the Competition and Markets Authority (CMA) as it prepares for new powers to oversee digital markets. The CMA has appointed specialists at the forefront of technological innovation, online competition and tackling the dominance of some of the world's most powerful firms.

This comes as the CMA prepares to be given new powers by government to tackle problems in online markets more rapidly, and ensure consumers benefit from free and fair competition. This means, for example, being able to set targeted rules which the most powerful firms must follow, rather than tackling problems after the harm has already been done.

NEWS

Teaching and Learning Technology Showcase success

In April, the Department of Computer Science hosted the first-ever Teaching and Learning Technology Showcase. The event was supported and attended by people from across the University, providing an opportunity to gain hands-on experience with various technologies that support and enhance teaching and learning. Regardless of their role, anyone interested in the theme was welcome to attend. The showcase featured examples of technology used to support reading and writing, such as eReaders, tablets, eInk monitors, and touch screen monitors. Additionally, there were demonstrations of the application of augmented reality (AR) and virtual reality (VR) in teaching, innovative translation technology for lectures and seminars, as well as presentation technology, including large touch-screen panels and our own lecture theatre technology. Three talks were also delivered, primarily focusing on tools and tips to improve academic productivity.

Feedback from the event was overwhelmingly positive. Attendees expressed their appreciation for the opportunity to test out the technologies, engage with experts, and network with like-minded individuals in the University. Many attendees visited the Department of Computer Science for the first time and were impressed by the quality of the facilities and the friendliness of the staff. The department was recognized as a hub of innovative people, eager to collaborate on projects. One attendee reached out after the event to convey their appreciation and said, 'I am enthusiastically looking forward to technological events hosted by the Department of Computer Science, such as the Showcase'.

It sounds like it is something that could turn into a regular event! If you are interested in being involved next time, please get in touch (email: jennifer.watson@cs.ox.ac.uk)

Research winners and commendations announced for MPLS Impact Awards 2023

We are delighted that the department's Sebastian Köhler was awarded an MPLS Early-Career Research Impact Award for his research on the security of electric vehicle charging.

Researchers across the Mathematical, Physical and Life Sciences Division (MPLS) have been recognised for their outstanding research impact at the annual MPLS Impact Awards.

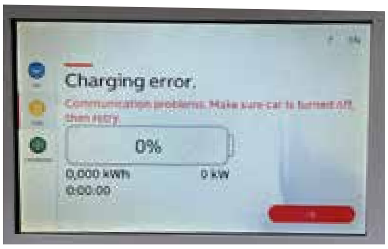
The awards celebrate the work of MPLS researchers who have made significant contributions to the economy or wider society at large through their research.

This year's winners were selected from nominations representing MPLS researchers at all career stages. The winners will each receive a £1,000 prize in recognition of their achievements.

In 2022, doctoral researcher Sebastian Köhler identified a significant vulnerability in the network protocol of the Combined Charging System (CCS), a widely adopted standard for electric vehicle (EV) charging in the US, Europe and Asia. Sebastian found that CCS charging could be disrupted from a certain distance away, using a remote signal, with the potential to impact approximately 20 million EVs and any electric battlefield, and emergency vehicles, buses, heavy trucks, boats, ferries, mining machinery or small aircraft using the CCS standard.

This vulnerability, known as 'Brokenwire', was assigned a Common Vulnerabilities and Exposures ID in April 2022 and has since been formally recognised by the National Institute of Standards and Technologies (US). To protect vehicles, machinery and aircraft using the standard against the malicious Brokenwire attacks that Sebastian's research scrutinised, adaptations to the CCS standard are required and hardware components in every affected car and charger will need to be updated.

Professor Leslie Goldberg, Head of the Department of Computer Science, said: 'This is a great example of the exciting and impactful work happening in our security research theme. Sebastian's research has immediate and significant implications for technology that is rapidly becoming a fundamental part of global infrastructure. An industry-wide response will be required to mitigate the vulnerability that Sebastian has identified and analysed.
'His discovery will help to make these technologies more robust and secure in the future, and he is moreover making an ongoing contribution to the coordinated industry and government response. Impact on this scale is truly impressive for a researcher at Sebastian's stage, and I'm very happy that it is being recognised through the MPLS Impact Awards.'



Charging error.

Communication problems. Make sure car is turned off, then retry

0%

0,000 kWh 0 kW
0:00:00

Error message shown by the charging station after the charging session has been disrupted

Andrew Ker recognised in this year's MPLS Teaching Awards

We are delighted that Andrew Ker is one of 10 academics who have been recognised with divisional teaching awards.

The Mathematical, Physical and Life Sciences (MPLS) divisional Teaching Awards scheme celebrates success and recognises and rewards excellence in innovative teaching practice. It is open to everyone who teaches, including graduate students, postdoctoral researchers, faculty and learning support staff.

Awards are made on merit, with winners selected by a cross-departmental panel. This year, entries closed in April and 108 members of staff were nominated, through 168 separate nominations. The panel met earlier this month to select ten winners, who will be recognised as part of a reception attended by senior leadership from across MPLS, in September 2023.

Andrew is a Lecturer and Tutorial Fellow in Computer Science. He received a student nomination praising the well-structured lectures he delivers, which help to support learning in areas which are not necessarily easy or favourites amongst his students. The Department of Computer Science endorsed the nomination and stated that other student feedback received for Andrew indicates that a number of other students might have wished to put forward a nomination for a Teaching Award, particularly since Andrew consistently receives high feedback scores for his lectures.

Professor Sam Howison, Head of the Mathematical, Physical and Life Sciences Division, said: 'Our Teaching Awards are an opportunity for us to shine a light on the commitment and innovation of our teaching staff at all levels, and across every department, who are so ably supporting the University's teaching mission and helping to inspire the next generation of leading scientists. How wonderful to see such a strong haul of entries, and I congratulate everyone on their nominations, especially our ten overall winners this year.'

Double-award success for research paper

The International Federation for Information Processing (IFIP) has given Nobuko Yoshida and colleagues Claudio Antares Mezzina and Francesco Tiezzi the DisCoTec (International Federated Conference on Distributed Computing Techniques) 2023 Best Paper Award for their paper Rollback Recovery in Session-based Programming. The same paper was in addition recognised with the COORDINATION 2023 (International Conference on Coordination Models and Languages) Best Paper Award.

The work detailed in the winning paper proposes a new reversible computing framework based on session types. The research enriches a language with programming facilities to commit session interactions, to roll back the computation to a previous commit point, and to abort the session. The correctness of the recovery is ensured at design-time (statically) by relying on a decidable compliance check at the type-level, implemented in MAUDE.

Both awards were announced at the DisCoTec 2023 conference, which took place in June 2023, in Portugal.

Inaugural PSS recognition awards

The first round of Computer Science Professional Services Staff Awards have been made. The Most Valuable Team Player Award is nominated termly, and recognises colleagues for outstanding work that has positively impacted their team and department. The Going Above and Beyond Award is nominated termly by management to recognise staff for showing initiative and demonstrating forward thinking that benefits their growth and the development of the department. The award winners were as follows:

Going Above & Beyond Award WINNER: Annette Vaneeden [below right], HIGHLY COMMENDED: Shannon Beesley

Most Valuable Team Player Award WINNER: Akiko Frellesvig [below left], HIGHLY COMMENDED: Grants Team – Ian Watts, Vernon Jenner, Christine Salmon and Claire Rugg

[L to R] Claire Rugg, Vernon Jenner and Ian Watts from the Grants Team.

Alumni Profile

Tristram Fenton-May – Vice President of Engineering at Deriv, one of the world's largest online brokers.

What course did you study here and when?

I matriculated in 1997, starting Maths and Computation, then switched to straight Computation after mods (first year exams) in the summer of 98. After completing my undergraduate degree I went on to a DPhil. The title of my thesis was 'Parallel and Hardware Implementation of Collision Detection Algorithms'.

What was your background before that?

Before Oxford I lived in South Wales, near Cardiff. I did my A-levels at Bishop of Llandaff in Cardiff.

What attracted you to studying Computer Science as a subject?

When I was a child my dad got an 8 bit computer, and wrote some simple games in BASIC which he gave to me for Christmas (on a tape) when I was about 5. Whilst I enjoyed playing them, learning how to write my own fascinated me more - one of my early memories is waking my dad up to ask how FOR loops work. School at that time didn't teach any programming until A-level – so I taught myself C and X86 assembly from books.

What aspects of the course you studied here did you particularly enjoy?

I really enjoyed being able to focus on Computer Science, and particularly enjoyed the architecture course where I learned the low level details of how computers worked. However, as interesting as the course was, one of the best aspects of studying in Oxford was meeting other people interested in computers. On the first day I met a fellow Computer Scientist and talked about the source code for Wolfenstein 3D, which had recently been released – we ended up doing a 3rd year project together implementing Wolfenstein 3D in hardware without a traditional load-fetch-execute microprocessor architecture. We are still good friends, and I made many life-long friends at university.

What did you do when you left Oxford?

Whilst I was finishing my DPhil I did a bit of consulting for a company who made solutions for truck companies. This included selling hardware for remote tracking using GPRS (a fairly new concept in 2003) and providing the data via SAAS. The contracting grew, and after I finally completed my DPhil I started working for them full-time.

It was a small company and I did a mixture of software and firmware for the devices we installed in vehicles. I was amazed when I was invited to observe a board meeting around 2005 and the first item on the CTO's agenda was 'migrate to linux' – something I had suggested to my boss a few weeks earlier.

How has the course you studied here helped you in your current career?

Whilst I have not written down an invariant very often, my theoretical computer science background has given me a great foundation to work from. During the course it seemed crazy that we learnt esoteric languages, but I rapidly learnt that specific languages come and go, and the understanding of the principles of language are much more important than any specific language. Many companies realise this, and when hiring people for the long term look more for deep understanding than specific buzzwords (of course buzzwords are important to HR and recruitment consultants!).

What advice would you give to current students on applying their knowledge in the workplace, when they leave university?

I think that a computer science background is a very useful foundation – I don't think you really need to go out of your way to use it. Once you learn the principles it is like riding a bike – every decision you make will be improved by the deep understanding of the problem space. In some cases, like implementing complex algorithms, or making architectural decisions this is obvious – but even when doing less obviously mathematical tasks, such as building front end websites, the foundations still show through in decisions about how to implement things.

What would the student you have thought about what you are currently doing – would you have been surprised, proud, amazed?

I have just left my role in the company that I contracted for during my doctorate (having been through a few mergers) and been the Vice President of Architecture and Chief Technology Officer. My new role which I will start shortly is in FinTech. I would never have imagined that my 'bit of contracting' would turn out to last over two decades!

GDPR: How a lack of enforcement came to undermine its ambition and reputation

By Doctoral Student Konrad Kollnig

Tracking, the collection and sharing of behavioural data about individuals, is widely used by app developers to analyse and optimise apps and to show ads. It also is a significant and ubiquitous threat in mobile apps, and often violates data protection and privacy laws.

Previously, our research group, led by Professor Sir Nigel Shadbolt, analysed 1 million Android apps from the Google Play Store from 2017. We found that about 90% of those apps could share data with Alphabet (the parent company of Google), and 40% with Facebook (now renamed 'Meta'). The data practices in children's apps were particularly worrisome, which is why our research group – in response – established a dedicated research strand on Kids Online Anonymity & Lifelong Autonomy (KOALA), led by Jun Zhao.

The much-debated General Data Protection Regulation (GDPR) came into force in the EU and UK in May 2018. This new law aimed to protect personal data better than its predecessor, the Data Protection Directive (DPD) from 1995.

Compared to the DPD, the GDPR introduces significant challenges for compliance in the context of tracking, particularly through higher potential fines (up to £17.5 million or 4% of global annual turnover), better regulatory alignment and enforcement and a higher bar for consent. Since there existed (and still exist) rather limited empirical insights into the extent to which the law achieved its intended aims, we set out to replicate the same analysis of Android apps several years later. Interestingly, compared to our previous study, our renewed work attracted much less public attention, which may underline that many of us have become used to invasive and illegal data practices by tech giants.

We analysed 2 million apps from the UK Google Play Store, 1 million apps from 2017 and 1 million from 2020. We performed an automated scan of apps' code to identify all domains that are known to belong to tracking companies, thereby characterising the companies that apps can potentially send personal data to. If there have been changes in the extent of tracking following the introduction of the GDPR, we should expect that they show up in our results. Crucially, we did not investigate what kinds of data were shared by apps, since machine learning approaches (commonly referred to as 'AI') nowadays make it possible to get detailed insights into the lives of individuals even from seemingly benign data.

Our results suggest that the GDPR has not had a large effect on the presence of tracking in apps on the UK Google Play Store. For instance, 85% of apps from 2017 could send data to Alphabet, compared to 89% in 2020. 43% could send data to Facebook in 2017, and 38% in 2020. Apps, on average, contain a similar number of trackers as before (5 companies in the median app). A consistent percentage of apps (15%) contain more than ten tracker companies.

Our analysis hints at a high level of concentration in the tracking market. Alphabet/Google and Meta/Facebook continue to dominate app tracking. This dominance might allow them to extract disproportionate profits from their digital advertising models, which would – in turn – increase the prices of consumer products. This dominance is currently subject to ongoing investigations by courts and authorities across the globe. At the same time, many relatively smaller companies are involved in app tracking. These smaller companies usually focus exclusively on mobile advertising, instead of having a broad portfolio of digital services like Alphabet/Google or Meta/Facebook. An important competitive advantage of these companies might be reduced public awareness and regulatory scrutiny, allowing them to compete with the market leaders at the expense of user privacy.

We further found that apps commonly shared with US-based companies. While one of the key aims of the GDPR is to facilitate the cross-border sharing of personal data between companies, the US also operate some of the most sophisticated intelligence agencies (like the NSA) and currently provide limited protections against surveillance by those agencies to non-US citizens. These practices were revealed by Edward Snowden in 2013. In light of this, the Court of Justice of the European Union repeatedly found that the US does not provide a similar level of data protection to EU citizens as the GDPR, and that personal data may usually not be sent to the US (Schrems II ruling). Apps' ubiquitous data sharing with US companies is thus problematic, if not illegal. While the GDPR is far-reaching, the law is not perfect. Apps continue to rely on invasive (and often illegal) tracking technologies. The law does not appear to have changed these incentive structures fundamentally.

Unfortunately, the GDPR remains rarely enforced in practice, which has led to a proliferation of illegal data practices online. These practices include the engagement in highly invasive data practices (including many, but not all, forms of tracking), frequent and insufficiently protected sending of personal data to the US, and also the wide adoption of annoying and ineffective consent banners.

Both the UK and EU are currently planning to revise their data protection laws. According to our broad body of research, the lack of enforcement is the central issue that needs to be addressed.

The self-regulation challenges of computing technology

By Ulrik Lyngs, Research Associate in the Human Centred Computing Group

Powerful, portable computing technologies like smartphones and laptops challenge our ability to control our attention and behaviour. Amidst constant notifications and AI-powered recommendation feeds that try to optimise ‘engagement’ with digital services, many people find it difficult to use their devices in the way they want without being distracted and wasting time. For example, our work might routinely take much longer than it should, because we keep interrupting ourselves to go to social media. Or we might fail to go to bed at the time we intended, because we get ‘sucked in’ by our devices.

There are two main reasons we can find it difficult to self-regulate digital device use. First, digital services are often deliberately designed to nudge us into behaving in ways that benefit the developer. This can take the form of manipulative ‘dark patterns’, such as pre-ticked boxes on shopping sites with deliberately ambiguous wording. It can also take the more ambivalent form of ‘grey patterns’ like infinitely scrolling feeds that, though they may surface genuinely valuable content, can act as attention sinks and make us forget our intentions for use.

Second, the sheer scale of instant access to information, connection, or entertainment can cause self-regulation difficulties. Psychological research shows that people who are better at self-regulating, are those who arrange their environments such that they have fewer temptations available. Therefore, if digital devices make everything always available with minimal effort, this should in itself make it difficult to self-regulate use.

Emerging design solutions

The good news is that in recent years, developers and researchers have experimented with design patterns that support self-regulation. On online stores for apps and browser extensions, hundreds of ‘digital self-control tools’ now provide interventions that help people stay in control, for example by blocking apps or hiding distracting website elements, tracking and visualising use, or providing rewards or punishments for how devices are used. Researchers have done controlled studies with such tools, and a wide range of interventions have been shown to both change behaviour and make people feel that they are better able to control their device use.

Different strategies have different likelihoods of success, and due to individual differences in goals, personal, and ecosystem of devices used. However, similarly to the findings of psychological research, our own research data from workshops with hundreds of students at the University of Oxford, suggests that the most generally useful strategies involve empowering users to control how much potentially distracting information they are exposed to.

For example, being able to hide recommended videos on YouTube, or use blocking tools to temporarily reduce the functionality of one’s devices to just those apps that are needed for a given task.

Platform conditions determine what solutions are available

The extent to which people are able to control the amount of information in their digital environments differs widely between platforms. The largest amount of freedom exists on the web, where internet browsers allow people to customise the look and functionality of any website via browser extensions. All major browsers now share the cross-browser Web Extensions API which allows users to, for example, tailor Facebook to their needs via injected stylesheets or scripts.

The situation is very different for mobile apps, however, where most people report self-regulation difficulties. No options currently exist to change the look or functioning of mobile apps to fit individual user needs beyond the options developers choose to provide.

Moreover, the tools available for blocking apps or tracking use differ dramatically between Android and iOS: on Android, a diverse ecosystem exists for apps that provide customised app blocking and usage visualisations, in addition to what Google provides in their ‘Digital Wellbeing’ app. On iOS, Apple takes a much more restrictive approach to developer permissions, which means that few options beyond Apple's own Screen Time tool are available for proper app blocking or time tracking.

Is there a role for regulation?

How might regulation support the availability of design patterns that support user self-regulation over digital device use? Let us consider the levels of (i) specific design patterns, (ii) user control over digital interfaces, and (iii) business models.

Banning specific design patterns?

Some attempts are currently being made at banning use of specific design patterns. For example, in relation to the GDPR, the design pattern of pre-ticked checked boxes has been ruled not to constitute valid consent. Similarly, EU's new Digital Services Act intends to ban "misleading interfaces known as 'dark patterns' and practices aimed at misleading users".

Banning specific patterns may apply to a small number of obviously manipulative practices. However, the much larger zone of 'grey patterns' commonly used to drive user engagement (for example, infinitely scrolling feeds, the selection algorithms that curate those feeds, autoplay), cannot be addressed through bans. Whereas these patterns may cause self-regulation difficulties as a side effect, they often serve important user needs.

Mandating increased control over digital interfaces?

An alternative is to mandate greater user control. An example is underway in the EU's Digital Services Act, which will require very large online platforms that use recommender systems to provide at least one option for algorithmic selection not based on personal profiling.

I wish to suggest a more far-reaching option: as mentioned, users already have the power on the web to control the appearance and functionality of websites via browser extensions. Mandating an analogue of this for mobile apps has the potential to radically empower users and allow the long tail of idiosyncratic 'online harms' to be addressed in a bottom-up fashion. The research project 'GreaseVision', from our Human Centred Computing group at the University of Oxford, is imagining what this might look like. GreaseVision is a technical system that allows end users to customise interfaces of their mobile apps to suit their needs, including elements connected with self-regulation challenges. For example, users can limit the amount of content on recommender feeds, change the colour of notification markers, add overlays to particular types of unwanted content, and so on.

Providing end users a 'right to tinker', via mandated provision of the equivalent of browser extensions for mobile apps, could be a highly impactful way to regulate.

Making non-exploitative business model more viable?

Regulators may also consider a more hands-off approach focused on business models. That is, instead of directly intervening in designs, regulators can make it more likely that apps which include design patterns that support self-regulation are viable on digital marketplaces.

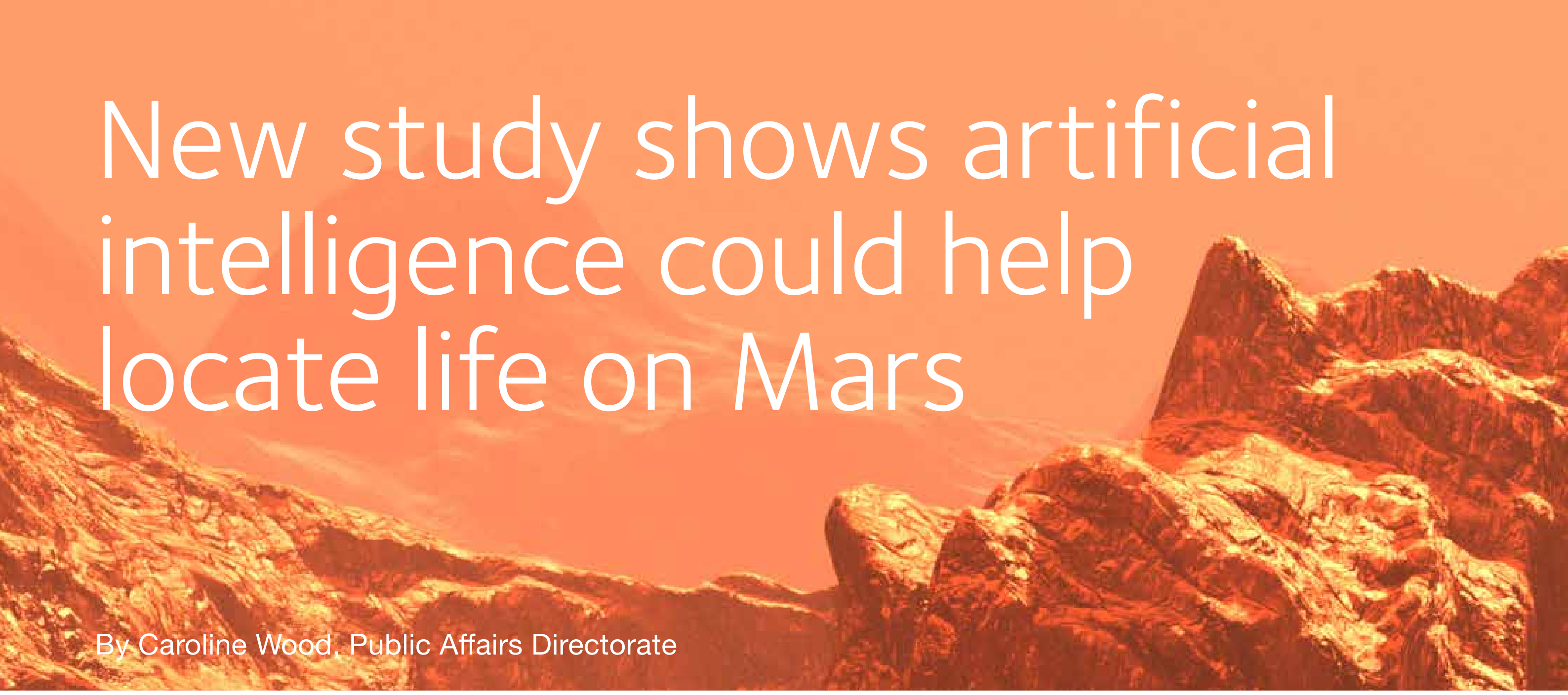
The underlying issue here is that many of the design patterns that cause self-regulation difficulties are a direct consequence of business models that rely on selling users' attention and data. Some have argued that digital platforms should be obliged to provide subscription options that do not rely on targeted advertisment. A gentler approach is to mandate disclosure of specific practices tied to business models. For example, Apple recently began to require that developers disclose on the App Store what user data apps collect. Such disclosures might make consumers more likely to pay for apps that do not rely on attention capture and targeted advertisement, analogous to how organic farming has been supported by labelling that makes consumers more willing to pay a premium.

A more radical approach is regulation that attempts to disrupt current dominant data infrastructures. Sir Tim Berners-Lee has proposed 're-decentralising the web' via interlinked 'Personal Online Data' (POD) stores kept securely on servers of the user's choice. People would decide which apps could access their data, and social networks like Facebook or Instagram be re-implemented as ways to interconnect individual PODs. Users' data would then not live in data stores owned by the companies, but in PODs under direct control of individual users who allow specific entities to read and interlink their data into specific services. A world where PODs were the basic nodes of the web might favour market competition around user interfaces and data interlinkages that effectively support self-regulation. However, the practical and regulatory obstacles involved in this transformation are likely to be substantial.

Conclusion

Digital devices like smartphones and laptops have revolutionised our world, but the instant connectivity they provide can lead to self-regulation failure that undermines productivity and wellbeing. This is especially true for individuals who are more attentionally vulnerable, for example children or adults with conditions like ADHD, making this an important concern for AI regulation.

On the web, browser extensions allow people to customise websites to their needs, and for example have YouTube without recommended videos. On mobile platforms, however, users have little control. To support self-regulation on mobile platforms, governmental regulation banning specific design patterns is likely to be useful only for obvious cases of manipulative dark patterns. Mandating greater user control over digital interfaces, however, could hold great potential to support a long tail of idiosyncratic user needs. This could take the form of a 'right to tinker' that provides users the same control over mobile interfaces as that which browser extensions enable on the web. Finally, regulators should require disclosure of data collection practices to make business models that are not based on attention capture more viable.



New study shows artificial intelligence could help locate life on Mars

By Caroline Wood, Public Affairs Directorate

A new study has found that artificial intelligence could accelerate the search for extra-terrestrial life by showing the most promising places to look. The findings have been published in *Nature Astronomy*.

In the search for life beyond Earth, researchers have few opportunities to collect samples from Mars or elsewhere. This makes it critical that these missions target locations that have the best chance of harbouring life. In this new study, researchers demonstrated that artificial intelligence (AI) and machine learning methods can support this by identifying hidden patterns within geological data that could indicate the presence of life.

Led by Dr Kimberley Warren-Rhodes at the SETI Institute and involving an international team of over 50 researchers, the first part of the study was an ecological survey of a 3 km^2 area in the Salar de Pajonales basin, at the boundary of the Chilean Atacama Desert and Altiplano in South America. This mapped the distribution of photosynthetic microorganisms and used techniques such as gene sequencing and infrared spectroscopy to reveal distinct markers of life, called 'biosignatures.' These data were then combined with aerial images captured by drones to train a machine learning model to predict which macro- and microhabitat types would be associated with biosignatures that could indicate life.

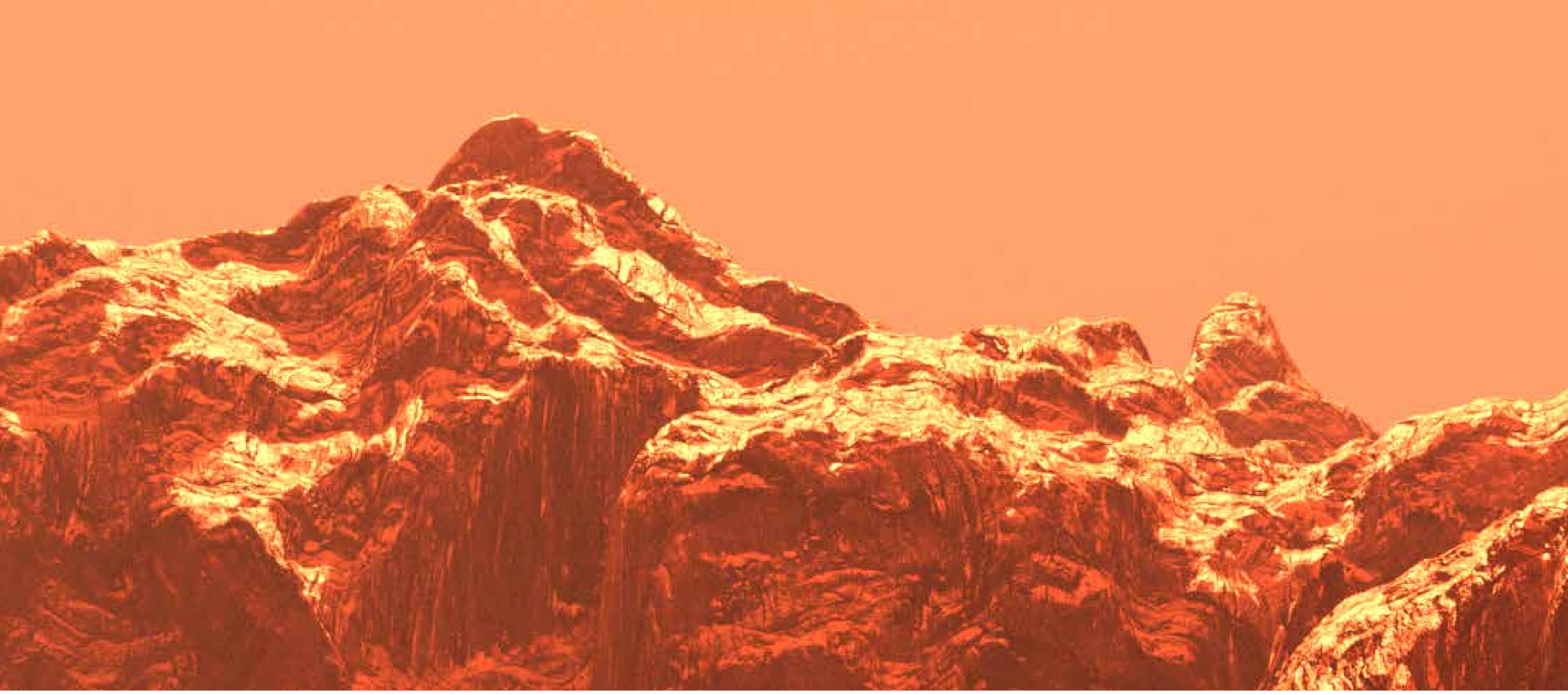
When tested using data on which it was not trained, the resulting model was capable of locating and detecting biosignatures up to 87.5% of the time (versus ≤10% by randomly searching). This decreased the search area required to find a positive result by up to 97%. Ultimately, similar models could be used to guide rovers exploring planets to the locations most likely to contain signs of life.

Freddie Kalaitzis, a Senior Researcher in the Department of Computer Science, led the application of machine learning methods to microhabitat data. He said: 'This work demonstrates an AI-guided protocol for searching for life on a Mars-like terrestrial analogue on Earth. This protocol is the first of its kind trained on actual field data, and its application can in principle generalise to other extreme life-harbouring environments. Our next steps will be to test this method further on Earth with the aim that it will eventually aid our exploration for biosignatures elsewhere in the solar system, such as Mars, Titan and Europa.'

'Our process combined statistical microbial ecology surveys, remote sensing from unmanned aerial vehicles, and machine learning to map, model, and predict the distribution of biosignatures in a Mars-relevant setting. The approach may also have applications for other astrobiology targets, such as the surface of Titan, the plumes of Enceladus, or the ice cover of Europa.' – Senior Research Fellow (Department of Computer Science) Freddie Kalaitzis

The Pajonales, a four-million-year-old lakebed, is one of the closest analogues to the Martian environment on Earth, and considered inhospitable to most forms of life. The high altitude (3,541 m) basin experiences exceptionally strong levels of ultraviolet radiation, hypersalinity and low temperatures – akin to the evaporitic basins of Mars.

continued on next page ▶



from previous page ▶

For the ecological survey, the researchers collected over 7,700 images and 1,150 samples, and used a variety of instruments to test for the presence of photosynthetic microbes living within the salt domes, rocks, and alabaster crystals that make up the basin's surface. Biosignature markers included carotenoid and chlorophyll pigments, which could be seen as orange-pink and green layers respectively.

The images recorded by the drones were combined with ground sampling data and 3D topographical mapping to classify regions into four macrohabitats (metre to kilometre scales) and six microhabitats (centimetre scale). Statistical analysis found that the microbial organisms across the study site were not distributed randomly, but clustered in distinct regions – despite the Pajonales having a near-uniform mineral composition. Follow-up experiments revealed that water availability is likely to be the crucial factor determining the position of these biological hotspots, rather than other environmental variables such as nutrient or light availability.

The combined dataset was used to train convolutional neural networks to predict which macro- and microhabitats were most strongly associated with biosignatures. 'For both the aerial images and ground-based centimetre-scale data, the model demonstrated high predictive capability for the presence of geological materials strongly likely to contain biosignatures' said Freddie. 'The results aligned well with ground-truth data, with the distribution of biosignatures being strongly associated with hydrological features.'

The research team now intend to test the model's ability to predict the location of similar yet different natural systems in the Pajonales basin, such as ancient stromatolite fossils. Going further, the model will be used to map other harsh ecosystems, including hot springs and permafrost soils. In time, the data from these studies will help inform and test hypotheses on the mechanisms that living organisms use to survive in extreme environments.

'Our study has once again demonstrated the power of machine learning methods to accelerate scientific discovery through its ability to analyse immense volumes of different data and identify patterns that would be indiscernible to a human being. Ultimately, we hope the approach will facilitate compilation of a databank of biosignature probability and habitability algorithms, roadmaps, and models that can serve as a guide for exploration on Mars.'

The paper 'Orbit-to-Ground Framework to Decode and Predict Biosignature Patterns in Terrestrial Analogues' has been published in Nature Astronomy.

The SETI Institute is a non-profit, multi-disciplinary research and education organization whose mission is to explore, understand, and explain the origin and nature of life in the universe and the evolution of intelligence.

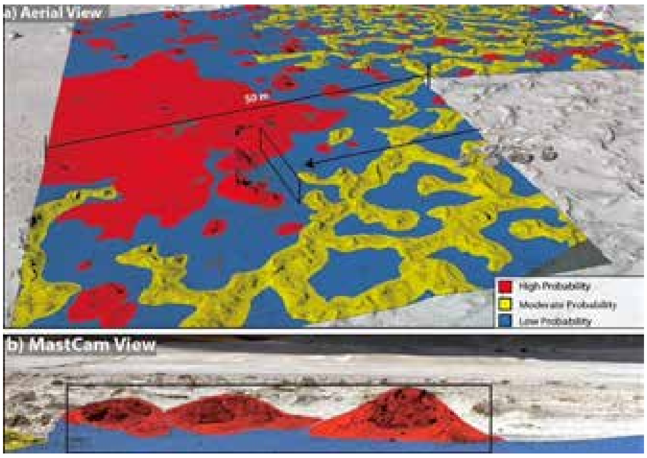
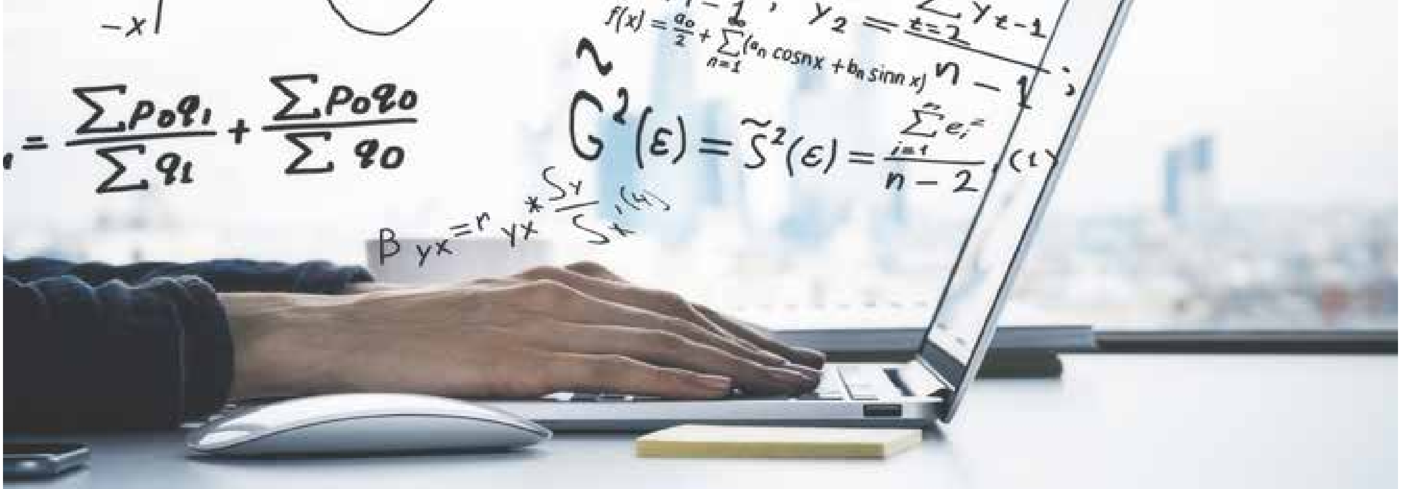


Figure 1: Aerial view (above) and ground view (below) from a rover of a biosignature probability map of the same area.
Image credit: M. Phillips, K. A. Warren-Rhodes & F. Kalaitzis.



ChatGPT: a mathematical assistant in the making?

By Doctoral Student Frieder Simont

The State of Affairs

The remarkable capabilities of ChatGPT and GPT-4 models gave new impetus to the goal of automating mathematics. Traditionally, the automated theorem-proving community had attempted to make math amenable to automation by formalizing it and searching for proofs. Unfortunately, this decades-old approach has not turned out to scale well to graduate-level mathematics and beyond. The rise of language models has led to new efforts to automate mathematics directly using natural language. The arrival of ChatGPT, in particular, ignited the Twittersphere, which abounds with anecdotal evidence of brilliant answers of ChatGPT – as well as with examples of math gone wrong.

Here is a particularly tricky one: What is the integral of 1/(2 + ln(3/2))? Tricky it is for ChatGPT indeed, as even the latest version from 3rd May 2023 starts its answer by claiming that 'the integral of 1/(2 + ln(3/2)) cannot be expressed in terms of elementary functions, so it is an example of a transcendental function', which -the author hopes- will amuse the readers. For even school children should find this integral to be anything but tricky; recognizing that there is no variable in this expression to integrate against means that arriving at the integral only amounts to adding a single variable, say "x" (and an arbitrary constant). ChatGPT's misstep is likely due to the rarity of such integrals in problem collections, and therefore in its training data.

GPT-4 on the other hand, whose performance on a large batch of standardized exams exceeds that of all previous versions of ChatGPT, manages to solve this integration task correctly! But this ray of optimism after ChatGPT's atrocious claim is immediately eclipsed by the slightly dubious reasoning GPT-4 provides, which we quote in full: "To find the integral of the function 1/(2 + ln(3/2)), we can first simplify the expression in the denominator. Let k = 2 + ln(3/2). Now the function becomes 1/k, where k is a constant. The integral of 1/k with respect to x is simply (1/k)x + C, where C is the constant of integration. So, the integral of 1/(2 + ln(3/2)) with respect to x is: (1/(2 + ln(3/2)))x + C".

There is still hope for mathematicians to continue their profession. Or is there? Examples such as these call for a rigorous evaluation of the mathematical abilities of ChatGPT.

A New Dataset for Natural-Language Mathematics

While there are several datasets for mathematical word problems, they lack the comprehensiveness of datasets in more established domains

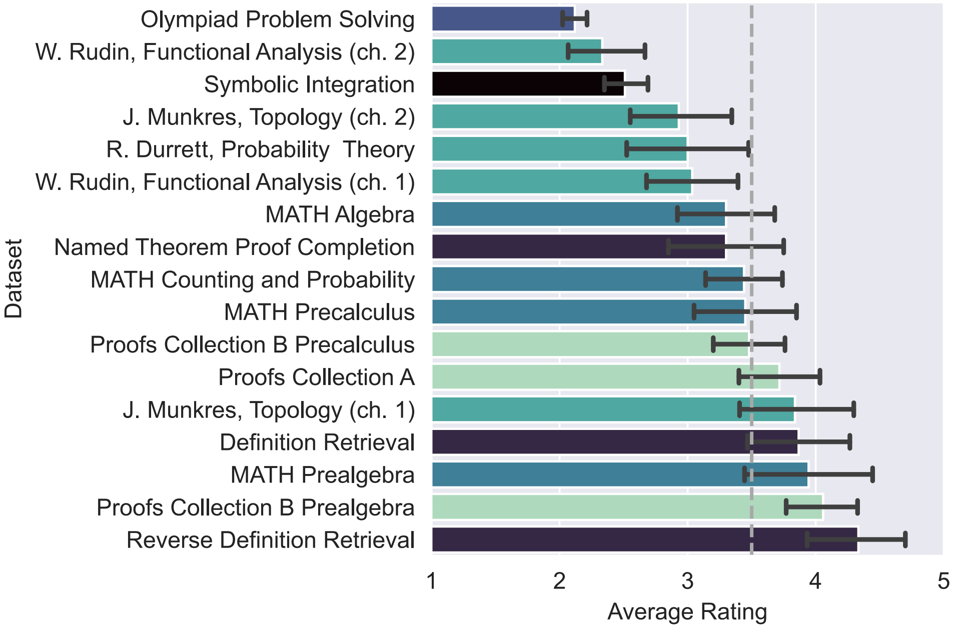


Figure 1: A description of various types of data included in the GHOSTS prompt dataset, plotted against the average rating that ChatGPT (version from 9 January 2023) achieved on that dataset. An average rating of 3.5 can be considered to be a reasonable passing grade. The vertical dotted line drawn at this rating indicates on which datasets ChatGPT would receive a passing grade.

continued on next page ▶

from previous page ▶

like computer vision or natural language processing. Their evaluation schemes rely typically on a simple correct-incorrect classification of the models' output. This is sufficient for common sense mathematical reasoning - but not enough to test advanced mathematical capabilities, which ChatGPT in some cases seemed to have.

In light of this, the author developed a new type of dataset and rating methodology and organized a team of researchers, comprised predominantly of mathematicians, to help and assist in measuring how well language models can do advanced mathematics. A preprint can be found at https://arxiv.org/abs/2301.13867.

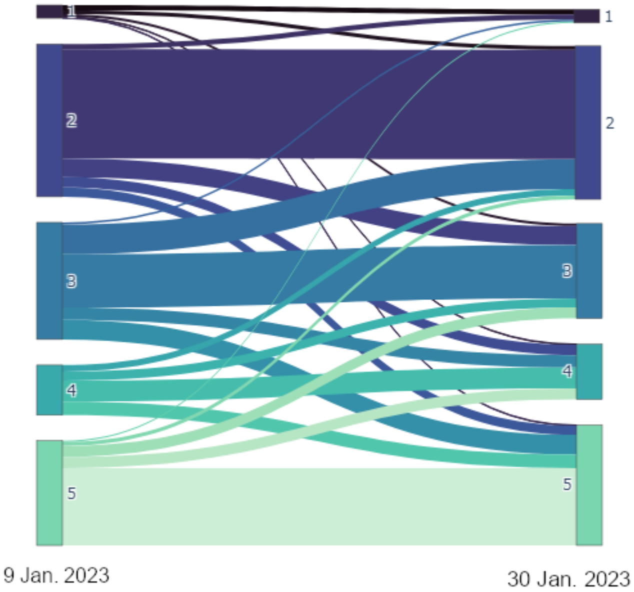


Figure 2: A Sankey diagram on the level of ratings (1 is the lowest, 5 is the best) that shows how the plots changed between the version of ChatGPT 9 January 2023 and the version from 30 January 2023, that according to the official release notes has "improved factuality and mathematical capabilities".

The new dataset, called preliminarily "GHOSTS", tests mathematical capabilities on more axes than any of the previous natural-language datasets: completing proofs, searching for theorems, and numerical reasoning abilities are just some aspects that are being tested. This is accompanied by a new methodology, that is more fine-grained than those accompanying previous datasets. To obtain a more nuanced understanding of the models' capabilities, we rated over 1500 responses on a scale of 1-5 and assigned various error codes and warning codes to flag certain undesirable reasoning behaviors of the model. These range from flagging whether the model got a bad rating because computations were wrong, proof steps were missing, or edge cases were ignored - to the models providing much more information than was asked for, or withholding names of well-known theorems. In a sense, the role between a compiler and a human has been reversed: A human is rating the output of a piece of software and providing error messages and warnings, for input that the software provides.

The 1500+ responses are split among different Mathematics Subject Classification (MSC) codes, levels of mathematical difficulty (ranging from elementary integration, as illustrated above, to questions similar to mathematical olympiads), levels of mathematical sophistication (we have included some graduate-level textbooks), whether the data was likely included in the data used to train the model or not - and many more categories.

This is just the first step. We are hard at work to add new datasets, that test further aspects of mathematical reasoning, to improve the methodology, and to evaluate further models in a more automated way. Because this undertaking is too much for a small set of researchers, we are working on letting the community submit datasets and opening up our dataset to let everyone contribute.

Results

ChatGPT works best as a mathematical search engine but falls short in proving theorems or performing numerical computations – though occasionally it does surprise! For the task of performing numerical computations, the Toolformer approach offers a solution. An instance of this is the collaboration between Wolfram Alpha and OpenAI to allow ChatGPT/GPT-4 to call directly Wolfram Alpha APIs, generate code, and integrate that answer in its output.

Searching for mathematical theorems is a non-trivial undertaking and the model manages to often correctly infer what is being sought (the author learned the name of 'vague topology' from ChatGPT). While on numerical or math word problem tasks ChatGPT didn't perform quite as well as other, older models that were specifically trained to solve these tasks, its performance was not abysmal either. See Figure 1 for a selection of domains from which our prompts were drawn, or self-devised and the corresponding rating of ChatGPT (the version from 9 January 2023).

Since updates to ChatGPT are released almost every month, one interesting thing to consider is how much better each model iteration gets. Comparing the same prompts on the 9 January 2023 version of ChatGPT to the 30 January version, which according to OpenAI's release notes had "improved factuality and mathematical capabilities", we see that reality is quite different: The Sankey diagram from Figure 2 shows that scores do not change in a consistent way between the versions, and largely remain similar.

While a more detailed discussion and further figures can be found at https://arxiv.org/abs/2301.13867, the overall conclusion is that there is promise in the language-model approach toward automating mathematics, as it is useable already today for selected use cases. There is research underway to unify the formal approach, championed by the automated theorem-proving community with the language model approach. From an educational perspective university teachers can rest assured that (for now) ChatGPT will not render them redundant or invalidate (sufficiently difficult) take-home assignments. Students are also not yet off the hook; if their goal is to use ChatGPT or GPT-4 to pass a university exam, it might just be safer to copy from their average peer!

Entangled Challenges: policy, publics, and UK plc

By Carolyn Ten Holter, Researcher in the Responsible Technology Institute

The rate of progress in quantum computing is accelerating, and as a result the commercial/ industrial sector is becoming more developed. With this commercial development, potential use-cases are becoming clearer, and thus preparation for the social impacts of quantum computing becomes increasingly important to delineate.

Within the Department of Computer Science, researchers have been involved with the UK's quantum computing programme since its inception, focused on the potential societal impacts of this game-changing technology. One of our previous projects on quantum computing, the NQT-RRI IAA project Questions of Responsibility, used qualitative research in academia, industry and policymaking to develop a roadmap for the creation of a dedicated hub for responsible innovation in quantum computing. Our team saw it as ever more crucial to deepen the conversation between the communities developing quantum computing, and wider society. Responsible innovation approaches, with their focus on inclusivity of diverse stakeholders in research, and anticipation of possible impacts, are viewed as one way to open up such a dialogue. Our new project in this area, Responsible Quantum Computing Communications (ResQCCom) seeks to engage with companies, policymakers, and citizens to discuss concerns and think collaboratively about the future. This process of unpacking and investigating the challenges and impacts – as well as potential unintended consequences – through an industry, policy, and social dialogue, can enable the forward-thinking aspects of responsible innovation to be turned into concrete action. Such an approach can help to ensure that society experiences not only the benefits of advanced quantum computing research, but that principles of good governance, transparency, and other aspects of responsible development are translated from the research environment into the commercial sector.

ResQCCom is focused on three key challenges:

1. Governance in technology transfer: at present, governance structures or 'responsible' approaches are not translated from the university research environment into the commercial world. Public confidence in quantum computing requires good governance, and citizens rightly expect that fledgling companies spinning out of university labs will be required to comply with industry standards, regulation etc – however such standards and regulation do not yet exist in quantum computing, placing a heavy burden on founders to ensure their work considers societal impacts. ResQCCom will create a toolkit for such companies to support business development.

2. Communication in policymaking: Policy around quantum computing is often bound up with perceptions of quantum technologies as matters of national strategic capability. There are concerns in the international quantum computing community that attempting to align development along 'national' lines is premature, unnecessary, and likely to hinder overall progress. At the same time, there may be insufficient recognition that quantum computing is an 'enabling' technology that should be of interest to multiple areas of government. ResQCCom will help to facilitate dialogue between different policymakers to ensure a cohesive approach.

3. Public engagement: Although the 2017 Public Dialogue in Quantum Computing was extremely valuable, it has not been followed up, despite the findings that the public were interested in quantum computing and invested in its development. Societal attitudes to, and familiarity with, quantum computing have therefore not been examined at scale for over five years, during which time quantum computing has developed far faster than anticipated. ResQCCom will conduct research into current public knowledge and attitudes, that can inform policy and research. ResQCCom will create publicly available, comprehensible explanations of the possible impacts of quantum computing including data visualisations, infographics, and an animated film for non-specialist audiences. ResQCCom will also work to connect up international researchers focused on the societal impact of quantum computing in order to share best practice, argue collectively for global RI approaches, and exchange developments in their respective countries.

Empowering children in the age of datafication: building a humane and autonomy-supportive future

AI is transforming the way children interact with the digital world, opening up exciting opportunities for their learning, playing and social networking. However, these experiences can come at a cost. Datafication – the practice of collecting, tracking, aggregating, analysing and profiling children's online data and digital footprints, raises increased concerns about children's data privacy, autonomy, and the potential of being subject to information manipulation. We have little understanding of how young children in the UK are affected by the practice of datafication and how much they are aware of the implications.

Through co-design workshops with 53 children, aged 7 to 14, Oxford Computer Science researchers (http://oxfordccai.org) explored how children recognise datafication practices around them and how they wish to be supported in coping with this. The findings provide crucial and timely insights for creating age-appropriate AI systems that nurture children's algorithmic literacy and advocate for a more humane and autonomy-supportive future.

Understanding Datafication and Children's Concerns:

The implications of datafication practices extend beyond data collection – they affect the online content children see, the promotion they are subject to, their digital autonomy, as well as their overall well-being. Such manipulations are often underpinned by intentional behavioural engineering, to prolong children's online engagement or influence their online social experiences and political opinions.

Our research has shown that children are not passive bystanders in this process. They genuinely care about their digital autonomy and are keen to seek for opportunities to take action and make changes. By actively involving children in co-design sessions over several months in late 2022, we gained first-hand insights into how young minds perceive and desire support regarding datafication.

Creating Age-Appropriate AI systems:

The study's findings underline the importance of tailored approaches to support children's coping with datafication. There is no one-size-fits-all solution; rather, individualised and age-appropriate strategies are critical. To begin with, children's digital literacy is a critical factor in empowering them to navigate the data-driven world effectively. To achieve this, national initiatives to foster children's digital literacy education and awareness must be developed, focusing on teaching children about their digital rights, privacy implications, data-based manipulations online and the potential risks and benefits of data sharing.

In addition to education, the study revealed that children desire technological interventions that prioritise their autonomy and well-being. They envision a future where the data ecosystem shifts towards a more 'humane-by-design' approach, respecting their rights and values, permitting for greater control over their personal information and greater transparency. This includes transparent consent processes, user-friendly privacy settings, and platforms that prioritise meaningful engagement over manipulative practices.

The Road Ahead - Towards a Humane AI Future:

As we move forward, it is vital to consider children's perspectives and involve them in shaping the digital landscape they inhabit. Academic researchers, industry leaders, policymakers, and educators must collaborate to create an environment that respects children's agency, protects their privacy and rights, and fosters their digital well-being.

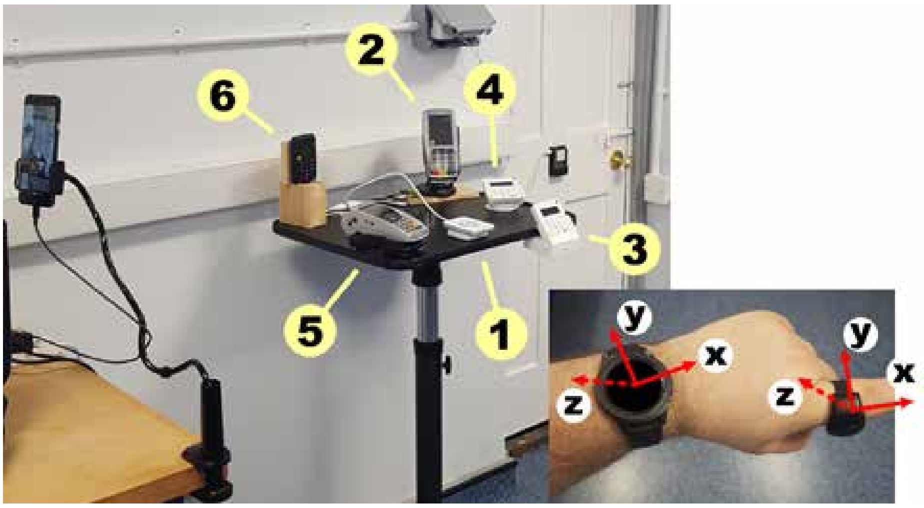
Building age-appropriate AI systems and tools requires interdisciplinary efforts that integrate insights from developmental psychology, computer science, learning sciences, and human-centered design. We must invest in research that explores children's evolving needs and aspirations regarding datafication. Only by doing so, can we develop interventions and policies that align with their expectations, ensuring that their rights are protected and their voices heard.

References

- *'Treat me as your friend, not a number in your database': Co-designing with Children to Cope with Datafication Online. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3544548.3580933*
- *'Don't make assumptions about me!': Understanding Children's Perception of Datafication Online. Proceedings of the ACM on Human-Computer Interaction. Volume 6. Issue CSCW2. Article No.: 419. pp 1–24. https://doi.org/10.1145/3555144*

Using Inertial Sensors to Authenticate Users

By Doctoral Students Jack Sturgess, Sebastian Köhler, Simon Birnbach and Professor Ivan Martinovic



Passwords are still the primary mechanism by which users authenticate themselves to digital services. Passwords must be long and random to resist guessing attacks and they must be often changed and not reused to resist dictionary attacks, all of which makes them burdensome to use effectively at scale. Biometrics, such as fingerprints or face geometry, provide a promising alternative. The initial barriers to adoption that biometric systems faced, such as high error rates and deployment costs, have evaporated in recent years as smartphone-based sensors have become readily available. But there are risks in using biometric authentication, as these characteristics can be captured by an attacker and used in impersonation—and, unlike passwords, they cannot be revoked or changed once compromised. To address this, behavioural biometrics have started to attract interest. Instead of measuring a physical characteristic, behavioural biometrics measure patterns of movement over time, such as gait or typing dynamics, and therefore can be collected unobtrusively without any effort on the part of the user and are more difficult to capture and impersonate. These systems were regarded as impractical due to the need for continuous measurement, but the widespread use of wearable devices (and therefore continuously worn sensors) has opened up new opportunities for implementation.

In the last issue of *Inspired Research*, we presented some of our recent work in mobile payment authentication in which we showed that smartwatch-users can be authenticated to the system using only the inertial sensors in the smartwatch. The movements of the arm and wrist that are performed by the user to make a payment, collectively called a tap gesture, are sufficiently unique to each user as to be capable of verifying identity. Our user study (n=31) also included an active attacker component, in which each participant watched video footage of other users making payments and attempted to impersonate them, and we showed that the system was able to identify and reject such attacks. We also showed that the tap gesture is sufficiently distinct from other arm and wrist movements made throughout the day such that we can recognise when a payment is made intentionally—as opposed to a payment that is accidental or initiated maliciously by a skimming attack, where an attacker rubs a rogue terminal against the payment device to trigger an unwanted payment response. The use of inertial sensors for user authentication and intent recognition come at no cost to the user (in terms of effort or delay) and can help to reduce payment fraud.

The miniaturisation of hardware has enabled the development of smart rings. Examples on the commercial market include Amazon Echo Loop, which uses a microphone and speaker to enable the user to interact with the Alexa virtual assistant and to make phonecalls through a paired smartphone and Blinq, which uses inertial sensors for fitness monitoring and provides a discreet panic button that triggers an application on a paired smartphone to send a help message containing geo-location information to contacts or on social media also Genki Wave, which uses inertial sensors to enable a musician to adjust the digital settings of an instrument by raising or lowering the finger. Smart keys (battery-less, NFC-enabled tokens that are powered via NFC when they are close to a terminal, such as contactless payment cards and key fobs) are also being cast in ring format and marketed as feature-poor smart rings. Examples of these include Tesla Ring, which can unlock and activate a vehicle, and Mastercard K-ring, which can make payments. As smart ring technology evolves and ring-based services start to require greater confidence in the identity of the user, they will require authentication capabilities.

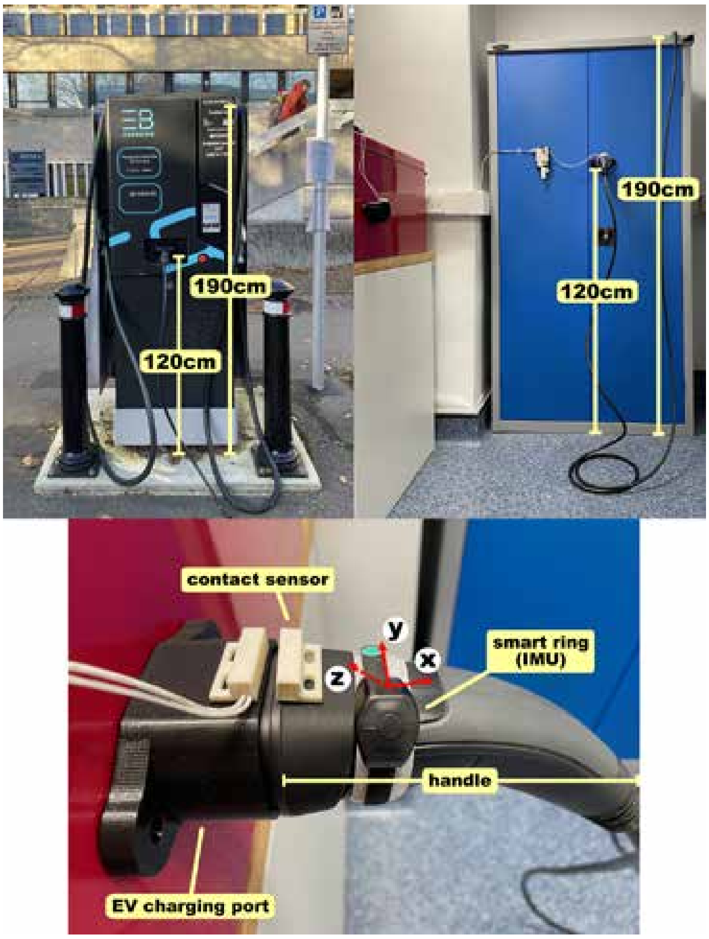
The tiny form factor of a smart ring restricts its input capabilities, which would make password- or PIN-based authentication user-unfriendly. To address this, we investigated the use of inertial sensors for ring-based systems. We conducted a user study (n=21) in which each participant wore a smart ring with inertial sensors embedded into it and performed some basic tasks while we collected their motion data. Firstly, we presented an array of six payment terminals and had each participant make payments using the smart

continued on next page ▶

from previous page ▶

ring by tapping it against the terminals. As with the smartwatch experiment, we also included an active attacker component and allowed participants to watch footage of and impersonate each other. We found that the ring-based sensors achieved equal error rates of 6% and were just as effective as the watch-based sensors at authenticating payments. Secondly, we had each participant knock on a closed door while wearing the smart ring to investigate the feasibility of using knock gestures for access control. Each participant knocked in sets of three, sets of five, and then in a 'secret knock' pattern of their choosing. To our surprise, we found that we were able to train a classifier that can distinguish between users based on their knocking data and that the knock gesture was suitable for use in user authentication. Throughout the study, we also had each participant wear a smartwatch on the same arm as the smart ring and simultaneously collected inertial sensor data from that as well. We found that the ring data could be used to authenticate tap and knock gestures made with the watch and vice versa, meaning that inertial sensors on either device could be used as a second factor to support the other. For the knocking experiment, we also attached inertial sensors to the door itself to see if a door-based system that would automatically grant access based on a knock was feasible—alas, the results for this were less promising.

Electric vehicles (EV) are gradually becoming more common. To recharge its battery, the user must park the vehicle next to a charging station, connect a charging cable, and make a payment to initiate the charge. A number of payment systems have been deployed by different station operators and include standard tap-and-pay systems, smart keys, and QR codes printed on the station that link to a payment system when scanned by a smartphone. EV batteries charge more slowly and must be recharged more often than the refuelling of vehicles with a combustion engine, so user-friendliness and convenience must be considered. As such, zero-interaction payment schemes, such as AutoCharge and Plug & Charge, are most desirable. These systems work by sending information from the vehicle to the charging station via the charging cable once it is connected; this information is linked to a payment instruction, typically administered by an application on a paired smartphone, to facilitate automatic billing. These schemes treat the vehicle as a token and authenticate the vehicle rather than the user. This could allow a thief to charge a stolen vehicle at the owner's expense or an attacker to capture information from a victim's vehicle and to inject it into the communication channel to charge a different vehicle at the victim's expense. To address this, the system requires some form of user authentication that does not inconvenience the user. We replicated a charging station and a Volkswagen ID.3 (a typical EV model) in our lab by 3D-printing some enclosures and fixing them in position to match the charging ports. We obtained an authentic charging cable and affixed inertial sensors to the handle and contact sensors on top so that we could timestamp when it was unhooked from or plugged into an enclosure. We conducted a user study (n=20) in which each participant unhooked the cable from its position in the charging station and plugged it in to the EV charging port and we collected the motion data. We segmented user gestures from around each timestamp, so as to disregard any variable travel time between the two actions, which would depend on how far away from the charging station the user had parked the EV, and we found that we were again able to authenticate each user using only inertial sensors. By tuning our model to favour security, we were able to add a layer of protection that rejected 82% of attacks without the user having to do anything. By tuning our model to favour usability, we were able to reduce the number of unnecessary authentication requests made by the application on the paired smartphone by 41% at no cost to security.



All of our user studies were approved by the department's research ethics committee. For more information, see our published papers entitled *WatchAuth*, *RingAuth*, and *CableAuth*.

Cohere: at the cutting edge of emerging AI

What is Cohere?

Cohere is an AI platform for the enterprise, allowing businesses everywhere to access cutting-edge large language models (LLMs) through a managed API. Founded in 2019, it has offices in Toronto, San Francisco, and London, with team members all over the globe. CEO and Co-founder Aidan Gomez is a graduate of our department, AI Leader Phil Blunsom is a current department academic and Ed Grefenstette* (another AI leader at Cohere) was also previously a member.

Aidan, who is also a co-author of the ground-breaking AI paper 'Attention is all you need' that introduced the Transformer architecture, says of his time at Oxford, 'I was surrounded by such a brilliant, diverse group of minds that I found myself in a constant state of inspiration – it felt like I learned something new every day. It was an incredibly impactful experience, both personally and academically, and one that played a tremendous role in shaping me into who I am today.'

Ed comments, 'I fondly remember my graduate years at Oxford's Department of Computer Science as the formative years in my career as a researcher. I was empowered to explore new topics, and formed friendships and collaborations which persist to this day.'

Cohere is just one company in a Generative AI revolution that has seen unprecedented consumer adoption, with OpenAI's ChatGPT registering 100 million active users within two months of its launch in November 2022.

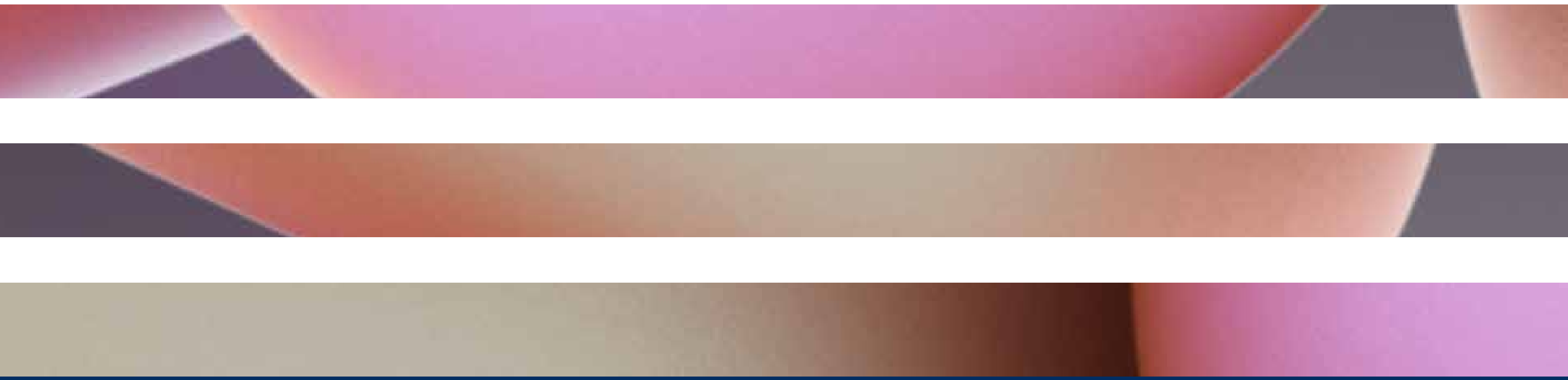
However, Cohere's approach has been quite different to that of other young companies in this area of development. Cohere's founders set out to create a world where any enterprise – even those without vast compute resources or computer science expertise – can access the transformative power of LLMs. Cohere takes on the massive upfront costs of building and training world-class foundational LLMs, and removes the need for enterprises to invest in the rare knowledge and resource-intensive development of their own.

This dramatically reduces the barriers to adoption, and allows enterprises of any kind to enhance their products and services with LLMs and drive business value. And because Cohere is designed specifically for the enterprise, its models are cloud-agnostic and can be deployed inside a customers' existing cloud environment, virtual private cloud (VPC), or on-site – able to adhere to the highest privacy and data security requirements.

Cohere has built a powerful, easy-to-deploy collection of APIs and tools designed for businesses that want to build products and experiences that read, understand, and respond with language. The company's founders believe that broadening access to LLMs will reduce the barriers to developing powerful product experiences rooted in language, and shape the future of how we interact with technology.

Geoffrey Hinton, Emeritus Professor of Computer Science at the University of Toronto (and previously

continued on next page ►



from previous page ▶

Engineering Fellow at Google) said of the company ‘Very large language models are now giving computers a much better understanding of human communication. The team at Cohere is building technology that will make this revolution in natural language understanding much more widely available.’

Research

Cohere For AI (C4AI) is a non-profit research lab that seeks to solve complex machine learning problems. The vision for the lab was to change how, where, and by whom research is done. Since that launch, the C4AI team has been hard at work publishing and promoting fundamental Machine Learning research, growing an inclusive and open research community, and creating programs designed to provide more entry points into ML research.

Key to the C4AI manifesto is the idea that ‘The best minds transcend borders and that discoveries are often made off the beaten path. We are committed to making meaningful progress in machine learning research through open collaboration. We believe that technology is powerful, and empowering different perspectives ensures responsible innovation. We see contributions to traditional conferences and publications in journals as an important part of our work, but also support efforts that go “beyond the research paper” and encourage scientific communication through different mediums.’

Scholars Program

The Cohere For AI (C4AI) Scholars Program provides the opportunity to work alongside C4AI Computer Scientists in an open, supportive environment that provides an alternative point of entry into NLP research. There is no insistence on prior degrees or formal experience working in a research lab. Instead, C4AI is attempting to identify emerging talent around the world.

Participants of the programme join a dedicated team of passionate researchers and industry experts and are paired with a project proposal. Participation is full-time, remote-first and paid. As part of the program, Scholars have access to a large-scale experimental framework, and help advance the company’s commitment to supporting responsible, fundamental research on machine learning topics while prioritizing good stewardship of open source scientific practices.

In this very fast moving area of AI development, Cohere has forged a clear identity, and an ethos that makes it stand out as innovative and ethical. This has enabled the company to attract huge investment. By summer 2023 its approximate value was over $2 billion. It’s exciting to see current and former members of our department involved in an enterprise at the forefront of the generative AI revolution.

*Since this article was written Ed Grefenstette has left Cohere to take up a new post as Director of Research at Google Deepmind.

What is the role of data recorders in the reconstruction of a psychological accident?

By Senior Researcher Pericle Salvini

Professor Marina Jirotka's Established Career Fellowship project, RoboTIPS (funded by EPSRC), is designing and testing an innovative design feature for social robots, the Ethical Black Box (EBB). This is a data recorder for robots, planned to function in a similar way to an aircraft's flight data recorder. We call it 'ethical' because we believe that in the future it will be considered unethical not to have one. In fact, we argue that these devices should be incorporated into all social robots – for example, robots capable of autonomous decision-making and interaction with human beings. The EBB collects data about a robot's actions in real time and in context - this can be drawn on following adverse incidents and accidents to understand how the event occurred and how it might be prevented from reccurring.

In order to evaluate the EBB's role in the reconstruction of accidents, we are devising a series of "mock accident" investigations with different robots in different domains. These simulations of accidents involve real robots and people, and simulate something going wrong - some sort of robot misbehaviour. We then hold an investigation with witnesses and investigators to reconstruct what happened. The investigators produce findings and recommendations just as a real investigatory board or inquiry might. In particular, we aim at answering the following research questions: (1) What does an investigation process for robot-related accidents look like? (2) What role does/could black box data from the robot play in this process? And (3) To what extent do human-robot interactions need to be logged, and how, in order to satisfactorily inform the accident investigation process? From these mock accident investigations, we will consider what new notions of responsibility might emerge through novel configurations of technology, society and governance, and develop models for how responsibility may be effectively distributed between people, institutions and computational agents.

Most recently we ran a pilot of our second mock accident investigation, which concerns a robot pet and its young owner. The goal of the pilot was to test the investigation procedure in preparation for the full event. The 'witnesses' parts were enacted by volunteers playing characters in a scenario written by the project researchers, drawing from expert input and related literature. Two investigators then took evidence and questioned the witnesses. The story is about the emotional attachment between a young girl and her robot pet, which becomes addictive and toxic due to a software virus that changes the robot's behaviour and turns it into an anthropomorphic and demanding artificial being. Its owner becomes more withdrawn and starts to exhibit challenging behaviours. The investigators were given the task of uncovering the sequence of events by questioning the witnesses and, alongside this, examining the robot's data logs – for example, the EBB. In the end, the investigators correctly deduced the role that the robot played in the accident, providing helpful feedback on how the scenario can be improved.

From a scientific perspective, this accident scenario was very challenging because it dealt with psychological safety in human-robot interactions. While it is relatively easy for a data recorder to provide data useful to reconstruct a physical accident, for instance data related to a collision or the malfunction of a robot component, it is still unknown whether and how data recorders could be useful in the reconstruction of accidents involving a person's cognitive, social, and emotional well-being. Considering the increasing presence of social robots in the market, in particular for therapy, education, and care/companionship, we believe it is highly relevant to find ways to improve psychological safety in human-robot interaction, and to provide ways in which to investigate incidents.

Our next step will be to run the official event by inviting expert, professional investigators to carry out inquiries into the incident. This will add to our growing body of evidence on how the EBB can be used in practice.

More details about the RoboTIPS project and research team are available at:https://bit.ly/3XQ3g3m