InspiredResearch

Winter 2020 Issue 17

News from the Department of Computer Science, University of Oxford

Lion's roar

Scientists discover the unique signature of a lion's roar using machine learning

Read more – p20

Police surveillance of Black Lives Matter:
The danger technology can pose to democracy – p12

From research to societal impact:
The work of the Oxford Research Software Engineering (OxRSE) – p14

CUBBITT – taking translation to a new level:
Advances in automatic machine translation – 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

Contributors

Joe Atherton, Emma Dunlop, Mihaela Duta, Catherine Eu Tong Jamie Frost, Marina Jirotka, Artem Kaznatcheev, Konrad Kollnig, Ivan Martinovic, Jason Nurse, Elisa Passini, Alex Rogers, Oliver Sampson, Anjuli Shere, Jakub Tomek, Ge Wang, Helena Webb, Standa Zivny

Cover picture by Charles J Sharp

Photographs in the newsletter are used for illustrative purposes and may have been taken before Covid-19 restrictions came into force

CONTENTS

Letter from the Head of Department

As the world went into lockdown in March this year, we all had to learn a new way of working. Tools like Zoom and Teams, which were perhaps rather disdained in some quarters before the pandemic, suddenly became an indispensable part of the global workplace. Without them, and without the sophisticated IT infrastructure that makes them possible (including of course the World-Wide Web, invented by our own Sir Tim Berners-Lee), the economic and social consequences of the pandemic would have been dramatically worse. Of course, you know all this, so why am I telling you? Because it is important to emphasise and re-emphasise just how crucial computing technology has been to getting us through the pandemic. Video conferencing tools like Zoom are based on a whole range of extremely sophisticated Computer Science techniques (image processing, real-time video compression, networking and streaming algorithms), which were the result of decades of research – they didn't come about by accident. All of which goes to show that investment in core computing techniques, in all these areas and many more, is money well-spent. The last year would have been unthinkable without the contributions of computer scientists.

When the lockdown began in March, I have to confess to thanking my lucky stars I was head of a Computer Science Department, rather than an experimental science department like physics or chemistry. Most researchers in our department require nothing more than a laptop and a whiteboard to do their research. (They already had laptops; we bought them whiteboards.) In this sense, although we all missed the face-to-face social interaction that is such a pleasure in our department, the actual transition to home working was relatively straightforward. I personally found that since I could no longer just drop into my students' offices, but had to book meetings in advance, I became more rigorous about meeting with them – a lesson I'll try to remember when we come through the other side. (You'll have to ask them whether they appreciate the extra attention, though.)

So, our research has continued apace, and while the experience is different, the results are every bit as impressive, as you will see in this edition of *Inspired Research*. And it is fascinating, as always, to see the extraordinary breadth of research that we undertake: from theoretical studies of algorithms for seemingly abstract mathematical problems like graph colouring, through to understanding how computer technology was used during the Black Lives Matter protests. I am also very excited about the DeepMind Professorship of AI – a new chair made possible by a philanthropic donation from DeepMind, the latest in their continued support of our department, for which we are very grateful. The chair is currently being advertised, and I hope to report on a successful appointment next spring.

Finally, you will surely be aware that Oxford has been at the centre of the response to the pandemic, and at the heart of this response has been the search for a vaccine. At the end of November Professor Andrew Pollard and his team were able to announce the efficacy of the Oxford vaccine, developed in collaboration with AstraZeneca. While this activity was obviously not led by Computer Science, advanced computing techniques were absolutely essential to its speedy development. It would have been impossible without them. Just as computing has proved to be essential to life in the age of the pandemic, so it has been essential in developing the vaccine that we all hope will release us from the pandemic..

Professor Michael Wooldridge
Head of Department of Computer Science
December 2020

How we prepared for returning to on-site in Michaelmas Term

Following the end of the initial Lockdown, it took hard work from a dedicated team of people led by Joe Atherton to prepare our buildings for the return of students and staff in Michaelmas term.

Preparing all Computer Science buildings for re-entry during the Covid-19 pandemic focussed on:

- Social distancing: Ensuring all signage relating to appropriate social distancing is visible to all staff and students in all areas.
- Office space: Ensuring all office desks were at least 2 metres apart. Where this was not possible we worked closely with HR and IT to create a re-entry rota system
- Teaching Spaces: all teaching spaces were measured and new seating layouts implemented. This has of course led to a significant reduction in numbers in certain offices and teaching rooms.

The team has installed over 70 sanitiser units in the department buildings. Regular cleaning schedules and detailed task lists for common areas help facilities teams focus on high-traffic areas and key touch-points.

Building access has been restricted so that only one entry point is available at the main reception door. All other doors are exit only. The door entry system has been adjusted to prevent entry outside of normal working hours. Joe's team also supported the induction of all staff and students to make sure they are aware of the procedures in place.

NEWS

News in brief

Professor Ian Horrocks has won a 'Test-of-Time' award for his 1998 paper 'Using an Expressive Description Logic: FaCT or Fiction?' The award was made at KR2020, the 17th International Conference on Principles of Knowledge Representation and Reasoning.

Professor Marta Kwiatkowska has been invited to join the Responsible AI Working Group of the Global Partnership on Artificial Intelligence (GPAI) initiative. Read more here: bit.ly/2I9FXyZ

Professor Georg Gottlob received an honorary doctorate from the University of Vienna. More information & video stream of award ceremony: bit.ly/34AHQML

Marcin Wrochna recipient of the 2020 Open Mind Prize in Combinatorics

Marcin Wrochna, a postdoctoral researcher at the Department of Computer Science, working with supervisor Professor Standa Zivny is the recipient of the 2020 Open Mind Prize. The prize is awarded biennially during the Polish Combinatorial Conference to a junior Polish researcher for outstanding research in combinatorics.

Marcin's research interests include parameterised algorithms, graph theory, and computational complexity. Read more: bit.ly/3jCHpHU

Wellcome Trust Award – improving access to software engineering skills and training for researchers

A recent national survey by the UK's Software Sustainability Institute found that 96% of biomedical researchers across the country use research software, and 69% believe their research would be impractical without it, results that were mirrored in a recent (2018) survey that we conducted at the University of Oxford. Despite this heavy reliance on research software, only a small fraction of researchers writing software have any formal software engineering training. To address these issues over half of the Russell Group universities have established Research Software Engineering (RSE) groups. These groups comprise professional Research Software Engineers who support the software development and training needs of their university on an as-needed basis. Through funding from the MPLS EPSRC Impact Acceleration Account (IAA), an Oxford RSE Group (OxRSE) was set up in early 2019. OxRSE is based in the Department of Computer Science and provides free research software support and training to projects that fall within the EPSRC remit.

Alumnus Jamie Frost wins international teaching award

Jamie Frost, an alumnus of the Department of Computer Science, has been recognised with an international teaching award. Jamie was a finalist in the Global Teaching Prize, (which was won by Indian teacher Ranjitsinh Disale), but he was instead awarded a special Covid Hero prize. Jamie has been recognised for the free maths tuition website he set up, DrFrostMaths.com. The website became an invaluable resource for teachers and parents while children were unable to attend school during the Covid-19 lockdown. Read more about Jamie in the Alumni Profile on page 11.

Innovate UK Soteria Project

The Soteria project is a cyber-security demonstrator for the e-commerce industrial market.

This project is led by THG Holdings plc (THG) and involves the universities of Oxford and Manchester. THG is a global online retailer and technology business with digital security at the heart of its operations. The University of Manchester will provide technical expertise on managed runtimes, smart network technologies, and business optimisation; and the University of Oxford will contribute expertise in formal verification. The research is aimed at reducing the potential impact of security breaches and attacks, and the costs required to secure digital businesses and services. The societal impacts are improved public perception of the benefits of technology, and less disruption to the daily activities of individuals due to security incidents, and service outages across public, government, and paid for digital services. Demonstrating the enhanced security capabilities and development of related products and technology strengthens the global and, especially, UK impact to online business, the public and government services and objectives. Finally, the proposed ingredients could constitute the basis for developing a software ecosystem capable of protecting against future cyber-attacks.

Professor Tom Melham from the Department of Computer Science comments, 'Soteria will demonstrate innovative extensions to the Morello capability architecture to toughen security protection for e-commerce and protect businesses. Oxford is delighted to being its deep experience and expertise in formal verification to this unique project, both to advance our research and to improve security assurance for the technology.'

Artem Kaznatcheev wins James S. McDonnell Foundation fellowship

Artem Kaznatcheev has won a prestigious James S. McDonnell Foundation fellowship in Dynamic & Multi-scale systems. It is a US $200k 'go-anywhere' postdoctoral fellowship.

Oxford University did particularly well in this round of fellowships, with Oxford students winning 3 of the 9 available awards. The other two awards went to DPhil students in the Maths and Biology departments.

Artem will be leaving us to start the fellowship in his chosen location, the Department of Biology at the University of Pennsylvania.

MSc Awards

The Hoare Prize for the best project submitted as part of the MSc in Computer Science has been awarded jointly to Raven Beutner and Maxence Draguent.

Raven's project 'On Termination of Higher-Order Recursive Programs with Continuous Distributions' was supervised by Professor Luke Ong.

Maxence's project 'Machine Learning in Particle Physics: Quark versus Gluon Jet' (joint with Physics) was supervised by Professor Alan Barr and Anisoara Calinescu.

The Hoare Prize for the best overall performance in the MSc in Computer Science 2020 has been awarded jointly to Raven Beutner and Julius Hense.

Professor Sadie Creese, the Chair of Examiners said, 'Congratulations and well done, this is a significant achievement which recognises your hard work. I wish you every success in your future careers.'

BCS Lovelace Medal 2020: awarded to Professors Ian Horrocks & Michael Wooldridge

The 2020 Lovelace Medal is to be shared between Ian Horrocks (Professor of Computer Science, University of Oxford), Nicholas Jennings (Professor of Artificial Intelligence and Vice-Provost for Research and Enterprise, Imperial College London) and Michael Wooldridge (Head of Department and Professor of Computer Science, University of Oxford, and Programme Director for AI at The Alan Turing Institute).

The award is presented annually to individuals who, in the opinion of BCS Academy Awards Committee, have made a significant contribution to the advancement of Information Systems.

Ian changed the whole approach to logic-based knowledge representation, demonstrating that a combination of sound theoretical foundations and insightful software engineering enabled the development of correct and scalable reasoning systems. (bit.ly/2UmfalE) Nicholas and Michael have been recognised for their major contribution as founders of the multi-agent systems (MAS) field. Their work on cooperating, coordinating and negotiating agents has been instrumental in transforming the nascent area from one studied by tens of people, to the global community it is today. (bit.ly/3lukIMi)

The Lovelace medal will be awarded to Ian, Nicholas and Michael at a ceremony at the Royal Society in London in autumn 2021.

The recipient of the Lovelace Medal is decided each year by a distinguished panel of academics, chaired in 2020 by Professor Anthony G Cohn.

Opening and closing dates for nominations for the 2021 award will be announced shortly.

Leverhulme Visiting Professorship

An award from the Leverhulme Trust is facilitating a 12-month visiting professorship to the Department by Henryk Michalewski. Henryk is currently a professor in the Department of Mathematics, Informatics and Mechanics at the University of Warsaw, Poland and has world-leading expertise in theorem proving and machine learning for game playing.

He has made important contributions to the theory of games and automata, relevant to both pure mathematics and Computer Science. He has also developed some of the highest performing game-playing systems.

Henryk will be hosted in the department by Professor Michael Benedikt. Henryk will present four Levenhulme lectures, will run a graduate seminar on the subject of machine learning for game playing and automated theorem-proving, as well as collaborating on research projects.

Whilst in the UK, Henryk plans to make short visits to several other universities and research institutes, including the Alan Turing Institute, both to initiate and to deepen collaborations with UK researchers who have used and contributed to machine learning and theorem proving.

News in brief

Professor Marina Jirotka was highly commended in the Policy Engagement category of the University of Oxford's Vice Chancellor's Awards 2020 for 'Putting out Digital Wildfires Before They Take Hold'. See: bit.ly/34FRldD

Professors Joel Ouaknine and James Worrell have been awarded the Salomaa Prize at the Development in Language Theory Symposium. See: bit.ly/2TQxS4U

Abdul Abdulrahim, a DPhil student at the Department of Computer Science will be part of this years' inaugural cohort of MPLS Equality, Diversity and Inclusion (ED&I) Fellows who have been selected to help advance ED&I work within the Mathematical Physical and Life Sciences division. Read more: bit.ly/3dO3OQp

As part of the Oxford at Home series for Oxford Sparks Live, DPhil student Maaike Zwaart took part in a question and answer titled 'Turning algebra upside down'. See the video here: bit.ly/3d9DQGC

Incoming undergraduate Maths and Computer Science student Abubakar Buwe worked on the KCL Covid System app. Read more: bit.ly/2I8SmDk

Department of Computer Science alumni Ashish Airon is on the Forbes '30 under 30 Asia' list in the Enterprise Technology category. Read more here: bit.ly/2BkWfC9

Read a 'People of ACM' interview with Head of Department Professor Michael Wooldridge here: bit.ly/36Dg1WB

The department's Professors Ian Horrocks, Boris Motik and Bernardo Cuenca Grau discuss two strains of Artificial Intelligence in this article: bit.ly/36LxYm2

DeepMind funds new post at Oxford University – the DeepMind Professorship of Artificial Intelligence

DeepMind, the leading British artificial intelligence company, has renewed and extended its commitment to supporting research in AI at the University of Oxford by funding a new post, the DeepMind Professorship of Artificial Intelligence.

The DeepMind Chair based at the Department of Computer Science, will enable a world-leading AI researcher to establish a new group within Oxford's thriving AI community, with the opportunity to pursue their own research interests. AI research at Oxford has strengths across the whole spectrum of contemporary AI, ranging from the philosophical and ethical foundations of the field through to reasoning, robotics, vision understanding, and deep learning.

The post-holder will benefit from being a Professorial Fellow of Exeter College. The College's diverse community of fellows and students from across the world include many engaged in research relevant to Artificial Intelligence.

DeepMind was established in London in 2010 and is now a world leader in artificial intelligence research and its application for positive impact. The DeepMind Chair is part of a wider initiative by the company to broaden participation in science and support wider research in the UK.

Demis Hassabis, co-founder and CEO, DeepMind, says: 'I'm delighted to expand our support of AI research at Oxford with the DeepMind Professorship of Artificial Intelligence. I look forward to seeing who the University appoints and where they decide to focus their research with the support of Oxford's world-class AI research community.'

Vice-Chancellor, Professor Louise Richardson, says: 'We are deeply grateful for DeepMind's continued support of our research and teaching, and tremendously excited about the prospects for this important new chair in an area in which Oxford is so active.'

Professor Michael Wooldridge, Head of Department of Computer Science, says: 'This is an outstanding opportunity for a world-class researcher to join a world-class AI research environment. We are tremendously excited about the possibilities for this post, and grateful for DeepMind's continued support of our work at Oxford.'

Professor Sir Rick Trainor, Rector (Head) of Exeter College, commented: 'Exeter College is delighted to be providing a Professorial Fellowship for the DeepMind Professor of Artificial Intelligence. The Chair relates directly to the interests of a number of Exeter Fellows, and strong intellectual synergy is expected.'

More details of the DeepMind Chair can be found here: bit.ly/3ep9Xmi

DeepMind video features alumna

A short film about department alumna Benedetta Mussati is included in a DeepMind blog about breaking down barriers to access to higher education through the DeepMind scholarship programme. Benedetta recently graduated with an MSc in Computer Science. In the film Benedetta talks about how the scholarship, and the course it enabled her to complete at Oxford have enabled her to fulfil her ambitions and grow in confidence.

Read the blog here: bit.ly/339urLN

Black History Month 2020

This year two of our alumna have featured prominently in Oxford University's Black History Month #BHM100 campaign. #BHM100 is a series celebrating the power of academic achievement and how Black people, past and present, have contributed to the calibre of Oxford University's reputation and understanding of Black experiences. The entries have their own unique value, and are not placed in any particular order.

Chao Mbogo is featured as number 85 in the #BHM100 and was also interviewed for the University's alumni pages. Chao is Dean of the School of Technology and Computer Science at the Kenya Methodist University and is the founder of a student mentoring scheme called KamiLimu. Read about Chao here: bit.ly/34Zntdx

Anne-Marie Imafidon (MBE) is number 71 in #BHM100. Anne-Marie

Chao Mbogo

is the founder of STEMettes, a social enterprise working to inspire and support young women into STEM careers. In 2017 Anne-Marie was awarded an MBE in the New Year's Honours 'for services to young women within STEM careers'. Read more about Anne-Marie here: bit.ly/3l0ZH6t

We are proud that these two alumna have founded organisations that help support access and encourage wider diversity in higher education.

News in brief

Alumna Chao Mbogo talks about her path breaking career in an interview conducted for the University of Oxford Science Blog: bit.ly/30K4S2u

Alumna Ann-Marie Imafidon (MBE) has been named as *Computer Weekly*'s Most Influential Woman in Tech. Read more: bit.ly/3nyuJo2

The annual The Times and Sunday Times Good University Guide has announced that the University of Oxford is ranked second in the UK for Computer Science, up two places from last year! Read more: bit.ly/3nqi7PL

Emeritus Professor Richard Bird and the department's Professor Jeremy Gibbons have released a book titled *Thinking Functionally with Haskell*. More information: bit.ly/2GC1Jeh

Once again Oxford tops the Computer Science league table in THE World University Rankings

Events for prospective undergraduates at the Department of Computer Science

Previously in-person events have been taking place online since March 2020, due to the Covid-19 pandemic

UNIQ
Annually, the department is involved in the University's multi-subject UNIQ spring and summer schools, providing the academic programme in Computer Science. UNIQ is designed as a five-day residential programme. Successful applicants normally stay in an Oxford college, attend academic and admissions sessions during the day and take part in social activities in the evenings.

The UNIQ spring school was postponed earlier this year and so the decision was made to combine the UNIQ spring school with the UNIQ summer school. The UNIQ virtual summer school in Computer Science ran from 28-31 July 2020 with just over 45 UK state school students in their penultimate year of secondary school. The students were taught by a number of our talented DPhil students, Tutors plus Professional Services Staff on topics such as: computing machines, recursion and iteration, ethics in technology, as well as some sessions on problem solving. The sessions were pre-recorded so that participants unable to join us 'live' were able to watch the video content at more suitable time. Each of the 'live' sessions was chaired by Computer Science student ambassadors, who coordinated the recorded content, and, alongside the content-creator, assisted the students with any set work from the sessions. The ambassadors were an integral part of the students' UNIQ experience. The ambassadors spent much of their time, not just on the academic programme but on other prepared talks including topics such as admissions preparation and student life.
More about UNIQ here: www.uniq.ox.ac.uk

Open days
Each year the Department of Computer Science takes part in the University-wide undergraduate open days. However, due to the pandemic, these events all took place online in 2020, and were rebranded by the University of Oxford as the 'Virtual Open Days'. The open days took place on 1 July, 2 July and 18 September 2020 and whilst they unfortunately could not be hosted in person, we had a number of live elements to our open days including a Question and Answer (Q&A) session with our Admissions Tutors as well as a Q&A with our current students.

In addition to the live elements content was also available during the open days, with a number of pre-recorded videos created by our current students on their life in Oxford, their experience of the application process, small introductions to interesting Computer Science topics, questions about the Oxford interviews, as well as information about making a competitive application. You can view those videos on the department's new YouTube channel here: bit.ly/3dl3icx

We hope that we will be able to see some students in-person next year but, of course, will be adhering to government guidelines to keep our all prospective applicants safe. We are currently working on plans to make our regular outreach offerings digital for the rest of the 2020-21 academic year.

In spite of the unforeseen circumstances this year, our student ambassadors, session leaders and staff have done a fantastic job of leading the online events this academic year in what has been a first for many across the Higher Education sector. Thanks to everyone that attended, staffed or created content for any of our online events this year, we could not have done it without you.

More information about the department's events throughout the year for prospective students and for school groups are available here: bit.ly/3dz31Tz

Science is Wonderful!

Between 22 and 24 September 2020, Elisa Passini from the Computational Cardiovascular Science group presented her research at the online exhibition 'Science is Wonderful', part of the European Research & Innovation Programme, attended by over 35,000 people.

The virtual booth displayed information about the use of computer modelling and simulations of the human heart for predictions of drug safety. Visitors could explore the booth and learn about the research, by reading posters and watching pre-recorded videos, and they could interact and ask live questions. Multiple Zoom talks were scheduled throughout the three days, both in English and Italian, to target different audiences: children, general public, and stakeholders.

This research is part of the European project TransQST (Translational Quantitative Systems Toxicology, transqst.org), to improve the understanding of the safety of medicines. The project focuses on four major systems (liver, kidney, cardiovascular and gastrointestinal-immune systems), and it includes academic and industrial partners from Europe and USA.

Industry Affiliate Group

The Department of Computer Science is proud to be affiliated with select industry partners who not only support and recruit our students but provide opportunities for collaborative work. We have now introduced the Industry Affiliate Group, which gives members improved ties with both the student body and academic community with a variety of different ways to get involved.

Funds raised from the Industry Affiliate Group are incredibly valued and are used for academic sponsorship to attract, reward, retain and develop our best students via bursaries and scholarships. We are also able to support students suffering from hardship with this important fund.

The industry scheme provides great opportunities for industry partners to work more closely with our students – read more about it here: bit.ly/2I6jyTw

For more information please email: industry@cs.ox.ac.uk

Oxford Women in Computer Science attend Grace Hopper Celebration 2020

The Department of Computer Science and Oxford Women in Computer Science Society (OxWoCS) co-funded four scholarships to support attendance at the Grace Hopper Celebration (GHC). The GHC is the world's largest gathering of women technologists from diverse backgrounds.

Yan Luo, Clara Pavillet, Lonie Sebagh, and Catherine Tong were awarded the scholarship to attend the virtual GHC, in the company of around 30,000 attendees from over 115 countries. DPhil students Catherine and Clara were particularly impressed by the keynote address from Serena Williams, who spoke about the need for a joint effort in using technology as a force of good. A great variety of workshops and company sessions were available. DPhil student Lonie was impressed by the organisers' ingenuity in keeping the virtual workshops engaging (eg an Escape Room format); she also particularly enjoyed the opportunity to interact with company representatives and learning about firms which she was previously unaware of.

Yan explains, 'At GHC, opportunities for spontaneous networking are plentiful and I had conversations with many brilliant women. For example, one lady talked about how she became a Microsoft Project Manager from a baker. After the GHC, we continued our conversation and friendship. This is the most valuable takeaway I have.'

More information here: ghc.anitab.org/

News in brief

While some of us may have found lockdown frustrating, Professor Bob Coecke decided to get creative and has been presenting talks and videos from his home, decorated with skeletons and lit like a stage set. It's certainly a different way of presenting quantum computing, and his videos have been viewed many times. bit.ly/36pi9R8

Oxford University creates its 200th spinout company PhishAR

The University of Oxford, through its research commercialisation arm Oxford University Innovation, has created its 200th company based on academic research, PhishAR.

PhishAR builds on the work of Professor Ivan Martinovic and Ivo Sluganovic, and employs augmented reality to crack down on 'phishing', an online scamming technique which fools a user into handing over their digital security credentials and other sensitive information, such as personal data or credit card details.

The cybercrime is typically achieved through utilisation of fake websites or emails that replicate the genuine article, such as an email from a bank or an employer's website. The fakes may appear as a near-perfect facsimile of the original, and may also use already obtained user information to form a sophisticated attack (spear phishing) or purposefully target senior members in an organisation for maximum gain (whaling).

PhishAR will help users identify fraudulent emails, websites and other online transactions by using artificial intelligence to scan and analyse what is presented on the screen in order to detect discrepancies such as style and formatting, used URLs, information requested from the user, and other security indicators. The company's software will run on mobile phones and augmented reality wearables to help mitigate the rapidly increasing impact of phishing, which costs individuals and organisations billions of dollars every year.

PhishAR's seed round was led by Kluz Ventures, and the company has been accepted onto the Mastercard Start Path Programme to accelerate its development.

This milestone comes at a significant point in Oxford's ongoing research commercialisation, with Oxford University spinout companies having raised £880.2m ($1.14bn) in external fundraising over the past financial year, eclipsing previous records. The innovation boom in Oxford driven by the creation and growth of these companies is fuelling an influx of talent and investment to Oxford's innovation ecosystem, known as the Oxford Cluster.

Ivan, co-Founder of PhishAR, said: 'Conventionally, service providers have been authenticating the end-user, which is a one-sided process, while the end-user struggles to know who they are communicating with, and this is what criminals are continuing to exploit, on an ever-increasing scale. PhishAR's innovation is in using augmented reality and artificial intelligence to see through their user's eyes and ensure that they are 'authenticating the authenticator'. Before any credentials are provided, PhishAR checks who is requesting the users' credentials. PhishAR is the only proven mechanism to allow the end user to authenticate and identify with whom they are communicating, and thereby dramatically reduce the risk from harmful phishing attacks while offering a seamless integration and minimal costs with existing authentication systems.'

Adam Stoten, Chief Operating Officer at Oxford University Innovation, said: 'PhishAR and its mission to safeguard millions of people ever year against cybercrime is a company worthy of Spinout #200, and we warmly welcome the team to the Oxford spinout family. While it took Oxford 55 years to get from Spinout #1 to Spinout #100, it took us a mere six years to reach 200. Consequently, we look forward to announcing our 300th spinout and many more companies in the near term.'

Chas Bountra, Pro-Vice Chancellor for Innovation at Oxford University, added: 'These 200 companies are testament to Oxford's research and its ability to have impact, and underline our status as the world's number one university. Through our innovation, we're tackling the world's biggest killers, we're tackling the biggest challenges of our era, and we're building the infrastructure for tomorrow's world – all while creating jobs and economic impact. Oxford may be an institution with its roots deep in the past, but we're demonstrating that we can also have a fundamental effect on shaping the future.'

Low-cost sensor makes tracking wildlife easier

Professor Alex Rogers is involved in the development of AudioMoth, a low-cost acoustic sensor that has been used by the Bat Conservation Trust in a trial to track rare wildlife in woodlands. The sensor has great potential to track rare creatures such as bats, and its low-cost means that many sensors to be used at multiple locations at the same time.
Read more here: bit.ly/2GWxv60

Alumni Profile

Jamie Frost : we talk to Jamie, whose career has taken an unusual turn as online teaching sensation @DrFrostMaths

What course did you study here and when?
I studied for a 4-year Masters in Computer Science between 2004-2008, and then subsequently a DPhil (Computational Linguistics group) from 2009-2012.

What was your background before that?
I went to a comprehensive school, Southborough School, from 1997-2002, followed by Tiffin School, a selective state school, from 2002-2004.

What attracted you to studying Computer Science as a subject?
I'd always been interested in programming from a young age, after my parents bought the first family computer when I was 10. At school I taught myself BASIC in order to program on their Acorn Computers. One particularly satisfying moment was when I was aged 13, and I tried to code an analogue clock; I couldn't work out how to find the coordinates of the endpoints of the hands. The day after I learnt about trigonometry in class, and rushed to a computer after school to finish my program! At age 17 I gained a recreational interest in Artificial Intelligence and built a Connect 4 player.

What aspects of the course you studied here did you particularly enjoy?
In my first year I particularly enjoyed the logic circuits course, and for my third year project I built a tool to construct and test circuits, including MIPS architecture components. I enjoyed the concurrent programming course too, and this software could output a circuit to CSP; I dug this out recently when making a GCSE Computer Science resource on the topic.

What did you do when you left Oxford?
Between my undergraduate degree and postgraduate studies, I was working for Morgan Stanley in bond trading technology. This involved working with traders and subsequently coding automated trading strategies.

After my DPhil I went into school teaching, having enjoyed my experience of departmental and college teaching at Oxford. I eventually built DrFrostMaths.com, which started off as a depository for the various teaching resources I was creating, but ended up as much more, including an online platform for students answering questions online, where teachers could also set and monitor work. While mostly for maths, there's also a Computer Science platform, with slides/worksheets for the GCSE syllabus, and an online testing-rig for mini-tasks. The platform overall is used by over 8000 schools across 133 countries, with 100 million questions answered online, 8 million resource downloads, and over 1 million hits a day. On the side I also do speaking engagements for both teacher and student groups.

How has the course you studied here helped you in your current profession?
A lot! The most demanding algorithm in my online platform was to take two expressions in LaTeX and compare them for mathematical equivalence, both on structural similarity and value equivalence (you can find a technical demo here: bit.ly/3paz4hO), using my knowledge of compilers/interpreters.

I've also had to create various bits of networked software for the platform, including a game where student devices connect with the teacher's device; and a virtual whiteboard where teachers can project exam questions to student screens, which the teachers can monitor student annotations in a single view.

When making teaching resources for GCSE Computer Science (and sometimes in Mathematics), I've made regular links to some of the things I studied at both undergraduate and postgraduate level, to give students a flavour of how they might eventually use this theory in practice.

What advice would you give to current students on applying their knowledge in the workplace, when they leave university?
At university I did all my coding in basic text editors, and my experience of refactoring and managing large amounts of code was pretty poor. It was therefore somewhat of a baptism by fire in my first internship in the summer between undergrad Year 3 and 4.

I'd also be wary of jobs with 'reputable employers' without scrutinising the nature of the work you're doing. My stint in the investment banking world was definitely not for me!

What would student you have thought about what you are currently doing?
I never thought I'd be able to blend my technological expertise with teaching in the way I have, nor achieve a level of niche B-Celebrity fame, where I regularly get asked for selfies at events by students and teachers! I'm just grateful that I have a job I love whilst also being able to keep my programming going.

Police surveillance of Black Lives Matter shows the danger technology poses to democracy

By Anjuli R Shere (Doctoral Researcher in Cyber Security, University of Oxford) and Jason Nurse (Assistant Professor in Cyber Security, University of Kent), previously Department of Computer Science, University of Oxford

US police forces have been turning to technology to track down Black Lives Matter protestors. Content from social media platforms and affiliated sites has been instrumental in the authorities being able to identify protestors based on photos of their faces, clothes and hair, or on the fact that they posted while at the protests. Meanwhile, drones have been added to the police's own means of capturing footage of the protests.

Making technology-driven state surveillance part of the police's response to democratic protest sets a dangerous precedent. There is a risk that the power this gives to police to target protestors could be abused and have a chilling effect on freedom of speech and assembly. This is particularly true in the case of Black Lives Matter, given alleged evidence of the infiltration of US law enforcement agencies by white supremacists.

What's more, the amount of data on people that is gathered by technology and potentially available to law enforcement is set to grow thanks to the rapid expansion of internet-connected devices (known as the Internet of Things, or IoT).

The Internet of Things could, if left unchecked, give authorities seemingly unlimited ways to mine for information on people, both users of the technology and bystanders. Voice operated assistants such as Amazon Alexa and Google Home record our conversations; smart watches and fitness trackers monitor our movements, and even many traditional home appliances now collect data on us, from smart fridges to washing machines. The growing prevalence and variety of these devices means a huge amount of data can be compiled on us by corporations in the name of improving user services or targeted advertising. But thanks to recent surveillance laws, state authorities can also request and gather a large amount of this data. And government bodies are already starting to capitalise on the new capabilities provided by the Internet of Things.

For example, some IoT technologies, such as internet-connected Amazon Ring doorbells that can record video footage, have become an informal addition to state surveillance infrastructure. Ring's partnerships with police forces gives them access to camera locations so they can request footage from specific device owners (and obtain it by warrant if they refuse).

Some deals have involved giving away the doorbells to the public for free. This effectively creates a cheap state monitoring network that has reportedly led to racial profiling among users.

Threat to protestors

IoT technology also could be used specifically against protestors, activists and journalists. Not only could collected data be used to identify or track people even more effectively than social media posts, but reliance on the technology could also leave people and groups vulnerable to cyber attacks.

For example, in Hong Kong we've seen attempts to disrupt the communication of protestors and force them to use less secure channels that can more easily be monitored. There's even a chance that the rise of hackable internet-connected cars could lead to more vehicular attacks on protests, as have occurred against anti-racism demonstrations in the US.

Despite these threats, our recent research shows journalists in particular are not generally aware of or protected from IoT technology being used to target them. What's more, having your data gathered by IoT devices might soon be unstoppable even if you don't own or use them. As part of our research, we surveyed 34 cyber security experts and found that 76.5% of them believe that it will not be possible for people to opt-out of interaction with the IoT within the next five years.

You might not be able to walk through a residential street without being filmed, or talk to a family member while in a doctor's waiting room without your conversation being recorded. For activists and protestors, this huge prevalence of technologies and databases that are accessible to the state means an ever-increasing risk of being identified, tracked and surveilled, as shown

continued on next page ▶

from previous page ▶

by the newly released Atlas of Surveillance.

With the growing threat of state surveillance through the IoT, activists are starting to take measures to protect themselves. More are becoming aware of the risks of taking a registered smartphone, which is essentially a personalised tracking device, on a protest. Others are following the example of protesters in Hong Kong, who recently adopted an informal all-black "uniform" complete with face masks to make it harder for authorities to identify individuals from online photos.

As well as providing secure, independent, encrypted messaging, the app Signal has responded to police forces' technological identification of protesters by creating a tool that blurs people's faces in photos. Although programs exists that can attempt to unblur pixelated photos, the fact that so much software isn't built with black people in mind could ironically make it worse at revealing the faces of people of colour.

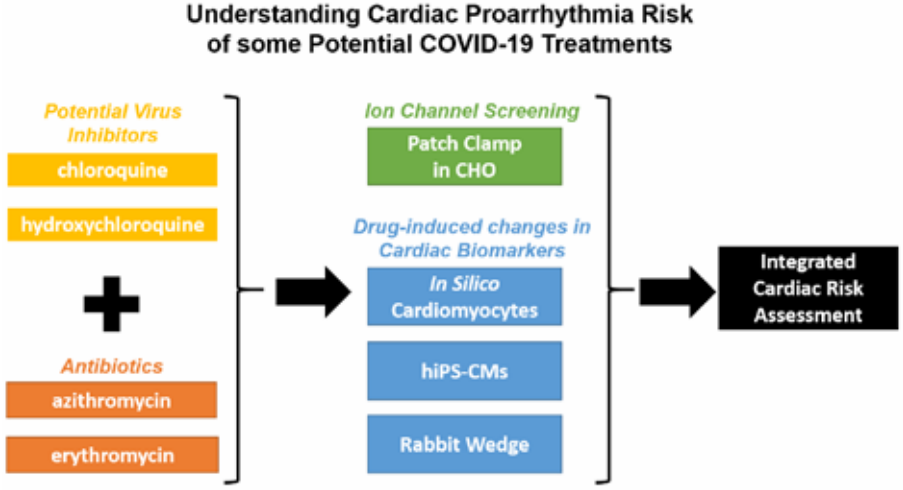
This issue reminds us that technology is never neutral, particularly when people exercising their right to protest have their data used against them. In this case, against people fighting against structural racism and police brutality against black and indigenous people.

Originally printed in *The Conversation*

An integrated cardiotoxicity assessment of potential Covid-19 therapies

Over the summer, Elisa Passini, Cristian Trovato, and Professor Blanca Rodriguez from the Computational Cardiovascular Science group contributed to a study aimed to assess the cardiac safety of potential Covid-19 therapies, eg the anti-malarial drugs chloroquine and hydroxychloroquine, taken alone or in combination without the macrolide antibiotics azithromycin and erythromycin.

These antimalarial drugs are clearly already marketed and used to treat patients. However, they were developed in the 1940s and 50s, a long time before the publication of the current ICH (The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use) safety and efficacy guidelines S7B and E14, which examine the potential for drugs to be associated with cardiac arrhythmias. There are non-clinical data available suggesting that chloroquine, hydroxychloroquine, azithromycin, and erythromycin all interact with a specific potassium channel in the heart, ie the hERG channel, which represents a key mechanism in drug-induced arrhythmia. There is also evidence of cardiac side effects for all four drugs. However, the drugs had not been formally tested before.



This study represents a very good example of the integration of modelling and simulations in the current pipelines for pre-clinical drug safety assessment. First of all, we performed the ion channel screening, using the patch clamp technique in a specific research cell line. This provided information about the effect of each of these drugs on specific cardiac ion channels. Then, using these data as inputs, we simulated the drug effect on computational model of the human heart, using our software Virtual Assay (v.3.2 © 2018 Oxford University Innovation Ltd. Oxford, UK). Finally, we performed two experimental assays: i) rabbit wedge preparations; ii) hiPS-CMs (Human-induced pluripotent stem cell-derived cardiomyocytes).

Drugs were tested both alone and in combinations, and the potential risk of drug-induced arrhythmias was evaluated based on a selection of specific biomarkers. Results were in agreement across the different techniques used, and this allowed us to classify the drugs as low, medium and high risk. The study was performed in collaborations with multiple industrial and academic partners, and it is described in a manuscript currently under review in a high impact journal.

From research to societal impact – the story of a software tool

The Oxford Research Software Engineering Group (OxRSE, www.rse.ox.ac.uk) is a Small Research Facility led by Professor David Gavaghan and Martin Robinson within the Department of Computer Science.

OxRSE offers software training and consultancy to researchers across the University. The group engages with researchers from all backgrounds and through interdisciplinary collaboration helps academics develop software to achieve goals that would otherwise be impossible. This approach is beginning to bear fruit, with some of our past projects achieving significant societal impact.

As part of a longstanding partnership with the Department of Education, OxRSE created the technical infrastructure for LanguageScreen, a language assessment which enables education professionals to assess young children's oral language skills. LanguageScreen is a key tool for the Nuffield Early Language Intervention (NELI) – which is a core part of the COVID-19 catch up package for primary schools in England (see announcement here: bit.ly/3k0CxMa).

Children are selected to participate in the NELI programme using LanguageScreen: a mobile app guides the examiner through a series of activities to assess a child's oral language skills and then uploads the assessment scores to a secure backend that generates an individualised assessment report. The concept behind the assessment activities was created in the Department of Education and their translation into an easy-to-use, production-grade software tool is a result of years of work and close communication between the researchers and research software engineers.

In July 2020, the UK Department of Education announced £1 billion of funding to support children and young people to catch up with education lost to COVID-19. The funds support a variety of programmes, including the Reception Year Early Language Programme, which incorporates NELI, a 20-week evidence-based intervention programme designed to improve the oral language skills of four- and five-year old children in need of language support.

This support will help schools deliver one-to-one and small-group support for children in Reception classes (students are usually aged four and five years old) whose spoken language skills may have suffered because of the pandemic. Every state-funded primary school with a Reception class in England is eligible to apply, with priority given to schools with a high proportion of disadvantaged pupils. Schools will receive resources and will be able to access online training from December 2020.

NELI was developed by a team led by Professor Charles Hulme (Department of Education) and Professor Maggie Snowling (Department of Experimental Psychology and St John's College) and was recently the subject of a large-scale effectiveness trial in 193 schools across England which tested the intervention programme and the LanguageScreen assessment in everyday school conditions.

An independent evaluation by the Education Endowment Foundation found that pupils who took part in the intervention programme made, on average, three months' more progress in language skills than a group of similar children who did not receive the intervention. The programme was also found to be an effective way of boosting language skills for children with English as an additional language.

continued on next page ▶

from previous page ▶

The NELI intervention requires schools to have reliable means of identifying children with language difficulties and for monitoring the development of children's oral language skills – which is where the LanguageScreen tool comes in. The LanguageScreen assessment tool consists of a cross-platform mobile app (available from Google Play and App Store) developed by Mihaela Duta (Department of Computer Science & Department of Experimental Psychology), which interfaces with a secure website (languagescreen.com) developed by the department's Abhishek Dasgupta, Fergus Cooper, Thibault Lestang and Graham Lee. The website backend stores and processes children's personal data alongside the LanguageScreen app assessment scores and automatically generates assessment reports for schools.

The reports contain individual and class level ranked evaluation of children's performance, with colour-coded alert labels to make it easy for education professionals to identify pupils in need of intervention. While conceptually straightforward, the software design and implementation of the integrated system posed several challenges.

The LanguageScreen app is aimed to be used by education professionals without any special expertise in language skills assessment, so it was important that the app be self-contained and useable 'straight out of the box', without any prior examiner training. Therefore, the LanguageScreen app guides the examiner through a series of activities to be done with the child via an intuitive interface that achieves the delicate balancing act between being both informative for the examiner and not off-putting for the child assessed.

The LanguageScreen assessment tool handles highly sensitive children's data therefore data security was a major consideration. The software design had to minimise the storage and transfer of identifiable personal data, while still enabling the reliable linkage between the assessment scores and the children's personal details. Since the LanguageScreen mobile app is designed to be run on the examiner's device, which could be either personal or school-owned, the device itself constitutes an important point of personal data vulnerability. It was therefore decided that the app would not store any identifiable information, but instead only read an encrypted QR code to identify the child and use that to label the uploaded data to the website backend, which handles the process of data linkage in order to produce the personalised reports. This meant that the backend had to provide a chain user interface and processing units that takes the children's personal data, process it to derive encrypted identification tokens, provide those tokens in the form of QR codes in a downloadable document, link the assessment scores uploaded by the app and produce the corresponding individualised reports.

This was a productive collaboration between OxRSE and researchers in Education and Experimental Psychology departments that has resulted in the creation of a software tool that will be used by thousands of schools. Working with researchers to create production quality software that also has societal impact is a key goal of the RSE team at Oxford, and we are happy that we have been able to achieve this. OxRSE continues to be involved with the project as a consultant. We look forward to work on similar projects in collaboration with other researchers across the University.

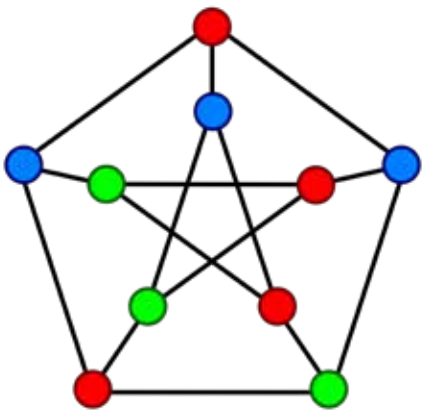
Read more about OxRSE here www.rse.ox.ac.uk

Approximate graph colouring: searching for approximate solutions to hard problems

by Professor Standa Živný

For many computational problems, finding an optimal answer is provably hard. In this article, we look at important problems for which finding an even approximate solution appears (and in some cases can be proved to be) hard.

The problem of graph colouring is one of the most studied computational problems: given an undirected graph G and k available colours, the task is to determine whether G admits a k-colouring; that is, whether each vertex of G can be assigned one of the k available colours in such a way that adjacent vertices are assigned different colours. Graph colouring belongs to the computational class NP, which means that if you are given an assignment of colours to the vertices it is easy to check whether the assignment is a proper colouring: for every edge, just check whether the two vertices of the edge are assigned different colours.



Colouring of the Peterson graph (source: Wikipedia).

Here is one of many possible interpretations of the problem. You are given a group of n people; these are the vertices of the graph. You're also told who is friends with whom. Here we assume that friendship is a symmetric relation; that is, if Alice is a friend with Bob then Bob is also a friend with Alice. The friendship relation is represented by the edges of the graph. Given an integer k, the k-colouring problem asks: can you partition the n people into k different groups so that in each group there are no friends?

The k-colouring problem is computationally easy if k=1 or k=2. If k=1 then there is no solution as long as there is at least one edge in the input graph. More interestingly, if k=2 then the k-colouring problem is equivalent to testing whether the vertices of the input graph can be partitioned into two parts such that all edges go between the two parts; this is a standard algorithmic task taught in first-year algorithms courses. For any fixed constant k≥3, the k-colouring problem is NP-complete, as shown by Karp in 1972 when he introduced the concept of NP-completeness. In particular, for any fixed k≥3, it is NP-hard to find a k-colouring of a given graph even if one is promised that a k-colouring exists.

As an aside, if you could provably show that k-colouring does or does not admit an efficient (polynomial-time) algorithm you would resolve the P vs NP problem, one of the Millennium Prize problems. This would come with a cheque for $1M and eternal fame.

The approximate graph colouring problem is a relaxation of the colouring problem, in which one is allowed to use more colours than is needed. For instance, given a 3-colourable graph G, can you find a 100-colouring of G? More formally, given two fixed integers k≤l, the (k,l)-colouring problem is as follows: given a graph G that admits a k-colouring (which is not given but promised to exist), find an l-colouring (which clearly exists, since a k-colouring exists and you could just not use the extra l-k colours).

As before, the problem is interesting only for k≥3. Since the work of Garey and Johnson in the 70s, it has been widely believed that for any fixed 3≤k≤l, the (k,l)-colouring problem is NP-hard. However, the progress has been very slow!

The state-of-the art, until 2019, was a result that showed that for every k≥3, (k,2k-2)-colouring is NP-hard. In 2019, a novel approach to the problem (in terms of techniques) showed that for any k≥3, (k,2k-1)-colouring is NP-hard. Lots of work to improve the hardness by just 1.

In 2020 Marcin Wrochna (also Department of Computer Science) and the author published a paper at the ACM-SIAM Symposium on Discrete Mathematics (SODA) that showed that for every k≥4, (k,b(k))-colouring is NP-hard, where b(k) is roughly $2^k/\sqrt{k}$.

Note that the result improves the previously known NP-hardness by an exponential factor. The proof of the result has several interesting features. Firstly, it is a relatively simple proof based on an old construction known as the arc digraph or line graph. As shown in the 70s, this construction has the remarkable property that while it can be computed in logarithmic space it can decrease the number of colours required in a controlled way. Secondly, the new result provably cannot be established by any previously used method. Despite this exciting improvement, a proof of the NP-hardness of

continued on next page ▶

from previous page ▶

(k,l)-colouring for all constant 3≤k≤l seems currently out of reach.

Looking back at the 3-colouring example, sometimes 3-colouring is really what is desired. Finding a 3-colouring of a 3-colourable graph is NP-hard, as Karp showed. What if the assumption is strengthened?

Well, given a graph that is promised to be 2-colourable, you can find a 3-colouring; in fact, you can find a 2-colouring! Using the notion of a circular chromatic number of a graph, one can make the following question precise: given a (2+ε)-colourable graph, can you find a 3-colouring? For ε=0, the problem is solvable efficiently. For ε=1, the problem is NP-hard. A recent result of Krokhin, Opršal (both from Durham University), Wrochna, and the author shows that for every $0<\varepsilon<1$, this problem is NP-hard. The proof of this theorem (and in fact of an even more general statement) involves ideas from algebraic topology, such as the winding number of a closed curve, and applying them to the discrete world of graphs.

All results discussed in this article are special cases of results established for a general framework of the promise constraint satisfaction problem (CSP). The CSPs include, in addition to graph colourings, many other fundamental computational problems and allow for devising mathematical theories that shed light on the computational complexity of the underlying combinatorial problems. On advantage of studying CSPs rather than problems in isolation is that the CSP framework abstracts out irrelevant issues specific to concrete problems and enables us to 'see the wood despite the trees'.

BigHeart: big data analysis and simulation of human heart function

The project 'BigHeart: big data analysis and simulation of human heart function' has been recently awarded to Professor Alfonso Bueno-Orovio by the EC Partnership for Advanced Computing in Europe (PRACE), under their first ICEI call.

Cardiovascular disease remains the major cause of death worldwide. The ageing of our population is further aggravating this societal burden, with a sustained increase of more than a million of deaths per year. These figures highlight the need of new Computational Medicine approaches, to complement clinical research in order to improve the understanding and treatment of heart disease.

BigHeart aims at generating the first large-scale dataset of patient-specific anatomical models and validated simulations of human cardiac function, in both health and disease. The results obtained during this project, through the exploitation of big clinical datasets such as the UK BioBank, will provide novel insights on mechanical dysfunction in human heart disease. These are expected to push the boundaries of in silico human clinical trials for drug and electrical therapy, and the realisation of the digital twin in clinical decision-making.

Alfonso says: 'The resources provided by BigHeart will realise our capabilities to integrate the processes of big data analysis, personalisation, and multiscale simulation of human cardiac function, enabling a further translation of our results to clinical research.'

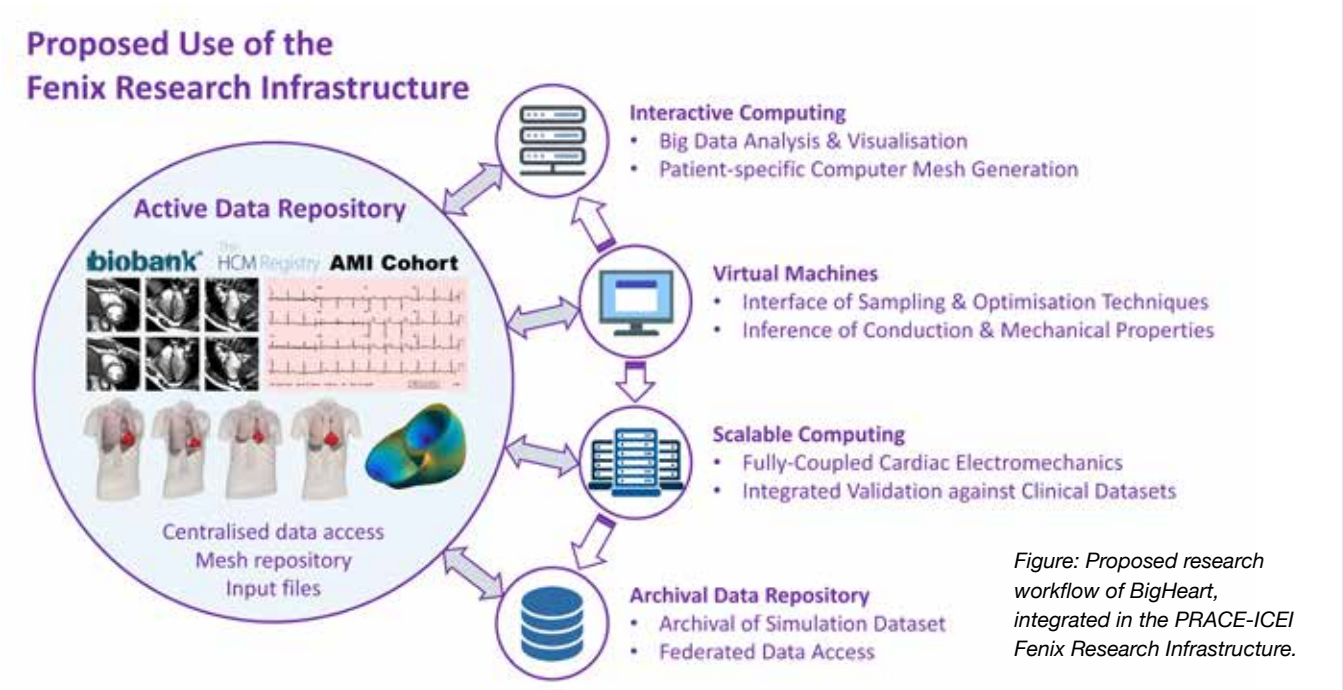


Figure: Proposed research workflow of BigHeart, integrated in the PRACE-ICEI Fenix Research Infrastructure.

The environmental impacts of digital technologies

Professor Marina Jirotka was a keynote speaker at this year's International Conference on ICT For Sustainability. She spoke about the potential for a 'responsible' revolution that will address the environmental impacts of digital technologies. Here are some extracts from her talk.

We are living through a revolution in information and communication technologies. Many of us will have found ubiquitous smartphones, broadband, and other IT infrastructure vital to help ameliorate the constraints of the COVID-19 pandemic. However, these technologies also have costs to the environment. ICT in general, and data centres and cloud computing in particular, have a heavy footprint, featuring high consumption of non-renewable energy, waste production and Co2 emissions.

A single data centre can use as much electricity as a medium-sized town. Most of these are powered by non-renewable energy sources. Many of the buildings also rely on 24/7 all-year-round air conditioning to prevent this equipment from overheating, which again consumes more electricity. This year, 3.5% of global emissions are expected to be produced by the data centre industry; this overtakes that of the aviation and shipping industries. Some researchers estimate that by 2025, one fifth of the planet's electricity will be consumed by data centres, amounting to 14% of global emissions in 2040.

Another example is Bitcoin. The energy use of Bitcoin is now calculated to be approximately 53 terawatts annually. This is roughly equivalent to the energy use of Switzerland. It is so bloated that Cambridge University's Business School has created a special tool to track it in real-time, scaling up to calculate annual usage. A team at MIT has estimated that the associated carbon emissions sit somewhere between levels produced by the nations of Jordan and Sri Lanka.

A final example is connected devices (internet of things, IoT). A recent report from the International Data Corporation forecasts that in 2025 there will be 41.6 billion connected devices enabled by 5G technologies and generating 79.4 zettabytes of data. The environmental costs of the generation and processing of such amounts of data through AI and smart devices go beyond CO2 emissions: they also include the unsustainable practices of mineral extraction and e-waste disposal – often conducted in informal settings in low and middle income countries and with exposure to toxic emissions for those working in these sites.

The environmental costs of ICT threaten to be very high indeed and combine with sustainability issues in other spheres to suggest a stark future in which (as described by the State of the Future report from the Millennium Project) challenges to humanity include radical climate change, increasing energy demands, population growth, and lack of access to clean water. Fortunately, we are already seeing some steps being taken to address these costs. Corporations are becoming 'greener', employing sustainability advocates, and increasing their renewable energy supply as a response to civil society organisations' criticism and public sensitiveness towards environmental questions.

Tech companies are well aware that they use vast quantities of energy, and, even if just for financial reasons, are trying to shift their energy demands to more sustainable models. Google already uses renewable energy to run its data centres, but in April it announced that it was moving its scheduled load-management – for example training AI models – to coincide with peaks in renewable energy supplies. So, when the wind blows and the sun shines, Google is running its energy-heavy video transcoding. Hopefully this drive for change will continue. It is slowly becoming undesirable, as a researcher, developer and designer, to ignore the societal and environmental consequences innovations.

continued on next page ▶

from previous page ▶

There is much that the research and innovation community can do to advance an organised and collective move towards digital sustainability. Research councils are funding research into sustainability in the digital economy and opening up other relevant funding streams. Our own research group, Human Centred Computing, hosted Oxford's first carbon neutral conference in September 2019: 100+ Brilliant Women in AI & Ethics (see *Inspired Research*, issue 15). We selected our print company on the basis of their own carbon-mitigation efforts and chose vegetarian conference food from sources that were as local as possible. We tried to use items that could be easily recycled and we calculated the carbon generated by the event (including delegate travel) so that it could then be offset.

Looking more broadly, our community can embrace the initiative known as Responsible Innovation (RI), which aims for science and society to work together. When adopting the RI approach, we ask questions such as 'what kind of society do we want?' and 'how can science and innovation help us to achieve this society?', and seek to bring in stakeholder perspectives across the stages of innovation. RI also involves anticipating the potential social, ethical and environmental impacts of innovation creating flexible and adaptive solutions to address their unintended but potentially harmful consequences. By identifying in advance the potential environmental impacts of innovation – in terms of energy use etc – we have the opportunity to bring them to the attention of developers so that they can be addressed during the development phase itself. We can thereby positively influence the trajectory of innovation.

We recently ran a workshop to explore some of these sustainability and RI issues further. This was led by Federica Lucivero from Ethox and the Wellcome Centre for Ethics and Humanities, at the University of Oxford, and Gabrielle Samuel from the Department of Global Health and Social Medicine at King's College London. We worked with interdisciplinary colleagues from various other universities to map out the conceptual landscape on the environmental impacts of data-driven technologies. One key area we identified was the importance of trade-offs when considering the environmental impacts of innovation. Data-driven environmental impacts need to be viewed alongside the social value of data-driven technologies. This is a complex balancing act. What benefits are more important than others and might make us willing to accept some environmental risks? Which are less important/less valuable? How do we decide which is better? Such questions are inherently complicated as there are different dimensions to take into account when discussing value trade-offs and balances.

Decisions about reducing environmental impacts can also be made at different levels of the data journey, from manufacturing to data use. Then, we also need to ask questions about responsibility: 'who is taking responsibility?' and 'who should be taking responsibility?' These questions are crucial and need to be addressed at different levels, including individual, corporate, and public/institutional responsibility. Furthermore, we need to encompass issues of values, interests, needs and inequalities. In this sense, to explore digital sustainability we need a holistic approach that takes into account moral, political, social, cultural, environmental, geographical and economic aspects together with technical and scientific ones. You can read a report on the workshop here: bit.ly/32lqXFB

In conclusion, the environmental impact of digital technologies is an urgent issue to be addressed. It something we all need to be aware of and that requires a response from multiple actors across society. We need to identify the impacts in an evidence-based manner so that we can then work to develop technical, social, business and governance-based solutions to promote change.

Scientists discover the unique signature of a lion's roar using machine learning

© William Warby

The roar of a lion is one of the most thrilling and captivating sounds of the wild. This characteristic call is typically delivered in a bout consisting of one or two soft moans followed by several loud, full-throated roars and a terminating sequence of grunts.

A team of scientists based in WildCRU at the University of Oxford, well-known for their research involving Cecil the Lion, has teamed up with colleagues in the Department of Computer Science to discover the precise ways in which each lion's roar is distinct, identifiable, and trackable.

Harnessing new machine learning techniques, the group designed a device, known as a biologger, which can be attached to an existing lion GPS collar to record audio and movement data. The biologgers allow the scientists to confidently associate each roar with the correct lion by cross-referencing movement and audio data through the large datasets of roar recordings collected. Video explainer: bit.ly/3p1jCEP

With the data collected by the biologgers, the scientists trained a pattern recognition algorithm to 'learn' each individual's roars and then tested the algorithm on sequences that it had not seen before to determine whether the shape of the contour as a whole is an important distinguishing feature.

Results, published in Bioacoustics, reveal that it is possible to classify roars according to individual identity with 91.5% accuracy. These findings suggest that the overall shape of the fundamental frequency (f0) of the full-throated roar contour is consistent within each individuals' roars and sufficiently different from other individuals to allow for accurate classification of individual identity.

Previous research has shown that lions can recognise the calls of other individuals, allowing them to locate distant companions and also to avoid potentially hostile neighbours. But little has been understood about how individuals convey identity information in the structure of their calls.

These new findings reveal a possible mechanism for individual vocal recognition amongst African lions. They indicate that individual lions may be able to learn the subtle variations in the fundamental frequency of other lions' roars and thereby associate particular variations with particular identities.

Andrew J. Loveridge, from WildCRU at the Department of Zoology, said: 'African lion numbers are declining and developing cost effective tools for monitoring, and ultimately better protecting, populations is a conservation priority. The ability to remotely evaluate the number of individual lions in a population from their roars could revolutionise the way in which lion populations are assessed.'

Professor Andrew Markham, from the Department of Computer Science, said: 'being able to accurately distinguish between individual roars using machine learning algorithms could facilitate the development of alternative techniques for assessing population density and tracking individual movements across the landscape.'

The scientists plan to develop their work by carrying out play-back experiments using modified calls. They hope that they will be able to determine whether the fundamental frequency alone conveys sufficient information on individual identity to enable vocal recognition. With rapid technological advances, automated acoustic monitoring of lion populations may not be far off.

Originally published in the Oxford University Science Blog.

© Benh LIEU SONG

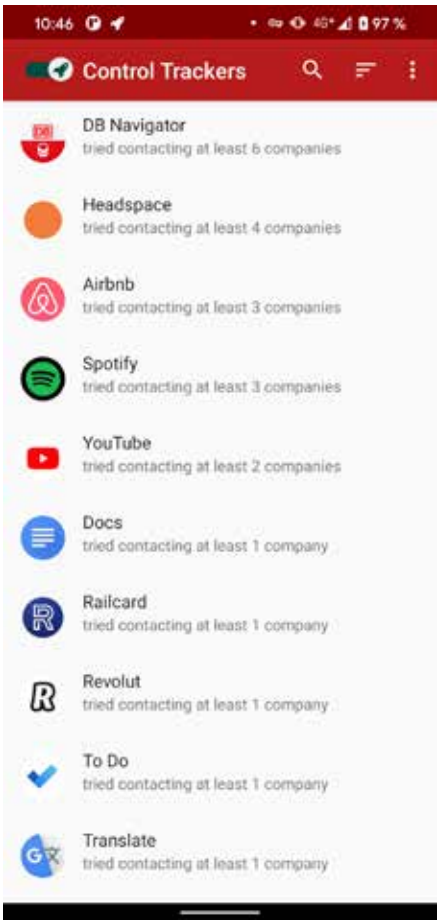
Take control of your phone, before it takes control of you

By Konrad Kollnig, a DPhil student in the Human Centred Computing research group

Many apps collect detailed data about your life. Some apps share this data with hundreds of companies from around the world. Should you be worried? To help you find answers, we at Human Centred Computing in the Department of Computer Science are developing TrackerControl, a new privacy app.

'I don't share much data with apps', you may object. The problem is that apps often collect data about you, without any notice or choice. This data can include your IP address, location, and sometimes even your phone number and those of your contacts. Most apps even finely record your app usage in an opaque practice known as tracking.

This data can make it surprisingly easy to monitor your life in great detail, by triangulating different information. Your last romantic date? Apps might know everything. Worse still, many apps share data with dozens – and sometimes hundreds – of companies. Whilst such data collection practices may not cause immediate harm, its future consequences are difficult to anticipate. This makes society vulnerable to a wide range of risks, some of which are yet to come into existence.



TrackerControl emerged as part of my MSc project in the department, together with Professor Max Van Kleek. While many privacy-enhancing apps already existed, none of them both (1) showed detailed insights into tracking and (2) enabled users to limit unwanted data transmissions. Only if users are given the necessary information and a choice, they can make more meaningful decisions about their data. In developing the app, a fundamental technical challenge was the balance between limiting data flows and preventing apps from breaking. If TrackerControl blocks too many data transmissions, it may impede app functionality. Often, data transmissions serve legitimate purposes, which can be difficult to untangle from unwanted tracking transmissions. Many existing privacy apps solve this challenge by blocking visible ads only. This is not enough, because invisible data transmissions can pose a similar privacy threat. TrackerControl instead blocks as much tracking as possible, whilst making it convenient for users to create exceptions from this rule.

The first version of the app was developed and released after a round of user interviews in November 2019. Since then, TrackerControl has attracted thousands of users, and appeared in various tech news outlets. TrackerControl users have formed a supportive and positive online community. They answer questions of new users, inform further app development, and have translated the app into nine languages. Working with the community has been a fruitful way to develop software alongside conducting research.

TrackerControl is part of wider efforts in our research group to scrutinise apps and support human flourishing in the digital world. One related project, called X-Ray, found that 90% of Android apps may share data with Google and 40% with Facebook, and that the power of companies over data might have tipped out of balance.

Tools such as TrackerControl can only be a partial solution. With data collection at scale, the privacy of individuals depends increasingly on the choices of other people. Many companies already have so much data that they need little to make accurate predictions about the individual. If we want to protect fundamental democratic rights in the digital age, such as the right to data protection and to a private, self-determined life, we need structural changes to the digital ecosystem, collective efforts, and possibly more effective regulation.

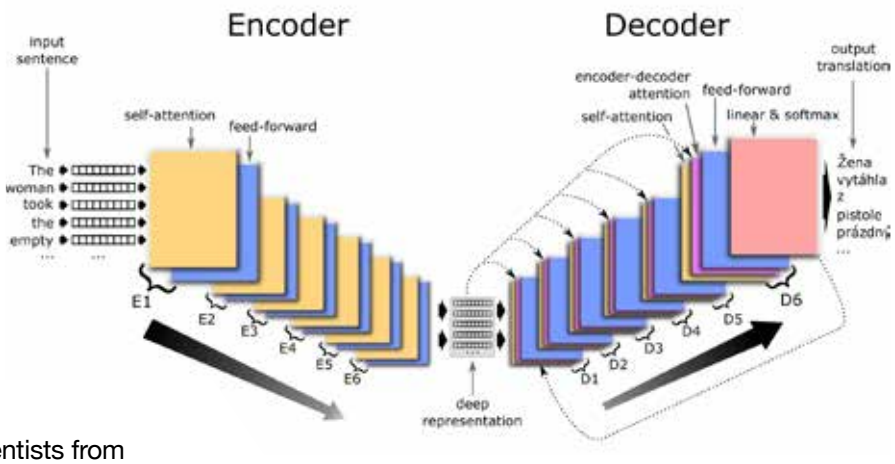
With TrackerControl, we showcase an exciting way to pursue and share research, as well as practical ways to improve privacy. You can help make a change, by downloading the app via trackercontrol.org, finding out how your apps transmit data, and sharing your experiences.

References:

- Binns, Lyngs, Van Kleek, Zhao, Libert, and Shadbolt (2018): 'Third Party Tracking in the Mobile Ecosystem' https://bit.ly/3eUf0f6
- Binns, Zhao, Van Kleek, and Shadbolt (2018): 'Measuring Third-party Tracker Power across Web and Mobile' https://bit.ly/3pjBEll

CUBBITT: taking translation to a new level

By Jakub Tomek from the Department of Physiology, Anatomy & Genetics

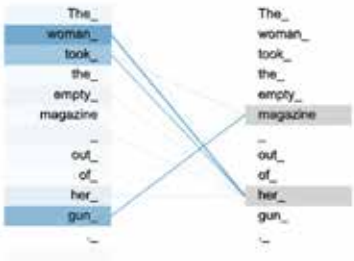


Automatic machine translation is an important application of artificial intelligence, with hundreds of billions of words translated on the internet every day. Although the quality of machine translation has improved greatly over the past decade, the view was still that the level of human translation was unattainable by artificial intelligence.

However, our collaboration between scientists from Charles University [1], Oxford University [2] and Google Brain [3] shows that human quality can not only be achieved, but even surpassed in some aspects [4]. CUBBITT, a newly developed deep learning system, was judged by human evaluators to be more accurate than the translation performed by a professional translation agency. CUBBITT makes fewer errors than a human translator in adding, removing, or shifting meaning, grammar, and spelling. While human translations were rated slightly better in terms of fluency of the target language, CUBBITT outperformed previous state-of-the-art systems in this respect as well. Moreover, CUBBITT was successful in the 'Translator's Turing Test', where, on a set of 100 independent sentences, most participants could not distinguish it significantly from human translations, while 94% of participants recognised translations from Google Translate as machine translations.

The success of CUBBITT is primarily due to two factors. The first is the concept of 'self-attention', which has proved enormously useful in the Transformer model from researchers at Google Brain [5]. Self-attention allows a network to learn the relationships between remote elements of a sentence and use them in translation. Therefore, the magazine in the picture below is correctly translated as gun magazine, and not as news magazine (which are two different words in Czech), because the neural network 'knows' that the magazine is linked to a gun in this sentence. At the same time, the use of attention allows for significantly better parallelisation than previous approaches and thus considerably faster learning on big data.

The second factor in CUBBITT's success is the clever use of monolingual data. Machine translation systems are primarily trained using pairs of translated sentences, and while long-term efforts by scientists have led to the creation of datasets containing tens of millions of such translated sentences for individual language pairs, this is still not enough. In contrast, monolingual data is abundant. A surprisingly effective strategy (called backtranslation) to get more data is to take this monolingual data, machine-translate it into a second language (for example, from Czech to English) and then use it for further training (for English to Czech translations). The standard approach is to train the system on randomly mixed machine-translated (artificial) and human-translated (authentic) sentences. However, the team from Charles University showed that better results can be achieved if the network receives artificial and authentic data in longer continuous blocks. Such an approach can better 'extract' knowledge and language structures in both types of text, combine them together, and apply them successfully in translation.



"Magazine" is translated as "zásobnik" (gun magazine), rather than "časopis" (news/lifestyle magazine) due to the association between itself and "gun"

The role of the Oxford part of the team (myself and Marketa Tomkova) was primarily in the design of human evaluation of translations, analysis of results and project management. We also managed to explain why it's better to submit data in longer blocks to the network, which is an interesting example of a problem where we see that a deep neural network does something well, but it's not at all obvious why.

Our study only covered news articles and not, for example, novels or poetry. Reference human translation was provided by a professional translation agency, not a specialist literature translator. Despite these 'buts', it is nonetheless clear that artificial intelligence has reached another major milestone. You can judge the current quality of the machine translation for yourself, given that this article was written in Czech and translated into English by CUBBITT.

[1] Charles University: Institute of formal and applied linguistics (Faculty of Mathematics and Physics) https://bit.ly/3mWJvUu [2] University of Oxford: Ludwig Cancer Research and Department of Computer Science [3] Google Brain: https://bit.ly/2HY4nff [4] Popel M, Tomkova M, Tomek J, Kaiser Ł, Uszkoreit J, Bojar O, Žabokrtský Z. Transforming machine translation: a deep learning system reaches news translation quality comparable to human professionals. Nature Communication. 2020. https://go.nature.com/3mRbnco [5] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I. Attention is all you need. In: Advances in Neural Information Processing Systems. 2017. bit.ly/32d5EFU

A view on the A-level results controversy

By Helena Webb, Senior Researcher, Human Centred Computing

Amongst the many dramatic events of 2020, the controversy surrounding the release of A-level results in England, Wales and Northern Ireland stands out as particularly memorable. When the examinations planned for the summer were cancelled due to COVID-19 it became necessary to find an alternative way to produce student grades. The schools regulator, the Office of Qualifications and Examinations Regulation (Ofqual) developed a standardisation algorithm to do this but when the resulting grades were released, they proved so unpopular that they were abandoned less than a week later. The episode serves to illustrate how unintended negative consequences can arise from automated decision-making processes.

The Ofqual algorithm was designed to moderate teacher-estimated grades whilst also limiting grade inflation. It was set to produce results that were similar (on a school and national level) to previous years. When the results were released, there was immediate outcry because large numbers of students – around 40% – received lower grades than their predictions. This left many fearful of losing University places and job opportunities. In addition, whilst the results did limit grade inflation on a national level, private (Independent) schools experienced a larger increase in the number of students getting top grades and many felt that state school students were more affected by downgraded marks than private school students.

It is possible to see how these problems arose. The algorithm was designed to moderate a ranked list of teacher-estimated grades and produce results that were similar to previous years for each school. This meant that, once a school had met its 'quota' for a certain grade, students with the same estimated grade but ranked below others in the school, were at risk of being downgraded. This had very severe impacts for large numbers of individual students. It also failed to acknowledge schools that had achieved improvements in performance from previous years. These rapidly improving schools were more likely to be in the state sector so this is one reason why the algorithm appeared to disadvantage state school students.

A further reason is that the algorithm gave greater weighting to teacher-estimated grades where class sizes were very small. This was designed to accommodate for fluctuations that would be more pronounced where only a small number of students were taking the exam. Small class sizes tend to be more common in the private sector and teacher-estimated grades tended to be more generous than the algorithm-moderated marks, so this was an additional way in which the algorithm had different consequences for different students.

Algorithmically-driven processes for automated decision-making are becoming increasingly prevalent in modern society and offer various efficiency benefits. However, as the A-level case vividly illustrates, they frequently have unintended consequences that can be to the detriment of individuals and communities. These consequences can be systematic and disadvantage particular demographic groups. Ofqual stated that the algorithm could not be biased as information about whether a school was private or state was not included; however other factors such class size and previous performance can serve as proxies for this information and be consequential for the outcomes delivered.

Amongst these negative concerns a more positive implication that can be drawn from the A-level case is the extent to which the general public engaged with it. There was massive media coverage at the time and a great of deal of public interest and discussion over the process through which the results had been generated. The word algorithm itself became more of a household term. This can be seen as a very positive step since wide-scale public engagement is necessary to foster a societal-level discussion about the role we want automated decision-making to play in our lives.

The Department has a research project looking further into aspects of algorithmic bias. Read about 'UnBias: Emancipating Users Against Algorithmic Biases for a Trusted Digital Economy' here: bit.ly/3ppGVrN

Parental control apps: protection or punishment?

By Ge Wang, a DPhil student in the Human Centred Computing research group

Children are now spending an unprecedented amount of time online each day via their smartphones and tablets. This unusual situation during 2020 has accelerated the need for children to spend time online, a previous report, 'Children's Media Lives: Life in Lockdown' (Ofcom, 2020) showed the ways the COVID-19 pandemic has impacted the structure of children's daily lives, shifting many their formerly face-to-face activities online.

While the Internet has brought significant new opportunities for children to learn and have fun, even during a pandemic, there have been significant concerns over the risks and dangers of children in online environments. In response to children's increasing exposure to both classes of risk, a new genre of apps, known as parental control apps, have emerged. These tools are designed to act as technical mediation support for parents to facilitate access to, and control over, their children's online activities as a means of protecting them online.

This rapid adoption and increasing reliance on parental control apps has raised corresponding questions about their efficacy, how such apps fit into existing parenting strategies, and the effects such apps are having on parents and their children. To understand where and how improvements might be achieved, we analysed 58 popular Android parenting control apps to examine how they regulate their parenting functionalities and how their functions effectively support or fail to support children and parents. As part of this process we also analysed both children's and parent's reviews of each of the apps.

From our studies, it is interesting to see that, instead of considering the apps as protection for them, children usually regard the apps as a form of punishment. Children reported these apps as 'socially damaging' or 'damaging to their mental health'. In some cases, they also stated that the apps are 'harmful' or 'detrimental to experiential learning'; and the apps are 'unnecessary' as 'why not just talk instead'. On the other hand, parents seem to see the apps as legitimate steps to ensure their children are safe online, an element which they struggled to make clear to their children. Almost all the functions we observed in parental controls apps were about either restriction (ie restricting children's activities online) or surveillance (ie monitoring children's activities). One of the major problems with restriction and surveillance is the extent to which these strategies reduce children's autonomy and violate their privacy.

How might apps be made to be more autonomy-supportive then? Here we offered two suggestions:

1) Our view is that the inherent problems with parental control apps is that they were focused exclusively on effective restriction and surveillance, instead of digital parenting. Rather than framing restriction and surveillance as protective measures to be used in perpetuity, for instance, a digital parenting app might reframe restriction and surveillance as a means of skills scaffolding. Starting with the most supportive scaffolding (maximum restrictions and surveillance) for the very young and most vulnerable, such apps could then align the gradual removal of restrictions and surveillance measures with milestones as children get older or demonstrate the ability to recognise and cope with online risks.

2) Another approach that might help is to frame restriction/surveillance as options of last resort. Primary autonomy-confirming strategies for older children might instead include, for example, simple verbal agreements about limits on online activity when requiring children to take responsibility for their actions. And a convincing rationale to justify the use of such apps might ensure that they are used only as necessary.

Read more about the department's research in Human Centred Computing here: bit.ly/3pIgPRb