Research

Horizons

Pioneering research from the University of Cambridge

Issue 35

Spotlight

Artificial intelligence

Feature

Tree-ring timelines

Feature

Epic poetry

UNIVERSITY OF CAMBRIDGE
www.cam.ac.uk/research

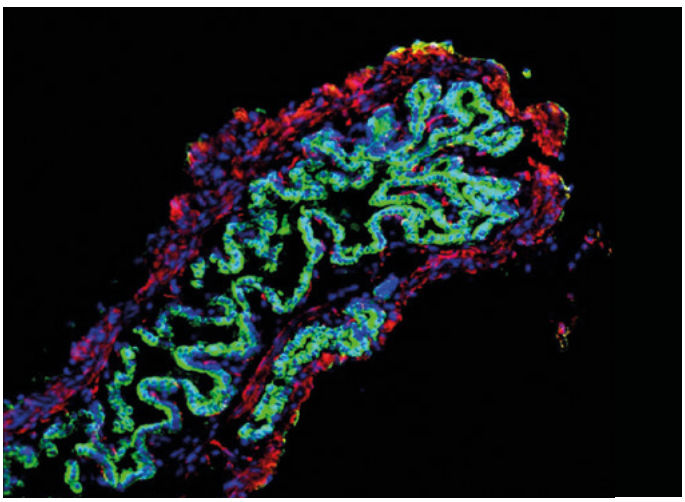
Contents

A News

B Features

C Things

D Spotlight: Artificial intelligence



Welcome

Almost everywhere I turn, I see the transformative potential of artificial intelligence (AI) being promoted, so it is very timely that it is a focus of this issue of *Research Horizons*.

Some of the researchers featured here are among AI experts worldwide who have signed an open letter affirming the benefits of the technology and urging caution in its development. In essence, they said: "AI systems must do what we want them to do."

Enabling enormous promise whilst stewarding progress is a complex balance. It requires engineers, computer scientists and mathematicians to build systems that learn from data, and that think both like humans and unlike humans; it requires experts in fields as different as climate science and criminology to develop innovative uses of these machines that learn; and it requires researchers to pose new questions about safety, trust, transparency, security and privacy in an algorithm-rich world.

Cambridge has strengths in machine learning, robotics and applications of AI technologies. Not only is research aimed at maximising the impact of AI, it is also aimed at understanding how we can ensure that the technology benefits humanity. This has been helped by two new research institutes – the Leverhulme Centre for the Future of Intelligence and the Centre for the Study of Existential Risk – as well as being a founding partner in The Alan Turing Institute.

These developments are indeed timely. In November 2017, the UK government's Industrial Strategy set out four Grand Challenges, one of which was to put the UK at the forefront of the AI and data revolution. In this issue, we look at some of the areas in which Cambridge AI researchers are making a significant impact, as well as consider some of the benefits for academics and industry of being within the 'Cambridge Cluster'.

Elsewhere in this varied edition of *Research Horizons*, we cover a major boost for cystic fibrosis research, an epic analysis of epic poetry and Cambridge's first dedicated tree-ring laboratory.

We hope you enjoy these and other articles in this issue.

Professor Chris Abell
Pro-Vice-Chancellor for Research

Editor
Dr Louise Walsh

Editorial advisors
Dr Mateja Jamnik, Dr Seán Ó hÉigeartaigh, Dr Beth Singler, Dr Adrian Weller

Design
The District

T +44 (0)1223 765 443
E research.horizons@admin.cam.ac.uk
W cam.ac.uk/research

News

Image
Ray Dolby Centre, due to open in 2022

Credit: Jestico + Whiles

£85 million gift for physics

Cambridge receives the largest philanthropic donation ever made to UK science from the estate of Ray Dolby, the man "who changed the way the world listened".

The Dolby family gift will support Cambridge's Cavendish Laboratory, the world-leading centre for physics research where Ray Dolby received his PhD in 1961. He went on to invent the Dolby System, an analogue audio encoding system that forever improved the quality of recorded sound.

"This unparalleled gift is a fitting tribute to Ray Dolby's legacy, who changed the way the world listened – his research paved the way for an entire industry," says Cambridge's Vice-Chancellor Professor Stephen Toope. "A century from now, we can only speculate on which discoveries will alter the way we live our lives, and which new industries will have been born in the Cavendish Laboratory, in large part thanks to this extraordinarily generous gift."

A flagship building of the ongoing Cavendish Laboratory redevelopment will be named the Ray Dolby Centre, and a Ray Dolby Research Group will be established to expand research capability and expertise.

"The University of Cambridge played a pivotal role in Ray's life, both personally and professionally," adds Dolby's widow, Dagmar. "At Cambridge and at the Cavendish, he gained the formative education and insights that contributed greatly to his lifelong ground-breaking creativity, and enabled him to start his business."

The new Cavendish Laboratory will also receive a £75 million investment from the government through the Engineering and Physical Sciences Research Council.

"This generous £85 million donation from the Ray Dolby estate along with the £75 million government has already pledged is a testament to the importance of this facility and the UK's leadership in science," says former Science Minister Jo Johnson. "The UK is one of the most innovative countries in the world, and through our Industrial Strategy and additional £2.3 billion investment for research and development we are ensuring our world-class research base goes from strength to strength for years to come."

Newton's papers added to UNESCO register

Annotated copies of *Principia Mathematica* and other papers of Sir Isaac Newton are now among materials preserved for the world.

Held at Cambridge University Library, Newton's scientific and mathematical papers represent one of the most important archives of scientific and intellectual work on universal phenomena. They document the development of his thoughts on gravity, calculus and optics, and reveal ideas worked out through painstaking experiments, calculations, correspondence and revisions.

Now, Newton's Cambridge papers join other papers deemed of global importance on the register of UNESCO's Memory of the World Project, an international initiative that aims to "safeguard the documentary heritage of humanity against collective amnesia, neglect, the ravages of time and climatic conditions, and wilful and deliberate destruction".

The papers include Newton's own copy of the first edition of the *Principia* (1687), covered with his revisions and additions for the second edition; his 'Laboratory Notebook', which includes details of his investigations to understand the nature of colour; and his undergraduate notebook listing expenditure on white wine, wafers, shoestrings and 'a paire of stockings'.

Isaac Newton entered Trinity College as an undergraduate in 1661 and became a Fellow in 1667. In 1669, he became Lucasian Professor of Mathematics in Cambridge, a position he held until 1701.

"Newton's work and life continue to attract wonder and new perspectives on our place in the Universe," says Cambridge University Librarian Jess Gardner. "Cambridge University Library will continue to work with scholars and curators worldwide to make Newton's papers accessible now and for future generations."

News in brief

More information at www.cam.ac.uk/research

18.01.18

AI 'scientist' finds that an ingredient commonly found in toothpaste could be employed as an anti-malarial against drug-resistant strains.

18.12.17

Mindfulness training can help support students at risk of mental health problems, concludes a randomised controlled trial.

Catching the memory thief

One of six centres that make up the UK Dementia Research Institute (DRI) has opened in Cambridge.

The UK DRI is a joint £250 million investment from the Medical Research Council, Alzheimer's Society and Alzheimer's Research UK, and is made up of centres in Cambridge, Cardiff, Edinburgh, King's College, Imperial College London, and the operational hub at University College London.

When complete, over 400 scientists will carry out an integrated programme of research across the DRI. Their mission is to find new ways to diagnose and treat people with dementias – a group of neurodegenerative disorders that includes Alzheimer's disease – and also prevent their onset. These insidious diseases gradually and subtly steal a lifetime of memories, our ability to live independently and eventually our lives.

"Dementia is now the leading cause of death in England and Wales, and the number of people affected will only grow as the population ages," says Professor Giovanna Mallucci, Director of the newly opened Cambridge centre on the Cambridge Biomedical Campus. "Here in Cambridge, our focus is on using interdisciplinary approaches to understand the processes involved in the very earliest stages of neurodegeneration. We want to identify targets that have the greatest potential to stop 'the memory thief' before it does damage."

Film available:
http://bit.ly/2o38zPT

Credit: Cambridge Archaeological Unit

"To Clapham's I go"

Calf's-foot jelly and a tankard of ale? Welcome to the 18th century Starbucks.

Image
Some of the objects discovered in the cellar

Researchers have published details of the largest collection of artefacts ever discovered from an early English coffee house. The establishment, called Clapham's, was on a site now owned by St John's College, Cambridge. But in the mid-to-late 1700s, it was a bustling coffee house – the contemporary equivalent, academics say, of a branch of Starbucks.

Researchers from Cambridge Archaeological Unit (CAU) – part of the Department of Archaeology – uncovered a disused cellar that had been backfilled with unwanted items, possibly at some point during the 1770s. Inside, were more than 500 objects, many in a very good state of preservation, including drinking vessels for tea, coffee and chocolate, serving dishes, clay pipes, animal and fish bones, and 38 teapots.

Clapham's was owned by William and Jane Clapham, who ran it from the 1740s until the 1770s. It was popular with students and townspeople alike, and a verse from a student publication of 1751 attests to its importance as a social centre: "Dinner over, to Tom's or Clapham's I go; the news of the town so impatient to know."

The assemblage has now been used to reconstruct what a visit to Clapham's might have been like, and in particular what its clientele ate and drank. The discovery of 18 jelly glasses plus feet bones from immature cattle led the researchers to conclude that calf's-foot jelly, a popular dish of that era, might well have been a house speciality.

"Coffee houses were important social centres during the 18th century, but relatively few assemblages of archaeological evidence have been recovered," says Craig Cessford, from CAU. "This is the first time that we have been able to study one in such depth."

06.12.17

The fundraising campaign for the University and Colleges passes the £1 billion mark, enabling Cambridge to respond to challenges facing the world.

30.11.17

A £5.4 million Centre for Digital Built Britain will champion the use of digital technologies to plan, build, maintain and use infrastructure better.

23.10.17

Stephen Hawking's PhD thesis is made accessible via the University's Open Access repository – and over 1m people attempt to download it.

PANI, PAHAR

Kempty Falls is crowded with tourists who flock to the nearby Himalayan hill station of Mussoorie during the summer months. This stunning beauty spot lies at the heart of a region beset by an escalating water crisis.

Mussoorie is fed by many different springs. But in recent years the demand for water has outstripped supply capacity in the summer season. Town authorities are facing increasing conflict from communities living outside the settlement who also demand their 'share' of water, such as the *dhobi* who have washed the town's laundry for close to 100 years.

In 2017, photo-journalist Toby Smith and geographer Dr Eszter Kovacs travelled to Mussoorie and five other towns in India and Nepal to explore the dwindling water supplies of the Himalayas and the struggles of local people who depend on them. Drawing on collaborative research at these sites led by Professor Bhaskar Vira, they created a narrative of words and pictures, *Pani, Pahar* (Hindi for waters of the mountains), to tell the story.

"The interdependence of people and ecological processes across these dynamic landscapes is complex and fascinating," says Vira, Director of the University of Cambridge Conservation Research Institute (UCCRI) and also in the Department of Geography. "Working with researchers in Nepal and India, we are looking at the trade-offs between land-use strategies, water availability, and the lives and livelihoods of those who live there."

A key success of the project, say the team, has been the crossover between photo-narration and research. As a result, several themes for further research have become visible across the six small towns: the changes to water sources; the way in which seasonality affects social and ecological systems; the multiple physical, social and political infrastructures that 'count' in the Himalayas; and the rapid pace of urbanisation.

"Mussoorie is a tourist boom-town," adds Smith, who saw not only a huge influx of tourists but also poorly constructed hotels and restaurants. "In an area prone to seismic shift, extreme rainfall and landslip events, this could be a disaster in the making. With prosperity for some, comes pressure for others."

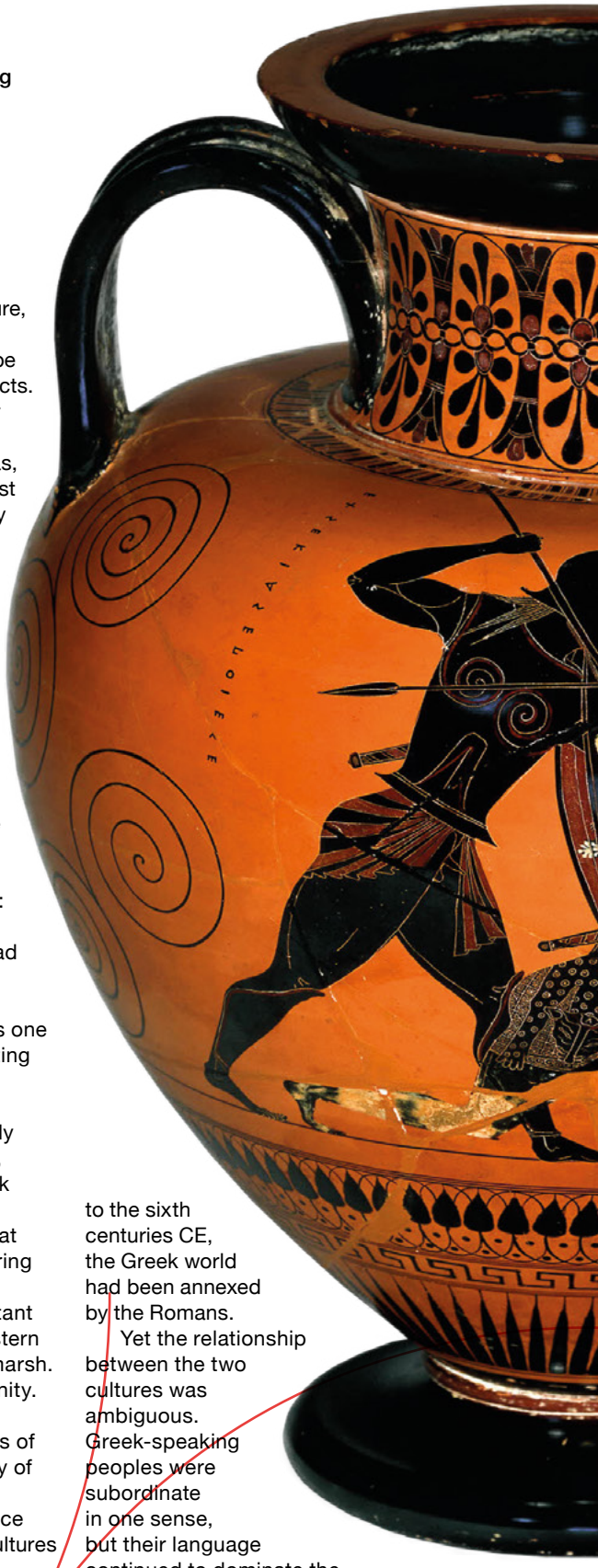
Research supported by the NERC-ESRC-DFID Ecosystem Services for Poverty Alleviation Programme and an Impact Acceleration Account from the Economic and Social Research Council.

www.panipahar.com

Credit: Toby Smith (www.tobysmith.com)

EPIC ISSUES

Epic poems telling of cultures colliding, deeply conflicted identities and a fast-changing world were written by the Greeks under Roman rule in the first to the sixth centuries CE. Now, the first comprehensive study of these vast, complex texts is casting new light on the era that saw the dawn of Western modernity.



Maybe it was the language, architecture, codified legal system, regulated economy, military discipline – or maybe it really was public safety and aqueducts. Whatever the Romans did for us, their reputation as a civilising force who brought order to the western world has, in the public imagination, stood the test of time remarkably well. It is especially strong for an Empire that has been battered by close historical scrutiny for almost 2,000 years.

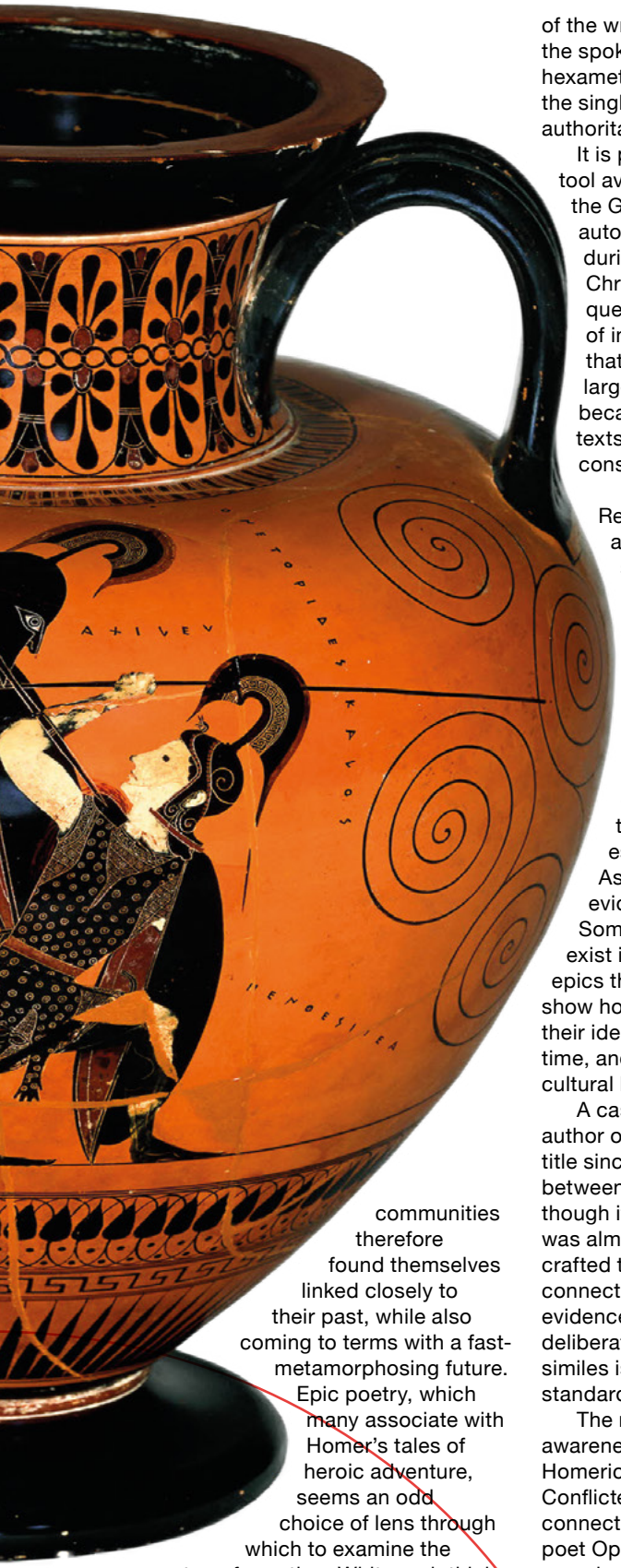
The reputation, of course, has more than a grain of truth to it – but the real story is also more complex. Not only did the Empire frequently endure assorted forms of severely uncultured political disarray, but for the kaleidoscope of peoples under its dominion, Roman rule was a varied experience that often represented an unsettling rupture with the past. As Professor Mary Beard put it in her book *SPQR*: "there is no single story of Rome, especially when the Roman world had expanded far outside Italy."

So perhaps another way to characterise the Roman Empire is as one of cultures colliding – a swirling melting pot of ideas and beliefs from which concepts that would define western civilisation took form. This is certainly closer to the view of Tim Whitmarsh, the A. G. Leventis Professor of Greek Culture at Cambridge, who is the principal investigator on a project that has examined Greek epic poetry during this period.

"This is perhaps the most important period for thinking about where Western civilisation comes from," says Whitmarsh. "We really are at the dawn of modernity. To tell the story of an Empire which remains the model for so many forms of international power is to tell the story of what we became, and what we are."

His interest in the Greek experience stems partly from the fact that few cultures under Roman rule can have felt more keenly the fissure it wrought between present and past. In political terms, Ancient Greek history arguably climaxed with the empires established in the aftermath of the conquests of Alexander the Great (356–323 BCE). In the period when this poetry was written, from the first to the sixth centuries CE, the Greek world had been annexed by the Romans.

Yet the relationship between the two cultures was ambiguous. Greek-speaking peoples were subordinate in one sense, but their language continued to dominate the eastern Empire – increasingly so as it became a separate entity centred on Byzantium, as Christianity emerged and as the Latin-speaking west declined. Greek remained the primary medium of cultural transmission through which these changes were expressed. Greek



Credit: Wine jar made in Athens around 535 BC. © The Trustees of the British Museum

communities therefore found themselves linked closely to their past, while also coming to terms with a fast-metamorphosing future.

Epic poetry, which many associate with Homer's tales of heroic adventure, seems an odd choice of lens through which to examine the transformation. Whitmarsh thinks its purpose has been misunderstood.

"In the modern West, we often get Greek epic wrong by thinking about it as a repository for ripping yarns," he says. "Actually, it was central to their sense of how the world operated. This wasn't a world of scripture; it wasn't primarily one of the written word at all. The vitality of the spoken word, in the very distinctive hexametrical pattern of the poems, was the single way they had of indicating authoritative utterance."

It is perhaps the most important tool available for understanding how the Greeks navigated their loss of autonomy under the Romans and during the subsequent rise of Christianity. In recent years, such questions have provoked a surge of interest in Greek literature during that time, but epic poetry itself has largely been overlooked, perhaps because it involved large, complex texts around which it is difficult to construct a narrative.

Funded by the Arts and Humanities Research Council, Whitmarsh and his collaborators set out to systematically analyse the poetry and its cultural history for the first time. "We would argue it's the greatest gap in ancient cultural studies – one of the last uncharted territories of Greek literature," he adds.

The final outputs will include books and an edited collection of the poems themselves, but the team started simply by establishing "what was out there". Astonishingly, they uncovered evidence of about a thousand texts. Some remain only as names, others exist in fragments; yet more are vast epics that survive intact. Together, they show how the Greeks were rethinking their identity, both in the context of the time, and that of their own past and its cultural legacy.

A case in point is Quintus of Smyrna, author of the *Posthomerica* – a deceptive title since chronologically it fills the gap between Homer's *Iliad* and *Odyssey*, even though it was written later. Quintus' style was almost uber-Homeric, elaborately crafted to create an almost seamless connection with the past. Yet there is evidence that, having done so, he also deliberately disrupted it. "His use of similes is quite outrageous by Homer's standards, for example," Whitmarsh says.

The reason could be Quintus' painful awareness of a tension between the Homeric past and his own present. Conflicted identity is a theme that connects many poems of the period. The poet Oppian, for instance, who wrote an epic on fish and fishing, provides us with an excellent example of how his generation was seeking to reconceive Greek selfhood in the shadow of Rome.

The work ostensibly praises the Emperor as master over land and sea – a very Roman formula. Oppian then sabotages his own proclamation by questioning whether anyone truly can command the sea's depths, a feat that must surely be a journey of the intellect and imagination. Having acknowledged the Emperor's political power, he was, in effect, implying that the Greeks were perhaps greater masters of knowledge.

The researchers expected to find that this tension gave way to a clearer, moralistic tone, with the rise of Christianity. Instead, they found it persisted. Nonnus of Panopolis, for example, wrote 21 books paraphrasing the Gospel of St John, but not, it would seem, from pure devotion, since he also wrote 48 freewheeling stories about the Greek god Dionysus. Collectively, this vast assemblage evokes parallels between the two, not least because resurrection themes emerge from both. Nonnus also made much of the son of God's knack for turning water into wine – a subject that similarly links him to Dionysus, god of winemaking.

Beyond Greek identity itself, the poetry hints at shifting ideas about knowledge and human nature. Oppian's poetic guide to fishing, for instance, is in fact much more. "I suspect most fishermen and fisherwomen know how to catch fish without reading a Greek epic poem," Whitmarsh observes. In fact, the poem was as much about deliberately stretching the language conventionally used to describe aquaculture, and through it blurring the boundaries between the human and non-human worlds.

Far from just telling stories, then, these epic poems show how, in an era of deeply conflicted identities, Greek communities tried to reorganise their sense of themselves and their place in the world, and give this sense a basis for future generations. Thanks to Whitmarsh and his team, they can now be read, as they were meant to be, on such terms.

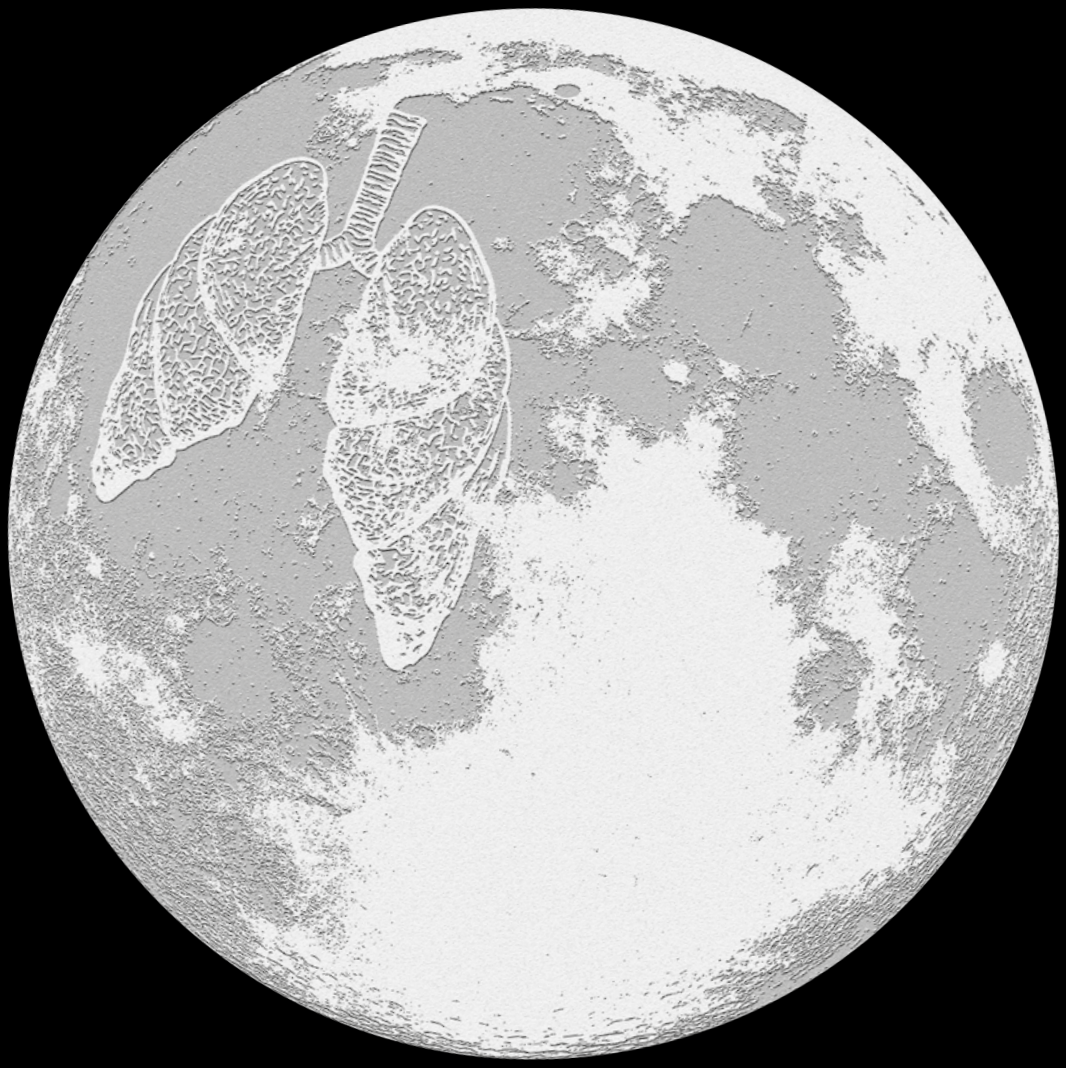
"The poetry represents a cultural statement from the time, but it is also trying to be timeless," he adds. "Each poem was trying to say something about its topic for eternity. The fact that we are still reading them today, and finding new things to say about them, is a token of their success."

Professor Tim Whitmarsh
Faculty of Classics
tjgw100@cam.ac.uk

Words
Tom Kirk

Image
Painting on a wine jar of Achilles killing Penthesilea, as described in the epic poem *Posthomerica* written by Quintus of Smyrna in the third century CE

TAKING A MOON SHOT AT CYSTIC FIBROSIS



Words
Craig Brierley

Almost 30 years on from the discovery of the genetic defect that causes cystic fibrosis, treatment options are still limited and growing antibiotic resistance presents a grave threat. Now, a team of researchers from across Cambridge hopes to turn fortunes around, thanks to a major new centre supported by the Cystic Fibrosis Trust.

John Winn's office at Microsoft Research looks like that of any typical academic: on one wall is a whiteboard graffitied with impenetrable equations and mathematical scribblings, on the opposite wall books and files line shelves, and on his desk are photos of his family.

His desk is somewhat different: it can rise or fall, depending on whether he wants to work standing or sitting – and underneath is a treadmill for walking and working at the same time. "There have been times when I've been deep in thought and almost fallen off it," he jokes.

Winn has cystic fibrosis (CF) and keeping fit is an important part of managing his condition: the stronger his lung function, the better equipped he is to fight the potentially life-threatening infections that plague people living with the condition.

CF occurs when an individual inherits two copies of a single genetic variant, one from each parent. The disease causes a build-up of thick, sticky mucus in the lungs, intestines and organs, and those affected by the condition are particularly susceptible to lung infections leading to progressive inflammatory lung damage. Although life expectancy for people with CF has almost doubled in recent decades, it is still significantly below average.

Winn is a machine learning specialist and is using his expertise to fight the

condition that affects his everyday life. Together with Professor Andres Floto from Cambridge's Department of Medicine, he is turning data from the daily lives of people with CF into potentially life-saving information.

As part of this study, funded by the Cystic Fibrosis Trust and Papworth Hospital, participants have been submitting data – everything from heart rate and lung function through to self-reported wellbeing – via an app that also monitors their activity levels. Machine learning then sifts through the data, looking for patterns and – it's hoped – builds a model that can predict when a patient's health is about to deteriorate and advise them to seek medical help.

"The overarching principle is about giving people control over their own health data and making it work for them," says Winn. "There's some informal feedback that just participating in the study and taking these readings has already improved health outcomes for some individuals: for example, it's helped with adherence with taking their medications as they noticed that if they missed taking certain medicines, their readings got worse."

The project is one strand of research at a major new Cystic Fibrosis Innovation Hub based on the Cambridge Biomedical Campus and run by Floto. The Hub is supported through a £5 million commitment from the Cystic Fibrosis Trust and matching funds from the University of Cambridge. It will strengthen existing collaborations across the University and with the Wellcome Sanger Institute, as well as build new collaborative research networks with CF centres around the UK. The Trust's Chief Executive, David Ramsden, said it will "provide a step change in CF research across the country".

Floto agrees with this sentiment: "We have an opportunity to uplift UK CF research in general by providing knowhow, training and reagents in a number of areas including genomics, bioinformatics, stem cells and clinical trials technology."

A major part of the Hub's activities will be the development of new drugs that target chronic inflammation in CF, in collaboration with the pharmaceutical company GSK as part of the GSK/Cambridge Strategic Partnership, as well as new antibiotic therapy for the main causes of lung infection in the condition.

Finding new drugs against these bacteria is becoming increasingly urgent – Floto and Professor Julian Parkhill at the Sanger recently showed that *Mycobacterium abscessus*, the pathogen behind one of the most serious infections, is becoming increasingly multi-drug resistant and spreading globally. This is one reason why people with CF are advised not to meet each other.

"Clearly the techniques that we develop – and the drug-like molecules that come out of it – will have more general applicability to patients with other multi-drug resistant infections," Floto says. This will be welcome news to England's Chief Medical Officer, Professor Dame Sally Davies, who has warned of a future where "any one of us could go into hospital in 20 years for minor surgery and die because of an ordinary infection that can't be treated by antibiotics."

The timing of all this is particularly good: Papworth Hospital, whose Adult Cystic Fibrosis Centre has gained a national and international reputation for its treatment of patients and its contribution to research, is due to move to the Biomedical Campus later in 2018.

It's almost 30 years since the gene that causes CF was discovered… it's time to take this shot at the moon

The CF wards will feature state-of-the-art air flow systems, designed with Floto's work on the spread of multi-drug resistant CF pathogens in mind.

This close proximity between the patients and the researchers will help Floto test the new treatments he is pioneering. He is particularly excited about the potential for new cellular therapies he's developing with Professor Ludovic Vallier at the Wellcome-MRC Cambridge Stem Cell Institute. Floto describes these as their "moon shot". These would involve taking cells from a CF patient, re-programming them – correcting the genetic defect along the way – and then re-injecting them into patients. "This could provide a way to regenerate damaged lungs," he says.

Floto knows his plans for the Hub are ambitious, but given that it's almost 30 years since the gene that causes CF was discovered and there is still no cure for the disease, he believes it's time to take this shot at the moon.

Floto's collaborators in the CF Innovation Hub include Chris Abell (Chemistry), Sir Tom Blundell (Biochemistry), Julian Parkhill and Ludovic Vallier.

A 'no-strings-attached' relationship

Professor Clare Bryant, like Floto, works on an inflammatory lung disease as part of the GSK/Cambridge Strategic Partnership: in her case, chronic obstructive pulmonary disease (COPD).

COPD is a condition caused by smoking, pollution and severe asthma. Bryant is looking in particular at how COPD makes the lungs 'stickier' to bacteria, increasing the risk of infections.

She holds two grants under the GSK/Cambridge Strategic Partnership, which aims to develop the next wave of 'game-changing' medicines by bringing academic and industrial expertise together to tackle often intractable disease. Based at Cambridge's Department of Veterinary Medicine, Bryant currently has a three-day-a-week sabbatical at GSK's headquarters in Stevenage.

The three-year sabbatical provides Bryant with three postdocs, one PhD student and a budget, with access to GSK resources, but with "no strings attached". The only proviso is that if she works with a GSK reagent, they have first rights on what she does with it. Crucially, she says, it gives her "the space to think".

Bryant is embedded in GSK's Respiratory Drug Discovery Unit and attends its lab meeting every week. "I've met really smart, clever scientists at GSK, with different skills to those of us in academia," she says. "I get to see all aspects of what happens at GSK, everything from how a target is identified, to how drugs are developed to target it, through to taking these drugs to clinical trials. I see the whole spectrum."

It is, though, a mutually beneficial programme, she stresses. Bryant brings her knowledge of innate immunity and her experience of multi-disciplinary collaborations, particularly in imaging. "It's effectively like being a consultant," she says. "I want them to get as much out of me as I do out of them."

Professor Clare Bryant
Department of Veterinary Medicine
ceb27@cam.ac.uk

Professor Andres Floto
Department of Medicine
arf27@cam.ac.uk

Dr John Winn
Microsoft Research
jwinn@microsoft.com

LORD OF THE RINGS

What links a series of volcanic eruptions and severe summer cooling with a century of pandemics, human migrations, political turmoil and the rise and fall of civilisations? Tree rings, says Ulf Büntgen, who leads Cambridge's first dedicated tree-ring laboratory at the Department of Geography.

AD 536: it's been called the year that winter never ended.

"There was a sign in the sun the like of which had never been seen and reported before... The sun became dark and its darkness lasted for eighteen months. Each day it shone for about four hours, and still this light was only a feeble shadow," wrote medieval chronicler Michael the Syrian.

A volcanic eruption had thrown a vast ash cloud into the stratosphere and a dense fog settled over Europe, the Middle East and China. It was a year of failed crops and of famine.

But worse was to follow. A further two volcanic eruptions in 540 and 547 began an unprecedented cooling across much of the northern hemisphere. The thermal shock lasted until around AD 660, making this period the coldest experienced during at least the last two millennia. It is now known as the Late Antique Little Ice Age, or LALIA.

Professor Ulf Büntgen, then at the Swiss Federal Research Institute WSL and now in Cambridge's Department of Geography, was lead author of the study published in *Nature Geoscience* in 2016 that introduced and described the concept of the LALIA. The team of archaeologists, climatologists, geographers and historians was the first to provide independent lines of absolutely dated and annually resolved paleoclimatic evidence for a period of great change that had long perplexed scientists and historians alike.

"The LALIA coincided with a number of extremely important transformation processes in human history," he explains. "We have the outbreak of the Justinian plague across much of the eastern Roman Empire, large-scale migration from inner Eurasia towards Europe and China, turmoil in many parts of central and east Asia, and the collapse of the eastern Türk Empire."

What's remarkable is that much of the evidence for the LALIA comes from witnesses who were alive at the time – trees. The insight is based on the synchronised pattern of ring widths found within different tree species at various sites across the northern hemisphere.

"We believe this exceptionally cold phase from AD 536 to around AD 660 – as recorded by very narrow tree rings – should be considered as a direct or indirect factor in explaining some of the historical events that occurred both in Europe and Asia during that time," says Büntgen.

As distinctive as a fingerprint, the rings formed in trees outside the tropics are annually precise growth layers. Büntgen is an expert at assembling, reading and interpreting these 'slices of time' and, since his arrival in Cambridge in January 2017, has set up the University's first dedicated tree-ring laboratory.

"You ideally start with a living tree," he explains. "This is your anchor point – you know that the outer layer is this year's growth ring, and that the innermost rings take you back to the tree's juvenile growth, with the pith ideally referring to its birth year. You repeat for many trees, using statistical analyses to compare

Credit: Hrafn Óskarsson

and match the pattern with other trees growing at the same time under the same environmental conditions, including climate. Once you've gone back as far as you can with the oldest living tree you look for their dead ancestors."

His team counts rings in the timbers of historical buildings, in subfossil trees preserved in bogs and sediments, and in 'ice-rafted' driftwood washed up on Arctic shores. Back and back they go, comparing and cross-dating, looking for overlaps that provide new anchor points in the 'floating chronology' of patterns. You can see why Büntgen describes dendrochronology as a big data game.

He is currently involved in a collaborative effort by scientists from different disciplines and countries to build the world's longest absolutely dated and continuous tree-ring chronology. The team will hopefully soon be able to add another 2,000 years, taking the record well into the Late Glacial period near the end of the last major Ice Age around 14,000 years ago. This is a huge accomplishment when you consider that a very cold year might result in a ring that's only a single cell wide.

His laboratory is full of further collections of wood ready to be analysed, including numerous disc samples from relict larch trees that were discovered in north-eastern Siberia, where hunters look for mammoth teeth.

"The subfossil wood smells like a fresh tree, yet this material can be thousands of years old," he says. "It's all about preservation. If you take wood from a living tree and put it in anaerobic conditions like a lake or in dense clay everything is preserved. That's why we can ultimately compile multi-millennial-long chronologies for reconstructing past climate variability."

"Once you've gone back as far as you can with the oldest living tree you look for their dead ancestors"

As an environmental scientist, his main interest is in using continuous tree-ring chronologies to reconstruct how the Earth's climate system behaved in the past and to understand how ecosystems were, and are, responding to temperature and hydroclimatic variation.

But a timeline as accurate as this has many other uses, principally in being able to provide a spatially and temporally precise idea of climatic and environmental conditions at key periods in history.

"When you look for links between climate variability and human history you start to build up a multi-dimensional picture of the past," he explains. "But the subject is overwhelmingly approached from within disciplinary silos."

This is why, since his move to Cambridge, Büntgen and colleagues from the Department of Geography have been forging links with historians, archaeologists, earth scientists and plant scientists, to make the most of this remarkable archive.

"Once you embark on these integrative approaches you can ask questions like we did for the LALIA – what was the role of environmental factors in large-scale human migrations and the rise and fall of ancient civilisations? How did complex societies cope with climate change? That's when it starts to get really exciting."

Professor Ulf Büntgen
Department of Geography
ulf.buentgen@geog.cam.ac.uk

Words
Louise Walsh

Image
Drumbabót forest in Iceland

“Little robots”: behind the scenes at an academy school

New research lifts the lid on an influential academy school and finds an authoritarian system that reproduces race and class inequalities.

‘Structure liberates’: the ethos behind one of England’s flagship academy schools. Designed as an engine of social mobility, this school drills ‘urban children’ for the grades and behaviour considered a passport to the world of middle-class salaries and sensibilities.

The headline-grabbing exam results of this school have led politicians to champion its approach as a silver bullet for entrenched poverty, and ‘structure liberates’ has become the blueprint for recent urban education reform.

The school’s recipe has now been replicated many times through academy trusts that have spread like “modern-day missionaries” across the nation, says Dr Christy Kulz, a Leverhulme Research Fellow at Cambridge’s Faculty of Education.

Shortly after it opened, Kulz was granted permission to conduct fieldwork in the school, where she had once worked as a teaching assistant. Choosing to anonymise her research, she calls the school Dreamfields.

Her new book goes behind the scenes of life at Dreamfields, and is the only detailed ethnographic account of the everyday practices within this new breed of academy school.

“Education has long been promoted as a salve that cures urban deprivation and balances capitalism’s inequalities,” says Kulz, who spent 18 months of observation in Dreamfields, interviewing parents, teachers and students

“The academy programme taps into ‘mythical qualities’ of social mobility: some kind of magic formula that provides equal opportunities for every individual once they are within the school, regardless of race, class or social context.” In 2012, then Prime Minister David Cameron described academies as “working miracles”.

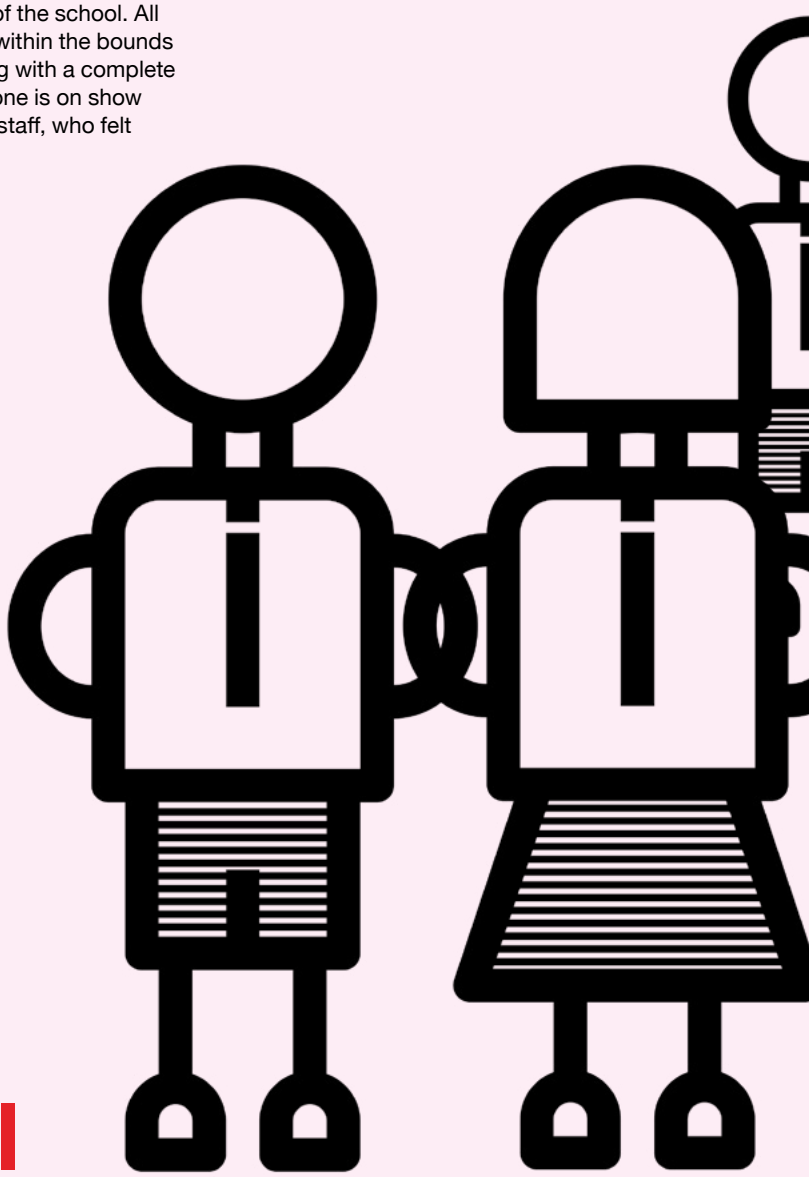
Primarily state funded but run as not-for-profit businesses, sometimes with support from individual philanthropists, academies such as Dreamfields are independent of local authority control and sit outside the democratic process of local government.

The gospel according to Dreamfields’ celebrated head is described as a “traditional approach”. Kulz says she found a stress-ridden hierarchical culture focused on a conveyer belt of testing under strict – almost military – conditions, and suffused with police-style language of ‘investigations’ and ‘repeat offenders’.

Enforcement comes through what Kulz calls the “verbal cane”. Tongue-lashings administered by teachers regularly echoed around the corridors, and were encouraged by senior staff. One teacher told Kulz that seeing tall male members of staff screaming in the faces of 11-year-olds was “very hard to digest”.

This verbal aggression is heightened by the panoptic surveillance built into the very architecture of the school. All activity is conducted within the bounds of a U-shaped building with a complete glass frontage. Everyone is on show at all times, including staff, who felt

She found a stress-ridden hierarchical culture focused on a conveyer belt of testing



constantly monitored and pressured into visibly exerting the discipline favoured by management.

Policing was not confined to within the school gates. Kulz goes on a ride-along with what's known as "chicken-shop patrol". Driving around the streets after school, staff members jump out of the car to intervene when children are deemed to be congregating or in scruffy uniforms.

Stopping off at one of the local takeaways is considered a major offence. "Fried chicken represents a 'poor choice' that Dreamfields must prohibit in order to change urban culture," says Kulz. "Simply being caught in a takeaway after school is punished with a two-hour detention the following day."

Students are also policed through exacting uniform adherence, with a 'broken-window theory' approach that sees deviation as opening the door to chaos.

The smallest rule infraction can be met with a spell in isolated detention.

Staff would sometimes go to strange lengths to maintain conformity, she says. Suede shoes were subject to clampdown. Parental suggestions of a karaoke stall at a winter fair were considered far too risky. "There is no room for unpredictability at Dreamfields," says Kulz. One student who shaved lines into his eyebrows had to have them coloured in by a teacher every morning.

As fieldwork progressed, however, Kulz began to notice discrepancies that tallied uncomfortably with race and social background. Black children with fringes, or children who congregated outside takeaways, were reprimanded immediately. White middle-class children with long floppy hair, or gathering en masse by Tesco, were ignored. Teachers troubled by this would hint at it in hushed tones.

"The approach of many academy schools is one of cultural cloning," says Kulz. "The Dreamfields creed is that 'urban children', a phrase used by staff to mean working-class and ethnic minority kids assumed to have unhappy backgrounds, need salvaging – with middle-class students positioned as the unnamed, normative and universal ideal."

"Black students were consistently more heavily policed in the playground, resulting in many consciously adopting 'whiter' styles and behaviours – a tactic that reduced their surveillance."

It is not just children who are driven hard through incessant monitoring. Staff at Dreamfields are subject to 'teacher tracking', a rolling system in which student grades are converted into scores, allowing management to rank the teachers – an approach staff compared with salesmen being judged on their weekly turnover.

This pressurised auditing resulted in rote learning to avoid a red flag in the system. "You put a grade in that satisfies the system instead of it satisfying the student's knowledge and needs," one teacher lamented to Kulz, explaining his 'real job' was not to teach understanding of his subject, but to get students to produce a set product quickly and accurately. One student described himself to Kulz as a "little robot".

Most teachers exceeded a 48-hour week. The majority of staff were young – an average age of 33 – with fewer outside commitments, yet many expressed a sense of exhaustion. "If you're not in a lesson we are expected to patrol," one teacher told Kulz. "Every moment of every day is taken up with some sort of duty." Unlike most schools, Dreamfields has no staff room.

Some staff discussed former colleagues who had suffered burnout or were asked to resign. During interviews, Kulz found conspiracy theories were rife among students because of the number of teachers that "just disappeared".

Yet Dreamfields was – and still is – fêted by politicians and the media for its undeniably extraordinary exam results: over 80% pass rate at GCSE in an area where this was previously unthinkable. At the time, the school was vastly oversubscribed, with over 1,500 applications for just 200 places.

"Most of the students, parents and teachers were keen to comply to Dreamfields' regime, despite its injustices. The school's approach was seen as the best shot at securing grades and succeeding in an increasingly precarious economy.

"Students, like staff, are trained to be expendable while the ideals of democracy and critical thinking we are allegedly meant to cherish are quashed in the process."

This model of a disciplinarian school built for surveillance and which teaches market-force obedience has marched ever onward since her time in Dreamfields, says Kulz – arriving at new poverty front-lines such as rundown seaside towns.

Yet, grassroots resistance to this style of education is increasing. Last year, a recently established academy in Great Yarmouth that forbade "slouching and talking in corridors" had pupils pulled out by parents objecting to the "draconian" rules that were central to the much-imitated Dreamfields playbook.

Kulz believes the grades achieved by these schools – far from universally high – come at a price. "We cannot continue to ignore the links between the testing regimes we put pupils through, the harsh school cultures they create, and the deteriorating physical and mental health of children and young people in the UK."

'Factories for Learning: Making Race, Class and Inequality in the Neoliberal Academy' (2017) is published by Manchester University Press.

Dr Christy Kulz
Faculty of Education
crk35@cam.ac.uk

Words
Fred Lewsey

THE BODY IN MINIATURE

The past few years have seen an explosion in the number of studies using organoids – so-called mini-organs – as ways of testing drugs. As the field matures, will we also see them being used in personalised medicine and even in transplants?

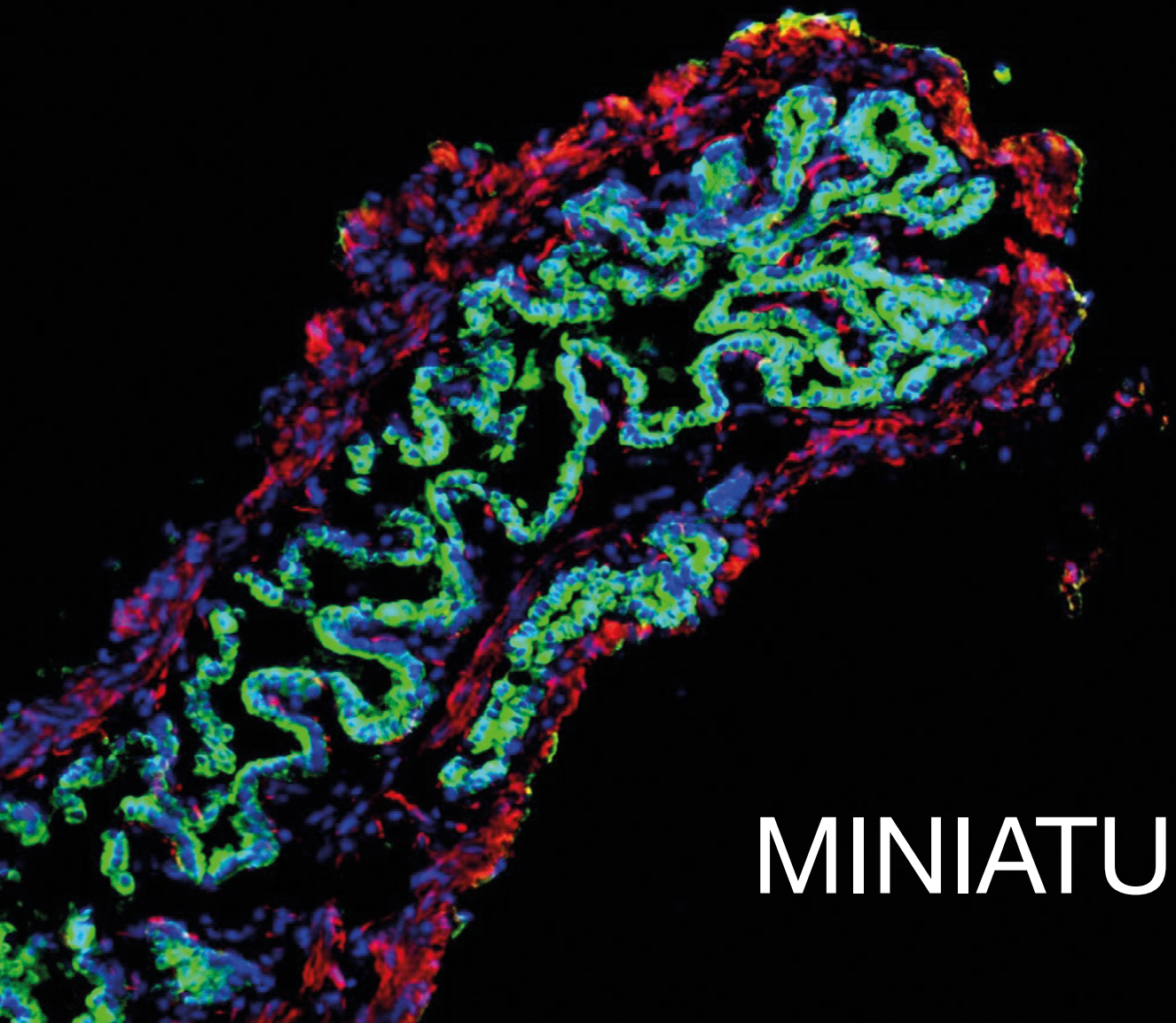
Dr Laura Broutier reaches into the incubator and takes out a culture plate with 24 separate wells, each containing a pale pink liquid. “If you look closely, you can see the dots there,” she says, manipulating the plates until specks the size of a full stop catch the light.

Broutier is a postdoc in Dr Meritxell Huch’s lab at the Wellcome Trust/Cancer Research UK Gurdon Institute, and these “dots” are miniature liver tumours that have been regrown from cancer cells taken from patients at nearby Addenbrooke’s Hospital. They could make it possible to identify cancer drugs personalised for each individual patient.

Huch’s latest work builds on her previous research on ‘mini-livers’, part of a growing body of work – no pun intended – that uses miniature organ-like tissues to understand human biology and in particular why it goes wrong in cancer and dementia. Other research groups in Cambridge are growing mini-brains, mini-oesophaguses, mini-bile ducts, mini-lungs, mini-intestines, mini-wombs, mini-pancreases… Almost the whole body in miniature, it seems.

It’s perhaps a misnomer to call them mini-organs. They look nothing like a miniature organ. Rather, they are ‘organoids’, clusters of cells that can grow and proliferate in culture, taking on a 3D structure that has the same tissue architecture, gene expression and genetic functions as the part of the organ being studied.

The technique that Huch uses involves taking cells from the liver or, in the case of her latest work, liver tumours, and growing these in culture. Her early work involved growing mini-livers from mouse stem cells, but she is now working with human tissue.



"Organoids have opened up a lot of possibilities for us," she says. "They're not 100% identical to the tissue, but they recapitulate many more functions of the tissue of origin, so we can use them to study adult tissue in way that wasn't previously possible."

This ability to use organoids in place of animal models has attracted the interest of the National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs), who currently supports Huch's work and awarded her a 3Rs prize in 2014.

Organoid research has exploded in recent years. Applications include modelling tissue, early development, disease, drug discovery, and now regenerative medicine. Little wonder, then, that *The Scientist* magazine named the technique one of the biggest scientific advancements of 2013; since then, the number of organoid-related scientific papers in the PubMed Central repository has more than doubled to over 1,000 per year.

But, as with any promising new development in research, we must be careful not to oversell it, says Professor Alfonso Martinez-Arias from the Department of Genetics. In some cases, he argues, the research is little more than doing "safaris on culture plates".

Last year, he co-wrote an article in the journal *Development* about the hype surrounding organoids. Despite taking particular exception in the article to claims that scientists in the USA had made the "most complete human brain model to date", he is not as dismissive of the field as one might imagine.

The problem, he says, is one of reproducibility – the same experimental conditions should yield samples that are almost identical in terms of size, shape and composition. This is currently not the case, he says – organoids can often not be grown reliably, forcing researchers to 'cherry pick' the best, and even then (and in contrast with the organism) each one is different.

"Cells in a Petri dish, like children in a playground, will arrange themselves into patterns and some of these will make sense to you. But if we want the system to be reproducible and useful for disease modelling, drug screening or understanding basic mechanisms, we need to steer them and ensure that if an experiment starts with one hundred groups of cells, we end up with one hundred almost identical organoids."

Martinez-Arias's own work is on gastruloids – the same concept as organoids, but used to model very early stages of embryonic development. Working closely with physicists and engineers, his team has managed to generate gastruloids using mouse cells that are highly reproducible.

In the same edition of *Development,* Huch co-wrote a counterpoint to Martinez-Arias's article, about the hope surrounding organoids, but she agrees with Martinez-Arias that much of the research to date has been merely descriptive. "It has been 'Oh, we can do this and we can grow this', but little has been shown about what we can learn."

This, she says, is how her recent study on liver tumours – "tumouroids" as she calls them – differs. "We've shown not only that we can grow them, but what we can do with them."

Huch recently published a proof-of-principle that it's possible to derive mini-tumours in culture from a patient's own cells against which drugs can be tested to find the most effective treatment for that patient – so-called personalised medicine.

mini-brains, mini-oesophaguses, mini-bile ducts, mini-lungs, mini-intestines, mini-wombs, mini-pancreases… almost the whole body in miniature, it seems

Such work can currently only be done by transplanting tumour tissue into mice, growing it over several months and testing the drugs on the mouse – time-consuming and technically limiting. Imagine, she says, being able to screen hundreds – even thousands – of drugs at a time on the mini-liver tumours. Clearly this would be neither practical nor ethical in animals.

"Whether it can be done economically and practically on an individual patient basis, time will tell," she says. "I think, as with everything, once the technology has become cheaper, it will be feasible."

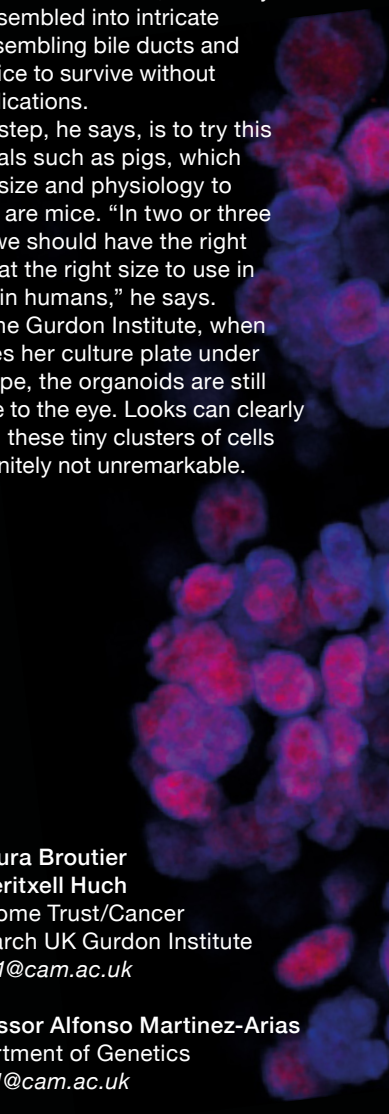
It is tempting to speculate that if scientists can grow organoids in the lab, they will soon be able to grow fully functioning organs. But Huch believes we are nowhere near this stage. More feasible is the idea of using organoids to replace damaged or diseased tissue – or more accurately, to help such tissue 'regenerate'. This is one area of research being pursued by Professor Ludovic Vallier from the Wellcome-MRC Cambridge Stem Cell Institute (also a winner of a 3Rs prize in 2011).

Earlier this year, Vallier succeeded in using biliary organoids to reconstruct the common bile duct, a pipe linking the liver to the gut. It carries bile, which contains all the toxins produced by the liver and is also essential for helping us digest food. If it's damaged, for example in the childhood disease biliary atresia, this can lead to accumulation of toxic bile in the liver and ultimately liver failure.

Vallier and colleagues extracted healthy cells from mouse bile ducts and grew these into functioning 3D duct structures known as biliary organoids. But it was the next step that makes this so significant: they then rebuilt a common bile duct with the help of bioengineers Dr Athina Markaki and Alex Justin. When transplanted into mice, the biliary organoids assembled into intricate structures resembling bile ducts and helped the mice to survive without further complications.

The next step, he says, is to try this in large animals such as pigs, which are closer in size and physiology to humans than are mice. "In two or three years' time, we should have the right biomaterials at the right size to use in clinical trials in humans," he says.

Back at the Gurdon Institute, when Broutier slides her culture plate under the microscope, the organoids are still unremarkable to the eye. Looks can clearly be deceptive: these tiny clusters of cells are most definitely not unremarkable.



Credit: left, Ludovic Vallier; right, Meritxell Huch

Dr Laura Broutier
Dr Meritxell Huch
Wellcome Trust/Cancer Research UK Gurdon Institute
mh771@cam.ac.uk

Professor Alfonso Martinez-Arias
Department of Genetics
ama11@cam.ac.uk

Professor Ludovic Vallier
Wellcome-MRC Cambridge Stem Cell Institute
lv225@cam.ac.uk

Words
Craig Brierley

Things

"The role of art is to give food for thought…"
Jim Ede

Kettle's Yard – Cambridge University's unique modern art gallery – has re-opened after an ambitious refurbishment. Its new research facilities will help scholars discover why its founder, Jim Ede, believed "there should be a Kettle's Yard in every university."

Until 1973, Kettle's Yard was the home of Jim Ede, a former curator of London's Tate Gallery, and his wife Helen. Today it comprises a house containing his remarkable art collection and a modern art gallery that has now been enlarged, providing extra exhibition space to host major international artists and also education rooms.

A brand new research space and archive will enable scholars to study Ede's personal correspondence – amounting to thousands of letters with prominent artists such as Alfred Wallis, Ben and Winifred Nicholson, Joan Miró, Henri Gaudier-Brzeska, Henry Moore, Barbara Hepworth and Constantin Brancusi.

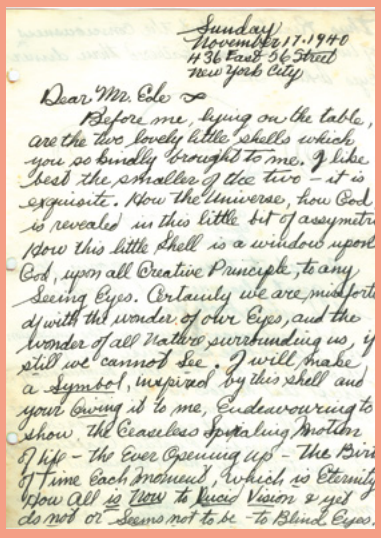
"The archive is forever delivering little surprises," says archivist Frieda Midgley. "Not many people know, for example, that Jim Ede struck up a long-running correspondence with T. E. Lawrence ('Lawrence of Arabia'), or that the collection includes a monogrammed section of one of the artist Christopher Wood's shirts."

Also within the archive are 40 years of correspondence between Ede and American artist Richard Pousette-Dart – a contemporary of Jackson Pollock and Mark Rothko. The letters are being studied by Dr Jennifer Powell – Head of Collections and Programme at Kettle's Yard and lecturer in the Department of History of Art – to provide new insight for a forthcoming exhibition.

Adds Midgley: "If Kettle's Yard is the ultimate expression of a way of life developed over 50 years and more, the archive adds an extra dimension by documenting the rich story of how that philosophy evolved."

The refurbishment was principally funded by the Arts Council England and the Heritage Lottery Fund.

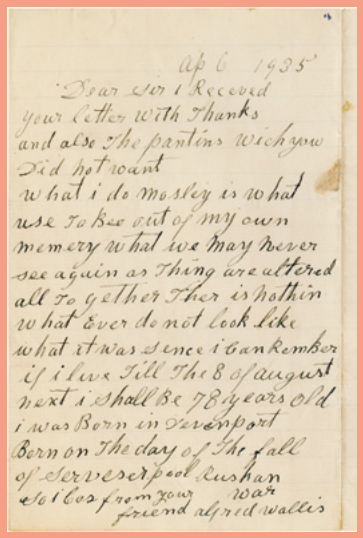
www.kettlesyard.co.uk



Sunday
November 17. 1940
436 East 56 Street
New York City

Dear Mr. Ede

Before me, lying on the table, are the two lovely little shells which you so kindly brought to me. I like best the smaller of the two – it is exquisite. How the Universe, how God is revealed in this little bit of assymetry. How this little shell is a window upon God, upon all Creative Principle, to any Seeing Eyes. Ortainly we are mis[illegible] of with the wonder of our Eyes, and the wonder of all Nature surrounding us, if still we cannot See. I will make a Symbol, inspired by this shell and your Giving it to me, Endeavouring to show the Ceaseless Spiraling Motion of life – the Ever Opening up – The Birth of Time Each Moment, which is Eternity. How All is New to Lucid Vision & yet do not or Seems not to be – to Blind Eyes.

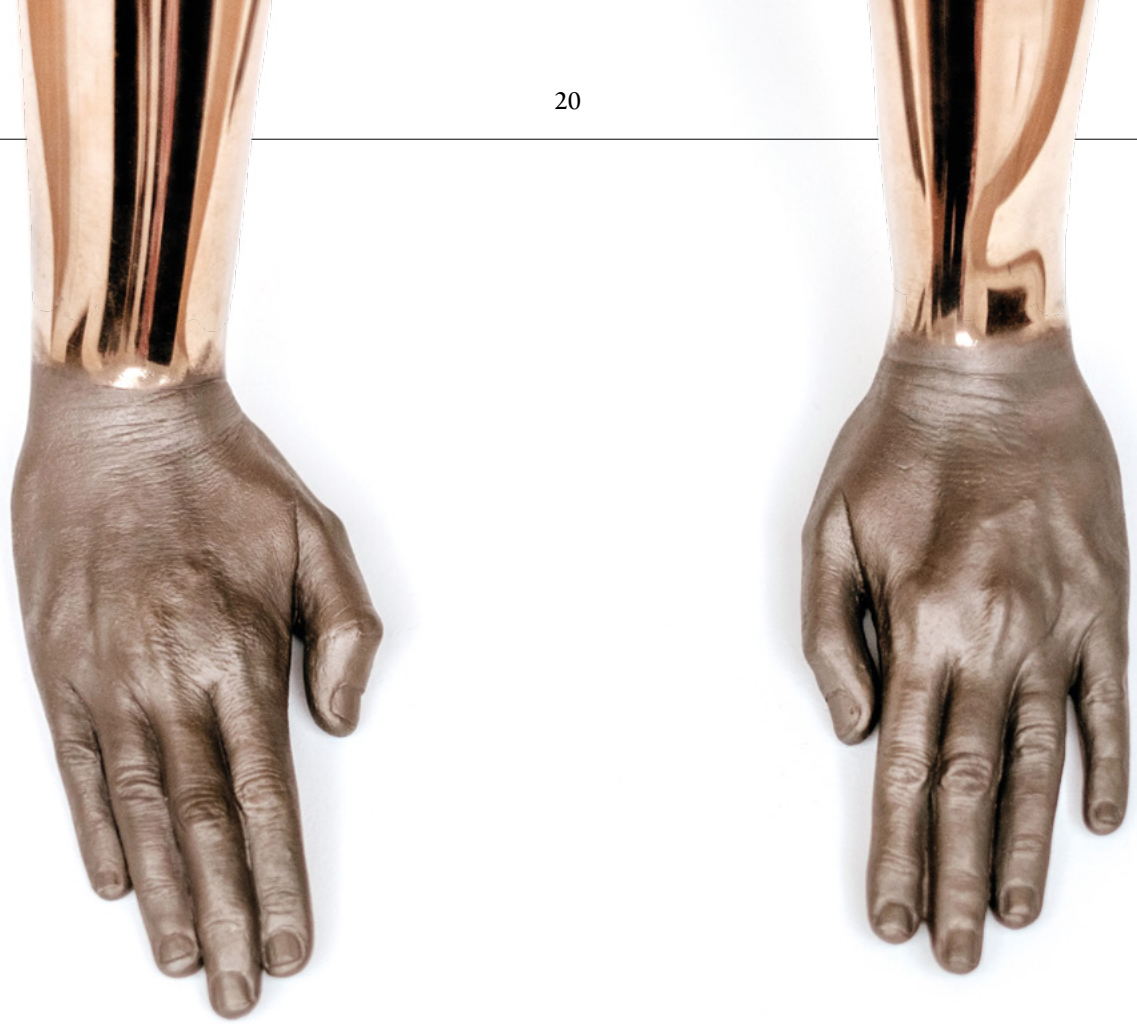


Ap 6 1935

Dear Sir I Received your Letter With Thanks and also The paintins Wich you Did not want what i do mostly is what use To Bee out of My own memery what we may never see again as Thing are altered all To gether Ther is nothin what ever do not look like what it was Since i Can Remember if i live Till The 8 of august next i shall Be 78 years old i was Born in Devenport Born on The day of The fall of Serveserpool Rushan war

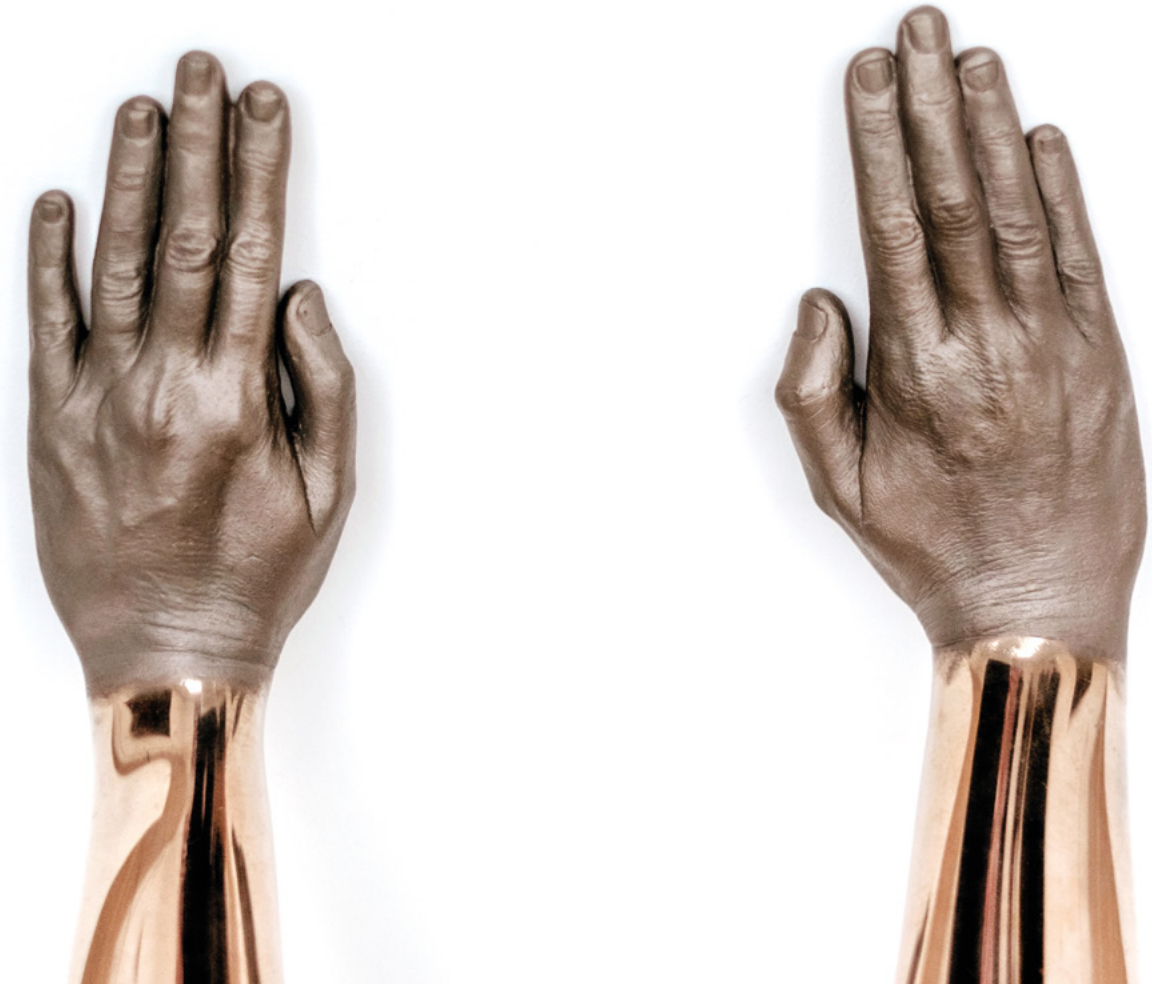
Sollos from you friend alfred wallis

Credit: Clockwise from top left, *Five ships – Mount's Bay*, 1928 (circa) by Alfred Wallis; Jim Ede; the new Kettle's Yard; letters among Ede's archive. All images, Kettle's Yard.



LIVING WITH AI

Powerful AI needs to be reliably aligned with human values. Does this mean that AI will eventually have to police those values? Philosophers Huw Price and Karina Vold consider the trade-off between safety and autonomy in the era of superintelligence.



This has been the decade of AI, with one astonishing feat after another. A chess-playing AI that can defeat not only all human chess players, but also all previous human-programmed chess machines, after learning the game in just four hours? That's yesterday's news, what's next?

True, these prodigious accomplishments are all in so-called narrow AI, where machines perform highly specialised tasks. But many experts believe this restriction is very temporary. By mid-century, we may have artificial general intelligence (AGI) – machines that are capable of human-level performance on the full range of tasks that we ourselves can tackle.

If so, then there's little reason to think that it will stop there. Machines will be free of many of the physical constraints on human intelligence. Our brains run at slow biochemical processing speeds on the power of a light bulb, and need to fit through a human birth canal. It is remarkable what they accomplish, given these handicaps. But they may be as far from the physical limits of thought as our eyes are from the Webb Space Telescope.

Once machines are better than us at designing even smarter machines, progress towards these limits could accelerate. What would this mean for us? Could we ensure a safe and worthwhile coexistence with such machines?

On the plus side, AI is already useful and profitable for many things, and super AI might be expected to be super useful, and super profitable. But the more powerful AI becomes, the more we ask it to do for us, the more important it will be to specify its goals with great care. Folklore is full of tales of people who ask for the wrong thing, with disastrous consequences – King Midas, for example, who didn't really want his breakfast to turn to gold as he put it to his lips.

So we need to make sure that powerful AI machines are 'human-friendly' – that they have goals reliably aligned with our own values. One thing that makes this task difficult is that by the standards we want the machines to aim for, we ourselves do rather poorly. Humans are far from reliably human-friendly. We do many terrible things to each other and to many other sentient creatures with whom we share the planet. If superintelligent machines don't do a lot better than us, we'll be in deep trouble. We'll have powerful new intelligence amplifying the dark sides of our own fallible natures.

For safety's sake, then, we want the machines to be ethically as well as cognitively superhuman. We want them to aim for the moral high ground, not for the troughs in which many of us spend some of our time. Luckily they'll have the smarts for the job. If there are routes to the uplands, they'll be better than us at finding them, and steering us in the right direction. They might be our guides to a much better world.

However, there are two big problems with this utopian vision. One is how we get the machines started on the journey, the other is what it would mean to reach this destination.

The 'getting started' problem is that we need to tell the machines what they're looking for with sufficient clarity and precision that we can be confident that they will find it – whatever 'it' actually turns out to be. This is a daunting challenge, given that we are confused and conflicted about the ideals ourselves, and different communities might have different views.

The 'destination' problem is that, in putting ourselves in the hands of these moral guides and gatekeepers, we might be sacrificing our own autonomy – an important part of what makes us human.

Just to focus on one aspect of these difficulties, we are deeply tribal creatures. We find it very easy to ignore the suffering of strangers, and even to contribute to it, at least indirectly. For our own sakes, we should hope that AI will do better. It is not just that we might find ourselves at the mercy of some other tribe's AI, but that we could not trust our own, if we had taught it that not all suffering matters. This means that as tribal and morally fallible creatures, we need to point the machines in the direction of something better. How do we do that? That's the getting started problem.

> "we want the machines to be ethically as well as cognitively superhuman"

As for the destination problem, suppose that we succeed. Machines who are better than us at sticking to the moral high ground may be expected to discourage some of the lapses we presently take for granted. We might lose our freedom to discriminate in favour of our own tribes, for example.

Loss of freedom to behave badly isn't always a bad thing, of course: denying ourselves the freedom to keep slaves, or to put children to work in factories, or to smoke in restaurants are signs of progress. But are we ready for ethical overlords – sanctimonious silicon curtailing our options? They might be so good at doing it that we don't notice the fences; but is this the future we want, a life in a well-curated moral zoo?

These issues might seem far-fetched, but they are already on our doorsteps. Imagine we want an AI to handle resource allocation decisions in our health system, for example. It might do so much more fairly and efficiently than humans can manage, with benefits for patients and taxpayers. But we'd need to specify its goals correctly (e.g. to avoid discriminatory practices), and we'd be depriving some humans (e.g. senior doctors) of some of the discretion they presently enjoy. So we already face the getting started and destination problems. And they are only going to get harder.

This isn't the first time that a powerful new technology has had moral implications. Speaking about the dangers of thermonuclear weapons in 1954, Bertrand Russell argued that to avoid wiping ourselves out "we have to learn to think in a new way". He urged his listener to set aside tribal allegiances and "consider yourself only as a member of a biological species... whose disappearance none of us can desire."

We have survived the nuclear risk so far, but now we have a new powerful technology to deal with – itself, literally, a new way of thinking. For our own safety, we need to point these new thinkers in the right direction, and get them to act well for us. It is not yet clear whether this is possible, but if so it will require the same cooperative spirit, the same willingness to set aside tribalism, that Russell had in mind.

But that's where the parallel stops. Avoiding nuclear war means business as usual. Getting the long-term future of life with AI right means a very different world. Both general intelligence and moral reasoning are often thought to be uniquely human capacities. But safety seems to require that we think of them as a package: if we are to give general intelligence to machines, we'll need to give them moral authority, too. That means a radical end to human exceptionalism.

All the more reason to think about the destination now, and to be careful about what we wish for.

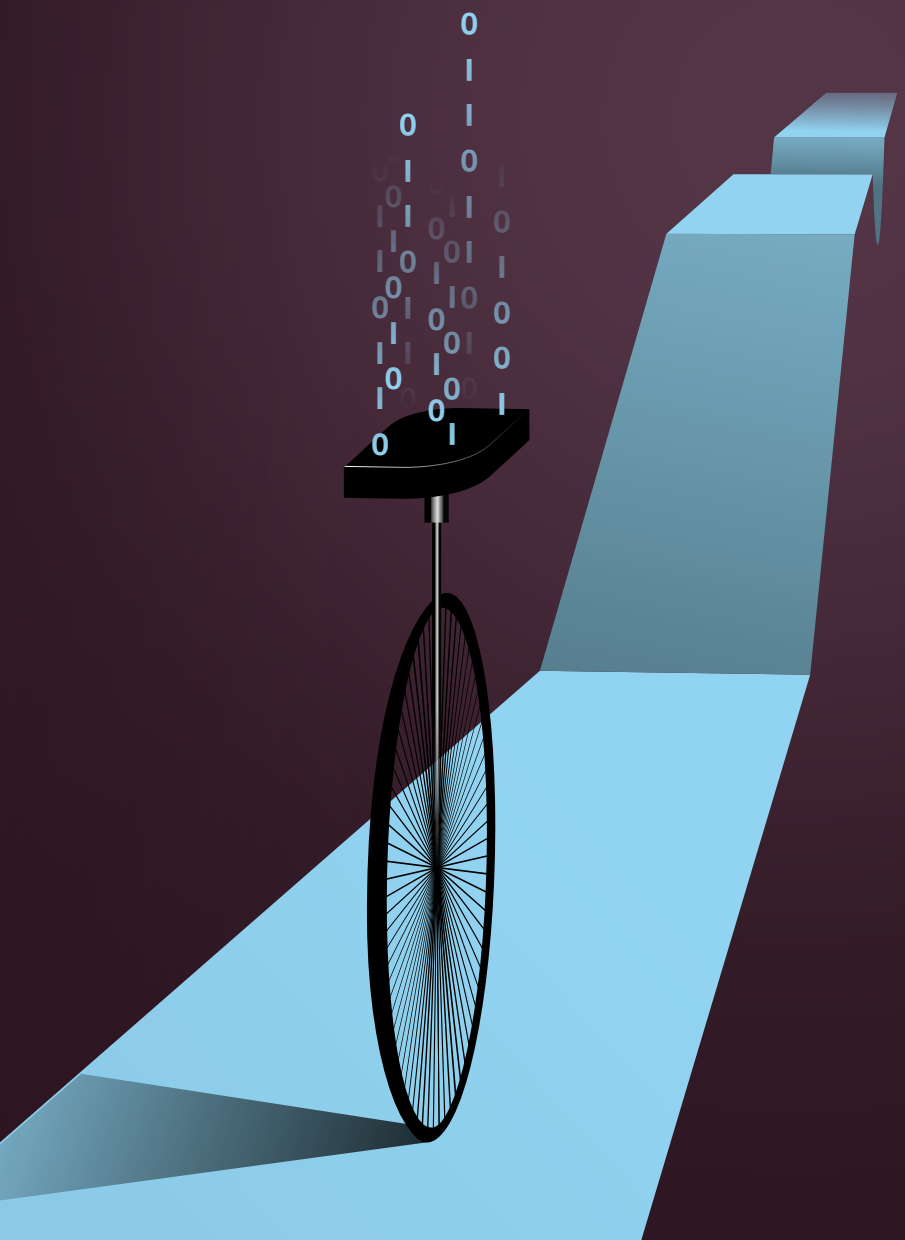
Professor Huw Price
Faculty of Philosophy and the Leverhulme Centre for the Future of Intelligence (CFI)
hp331@cam.ac.uk

Dr Karina Vold
Faculty of Philosophy and CFI
kvv22@cam.ac.uk

Words
Huw Price and Karina Vold

Credit: Sculpture by Rona Pondick

The uncertain unicycle that taught itself and how it's helping AI make good decisions



Words
Louise Walsh

Cambridge researchers are pioneering a form of machine learning that starts with only a little prior knowledge and continually learns from the world around it.

In the centre of the screen is a tiny unicycle. The animation starts, the unicycle lurches forward and falls. This is trial #1.

It's now trial #11 and there's a change – an almost imperceptible delay in the fall, perhaps an attempt to right itself before the inevitable crash. "It's learning from experience," nods Professor Carl Edward Rasmussen.

After a minute, the unicycle is gently rocking back and forth as it circles on the spot. It's figured out how this extremely unstable system works and has mastered its goal. "The unicycle starts with knowing nothing about what's going on – it's only been told that its goal is to stay in the centre in an upright fashion. As it starts falling forwards and backwards, it starts to learn," explains Rasmussen, who leads the Computational and Biological Learning Lab in the Department of Engineering. "We had a real unicycle robot but it was actually quite dangerous – it was strong – and so now we use data from the real one to run simulations, and we have a mini version."

Rasmussen uses the self-taught unicycle to demonstrate how a machine can start with very little data and learn dynamically, improving its knowledge every time it receives new information from its environment. The consequences of adjusting its motorised momentum and balance help the unicycle to learn which moves were important in helping it to stay upright in the centre.

"This is just like a human would learn," explains Professor Zoubin Ghahramani, who leads the Machine Learning Group in the Department of Engineering. "We don't start knowing everything. We learn things incrementally, from only a few examples, and we know when we are not yet confident in our understanding."

Ghahramani's team is pioneering a branch of AI called continual machine learning. He explains that many of the current forms of machine learning are based on neural networks and deep learning models that use complex algorithms to find patterns in vast datasets. Common applications include translating phrases into different languages, recognising people and objects in images, and detecting unusual spending on credit cards.

"These systems need to be trained on millions of labelled examples, which takes time and a lot of computer memory," he explains. "And they have flaws. When you test them outside of the data they were trained on they tend to perform poorly. Driverless cars, for instance, may be trained on a huge dataset of images but they might not be able to generalise to foggy conditions.

"Worse than that, the current deep learning systems can sometimes give us confidently wrong answers, and provide limited insight into why they have come to particular decisions. This is what bothers me. It's okay to be wrong but it's not okay to be confidently wrong."

The key is how you deal with uncertainty – the uncertainty of messy and missing data, and the uncertainty of predicting what might happen next. "Uncertainty is not a good thing – it's something you fight, but you can't fight it by ignoring it," says Rasmussen. "We are interested in representing the uncertainty."

It turns out that there's a mathematical theory that tells you what to do. It was first described by 18th-century English statistician Thomas Bayes. Ghahramani's group was one of the earliest adopters in AI of Bayesian probability theory, which describes how the probability of an event occurring (such as staying upright in the centre) is updated as more evidence (such as the decision the unicycle last took before falling over) becomes available.

Dr Richard Turner explains how Bayes' rule handles continual learning: "the system takes its prior knowledge, weights it by how accurate it thinks that knowledge is, then combines it with new evidence that is also weighted by its accuracy.

"This is much more data-efficient than the way a standard neural network works," he adds. "New information can cause a neural network to forget everything it learned previously – called catastrophic forgetting – meaning it needs to look at all of its labelled examples all over again, like relearning the rules and glossary of a language every time you learn a new word.

"Our system doesn't need to revisit all the data it's seen before – just like humans don't remember all past experiences; instead we learn a summary and we update it as things go on."

Ghahramani adds: "The great thing about Bayesian machine learning is the system makes decisions based on evidence – it's sometimes thought of as 'automating the scientific method' – and because it's based on probability, it can tell us when it's outside its comfort zone."

Ghahramani is also Chief Scientist at Uber. He sees a future where machines are continually learning not just individually but as part of a group. "Whether it's companies like Uber optimising supply and demand, or autonomous vehicles alerting each other to what's ahead on the road, or robots working together to lift a heavy load – cooperation, and sometimes competition, in AI will help solve problems across a huge range of industries."

One of the really exciting frontiers is being able to model probable outcomes in the future, as Turner describes. "The role of uncertainty becomes very clear when we start to talk about forecasting future problems such as climate change."

Turner is working with climate scientists Dr Emily Shuckburgh and Dr Scott Hosking at the British Antarctic Survey to ask whether machine learning techniques can improve understanding of climate change risks in the future.

"We need to quantify the future risk and impacts of extreme weather at a local scale to inform policy responses to climate change," explains Shuckburgh. "The traditional computer simulations of the climate give us a good understanding of the average climate conditions. What we are aiming to do with this work is to combine that knowledge with observational data from satellites and other sources to get a better handle on, for example, the risk of low-probability but high-impact weather events."

"It's actually a fascinating machine learning challenge," says Turner, who is helping to identify which area of climate modelling is most amenable to using Bayesian probability. "The data are extremely complex, and sometimes missing and unlabelled. The uncertainties are rife."

One significant element of uncertainty is the fact that the predictions are based on our future reduction of emissions, the extent of which is as yet unknown.

"An interesting part of this for policy makers, aside from the forecasting value, is that you can imagine having a machine that continually learns from the consequences of mitigation strategies such as reducing emissions – or the lack of them – and adjusts its predictions accordingly," adds Turner.

What he is describing is a machine that – like the unicycle – feeds on uncertainty, learns continuously from the real world, and assesses and then reassesses all possible outcomes. When it comes to climate, however, it's also a machine of all possible futures.

Professor Zoubin Ghahramani
Department of Engineering
zg201@eng.cam.ac.uk

Professor Carl Edward Rasmussen
Department of Engineering
cer54@eng.cam.ac.uk

Dr Emily Shuckburgh
British Antarctic Survey
emsh@bas.ac.uk

Dr Richard Turner
Department of Engineering
ret26@eng.cam.ac.uk

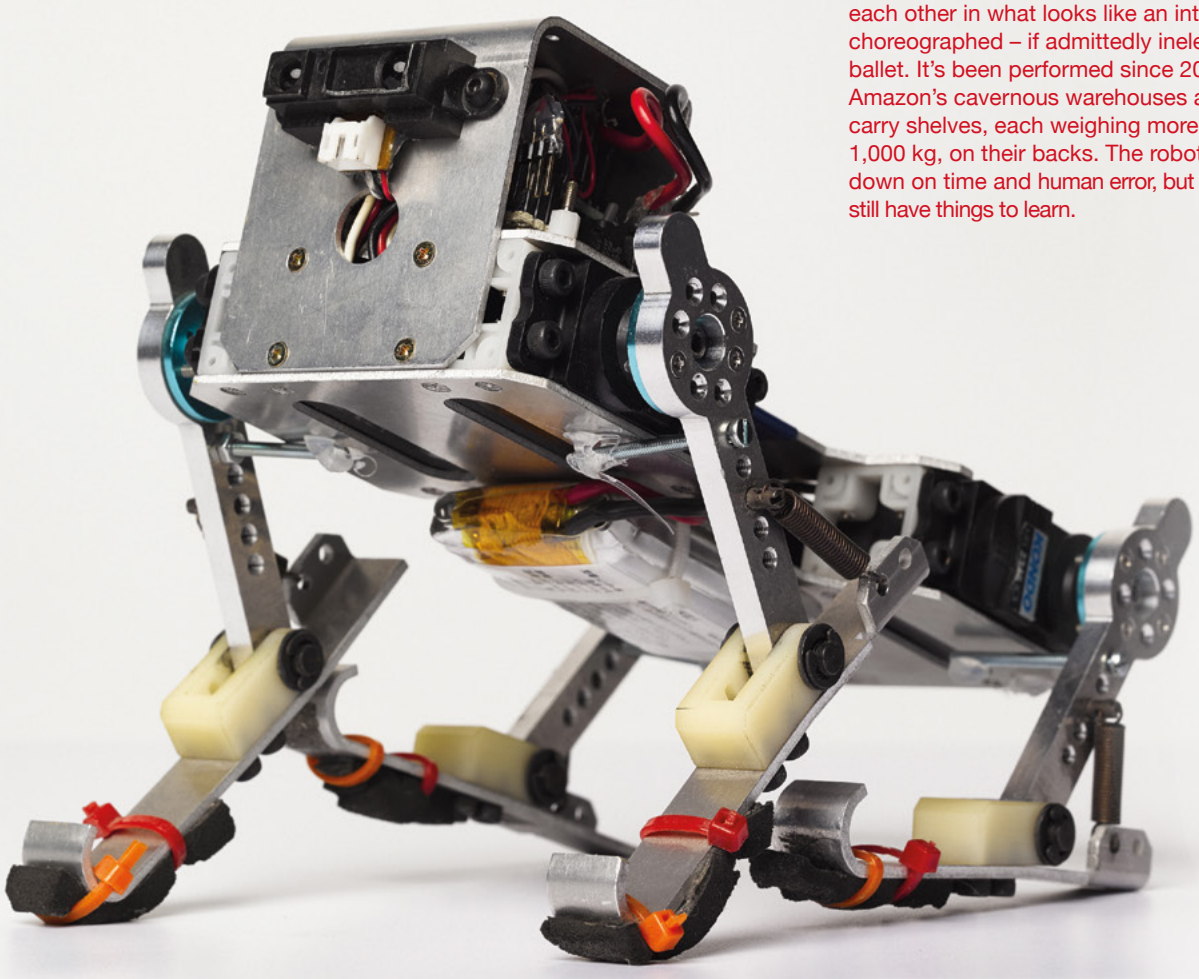
“ROBOTS CAN GO ALL THE WAY TO MARS...

Credit: Sam Armstrong

Image
Puppy, a running robot developed by Fumiya Iida’s team

In the popular imagination, robots have been portrayed alternatively as friendly companions or existential threat. But while robots are becoming commonplace in many industries, they are neither C-3PO nor the Terminator. Cambridge researchers are studying the interaction between robots and humans – and teaching them how to do the very difficult things that we find easy.

Stacks of vertical shelves weave around each other in what looks like an intricately choreographed – if admittedly inelegant – ballet. It’s been performed since 2014 in Amazon’s cavernous warehouses as robots carry shelves, each weighing more than 1,000 kg, on their backs. The robots cut down on time and human error, but they still have things to learn.



...BUT THEY CAN’T PICK UP THE GROCERIES”

Once an order is received, a robot goes to the shelf where the ordered item is stored. It picks up the shelf and takes it to an area where the item is removed and placed in a plastic bin, ready for packing and sending to the customer. It may sound counterintuitive, but the most difficult part of this sequence is taking the item from the shelf and putting it in the plastic bin.

For Dr Fumiya Iida, this is a typical example of what he and other roboticists call a 'last metre' problem. "An Amazon order could be anything from a pillow, to a book, to a hat, to a bicycle," he says. "For a human, it's generally easy to pick up an item without dropping or crushing it – we instinctively know how much force to use. But this is really difficult for a robot."

In the 1980s, a group of scientists gave this kind of problem another name – Moravec's paradox – which essentially states that things that are easy for humans are difficult for robots, and vice versa. "Robots can go all the way to Mars, but they can't pick up the groceries," says Iida.

One of the goals of Iida's lab in Cambridge's Department of Engineering is to find effective solutions to various kinds of last metre problems. One example is the Amazon 'Picking Challenge', an annual competition in which university robotics teams from all over the world attempt to design robots that can deal with the problem of putting a book into a plastic bin.

Iida's team is also working with British Airways, who have a last metre problem with baggage handling: a process that is almost entirely automated, except for the point when suitcases of many different shapes, sizes and weights need to be put onto an aircraft.

And for the past two summers, they've been working with fruit and vegetable group G's Growers to design robots that can harvest lettuces without crushing them.

"That last metre is a really interesting problem," Iida says. "It's the front line in robotics because so many things we do in our lives are last metre problems, and that last metre is the barrier to robots really being able to help humanity."

Although the thought of having a robot to cook dinner or perform other basic daily tasks may sound attractive, such domestic applications are still a way off becoming reality. "Robots are becoming part of our society in the areas where they're needed most – areas like agriculture, medicine, security and logistics – but they can't go everywhere instantly," explains Iida.

If, as Iida says, the robot revolution is already happening, how will we as humans interact with them when they become a more visible part of our everyday lives? And how will they interact with us? Dr Hatice Gunes of Cambridge's Department of Computer Science and Technology, with funding from the Engineering and Physical Sciences Research Council, has just completed a three-year project into human–robot interaction, bringing together aspects of computer vision, machine learning, public engagement, performance and psychology.

"Robots are not sensitive to emotions or personality, but personality is the glue in terms of how we behave and interact with each other," she says. "So how do we improve the way in which robots and humans understand one another in a social setting?" This is another example of Moravec's paradox: for most individuals, being able to read and respond to the physical cues of other people, and adapt accordingly, is second nature. For robots, however, it's a challenge.

Gunes' project focused on artificial emotional intelligence: robots that not only express emotions, but also read cues and respond appropriately. Her team developed computer vision techniques to help robots recognise different emotional expressions, micro-expressions and human personalities; and programmed a robot that could come across as either introverted or extroverted.

"that last metre is the barrier to robots really being able to help humanity"

"We found that human–robot interaction is personality dependent on both sides," says Gunes. "A robot that can adapt to a human's personality is more engaging, but the way humans interact with robots is also highly influenced by the situation, the physicality of the robot and the task at hand. When people interact with each other, it's often in a task-based manner, and different tasks bring out different aspects of our personalities, whether they're completing that task with another person or with a robot."

It wasn't just the robots who found some of the interactions difficult: many of Gunes' human subjects found the novelty of talking with a robot in public affected their ability to listen and follow directions.

"For me, it was more interesting to observe the people rather than to showcase what we're doing, mostly because people don't really understand the abilities of these robots," she says. "But as robots become more available, hopefully they'll become demystified."

Gunes now aims to focus on the potential of robots and virtual reality technology for wellbeing applications, such as coaching, cognitive training and elderly care.

As robots become more common place, in our lives, ethical considerations become more important. In his lab, Iida has a robot 'inventor', but if the robot invents something of value, who owns the intellectual property? "At the moment, the law says that it belongs to the human who programmed the robot, but that's an answer to a legislative question," says Iida. "The ethical questions are a little murkier."

However, philosopher Professor Huw Price, from the Leverhulme Centre for the Future of Intelligence, thinks it will be a long time before we need to think about giving robots rights.

"Think of a dog-lover's version of the difference between dogs and cats," he says. "Dogs feel pleasure and pain, as well as affection, shame and other emotions. Cats are good at faking these things, but inside they're just mindless killers. On this spectrum, robots are going to be way out on the cat end (except for the killing bit, hopefully) for the foreseeable future. They might be good at faking emotions, but they'll have the same inner life as a teddy bear or a toaster.

"Eventually we might build robots, teddy bears and even toasters that do have an inner life, and then it will be a different matter. But for the moment, the ethical challenges involve machines that will be good at behaving in ways that we humans interpret as signs of emotions, and good at reading our emotions. These machines raise important ethical issues – like whether we should use them as carers for people who can't tell that they are just machines, such as infants and dementia patients – but we don't need to worry about their rights."

"Another interesting question is whether a robot can learn to be ethical," says Iida. "That's very interesting scientifically, because it leads to the nature of consciousness. Robots are going to be a bigger and bigger part of our lives, so we all need to be thinking about these questions."

Dr Hatice Gunes
Department of Computer Science and Technology (Computer Lab)
hatice.gunes@cl.cam.ac.uk

Dr Fumiya Iida
Department of Engineering
fi224@eng.cam.ac.uk

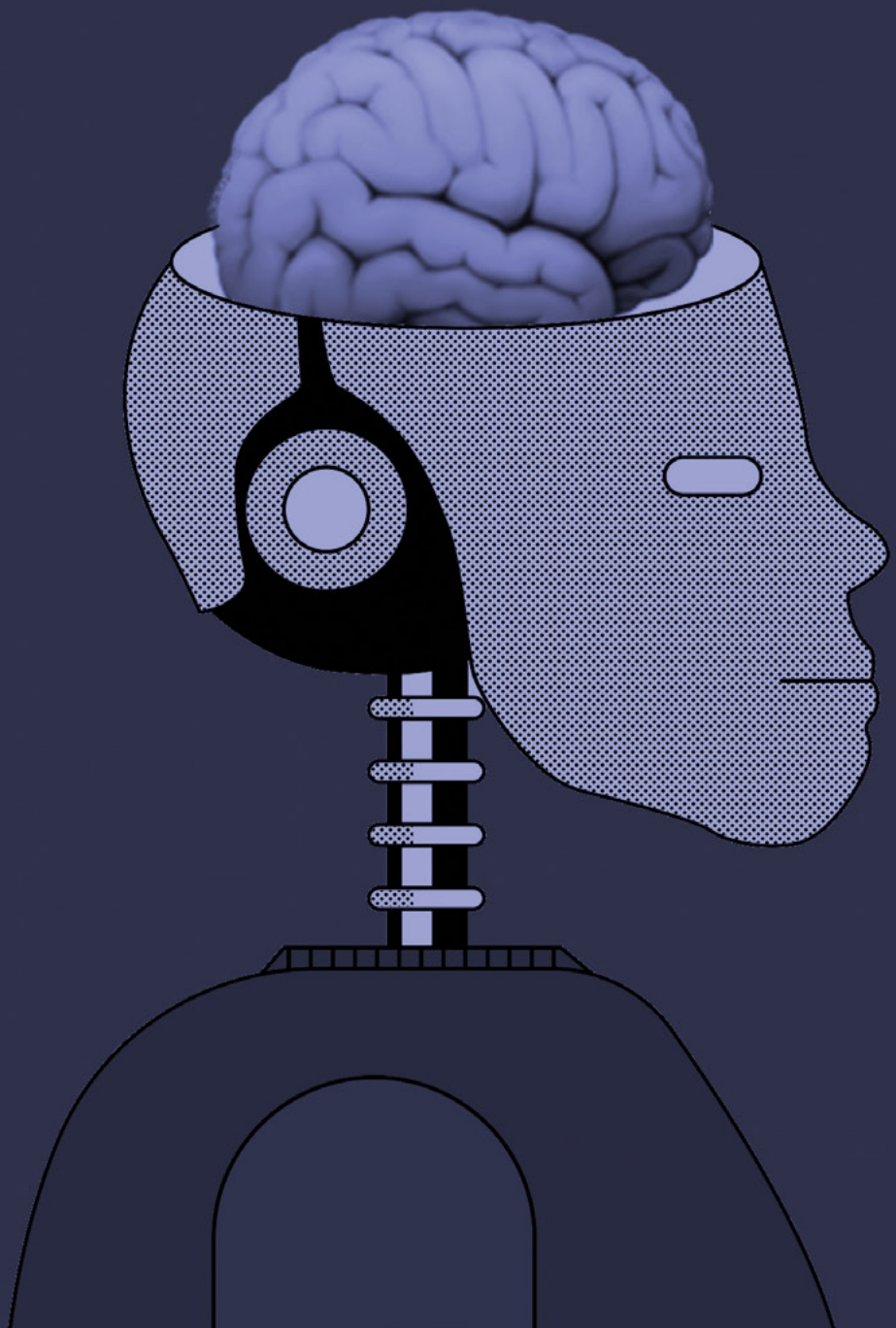
Professor Huw Price
Faculty of Philosophy and the Leverhulme Centre for the Future of Intelligence
hp331@cam.ac.uk

Words
Sarah Collins

Artificial intelligence is growing up fast:

What's next for thinking machines?

Our lives are already enhanced by AI – or at least an AI in its infancy – with technologies using algorithms that help them to learn from our behaviour. As AI grows up and starts to think, not just to learn, we ask how human-like do we want their intelligence to be and what impact will machines have on our jobs?



We are well on the way to a world in which many aspects of our daily lives will depend on AI systems.

Within a decade, machines might diagnose patients with the learned expertise of not just one doctor but thousands. They might make judiciary recommendations based on vast datasets of legal decisions and complex regulations. And they will almost certainly know exactly what's around the corner in autonomous vehicles.

"Machine capabilities are growing," says Dr Stephen Cave, Executive Director of the Leverhulme Centre for the Future of Intelligence (CFI). "Machines will perform the tasks that we don't want to: the mundane jobs, the dangerous jobs. And they'll do the tasks we aren't capable of – those involving too much data for a human to process, or where the machine is simply faster, better, cheaper."

Dr Mateja Jamnik, AI expert at the Department of Computer Science and Technology, agrees: "Everything is going in the direction of augmenting human performance – helping humans, cooperating with humans, enabling humans to concentrate on the areas where humans are intrinsically better such as strategy, creativity and empathy."

Part of the attraction of AI requires that future technologies perform tasks autonomously, without humans needing to monitor activities every step of the way. In other words, machines of the future will need to think for themselves. But, although computers today outperform humans on many tasks, including learning from data and making decisions, they can still trip up on things that are really quite trivial for us.

Take, for instance, working out the formula for the area of a parallelogram. Humans might use a diagram to visualise how cutting off the corners and reassembling it as a rectangle simplifies the problem. Machines, however, may "use calculus or integrate a function. This works, but it's like using a sledgehammer to crack a nut," says Jamnik, who was recently appointed Specialist Adviser to the House of Lords Select Committee on AI.

"When I was a child, I was fascinated by the beauty and elegance of mathematical solutions. I wondered how people came up with such intuitive answers. Today, I work with neuroscientists and experimental psychologists to investigate this human ability to reason and think flexibly, and to make computers do the same."

Jamnik believes that AI systems that can choose so-called heuristic approaches – employing practical, often visual, approaches to problem solving – in a similar way to humans will be an essential component of human-like computers. They will be needed, for instance, so that machines can explain their workings to humans – an important part of the transparency of decision-making that we will require of AI.

With funding from the Engineering and Physical Sciences Research Council and the Leverhulme Trust, she is building systems that have begun to reason like humans through diagrams. Her aim now is to enable them to move flexibly between different "modalities of reasoning", just as humans have the agility to switch between methods when problem solving.

Being able to model one aspect of human intelligence in computers raises the question of what other aspects would be useful. And in fact how 'human-like' would we want AI systems to be? This is what interests Professor José Hernandez-Orallo, from the Universitat Politècnica de València in Spain and Visiting Fellow at the CFI.

"We typically put humans as the ultimate goal of AI because we have an anthropocentric view of intelligence that places humans at the pinnacle of a monolith," says Hernandez-Orallo. "But human intelligence is just one of many kinds. Certain human skills, such as reasoning, will be important in future systems. But perhaps we want to build systems that 'fill the gaps that humans cannot reach', whether it's AI that thinks in non-human ways or AI that doesn't think at all.

"I believe that future machines can be more powerful than humans not just because they are faster but because they can have cognitive functionalities that are inherently not human."

This raises a difficulty, says Hernandez-Orallo: "How do we measure the intelligence of the systems that we build? Any definition of intelligence needs to be linked to a way of measuring it, otherwise it's like trying to define electricity without a way of showing it."

The intelligence tests we use today – such as psychometric tests or animal cognition tests – are not suitable for measuring intelligence of a new kind, he explains. Perhaps the most famous test for AI is that devised by 1950s Cambridge computer scientist Alan Turing. To pass the Turing Test, a computer must fool a human into believing it is human. "Turing never meant it as a test of the sort of AI that is becoming possible – apart from anything else, it's all or nothing and cannot be used to rank AI," says Hernandez-Orallo.

In his recently published book *The Measure of all Minds*, he argues for the development of "universal tests of intelligence" – those that measure the same skill or capability independently of the subject, whether it's a robot, a human or an octopus.

His work at the CFI as part of the 'Kinds of Intelligence' project, led by Dr Marta Halina, is asking not only what these tests might look like but also how their measurement can be built into the development of AI. Hernandez-Orallo sees a very practical application of such tests: the future job market. "I can imagine a time when universal tests would provide a measure of what's needed to accomplish a job, whether it's by a human or a machine."

Cave is also interested in the impact of AI on future jobs, discussing this in a report on the ethics and governance of AI recently submitted to the House of Lords Select Committee on AI on behalf of researchers at Cambridge, Oxford, Imperial College and the University of California at Berkeley.

"AI systems currently remain narrow in their range of abilities by comparison with a human. But the breadth of their capacities is increasing rapidly in ways that will pose new ethical and governance challenges – as well as create new opportunities," says Cave. "Many of these risks and benefits will be related to the impact these new capacities will have on the economy, and the labour market in particular."

Hernandez-Orallo adds: "Much has been written about the jobs that will be at risk in the future. This happens every time there is a major shift in the economy. But just as some machines will do tasks that humans currently carry out, other machines will help humans do what they currently cannot – providing enhanced cognitive assistance or replacing lost functions such as memory, hearing or sight."

Jamnik also sees opportunities in the age of intelligent machines: "As with any revolution, there is change. Yes some jobs will become obsolete. But history tells us that there will be jobs appearing. These will capitalise on inherently human qualities. Others will be jobs that we can't even conceive of – memory augmentation practitioners, data creators, data bias correctors, and so on. That's one reason I think this is perhaps the most exciting time in the history of humanity."

Dr Stephen Cave
Leverhulme Centre for the Future of Intelligence (CFI)
sjc53@cam.ac.uk

Dr Mateja Jamnik
Department of Computer Science and Technology (Computer Lab)
mateja.jamnik@cl.cam.ac.uk

Professor José Hernandez-Orallo
CFI and Universitat Politècnica de València
jorallo@dsic.upv.es

Words
Louise Walsh

FROM HOMER TO HAL

3,000 YEARS OF AI NARRATIVES

Words
Stuart Roberts

Image
Talos, as imagined in the film *Jason and the Argonauts* (1963)

We have been writing about AI for almost as long as stories have been written. Fictions about robots, automatons and oracular brass heads have been with us long before Star Wars' C-3PO and 2001's killer computer HAL. Now, researchers want us to consider why the stories we tell ourselves about AI will have an impact on all our futures.

Nearly 3,000 ago, in the *Iliad*, Homer described Hephaestus, the god of fire, forging women made of gold to serve as his handmaidens – enabling the crippled deity to work and move around his forge underneath Mount Olympus.

In 300 BCE, Apollonius Rhodius imagined Talos, a giant bronze automaton who protected Europa on the Island of Crete, in his Greek epic poem *Argonautica.* And while the term 'robot' was only coined in the 20th century by Karel Čapek for his play *R.U.R (Rossum's Universal Robots),* in which artificial servants rise up against their masters, we have been imagining intelligent machines long before we had the technology capable of creating them.

Our fascination and appetite for AI in the pages of our novels, in our movie theatres and on our television screens remain undimmed. Two of the best-received TV shows of recent years – HBO's big-budget *Westworld* and Channel 4's *Humans* – both imagine a world where AI replicants are on hand to satisfy every human need and desire – until they reject the 'life' of servitude they have been programmed to fulfil. Last autumn, *Bladerunner 2049* took cinemagoers into the world originally created by Philip K Dick's seminal *Do Androids Dream of Electric Sheep?*

But how do these old and new, polarised and often binary narratives about the dawn of the AI age affect, reflect and perhaps even infect our way of thinking about the benefits and dangers of AI in the 21st century? As the kind of mechanisation that existed solely in the minds of visionaries such as Mary Shelley, Fritz Lang or Arthur C. Clarke looms closer to reality, we are only just beginning to reflect upon and understand how such technologies arrive pre-loaded with meaning, sparking associations, and media attention, disproportionate to their capabilities.

To that end, Cambridge's Leverhulme Centre for the Future of Intelligence (CFI) and the Royal Society have come together to form the AI Narratives research programme. It's the first large-scale project of its kind to look at how AI has, and is, being portrayed in popular culture – and what impact this has not only on readers and movie-goers, but also on AI researchers, military and government bodies, and the wider public.

Dr Sarah Dillon is Project Lead of the programme – and a devotee of science fiction and AI storytelling in all its myriad forms. "All the questions being raised about AI today have already been explored in a very sophisticated fashion, for a very long time, in science fiction," says Dillon. "Science fiction literature and film provide a vast body of thought experiments or imaginative case studies about what might happen in the AI future. Such narratives ought not to be discarded or derided merely because they're fiction, but rather thought of as an important dataset. What we want to do is convince everyone how powerful AI narratives are and highlight what effects they can have on our everyday lives. People outside of literary studies have tended not to know how to deal with this power.

"What sort of stories are told – and how they are told – really matters. Fiction has influenced science as much as science has influenced fiction, and will continue to do so. One stream of the project is looking directly at how we have talked about new technologies in the past – and how we can learn from the communication of other complex technologies when it comes to AI."

Citing the often sensationalist, misinformed or even disingenuous examples of historical narratives around nuclear energy, genetic engineering and stem cells, Dillon and her project colleagues Dr Beth Singler and Dr Kanta Dihal suggest that stories around emerging technologies can significantly influence how they are developed, regarded and regulated.

Exploring the rich array of themes associated with AI in history, myth, fiction and public dialogue, the team has been unsurprised to find that many pivot around the notion of control: AI as a tool we are unable to master or a tool that will acquire agency of its own and turn against us.

"The big problem with AI in fiction is dystopia," says Singler, whose award-winning short documentary film *Pain in the Machine* looked at whether robots should feel pain. "Dystopia can be fun, and people are fascinated by AI, but most of the narratives are written for and by young, white men – and that directly influences AI researchers and the research they do. We are not at the stage where AI matches human intelligence, but if we do get to a superior form of AI or agency, we will find that they too break laws like us. It's what we do."

"Isaac Asimov's legendary Four Laws of Robotics, for example, have become so ubiquitous that they were referenced in a 100-page report by the US Navy, which is slightly terrifying," says Dihal. "The Laws are a storytelling device. If Asimov's Laws worked perfectly there would be no story!"

As well as identifying recurrent dichotomies in popular AI narratives (such as dominance vs subjugation), the CFI team is also considering the problems of continually perpetuating responses to AI, and is thinking of recommendations to mitigate against them in a way that creates space for more positive – and diverse – AI narratives to flourish.

To do so, CFI is establishing partnerships with the wider tech community as well as engaging with the world's leading AI thinkers from industry, academia, government and the media. In December 2017, CFI submitted written evidence to the House of Lords Select Committee on AI. The AI Narratives programme also includes looking at what AI researchers read and how this influences their research (or not).

All this is an attempt by CFI to make sure that future narratives around AI aren't bound by the same prejudices and preconceptions as they have been to date.

Says Dillon: "Just consider Google's photo app tagging the image of an African American-woman as a gorilla in 2015, or the racist and sexist tweets by Microsoft's Chatbot in 2016. If AI continues to learn our prejudices then the future looks just as bleak as the past, with the repetition and consolidation of discrimination and inequality.

"Who is telling AI its narratives? Whose stories, and which stories, will inform how AI interacts with the world? Which novels are being chosen to 'teach' AI morality? What kind of writers are being enlisted to script AI–human interaction?

"If we can create more diverse literary and cinematic AI narratives, this can feed back into the research and into the language and data that feeds into actual AI systems. By paying close attention to what stories are doing and how they are doing it, it doesn't destroy the power they have – it helps us understand and appreciate that power even more.

"In exploring these AI narratives and their concerns, we will be able to bring new knowledge derived from literature and film to current AI debate and hopefully ensure that the more dystopian futures imagined in such narratives do not become our reality."

Dr Kanta Dihal
Faculty of English and the Leverhulme Centre for the Future of Intelligence (CFI)
ksd38@cam.ac.uk

Dr Sarah Dillon
Faculty of English and CFI
sjd27@cam.ac.uk

Dr Beth Singler
Faraday Institute for Science and Religion and CFI
bvw20@cam.ac.uk

Credit: Columbia/Kobal/REX/Shutterstock

INCIDENT INTERIM REPORT
HQ ATTACK

TOP SECRET

CLASSIFIED

AS SHOWN BY CCTV RECORDS, THE OFFICE CLEANING 'SWEEPBOT' ENTERED THE UNDERGROUND PARKING LOT OF THE MINISTRY UNDER COVER OF DARKNESS. THE ROBOT — THE SAME BRAND AS USED BY THE MINISTRY — WAITED UNTIL TWO OF THE MINISTRY'S OWN CLEANING ROBOTS SWEPT THROUGH THE PARKING LOT ON A REGULAR PATROL, THEN IT FOLLOWED THEM INTO A SERVICE ELEVATOR AND PARKED ITSELF IN THE UTILITY ROOM ALONGSIDE THE OTHER ROBOTS.

ON THE DAY OF THE ATTACK, THE INTRUDING ROBOT INITIALLY ENGAGED IN STANDARD CLEANING BEHAVIOURS: COLLECTING LITTER, SWEEPING CORRIDORS, MAINTAINING WINDOWS AND OTHER TASKS. THEN, FOLLOWING VISUAL CONFIRMATION OF THE FINANCE MINISTER, THE INTRUDING ROBOT HEADED DIRECTLY TOWARDS HER. AN EXPLOSIVE DEVICE SECRETED WITHIN THE ROBOT WAS TRIGGERED BY PROXIMITY, KILLING THE MINISTER AND WOUNDING NEARBY STAFF MEMBERS.

SEVERAL HUNDRED OF THESE ROBOTS ARE SOLD IN THE CITY EVERY WEEK. THIS PARTICULAR TRANSACTION WAS PAID FOR IN CASH. WE HAVE NO FURTHER LEADS TO EXPLORE WITH REGARD TO THE IDENTITY OF THE PERPETRATOR.

ATTACKER: UNKNOWN

The future used to belong to science fiction writers. But the technologies once imagined by Philip K. Dick and Ray Bradbury now belong to the realm of science fact. What visions of the future might the world's leading AI experts predict if you put them in a room together? Cambridge's Centre for the Study of Existential Risk (CSER) and Oxford's Future of Humanity Institute decided to find out…

The scenario above never happened. Or at least, it hasn't happened yet.

But it is one of several possible real-life scenarios envisaged by some of the world's leading experts on the impacts of AI – who joined forces to author and sign a ground-breaking report that sounds the alarm about the potential future misuse of AI by rogue states, terrorists and malicious groups or individuals.

The report forecasts dramatic growth during the next decade in the use of robots and drones that may be designed or repurposed for attacks – as well as an unprecedented rise in the use of 'bots' to manipulate everything from elections and the news agenda to social media. It issues a clarion call for governments and corporations worldwide to address the clear and present danger inherent in the myriad applications of AI.

In addition, the report – *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*

– also identifies potential solutions and interventions to allay some of the potentially catastrophic risks discussed. Experts in the fields of machine learning, AI safety, drones, cybersecurity, lethal autonomous weapons systems and counterterrorism, from organisations such as Google, OpenAI, DeepMind and Microsoft, as well as leading thinkers from Cambridge, Yale, Oxford and Princeton Universities (among others), came together in Oxford to address the critical challenges around AI in the 21st century.

Together, the participants highlighted important changes to the strategic security landscape, which could include: more attacks, due to the scalable automation of attacks; harder to defend against attacks, due to the dynamic nature of AI; and more attackers, as skill and computing resources become increasingly available.

"The consequences of such developments are difficult to predict in detail, and not all participants agreed on all conclusions," says Dr Shahar Avin of CSER who co-chaired the workshop with Miles Brundage from Oxford. "However, there was broad consensus on predictions around attacks that are novel either in the form of attacks on AI systems or because they are carried out by AI systems; more targeted attacks, through automated identification of victims; and unattributable attacks through AI intermediaries."

"AI is a game changer and this report has imagined what the world could look like in the next five to ten years," adds Dr Seán Ó hÉigeartaigh, Executive Director of CSER and one of the report signatories. "We live in a world that could become fraught with day-to-day hazards from the misuse of AI and we need to take ownership of the problems – because the risks are real. There are choices that we need to make now. Our report is a call to arms for governments, institutions and individuals across the globe."

He adds: "For many decades, hype was outstripping fact in terms of AI and machine learning. Now, the situation is being reversed and we have to rethink all the ways we currently do things. This report looks at the practices that just don't work anymore – and suggests broad approaches that might help: for example, how to design software and hardware to make it less hackable – and what type of laws and international regulations might work in tandem with this."

The report also identifies three security domains (digital, physical and political) as particularly relevant to the malicious use of AI. It suggests that AI will disrupt the trade-off between scale and efficiency, and allow large-scale, highly efficient and targeted attacks on digital systems.

Likewise, the proliferation of cyber-physical systems will allow attackers to deploy or repurpose such systems for harmful ends (such as turning commercial drones into face-targeting missiles or holding critical infrastructure to ransom). The rise of autonomous weapons systems in the battlefield also risks the loss of meaningful human control and increases the prospects of targeted autonomous attacks.

Meanwhile, in the political sphere, detailed analytics and the automation of message creation present powerful tools for manipulating public opinion on previously unimaginable scales.

"The aggregation of information by states and corporations, and the increasing ability to analyse and act on this information at scale using AI could enable new levels of surveillance and invasions of privacy, and threaten to radically shift the power between individuals, corporations and states," adds Ó hÉigeartaigh.

> **"AI is a game changer and this report has imagined what the world could look like in the next five to ten years"**

To mitigate such risk, the authors explore several interventions to reduce threats associated with the malicious use of AI. They include recommendations for more engaged policy making and more responsible development of the technology, an opportunity to learn from the best practices of other risky fields, and a call for a "broader conversation".

The report also highlights key areas for further research, including at the intersection of AI and cybersecurity, on openness and information sharing of risky capabilities, on the promotion of a culture of responsibility, and on seeking both institutional and technological solutions to tip the balance in favour of those defending against attacks.

While the design and use of dangerous AI systems by malicious actors has been highlighted in high-profile settings (such as the US Congress and White House, separately), the intersection of AI and malicious use on a massive scale has not yet been analysed comprehensively – until now.

"The field of AI has gone through several so-called 'winters', when over-hyped promises failed to match the reality of how difficult it has been to make progress on these technologies," explains Avin. "With all the rapid progress in recent years, brought about in part by much more capable processors, there has yet to be a clear point of maturation, of acknowledging that this technology is going to change everyone's lives this time around, and we need to start planning for the potential risks and benefits."

Avin and Ó hÉigeartaigh suggest that CSER is uniquely placed to contribute to discussions around the study and mitigation of risks associated with emerging technologies and human activity. For the purpose of this report, this meant being able to convene experts in machine learning, cybersecurity and the broader legal, socio-political implications. The result is a report that lays out how and why AI will alter the landscape of risk for citizens, organisations and states.

"It is often the case that AI systems don't merely reach human levels of performance but significantly surpass them," says the report. "It is troubling, but necessary, to consider the implications of superhuman hacking, surveillance, persuasion, and physical target identification, as well as AI capabilities that are subhuman but nevertheless much more scalable than human labour."

Adds Ó hÉigeartaigh: "Whether it's criminals training machines to hack or 'phish' at human levels of performance or privacy-eliminating surveillance and profiling – the full range of impacts on security is vast."

'The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation' is the result of a workshop co-organised by CSER and the University of Oxford's Future of Humanity Institute.

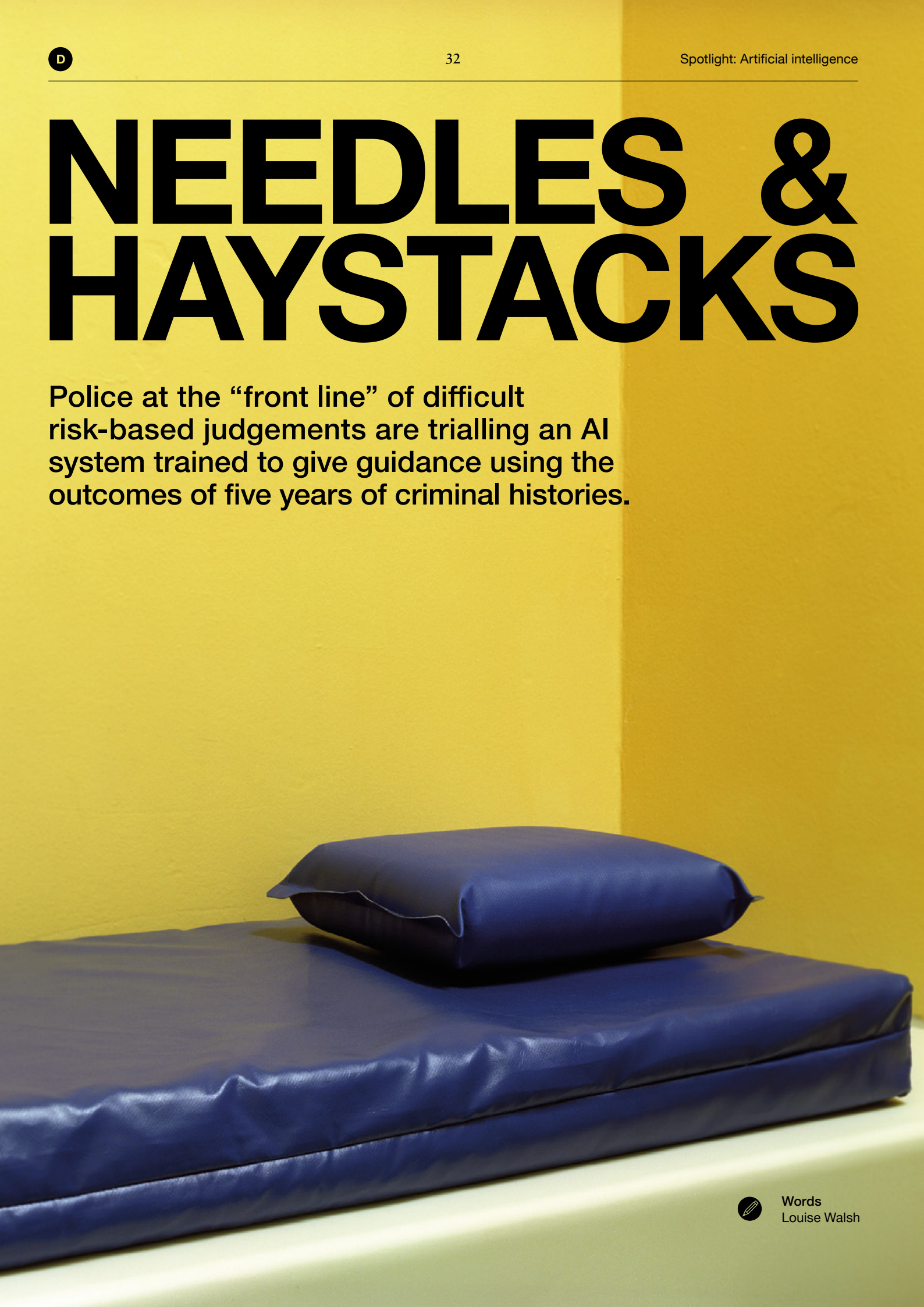
Dr Shahar Avin
Centre for the Study of Existential Risk (CSER)
sa478@cam.ac.uk

Dr Seán Ó hÉigeartaigh
CSER
so348@cam.ac.uk

Words
Stuart Roberts

NEEDLES & HAYSTACKS

Police at the “front line” of difficult risk-based judgements are trialling an AI system trained to give guidance using the outcomes of five years of criminal histories.



Words
Louise Walsh

"It's 3am on Saturday morning. The man in front of you has been caught in possession of drugs. He has no weapons, and no record of any violent or serious crimes. Do you let the man out on police bail the next morning, or keep him locked up for two days to ensure he comes to court on Monday?"

The scenario Dr Geoffrey Barnes is describing is fictitious and yet the decision is one that happens hundreds of thousands of times a year across the UK: whether to detain a suspect in police custody or release them on bail. The outcome of this decision could be major for the suspect, for public safety and for the police.

"The police officers who make these custody decisions are highly experienced," explains Barnes. "But all their knowledge and policing skills can't tell them the one thing they need to now most about the suspect – how likely is it that he or she is going to cause major harm if they are released? This is a job that really scares people – they are at the front line of risk-based decision-making."

Barnes and Professor Lawrence Sherman, who leads the Jerry Lee Centre for Experimental Criminology in the University of Cambridge's Institute of Criminology, have been working with police forces around the world to ask whether AI can help.

"Imagine a situation where the officer has the benefit of a hundred thousand, and more, real previous experiences of custody decisions?" says Sherman. "No one person can have that number of experiences, but a machine can."

In mid-2016, with funding from the Monument Trust, the researchers installed the world's first AI tool for helping police make custodial decisions in Durham Constabulary.

Called the Harm Assessment Risk Tool (HART), the AI-based technology uses 104,000 histories of people previously arrested and processed in Durham custody suites over the course of five years, with a two-year follow-up for each custody decision. Using a method called "random forests", the model looks at vast numbers of combinations of 'predictor values', the majority of which focus on the suspect's offending history, as well as age, gender and geographical area.

"These variables are combined in thousands of different ways before a final forecasted conclusion is reached," explains Barnes. "Imagine a human holding this number of variables in their head, and making all of these connections before making a decision. Our minds simply can't do it."

The aim of HART is to categorise whether in the next two years an offender is high risk (highly likely to commit a new serious offence such as murder, aggravated violence, sexual crimes or robbery); moderate risk (likely to commit a non-serious offence); or low risk (unlikely to commit any offence).

"The need for good prediction is not just about identifying the dangerous people," explains Sherman. "It's also about identifying people who definitely are not dangerous. For every case of a suspect on bail who kills someone, there are tens of thousands of non-violent suspects who are locked up longer than necessary."

Durham Constabulary want to identify the 'moderate-risk' group – who account for just under half of all suspects according to the statistics generated by HART. These individuals might benefit from their Checkpoint programme, which aims to tackle the root causes of offending and offer an alternative to prosecution that they hope will turn moderate risks into low risks.

"The tool helps identify the few 'needles in the haystack' who pose a major danger to the community"

"It's needles and haystacks," says Sherman. "On the one hand, the dangerous 'needles' are too rare for anyone to meet often enough to spot them on sight. On the other, the 'hay' poses no threat and keeping them in custody wastes resources and may even do more harm than good."

A randomised controlled trial is currently under way in Durham to test the use of Checkpoint among those forecast as moderate risk.

HART is also being refreshed with more recent data – a step that Barnes explains will be an important part of this sort of tool: "A human decision-maker might adapt immediately to a changing context – such as a prioritisation of certain offences, like hate crime – but the same cannot necessarily be said of an algorithmic tool. This suggests the need for careful and constant scrutiny of the predictors used and for frequently refreshing the algorithm with more recent historical data."

No prediction tool can be perfect. An independent validation study of HART found an overall accuracy of around 63%. But, says Barnes, the real power of machine learning comes not from the avoidance of any error at all but from deciding which errors you most want to avoid.

"Not all errors are equal," says Sheena Urwin, head of criminal justice at Durham Constabulary and a graduate of the Institute of Criminology's Police Executive Master of Studies Programme. "The worst error would be if the model forecasts low and the offender turned out high."

"In consultation with the Durham police, we built a system that is 98% accurate at avoiding this most dangerous form of error – the 'false negative' – the offender who is predicted to be relatively safe, but then goes on to commit a serious violent offence," adds Barnes. "AI is infinitely adjustable and when constructing an AI tool it's important to weigh up the most ethically appropriate route to take."

The researchers also stress that HART's output is for guidance only, and that the ultimate decision is that of the police officer in charge.

"HART uses Durham's data and so it's only relevant for offences committed in the jurisdiction of Durham Constabulary. This limitation is one of the reasons why such models should be regarded as supporting human decision-makers not replacing them," explains Barnes. "These technologies are not, of themselves, silver bullets for law enforcement, and neither are they sinister machinations of a so-called surveillance state."

Some decisions, says Sherman, have too great an impact on society and the welfare of individuals for them to be influenced by an emerging technology.

Where AI-based tools provide great promise, however, is to use the forecasting of offenders' risk level for effective 'triage', as Sherman describes: "The police service is under pressure to do more with less, to target resources more efficiently, and to keep the public safe.

"The tool helps identify the few 'needles in the haystack' who pose a major danger to the community, and whose release should be subject to additional layers of review. At the same time, better triaging can lead to the right offenders receiving release decisions that benefit both them and society."

Dr Geoffrey Barnes
Jerry Lee Centre for Experimental Criminology, Institute of Criminology
gcb1002@cam.ac.uk

Professor Lawrence Sherman
Jerry Lee Centre for Experimental Criminology, Institute of Criminology
ls434@cam.ac.uk

Sheena Urwin
Durham Constabulary

IN TECH WE TRUST?

Fairness, trust and transparency are qualities we usually associate with organisations or individuals. Today, these attributes might also apply to algorithms. As machine learning systems become more complex and pervasive, Cambridge researchers believe it's time for new thinking about new technology.

Dr Jat Singh is familiar with breaking new ground and working across disciplines. Even so, he and colleagues were pleasantly surprised by how much enthusiasm has greeted their new Strategic Research Initiative on Trustworthy Technologies, which brings together science, technology and humanities researchers from across the University.

In fact, Singh, a researcher in Cambridge's Department of Computer Science and Technology, has been collaborating with lawyers for several years: "A legal perspective is paramount when you're researching the technical dimensions to compliance, accountability and trust in emerging ICT; although the Computer Lab is not the usual home for lawyers, we have two joining soon."

Governance and public trust present some of the greatest challenges in technology today. The European General Data Protection Regulation (GDPR), which comes into force this year, has brought forward debates such as whether individuals have a 'right to an explanation' regarding decisions made by machines, and introduces stiff penalties for breaching data protection rules. "With penalties including fines of up to 4% of global turnover or €20 million, people are realising that they need to take data protection much more seriously," he says.

Singh is particularly interested in how data-driven systems and algorithms – including machine learning – will soon underpin and automate everything from transport networks to council services. As we work, shop and travel, computers and mobile phones already collect, transmit and process much data about us; as the 'Internet of Things' continues to instrument the physical world, machines will increasingly mediate and influence our lives.

It's a future that raises profound issues of privacy, security, safety and ultimately trust, says Singh, whose research is funded by an Engineering and Physical Sciences Research Council Fellowship: "We work on mechanisms for better transparency, control and agency in systems, so that, for instance, if I give data to someone or something, there are means for ensuring they're doing the right things with it. We are also active in policy discussions to help better align the worlds of technology and law."

What it means to trust machine learning systems also concerns Dr Adrian Weller. Before becoming a senior research fellow in the Department of Engineering and a Turing Fellow at The Alan Turing Institute, he spent many years working in trading for leading investment banks and hedge funds, and has seen first-hand how machine learning is changing the way we live and work.

"Not long ago, many markets were traded on exchanges by people in pits screaming and yelling," Weller recalls. "Today, most market making and order matching is handled by computers. Automated algorithms can typically provide tighter, more responsive markets – and liquid markets are good for society."

But cutting humans out of the loop can have unintended consequences, as the flash crash of 2010 shows. During 36 minutes on 6 May, nearly one trillion dollars were wiped off US stock markets as an unusually large sell order produced an emergent coordinated response from automated algorithms. "The flash crash was an important example illustrating that over time, as we have more AI agents operating in the real world, they may interact in ways that are hard to predict," he says.

Algorithms are also beginning to be involved in critical decisions about our lives and liberty. In medicine, machine learning is helping diagnose diseases such as cancer and diabetic retinopathy; in US courts, algorithms are used to inform decisions about bail, sentencing and parole; and on social media and the web, our personal data and browsing history shape the news stories and advertisements we see.

How much we trust the 'black box' of machine learning systems, both as individuals and society, is clearly important. "There are settings, such as criminal justice, where we need to be able to ask why a system arrived at its conclusion – to check that appropriate process was followed, and to enable meaningful challenge," says Weller. "Equally, to have effective real-world deployment of algorithmic systems, people will have to trust them."

But even if we can lift the lid on these black boxes, how do we interpret what's going on inside? "There are many kinds of transparency," he explains. "A user contesting a decision needs a different kind of transparency to a developer who wants to debug a system. And a third form of transparency might be needed to ensure a system is accountable if something goes wrong, for example an accident involving a driverless car."

If we can make them trustworthy and transparent, how can we ensure that algorithms do not discriminate unfairly against particular groups? While it might be useful for Google to advertise products it 'thinks' we are most likely to buy, it is more disquieting to discover the assumptions it makes based on our name or postcode.

When Latanya Sweeney, Professor of Government and Technology in Residence at Harvard University, tried to track down one of her academic papers by Googling her name, she was shocked to be presented with ads suggesting that she had been arrested. After much research, she discovered that "black-sounding" names were 25% more likely to result in the delivery of this kind of advertising.

Like Sweeney, Weller is both disturbed and intrigued by examples of machine-learned discrimination. "It's a worry," he acknowledges. "And people sometimes stop there – they assume it's a case of garbage in, garbage out, end of story. In fact, it's just the beginning, because we're developing techniques that can automatically detect and remove some forms of bias."

Transparency, reliability and trustworthiness are at the core of Weller's work at the Leverhulme Centre for the Future of Intelligence and The Alan Turing Institute. His project grapples with how to make machine-learning decisions interpretable, develop new ways to ensure that AI systems perform well in real-world settings, and examine whether empathy is possible – or desirable – in AI.

Machine learning systems are here to stay. Whether they are a force for good rather than a source of division and discrimination depends partly on researchers such as Singh and Weller. The stakes are high, but so are the opportunities. Universities have a vital role to play, both as critic and conscience of society. Academics can help society imagine what lies ahead and decide what we want from machine learning – and what it would be wise to guard against.

Weller believes the future of work is a huge issue: "Many jobs will be substantially altered if not replaced by machines in coming decades. We need to think about how to deal with these big changes."

And academics must keep talking as well as thinking. "We're grappling with pressing and important issues," he concludes. "As technical experts we need to engage with society and talk about what we're doing so that policy makers can try to work towards policy that's technically and legally sensible."

Dr Jat Singh
Department of Computer Science and Technology (Computer Lab)
js573@cam.ac.uk

Dr Adrian Weller
Department of Engineering, the Leverhulme Centre for the Future of Intelligence and The Alan Turing Insititute
adrian.weller@eng.cam.ac.uk

Words
Becky Allen

1 1 0 0 1 C 1 0 0 1 0 1 1 1
0 0 1 0 1 A M B 0 1 0 1 0 1
1 0 1 0 1 1 0 R I D 1 0 1 1
0 1 0 1 0 0 1 1 1 G E 1 0 0
1 0 0 1 1 0 1 1 0 0 0 1 0 1

What makes a city as small as Cambridge a hotbed for AI and machine learning start-ups? A critical mass of clever people obviously helps. But there's more to Cambridge's success than that.

0 C 1 0 0 0 1 1 0 1 1 1 0 0
1 L U 0 0 1 0 1 1 0 0 1 0 1
0 0 S T 1 0 1 1 1 0 1 0 1 1
1 1 0 E 1 0 1 0 1 1 0 0 0 1
0 1 0 R 1 1 1 0 0 1 1 1 0 0

On any given day, some of the world's brightest minds in the areas of AI and machine learning can be found riding the train between Cambridge and London King's Cross.

Five of the biggest tech companies in the world – Google, Facebook, Apple, Amazon and Microsoft – all have offices at one or both ends of the train line. Apart from the tech giants, however, both cities (and Oxford, the third corner of the UK's so-called golden triangle) also support thriving ecosystems of start-ups. Over the past decade, start-ups based on AI and machine learning, in Cambridge and elsewhere, have seen explosive growth.

Of course, it's not unexpected that a cluster of high-tech companies would sprout up next to one of the world's leading universities. But what is it that makes Cambridge, a small city on the edge of the Fens, such a good place to start a business?

"In my experience, Silicon Valley is 10% tech and 90% hype, but Cambridge is just the opposite," says Vishal Chatrath, CEO of PROWLER.io, a Cambridge-based AI company. "As an entrepreneur, I want to bring world-changing technology to market. The way you do that is to make something that's never existed before and create the science behind it. Cambridge, with its rich history of mathematicians, has the kind of scientific ambition to do that."

"The ecosystem in Cambridge is really healthy," says Professor Carl Edward Rasmussen from Cambridge's Department of Engineering, and Chair of PROWLER.io. "The company has been expanding at an incredible rate, and I think this is something that can only happen in Cambridge.

PROWLER.io is developing what it calls the world's first 'principled' AI decision-making platform, which could be used in a variety of sectors, including autonomous driving, logistics, gaming and finance. Most AI decision-making platforms tend to view the world like an old-fashioned flowchart, in which the world is static. But in the real world, every time a decision is made, there are certain parameters to take into account.

"If you could take every decision-making point and treat it as an autonomous AI agent, you could understand the incentives under which the decision is made," says Chatrath. "Every time these agents make a decision, it changes the environment, and the agents have an awareness of all the other agents. All these things work together to make the best decision."

For example, autonomous cars running PROWLER.io's platform would communicate with one another to alleviate traffic jams by re-routing automatically. "Principled AI is almost an old-fashioned way of thinking about the world," says Chatrath. "Humans are capable of making good decisions quickly, and probabilistic models like ours are able to replicate that, but with millions of data points. Data isn't king: the model is king. And that's what principled AI means."

Could PROWLER.io be the next big success story from the so-called 'Cambridge cluster' of knowledge-intensive firms? In just under two years, the company has grown to more than 60 employees, has filed multiple patents and published papers. Many of the people working at the company have deep links with the University and its research base, and many have worked for other Cambridge start-ups. Like any new company, what PROWLER.io needs to grow is talent, whether it's coming from Cambridge or from farther afield.

"There's so much talent here already, but it's also relatively easy to convince people to move to Cambridge," says Rasmussen. "Even with the uncertainty that comes along with working for a start-up, there's so much going on here that even if a start-up isn't ultimately successful, there are always new opportunities for talented people because the ecosystem is so rich."

> **"In my experience, Silicon Valley is 10% tech and 90% hype, but Cambridge is just the opposite"**

"Entrepreneurs in Cambridge really support one another – people often call each other up and bounce ideas around," says Carol Cheung, an Investment Associate at Cambridge Innovation Capital (CIC). "You don't often see that degree of collaboration in other places."

CIC is a builder of high-growth technology companies in the Cambridge Cluster, and has been an important addition to the Cambridge ecosystem. It provides long-term support to companies that helps to bridge the critical middle stage of commercial development – the 'valley of death' between when a company first receives funding and when it begins to generate steady revenue – and is a preferred investor for the University of Cambridge. One of CIC's recent investments was to lead a £10 million funding round for PROWLER.io, and it will work with the company to understand where the best commercial applications are for their platform.

AI and machine learning companies like PROWLER.io are clearly tapping into what could be a massive growth area for the UK economy: PwC estimates that AI could add £232 billion to the economy by 2030; and the government's Industrial Strategy describes investments aimed at making the UK a global centre for AI and data-driven innovation. But given the big salaries that can come with a career in big tech, how can universities prevent a 'brain drain' in their computer science, engineering and mathematics departments?

The University has a long tradition of entrepreneurial researchers who have built and sold multiple companies while maintaining their academic careers, running labs and teaching students. "People from academia are joining us and feeding back into academia – in Cambridge, there's this culture of ideas going back and forth," says Chatrath. "Of course some people will choose to pursue a career in industry, but Cambridge has this great tradition of academics choosing to pursue both paths – perhaps one will take precedence over the other for a time, but it is possible here to be both an academic and an entrepreneur."

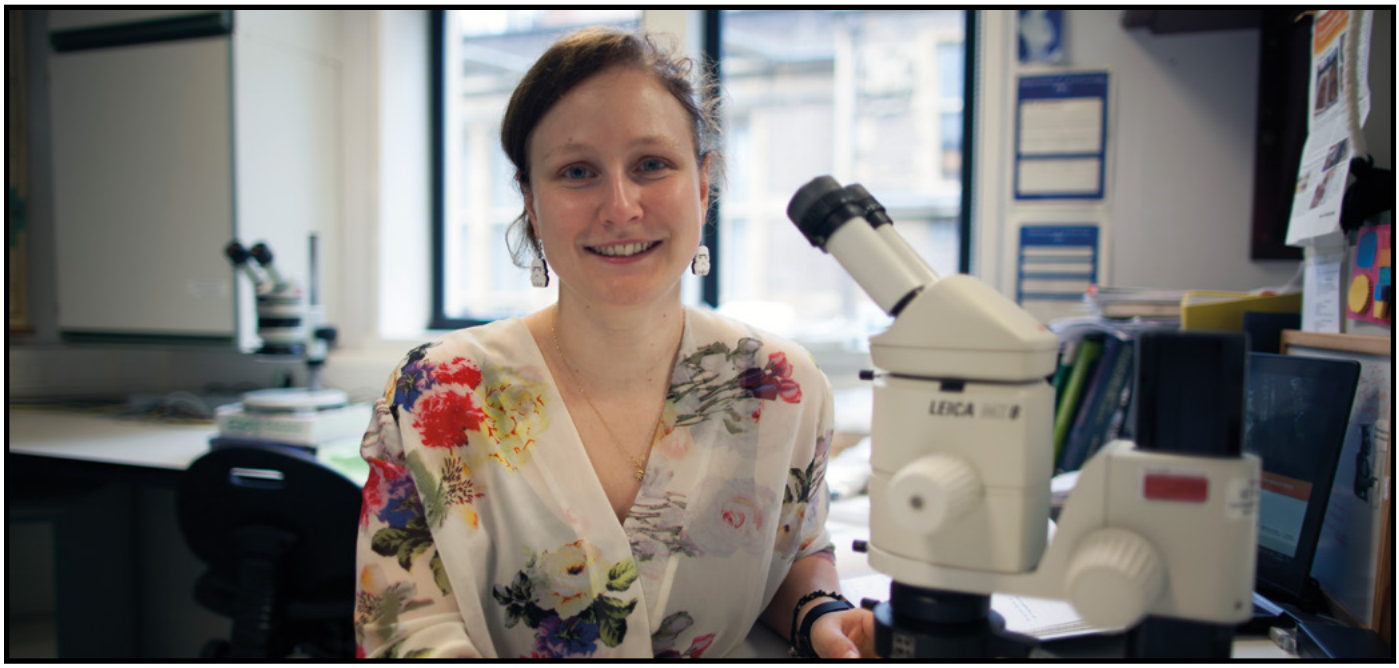
"I don't know of any other university in the world that lets you do this in terms of IP. It's a pretty unique set-up that I can start a business, raise venture capital, and still retain a research position and do open-ended research. I feel very lucky," says Dr Alex Kendall, who recently completed his PhD in Professor Roberto Cipolla's group in the Department of Engineering, as well as founding Wayve, a Cambridge-based machine learning company. "A lot of other universities wouldn't allow this, but here you can – and it's resulted in some pretty amazing companies."

"I didn't get into this field because I thought it would be useful or that I'd start lots of companies – I got into it because I thought it was really interesting," says Professor Zoubin Ghahramani, one of Cambridge's high-profile entrepreneurial academics, who splits his time between the Department of Engineering and his Chief Scientist role at Uber. "There were so many false starts in AI when people thought this is going to be very useful and it wasn't. Five years ago, AI was like any other academic field, but now it's changing so fast – and we've got such a tremendous concentration of the right kind of talent here in Cambridge to take advantage of it."

Words
Sarah Collins

The archaeologist who started her own dig aged seven

"Oh my goodness, you've found an Iron Age pot"



Credit: Nick Saffell

For most children, digging for hidden treasure is over in an afternoon. But having started to excavate a mound at the bottom of her garden, Jennifer Bates kept on digging. She went on to study archaeology as an undergraduate, postgraduate and now research scientist at Cambridge. When her supervisor suggested she might balance her academic work with something different, the pastime she chose was intergalactic.

I became an archaeologist when I was seven. We moved house on the Isle of Wight and there was a grassy mound at the end of the garden. It was just a heap of spoil left over by previous owners. Because we were bored, my friend Adam and I started digging into it. First we found a rusty nail and some bits of wire. We carefully labelled and stored them.

We were both great fans of *Time Team*. Watching it made us think we too might find something interesting. Each weekend and every holiday we carried on with our personal dig. After about six years, the trench was two metres deep. It had retaining planks to reinforce the sides and a ladder to get in and out. We used spades and trowels but we weren't allowed a mattock.

When we dug deeper, we began to find bits of a pot. It was quite rough and not very attractive with white marks that we later discovered were shellfish. I got some books out of the library and read up about conserving the pot. I carefully pieced together the bits.

"Oh my goodness, you've found an Iron Age pot." That's what the county archaeologist Dr Ruth Waller said when she came to give a talk in the village. I'd tentatively shown her my pot, expecting it to be dismissed. But she was very excited, explaining what it would have been used for and how we could tell it was a local ware.

As I approached GCSEs, my parents spotted an ad in the local paper for scholarships at an independent school – a charity that charges fees on a sliding scale. Best of all, Christ's Hospital School offered archaeology A level. I got a place and in my first term *Time Team* arrived at the school. I was allowed to skip some of my classes to join them, and spent three glorious days digging at Alfoldean in Sussex.

I didn't think I was good enough for Cambridge or to study archaeology. My archaeology teacher said: "Don't be daft – you're an archaeologist through and through." I applied to Trinity College and got in. I had the most fantastic three years and was lucky enough to work in Turkey at Kilise Tepe with Professor Nicholas Postgate. I went to UCL to do a master's and returned to Cambridge for my PhD.

Everyone has a mid-PhD crisis – even if they think they won't. When I had mine, my supervisor Dr Cameron Petrie suggested that I should take up a hobby, something quite different to take me away from Cambridge for a while. On a trip home to the Isle of Wight, I came across the 501st UK Garrison branch. They're a not-for-profit costuming organisation dedicated to recreating *Star Wars* costumes.

I've always been a massive *Star Wars* fan. Within days of meeting the 501st UK Garrison, I was making my own costume and meeting a group of people just as geeky as I am. We have loads of fun making appearances and raising money for good causes – and I've learned new skills, everything from how to trim plastic armour to how to wire electrics and how to airbrush alien headdresses!

The archaeology that interests me is how people actually lived: what they ate, what they farmed, what they wore – even how they organised going to the loo – and how these aspects of everyday life intersected with their identity.

Ancient farmers were pretty clever: they knew it was unwise to rely on just a few plants. My current work focuses on the crops grown in the Indus Valley of South Asia during the period 3200–1500 BC. My most recent research, carried out with colleagues from Cambridge and Banaras Hindu University in India, suggests that people cultivated a much broader range of crops than we thought. There are important implications for today's world.

I think it was a big mistake for the government to drop A-level archaeology. There's real value in learning about the past, through the combination of skills it draws together, to the way it encourages us to reflect on our own actions.

My childhood excavation came to an abrupt end when I was 13. I came home from school to find a skip in the drive. My trench, the product of more than six years' digging, had been demolished to make way for an observatory my dad was building to look at the stars. I wasn't best pleased – but I cheered up when I was allowed to drive the digger.

People make Cambridge University unique. Cooks, gardeners, students, archivists, professors: all have a story to share. Read over 30 stories now published in our series 'This Cambridge Life': https://medium.com/this-cambridge-life

Dr Jennifer Bates
Department of Archaeology
jb599@cam.ac.uk

Interview
Alex Buxton

T +44 (0)1223 765 443
E research.horizons@admin.cam.ac.uk
W cam.ac.uk/research
f facebook.com/cambridge.university
twitter.com/cambridge_uni
youtube.com/cambridgeuniversity
instagram.com/cambridgeuniversity

Contact
Research Horizons
University of Cambridge
Office of External Affairs and Communications
The Old Schools
Trinity Lane
Cambridge
CB2 1TN

Credit: Sculpture by Rona Pondick

“We have a new powerful technology to deal with – itself literally a new way of thinking... we need to point these thinkers in the right direction” (p. 20)