TRUST & TECHNOLOGY INITIATIVE

Exploring the dynamics of trust and distrust around internet technologies, society and power.

www.trusttech.cam.ac.uk
www.twitter.com/CamTrustTech

UNIVERSITY OF CAMBRIDGE

Contents

Introducing the Initiative

The Trust & Technology Initiative brings together and drives forward interdisciplinary research from Cambridge and beyond to:

- Explore the dynamics of trust and distrust in relation to internet technologies, society and power
- Better inform trustworthy design and governance of next generation tech at the research and development stage
- Promote informed, critical, and engaging voices supporting individuals, communities and institutions in light of technology's increasing pervasiveness in societies.

The Initiative is unique in considering the interplays and feedback loops between technology fundamentals, societal impact and governance of next generation systems at the research and development stage. Our particular ability to connect cutting edge deep technology with social science and humanities expertise enables dynamic exploration of emergent use cases, and for us to envisage and experiment with realistic future scenarios.

A network around trust, technology, society and power

As a network, the Initiative is a 'big tent,' bringing people together, facilitating collaboration, and engaging industry, civil society, government, and the public, across:

- Relationships and interplays between technology and society; the legal, ethical and political frameworks impacting both trust and technology, and innovative governance, in areas such as transport, critical infrastructure, identity, manufacturing, healthcare, financial systems and networks, communications systems, internet of things
- The nature of trust and distrust; trust in technology, and trust through technology; the many dimensions of trust at individual, organisational and societal levels

- Rigorous technical foundations, for resilient, secure and safe computer systems, including data and communications platforms, artificial intelligence, and robotics

What the Trust & Technology Initiative does

- Connects the research community around trust and technology
- Catalyses new collaborative projects and activities
- Builds capacity and strengthens knowledge transfer
- Influences national and international research and policy agendas
- Acts as a helpful gateway to Cambridge for external partners

How we work

The Trust & Technology Initiative team proactively engages researchers and partners, and uses creative ways to bring together diverse participants and enable effective discussion and collaboration. We help interdisciplinary research ideas to emerge, and can support proposal development and securing resources. The Initiative also creates content to bridge between disciplines and sectors, and seeks new ways to connect researchers and enable prototyping and testing of ideas.

We are interested in more than just research collaborations, and are exploring what value the Initiative can offer potential partner organisations, including networking and brokering support, workshops and roundtables, strategic reports, and other services.

Get in touch

- Website: www.trusttech.cam.ac.uk
- Twitter: @CamTrustTech
- Email: admin@trusttech.cam.ac.uk

Get involved

We're developing a variety of ways to get more involved with the Trust & Technology Initiative. If you'd like to work with us in some way, please email us at admin@trusttech.cam.ac.uk, including a few sentences about your research and interests and how they relate to Trust & Technology, and whether you'd like to play an active part in our work (for instance, organising events, writing blog posts, etc). We'll be in touch to discuss options.

All are welcome to contribute to the Trust & Technology Initiative's Zotero library of interesting papers and articles. Find the library here: http://bit.ly/camtrusttechlibrary.

Executive team

Prof. Simon Moore, Co-Chair

Department of Computer Science and Technology

Simon Moore is a Professor of Computer Engineering at the University of Cambridge Computer Laboratory in England, where he undertakes research and teaching in the general area of computer design with particular interests in secure and rigorously-engineered computer architecture. Professor Moore is the senior member of the Computer Architecture research group.

Dr Jatinder Singh, Co-Chair

Department of Computer Science and Technology

Jat is an EPSRC Research Fellow at the Department of Computer Science & Technology, where he leads the newly formed "Compliant and Accountable Systems" research group. His research concerns the

technical and legal aspects of emerging technology, taking an interdisciplinary approach towards issues of security, privacy, transparency, accountability and agency as it relates to ICT. Jat is a Fellow at the Alan Turing Institute, and is active in the tech-policy space, serving on advisory councils for the UK Government and Financial Conduct Authority.

Dr Jennifer Cobbe, Coordinator

Department of Computer Science and Technology

Jennifer Cobbe is the Coordinator of the Cambridge Trust & Technology Initiative and a researcher in the Department of Computer Science and Technology at Cambridge. Jennifer holds PhD and master's degrees in Law from Queen's University, Belfast. For her PhD, she studied commercial and state internet surveillance, data protection, and privacy. She researches and writes on law, technology, and society.

Contact Jennifer at: jennifer.cobbe@cl.cam.ac.uk

Dr Laura James, Entrepreneur in Residence

Department of Computer Science and Technology

Laura supports the Initiative part time, alongside other ventures working with emerging internet technologies in different contexts. She has worked extensively in technology and leadership roles in R&D, startups, civil society, humanitarian relief, co-operatives and more. Laura holds Masters and PhD degrees in Engineering from the University of Cambridge, received the Royal Academy of Engineering Leadership Award and a NESTA Crucible Fellowship, and is a Chartered Engineer.

Contact Laura at: laura.james@cl.cam.ac.uk

Dr Ella McPherson

Department of Sociology

Dr Ella McPherson is the Department of Sociology's Lecturer in the Sociology of New Media and Digital Technology as well as the Anthony L. Lyster Fellow in Sociology at Queens' College. She is also Co-Director of the Centre of Governance and Human Rights, where she leads the research theme on human rights in the digital age. Ella's research focuses on symbolic struggles surrounding the media in times of transition, whether democratic or digital. She is particularly interested in the implication of these struggles for the formation, evaluation and contestation of truth-claims. Her current research, which has been funded by an ESRC Future Research Leader fellowship as well as by the Isaac Newton Trust, is on human rights fact-finding in the digital age.

Ella also leads The Whistle, an academic startup supported by an EU Research and Innovation Horizon 2020 grant, which aims to support the collection and verification of human rights information for evidence.

Steering Committee

Dr Anne Alexander

Centre for Research in the Arts, Social Sciences and Humanities (CRASSH)

Anne Alexander is the co-ordinator of the Cambridge Digital Humanities Network, a network of researchers at the University of Cambridge who are interested in how the use of digital tools is transforming scholarship in the humanities and social sciences. This transformation spans both the content and practice of humanities research, as the diffusion of digital technologies opens up new fields of study and generates research questions which breach traditional disciplinary boundaries.

Dr Richard Clayton

Department of Computer Science and Technology

Dr Richard Clayton is a software developer by trade. In the 1980s his company created the system software for the best-selling Amstrad CPC and PCW computers, and then developed "Turnpike" in the 1990s -- one of the first Internet access packages for Windows. The company was sold to Demon Internet, then the UK's largest ISP in 1995, and Richard worked at Demon until in 2000 he was given the opportunity to study for a PhD at the University of Cambridge, He remains an academic ("because it's more fun than working"), doing research into email spam, fake bank "phishing" websites, and other Internet wickedness. As an expert in these areas, he is a regular speaker and media commentator. He has also assisted the APIG and APComms all-party groups of MPs in their inquiries into Internet issues, and he acted as the "specialist adviser" for the House of Lords Science and Technology Committee's two inquiries into "Personal Internet Security".

Dr Rob Doubleday

Centre for Science and Policy

Rob Doubleday has been Executive Director of the Centre for Science and Policy at the University of Cambridge since 2012. Previously Rob established CSaP's research programme. His research interests include the role of science, evidence and expertise in contemporary societies, in particular the relationship between scientific advice, public policy and democracy. In 2010 Rob spent a year on secondment to the Government Office for Science, working on policies to promote engagement between academia and government. Prior to this Rob was the principal investigator of a three-year Wellcome Trust funded project that studied the policy and public dimensions of nanotechnologies. Rob has degrees in Chemistry (Imperial College, London) and Science and Technology Policy (SPRU, University of Sussex). He has a PhD in Geography and Science & Technology Studies from University College London and studied at the Harvard Kennedy School on a Fulbright Scholarship. Rob is

also a Senior Research Associate in the Department of Geography at Cambridge.

Dr David Erdos

Centre for Intellectual Property and Information Law

Dr David Erdos is Deputy Director of the Centre for Intellectual Property and Information Law (CIPIL) and University Senior Lecturer in Law and the Open Society in the Faculty of Law. He is also WYNG Fellow in Law at Trinity Hall. David's principal research interests focuses on data protection looking especially at how to reconcile this with freedom of expression in the areas of journalism, academic scholarship and social media. This work intersects with debates on internet governance generally including, in particular, the liability and responsibility of "intermediary" actors such as Facebook and Google. David's work has been published widely in leading legal and socio-legal journals including the Cambridge Law Journal, the Common Market Law Review, Public Law and the Journal of Law and Society.

Dr Julian Huppert

Intellectual Forum, Jesus College

Julian is the Founding Director of the Intellectual Forum, which is aimed at covering the widest range of academic interests across the College. His background is as a scientist, working on unusual structures of DNA. In particular, DNA of particular sequences can form four-stranded knot-like structures called G-quadruplexes, which can function as genomic switches, turning genes on and off. His work used biophysical and computational methods to predict the formation of these structures, and has led to the identification of a large number of possible anti-cancer drug targets.

After five years away as the MP for Cambridge, his research focus changed to look at science and technology policy, including the

challenges of privacy in the digital age. He has also worked on how to best use evidence in public policy making – a perennial challenge.

Prof Adrian Kent

Department of Applied Mathematics and Theoretical Physics

Adrian Kent is Professor of Quantum Physics in the Department of Applied Mathematics and Theoretical Physics and a Distinguished Visiting Research Chair at Perimeter Institute for Theoretical Physics. His research interests span the foundations of physics and technological applications of quantum information. He pioneered the use of relativistic signalling constraints in cryptography, and co-authored research that sparked the field of "device-independent" quantum cryptography, which gives users security guarantees even when their devices may have been designed by a malicious supplier. More recently, he has developed "supermoney", a form of token that gives users privacy and issuers security against fraud and is faster and more flexible than any existing technology. He has a strong interest in how we most effectively channel science and technological developments to shape our future in positive directions and to reduce catastrophic threats, and is a member of the scientific advisory board of the Cambridge Centre for the Study of Existential Risk.

Prof Dame Theresa Marteau

Behaviour and Health Research Unit

Professor Dame Theresa Marteau is Director of the Behaviour and Health Research Unit at the University of Cambridge. She is also Director of Studies for Psychological and Behavioural Sciences at Christ's College. Her research focuses on the development and evaluation of interventions to change behaviour (principally diet, physical activity, tobacco and alcohol consumption) to improve population health and reduce health inequalities, with a particular focus on targeting non-conscious processes. She also researches the acceptability to publics and policy-makers of government intervention to change behaviour.

Prof John Naughton

CRASSH

Professor John Naughton is a Senior Research Fellow at CRASSH, Emeritus Professor of the Public Understanding of Technology at the Open University, Director of the Wolfson Press Fellowship Programme and the Technology columnist of the London *Observer*. His most recent book, *From Gutenberg to Zuckerberg: what you really need to know about the Internet,* is published by Quercus. He was co-director of the *Technology and Democracy* and *Conspiracy and Democracy* research projects at CRASSH. His most recent work and publications have focussed on surveillance capitalism and the power and responsibilities of technology corporations.

Prof Daniel Ralph

Judge Business School

Daniel Ralph is Professor of Operations Research at Cambridge Judge Business School, and is part of the School's Operations & Technology Management subject group. Professor Ralph is a member of the Australian Mathematical Society, INFORMS, the Mathematical Optimization Society and SIAM. He was Editor-in-Chief of *Mathematical Programming (Series B)* from 2007-2013 and has served on the editorial boards of *Mathematics of Operations Research* and the *SIAM Journal on Optimization*

Dr Manj Sandhu

Department of Public Health and Primary Care

Manj Sandu's research focuses on the integration of principles and procedures underlying population genetics and epidemiology. Together with current and emerging genome-wide technologies, this approach provides unparalleled opportunities to identify the biological mechanisms underlying the development of complex diseases and traits. His work has largely centred on the genetic basis of

cardiometabolic traits and diseases, particularly lipid metabolism and coronary artery disease, and the use of genetic tools for causal inference. More recently, he has begun developing epidemiological resources to explore genomic diversity and its impact on infectious and cardiometabolic risk factors and diseases in Sub-Saharan African populations, as part of a public health and epidemiological research programme.

Dr Phillip Stanley-Marbell

Department of Engineering

My research focuses on designing hardware architectures, algorithms, and programming language constructs that use an understanding of the physical world and the flexibility of sensing systems to improve the efficiency of computing systems that interact with nature. My research results range from fundamental theory, to algorithms, programming languages, and compiler tools. I frequently build printed circuit board and FPGA prototypes to validate concepts.

Dr Adrian Weller

Department of Engineering

Adrian Weller is Programme Director for AI at The Alan Turing Institute, the national institute for data science and AI, where he is also a Turing Fellow leading a group on Fairness, Transparency and Privacy. He is a Senior Research Fellow in Machine Learning at the University of Cambridge, and at the Leverhulme Centre for the Future of Intelligence (CFI) where he leads a project on Trust and Transparency. He is very interested in all aspects of AI, its commercial applications and how it may be used to benefit society. He advises several companies and charities. Previously, Adrian held senior roles in finance. He received a PhD in computer science from Columbia University, and an undergraduate degree in mathematics from Trinity College, Cambridge.

Cambridge perspectives on Trust & Technology

In the run up to our launch event, we asked researchers from Cambridge to give us their thoughts on trust and technology. This is what they said.

Full versions of articles, with references where appropriate, are available on our website.

Excerpts from the Reith Lectures, 2002

Baroness Onora O'Neill

Faculty of Philosophy

Republished with permission; full lectures available at https://www.bbc.co.uk/radio4/reith2002

'Without Trust We Cannot Stand'

Confucius told his disciple Tsze-kung that three things are needed for government: weapons, food and trust. If a ruler can't hold on to all three, he should give up the weapons first and the food next. Trust should be guarded to the end: "without trust we cannot stand". Confucius' thought still convinces. Weapons did not help the Taliban when their foot soldiers lost trust and deserted. Food shortages need not topple governments when they and their rationing systems are trusted, as we know from WWII.

It isn't only rulers and governments who prize and need trust. Each of us and every profession and every institution needs trust. We need it because we have to be able to rely on others acting as they say that they will, and because we need others to accept that we will act as we say we will. The sociologist Niklas Luhman was right that 'A complete absence of trust would prevent [one] even getting up in the morning.'

The Crisis of Trust

We may need trust, but trusting often seems hard and risky. Every day we read of untrustworthy action by politicians and officials, by hospitals and exam boards, by companies and schools. We supposedly face a deepening crisis of trust. Everyday we also read of aspirations and attempts to make business and professionals, public servants and politicians more accountable in more ways to more stakeholders. But can a revolution in accountability remedy our crisis of trust?

The experts and exponents of the crisis of trust are mainly sociologists and journalists: they've tried to find out whom we do and don't trust, in particular whom we say we do and don't trust. They have produced a lot of dispiriting evidence. Remedies are proposed on all sides: politicians and campaigning groups, academics and journalists advocate greater respect for human rights, higher standards of accountability and greater transparency. If these are really the remedies for the crisis of trust, we should surely be seeing some results by now. On the contrary, the accusations mount.

I shall look at trust from a more philosophical but also (I hope) more practical standpoint: these (I believe) go together quite naturally. What does it take for us to place trust in others? What evidence do we need to place it well? Does the revolution in accountability support or possibly undermine trust?

The common ground from which I begin is that we cannot have guarantees that everyone will keep trust. Elaborate measures to ensure that people keep agreements and do not betray trust must, in the end, be backed by --trust. At some point we just have to trust. There is, I think, no complete answer to the old question: 'who will guard the guardians?'. On the contrary, trust is needed precisely because all guarantees are incomplete.

Where we have guarantees or proofs, we don't need to trust.

Since trust has to be placed without guarantees, it is inevitably sometimes misplaced: others let us down and we let others down. When

this happens trust and relationships based on trust are both damaged. Trust, it is constantly observed, is hard earned and easily dissipated. It is valuable social capital and not to be squandered.

If there are no guarantees to be had, we need to place trust with care. This can be hard. The little shepherd boy who shouted 'Wolf! Wolf!' eventually lost his sheep, but we note not before his false alarms had deceived others time and again. Deception and betrayal often work. Traitors and terrorists, embezzlers and con artists, forgers and plagiarists, false promisers and free riders cultivate then breach others' trust. They often get away with it.

We take elaborate steps to deter and prevent deception and fraud: we set and enforce high standards. Human rights requirements are imposed on the law, on institutions, on all of us. Contracts clarify and formalise agreements and undertakings with ever-greater precision. Professional codes define professional responsibilities with ever-greater accuracy.

Huge efforts also go into ensuring trustworthy performance. Auditors scrutinise accounts (but are they trustworthy?). Examiners control and mark examinees (but are they trustworthy?). The police investigate crimes (but are they trustworthy?). Increasingly sophisticated technologies are deployed to prevent and detect breaches of trust, ranging from locks and safes, passwords and identity cards, to CCTV cameras and onto the most elaborate encryption. The efforts to prevent abuse of trust are gigantic, relentless and expensive; and inevitably their results are always less than perfect.

Have these countermeasures begun to restore trust, or to reduce suspicion? Sociologists and journalists report few signs. They claim that we are in the grip of a deepening crisis of public trust directed even at our most familiar institutions and office-holders. Mistrust, it seems is now directed not just at those clearly in breach of law and accepted standards, not just at crooks and wide boys. Mistrust and suspicion have spread across all areas of life, and supposedly with good reason. Citizens, it is said, no longer trust governments, or politicians, or ministers, or the police, or the courts, or the prison service. Consumers, it is said, no longer trust business, especially big business, or their

products. None of us, it is said, trusts banks, or insurers, or pension providers. Patients, it is said, no longer trust doctors (think of Dr Shipman!), and in particular no longer trust hospitals or hospital consultants. 'Loss of trust' is in short, a cliché of our times.

How good is the evidence for this crisis of trust? A lot of the most systematic evidence for the UK can be found in public opinion polls and similar academic research. The pollsters ask carefully controlled cross-sections of the public whether they trust certain professions or office-holders. The questions aren't easy to answer. Most of us would want to say that we trust some but not other professionals, some but not other office-holders, in some matters but not in others. I might trust a schoolteacher to teach my child arithmetic but not citizenship. I might trust my GP to diagnose and prescribe for a sore throat, but not for a heart attack. I might trust my bank with my current account, but not with my life savings. In answering the pollsters we suppress the complexity of our real judgements, smooth out the careful distinctions we draw between different individuals and institutions, and average our judgements about their trustworthiness in different activities.

We depend on journalists for our knowledge of the results of these polls and the levels of reported public trust. There is some irony in this, since these polls repeatedly show that no profession is less trusted in the UK than journalism. Often newspaper reports of public opinion highlight the most dramatic statistic, typically the one that suggests the most extreme distrust. They seldom comment on the ambiguities of the questions or the categories, or linger on cases where trust is average or above average or high.

Active Trust

The polls supposedly show that in the UK public trust in office-holders and professionals of many sorts is low and declining. They certainly reveal a mood of suspicion. But do they show anything more? Are the opinions we divulge to pollsters backed up by the ways in which we actively place our trust in others, and specifically by the way that we place it, or refuse to place it, in public servants, or professionals and institutions?

Much of the evidence of the way we actively place our trust seems to me to point in quite different directions. We constantly place trust in others, in members of professions and in institutions. Nearly all of us drink water provided by water companies and eat food sold in supermarkets and produced by ordinary farming practices. Nearly all of us use the roads (and, even more rationally, the trains!). Even if we have some misgivings, we go on placing trust in medicines produced by the pharmaceutical industry, in operations performed in NHS hospitals, in the delivery of letters by the post office, and in roads that we share with many notably imperfect drivers. We constantly place active trust in many others.

Does action speak louder than words? Are the ways we actually place our trust a more accurate gauge of trust than our comments to pollsters? If we were really as mistrusting as some of us tell the pollsters, would we behave like this? We might do so if we had no options. Perhaps the fact of the matter is that we simply have to rely on institutions and persons although we don't really trust them. In many of these examples, it may seem, we have little choice. How can we avoid tap water, even if we mistrust the water companies, since it is the only ready source of supply? How can we avoid conventional medicines, even if we mistrust the pharmaceutical industry, since there are no effective and available alternatives?

But are these thoughts really convincing? Those who seriously mistrust producers and suppliers of consumer goods can and do refuse to rely on them. Those who really mistrust the tap water drink bottled water, or boil it, or use water purification tablets: where water supplies are seriously questionable people do so. Those who really mistrust the pharmaceutical industry and its products can refuse them and choose to rely on alternative, more natural, remedies and some people do so, but not many. Those who really mistrust the standards of food safety of conventional agriculture, food processing, shops and restaurants can eat organically grown food: it may cost more, but is less expensive than convenience foods and eating out. Where people have options we can tell whether they really mistrust by seeing whether they put their money where they put their mouths. The evidence suggests that we still

constantly place trust in many of the institutions and professions that we profess to not to trust.

Evidence for trust or mistrust is less clear when opting out is hard or impossible. There is no way of opting out of public goods-or public harms. It seems to me that where people have no choice, their action provides poor evidence that they trust-and poor evidence that they mistrust.

Where we have no choice, the only evidence of mistrust is what people say. But we know from cases where they have choice that this can be unreliable evidence. If what we say is unreliable evidence when we have choices, why should we think it reliable evidence when we have no choices? Expressions of mistrust that are divorced from action come cheap: we can assert and rescind, flaunt or change, defend or drop attitudes and expressions of mistrust without changing the way we live. This may show something about indeed rather a lot attitudes of suspicion, but little or nothing about where we actually place our trust.

Is trust failing?

A standard account of the supposed 'crisis of public trust' is that the public rightly no longer trusts professionals and public servants because they are less trustworthy. But is this true? A look at past news reports would show that there has always been some failure and some abuse of trust; other cases may never have seen the light of day. Since we never know how much untrustworthy action is undetected, we can hardly generalise. Growing mistrust would be a reasonable response to growing untrustworthiness: but the evidence that people or institutions are less trustworthy is elusive.

In fact I think there isn't even very good evidence that we trust less. There is good evidence that we say we trust less: we tell the pollsters, they tell the media, and the news that we say we do not trust is then put into circulation. But saying repeatedly that we don't trust no more shows that we trust less, than an echo shows the truth of the echoed words; still less does it show that others are less trustworthy.

Could our actions provide better evidence than our words and show that we do indeed trust less than we used to? Curiously I think that our action often provides evidence that we still trust. We may say we don't trust hospital consultants, and yet apparently we want operations -- and we are pretty cross if they get delayed. We may say that we don't trust the police, but then we call them when trouble threatens. We may say that we don't trust scientists and engineers, but then we rely on hi-tech clinical tests and medical devices. The supposed 'crisis of trust' may be more a matter of what we tell inquisitive pollsters than of any active refusal of trust, let alone of conclusive evidence of reduced trustworthiness. The supposed 'crisis of trust' is, I think, first and foremost a culture of suspicion.

More Perfect Accountability?

The diagnosis of a crisis of trust may be obscure: we are not sure whether there is a crisis of trust. But we are all agreed about the remedy. It lies in prevention and sanctions. Government, institutions and professionals should be made more accountable. And in the last two decades, the quest for greater accountability has penetrated all our lives, like great draughts of Heineken's, reaching parts that supposedly less developed forms of accountability did not reach.

For those of us in the public sector the new accountability takes the form of detailed control. An unending stream of new legislation and regulation, memoranda and instructions, guidance and advice floods into public sector institutions. Central planning may have failed in the former Soviet Union but it is alive and well in Britain today. The new accountability culture aims at ever more perfect administrative control of institutional and professional life.

The new legislation, regulation and controls are more than fine rhetoric. They require detailed conformity to procedures and protocols, detailed record keeping and provision of information in specified formats and success in reaching targets. Detailed instructions regulate and prescribe the work and performance of health trusts and schools, of universities and research councils, of the police force and of social workers. And beyond the public sector, increasingly detailed legislative and regulatory

requirements also bear on companies and the voluntary sector, on self-employed professionals and tradesmen. All institutions face new standards of recommended accounting practice, more detailed health and safety requirements, increasingly complex employment and pensions legislation, more exacting provisions for ensuring non-discrimination and, of course, proliferating complaint procedures.

The new accountability has quite sharp teeth. Performance is monitored and subjected to quality control and quality assurance. The idea of audit has been exported from its original financial context to cover ever more detailed scrutiny of non-financial processes and systems. Performance indicators are used to measure adequate and inadequate performance with supposed precision. This audit explosion, as Michael Power has so aptly called it, has often displaced or marginalised older systems of accountability. In the universities external examiners lost influence as centrally planned teaching quality assessment was imposed; in the health services professional judgement is constrained in many ways; in schools curriculum and assessment of pupils is controlled in pretty minute detail. Schools, hospitals and universities are then all judged and funded by their rankings in league tables of performance indicators.

Managerial accountability for achieving targets is also imposed on institutions although they are given little institutional freedom. Hospital Trusts may be self-governing, but they do not choose which patients to admit or what standards of care to provide. School governors and head teachers have few discretionary powers: they may not select their pupils or expel those whose exam performance will damage their rankings. Universities are supposedly still autonomous, but they have little choice but to cut or close departments with lower research ratings who lose their funding. We are supposedly on the high road towards ever more perfect accountability. Well, I wonder.

Accountability and Mistrust

Have these instruments for control, regulation, monitoring and enforcement worked? Their effects are certainly pretty evident in the daily lives of conscientious professionals and administrators. Professionals have to work to ever more exacting-if changing-standards

of good practice and due process, to meet relentless demands to record and report, and they are subject to regular ranking and restructuring. I think that many public sector professionals find that the new demands damage their real work. Teachers aim to teach their pupils; nurses to care for their patients; university lecturers to do research and to teach; police officers to deter and apprehend those whose activities harm the community; social workers to help those whose lives are for various reasons unmanageable or very difficult. Each profession has its proper aim, and this aim is not reducible to meeting set targets following prescribed procedures and requirements.

If the new methods and requirements supported and didn't obstruct the real purposes of each of these professions and institutions, the accountability revolution might achieve its aims. Unfortunately I think it often obstructs the proper aims of professional practice. Police procedures for preparing cases are so demanding that fewer cases can be prepared, and fewer criminals brought to court. Doctors speak of the inroads that required record-keeping makes into the time that they can spend finding out what is wrong with their patients and listening to their patients. Even children are not exempt from the new accountability: exams are more frequent and time for learning shrinks. We are heading towards defensive medicine, defensive teaching and defensive policing.

The new accountability is widely experienced not just as changing but I think as distorting the proper aims of professional practice and indeed as damaging professional pride and integrity. Much professional practice used to centre on interaction with those whom professionals serve: patients and pupils, students and families in need. Now there is less time to do this because everyone has to record the details of what they do and compile the evidence to protect themselves against the possibility not only of plausible, but of far-fetched complaints. Professionals and public servants understandably end up responding to requirements and targets and not only to those whom they are supposed to serve.

But I'd like to suggest that the revolution in accountability be judged by the standards that it proposes. If it is working we might expect to see indications -- performance indicators!-- that public trust is reviving. But

we don't. In the very years in which the accountability revolution has made striking advances, in which increased demands for control and performance, scrutiny and audit have been imposed, and in which the performance of professionals and institutions has been more and more controlled, we find in fact growing reports of mistrust. In my view these expressions of mistrust suggest that just possibly we are imposing the wrong sorts of accountability. The new systems of control may have aims and effects that are quite distinct from the higher standards of performance, monitoring and accountability that are their ostensible, publicly celebrated aims. We can see this by asking to whom the new audit culture makes professionals and institutions accountable, and for what it makes them accountable.

In theory the new culture of accountability and audit makes professionals and institutions more accountable to the public. This is supposedly done by publishing targets and levels of attainment in league tables, and by establishing complaint procedures by which members of the public can seek redress for any professional or institutional failures. But underlying this ostensible aim of accountability to the public the real requirements are for accountability to regulators, to departments of government, to funders, to legal standards. The new forms of accountability impose forms of central control-quite often indeed a range of different and mutually inconsistent forms of central control.

Some of the new modes of public accountability are in fact internally incoherent. Some of them set targets that cannot be combined without fudging: for example, universities are soon to be told to admit 50% of the age group, but also to maintain current standards. Others are incoherent because they require that targets be achieved by following processes that do not dovetail with targets and can't be made to dovetail with those targets. Again, universities are to treat each applicant fairly on the basis of ability and promise: but they are supposed also to admit a socially more representative intake. There's no guarantee that the process meets the target. Hospitals are to treat each patient on a basis of need and prioritise emergencies, but they are going to be criticised if they postpone non-urgent operations. That might be legitimate grounds for criticism if they could build in spare capacity and do the non-urgent as well as the urgent operations. But the NHS has been made tightly

efficient in its use of resources, so it cannot build in spare capacity on the necessary scale. Incompatible or barely compatible requirements invite compromises and evasions; they undermine both professional judgement and institutional autonomy.

In theory again the new culture of accountability and audit makes professionals and institutions more accountable for good performance. This is manifest in the rhetoric of improvement and rising standards, of efficiency gains and best practice, of respect for patients and pupils and employees. But beneath this admirable rhetoric the real focus is on performance indicators chosen for ease of measurement and control rather than because they measure accurately what the quality of performance is. Most people working in the public service have a reasonable sense not only of the specific clinical, educational, policing or other goals for which they work, but also of central ethical standards that must meet. They know that these complex sets of goals may have to be relegated if they are required to run in a race to improve performance indicators. Even those who devise the indicators know that they are at very best surrogates for the real objectives. Nobody after all seriously thinks that numbers of exam pass levels are the only evidence of good teaching, or that crime clear up rates the only evidence of good policing. Some exams are easier, others are harder, some crimes are easier to clear up, others are harder. However the performance indicators have a deep effect on professional and institutional behaviour. If a certain 'A' level board offers easier examinations in a subject, schools have reason to choose that syllabus even if it is educationally inferior. If waiting lists can be reduced faster by concentrating on certain medical procedures, hospitals have reason so to do, even if medical priorities differ. Perverse incentives are real incentives. I think we all know that from our daily lives. Much of the mistrust and criticism now directed at professionals and public institutions complains about their diligence in responding to incentives to which they have been required to respond rather than pursuing the intrinsic requirements for being good nurses and teachers, good doctors and police officers, good lecturers and social workers. But what else are they do under present regimes of accountability?

In the end, the new culture of accountability provides incentives for arbitrary and unprofessional choices. Lecturers may publish prematurely

because their department's research rating and its funding requires it. Schools may promote certain subjects in which it is easier to get 'As' in public examinations in those subjects. Hospital trusts have to focus on waiting lists even where these are not the most significant measures of medical quality. To add to their grief, the Sisyphean task of pushing institutional performance up the league tables is made harder by constantly redefining and adding targets and introducing initiatives, and of course with no account taken of the costs of competing for initiative funding.

In the New World of accountability, conscientious professionals often find that the public claim to mistrust them-but the public still demand their services. Claims of mistrust are poor reward for meeting requirements that allegedly embody higher standards of public accountability. In ancient Troy the prophetess Cassandra told the truth, but she wasn't believed. Like Cassandra, professionals and institutions doing trustworthy work today may find that the public say that they do not trust them-- but (unlike Cassandra) their services are still demanded. The pursuit of ever more perfect accountability provides citizens and consumers, patients and parents with more information, more comparisons more complaints systems; but it also builds a culture of suspicion, low morale and may ultimately lead to professional cynicism, and then we would have grounds for public mistrust.

Real Accountability?

Perhaps the present revolution in accountability will make us all trustworthier. Perhaps we shall be trusted once again. But I think that this is a vain hope -- not because accountability is undesirable or unnecessary, but because currently fashionable methods of accountability damage rather than repair trust. If we want greater accountability without damaging professional performance we need intelligent accountability. What might this include?

Let me share my sense of some of the possibilities. Intelligent accountability, I suspect, requires more attention to good governance and fewer fantasies about total control. Good governance is possible only if institutions are allowed some margin for self-governance of a

form appropriate to their particular tasks, within a framework of financial and other reporting. Such reporting, I believe, is not improved by being wholly standardised or relentlessly detailed, and since much that has to be accounted for is not easily measured it cannot be boiled down to a set of stock performance indicators. Those who are called to account should give an account of what they have done and of their successes or failures to others who have sufficient time and experience to assess the evidence and report on it. Real accountability provides substantive and knowledgeable independent judgement of an institution's or professional's work.

Serious and effective accountability, I believe, needs to concentrate on good governance, on obligations to tell the truth and needs to seek intelligent accountability. I think it has to fantasise much less about Herculean micro-management by means of performance indicators or total transparency. If we want a culture of public service, professionals and public servants must in the end be free to serve the public rather than their paymasters.

Trust and Information

Socrates did not want his words to go fatherless into the world, transcribed onto tablets or into books that could circulate without their author, to travel beyond the reach of discussion and questions, revision and authentication. So he talked and chatted and argued with others on the streets of Athens, but he wrote and published nothing. (Plato disregarded his teacher's worry and published Socrates' thoughts and conversations anyhow - to our benefit). The problems to which Socrates pointed are acute in an age of recirculated 'news', public relations, global gossip and Internet publication. How can we tell which claims and counterclaims, reports and supposed facts are trustworthy when so much information swirls around us. It is hard to distinguish rumour from report, fact from fiction, reliable source from disinformant, truth-teller from deceiver?

A crisis of trust cannot be overcome by a blind rush to place more trust. Our ambition is not to place trust blindly, as small children do, but with good judgement. In judging whether to place our trust in others' words or

undertakings, or to refuse that trust, we need information and we need the means to judge that information. To place trust reasonably we need to discover not only which claims or undertakings we are invited to trust, but what we might reasonably think about them.

Reasonably placed trust requires not only information about the proposals or undertakings that others put forward, but also information about those who put them forward. Gullible people who put their trust eagerly in blind dates, or pyramid selling schemes, or snake oil merchants and other unlikely purveyors of sure-fire magic do so on the basis of patently inadequate evidence about those who make the proposals they accept; they get our pity or derision but hardly our sympathy. We reserve that sympathy for people who cannot judge those who deceived them. If we are to place trust with assurance we need to know what we are asked be believe or accept, and who is soliciting our trust. Here, it may seem, we are in plenty of luck.

There has never been more abundant information about the individuals and institutions whose claims we have to judge. Openness and transparency are now possible on a scale of which past ages could barely dream. We are flooded with information about government departments and government policies, about public opinion and public debate, about school, hospital and university league tables. We can read facts and figures that supposedly demonstrate financial and professional accountability, cascades of rebarbative semi-technical detail about products and services on the market, and lavish quantities of information about the companies that produce them. At the click of a mouse those with insatiable appetites for information can find out who runs major institutions, look at the home pages and research records of individual scientists, inspect the grants policies of research councils and major charities, down-load the annual reports and the least thrilling press releases of countless minor public, professional and charitable organisations, not to mention peruse the agenda and the minutes of increasing numbers of public bodies. It seems no information about institutions and professions is too boring or too routine to remain unpublished. So if making more information about more public policies, institutions and professionals more widely and freely available is the key to building trust, we must be well on the high road towards an ever more

trusting society.

This high road is built on new technologies that are ideal for achieving transparency and openness. It has become cheap and easy to spread information, indeed extraordinarily hard to prevent its spread. Secrecy was technically feasible in the days of words on paper. But it is undermined by easy, instantaneous, multiple replication-and endless possibilities for subtle or less-than-subtle revision. We may still speak quaintly of 'leaks' of sensitive information, as if information could be sealed in watertight compartments and archives. But in fact we live amid electronic networks through which information travels with ease, at almost no cost in time, skill or money. Special regimes for data protection and for penalising breaches of commercial or other specific sorts of confidentiality are needed to retain even limited areas of secrecy. Openness or transparency is now all too easy: if they can produce or restore trust, trust should surely be within our grasp.

Yet this high enthusiasm forever more complete openness and transparency has done little to build or restore public trust. On the contrary, trust seemingly has receded as transparency has advanced. Perhaps on reflection we should not be wholly surprised. It is quite clear that the very technologies that spread information so easily and efficiently are every bit as good at spreading misinformation and disinformation. Some sorts of openness and transparency may be bad for trust.

In fact, our clearest images of trust do not link it to with openness or transparency at all. Family life is often based on high and reciprocal trust, but close relatives do not always burden one another with full disclosure of their financial or professional dealings, let alone with comprehensive information about their love lives or health problems; and they certainly do not disclose family information promiscuously to all the world. Similarly, in trusting doctor-patient relationships (that's the sort we supposedly no longer enjoy) medically relevant information was disclosed under conditions of confidence. Mutual respect precludes rather than requires across-the-board openness between doctor and patient, and disclosure of confidential information beyond the relationship is wholly unacceptable. I may trust my friends, colleagues

and neighbours whole-heartedly, without any wish, or need, to know everything about their private lives - or to have them know everything about mine.

Perhaps it is not then surprising that public distrust has grown in the very years in which openness and transparency have been so avidly pursued. Transparency certainly destroys secrecy: but it may not limit the deception and deliberate misinformation that undermine relations of trust. If we want to restore trust we need to reduce deception and lies rather than secrecy. Some sorts of secrecy indeed support deception, others do not. Transparency and openness may not be the unconditional goods that they are fashionably supposed to be. By the same token, secrecy and lack of transparency may not be the enemies of trust.

Deception and Misinformation

I think that deception is the real enemy of trust. Deception is not just a matter of getting things wrong. It can be pretty irritating to be misled by somebody's honest mistake, but it is not nearly as bad as being their dupe. The passer-by who in all honesty provides a stranger with inaccurate directions for reaching a destination or the club secretary who in all honesty sends out notices for a meeting on the wrong date are not deceivers. Nor, irritating as they may be, are those who in all honesty undertake to perform tasks that are beyond their competence. Deceivers by contrast mislead intentionally, and it is because their falsehood is deliberate, and because it implies a deliberate intention to undermine, damage or distort others' plans and their capacities to act, that it damages trust and future relationships.

Deception is not a minor or a marginal moral failure. Deceivers do not treat others as moral equals; they exempt themselves from obligations that they rely on others to live up to. Deception lies at the heart of many serious crimes, including fraud and embezzlement, impersonation and obtaining goods by false pretences, forgery and counterfeiting, perjury and spying, smuggling and false accounting, slander and libel. Deception is also part of nearly all theft and most crimes of violence and coercion: burglars enter houses surreptitiously; spies and terrorists establish bogus credentials, live under assumed names, conduct spurious

businesses and form deceptive friendships; murderers often lull their victims into false security and lure them to their deaths. Deceptions may amount to treachery or betrayal. Soviet historians lyingly misrepresented the massacre of Polish officers at Katyn as a German rather than a Soviet war crime; Judas Iscariot falsely played the part of the faithful disciple; Macbeth falsely acted the part of Duncan's faithful vassal. Wolves who wear sheep's (or grandmothers') clothing are not just making mistakes. Nor are card cheats and plagiarists, those who promote false history or scientific fraud, those who write false references for friends (or for colleagues whom they want to shed) or those who corruptly swing contracts, jobs or other favours in the direction of their cronies. Nor are those who hide their conflicts of interest, who promise commitments they have no intention of honouring, or who two-time their partners.

If we want to increase trust we need to avoid deception rather than secrecy. Although some ways of increasing transparency may indirectly reduce deception, many do not. Unless there has been prior deception, transparency does nothing to reduce deception; and even if there has been deception, openness is not a sure-fire remedy. Increasing transparency can produce a flood of unsorted information and misinformation that provides little but confusion unless it can be sorted and assessed. It may add to uncertainty rather than to trust. And unless the individuals and institutions who sort, process and assess information are themselves already trusted, there is little reason to think that transparency and openness are going to increase trust. Transparency can encourage people to be less honest, so increasing deception and reducing reasons for trust: those who know that everything they say or write is to be made public may massage the truth. Public reports may underplay sensitive information; head teachers and employers may write blandly uninformative reports and references; evasive and uninformative statements may substitute for truth-telling.

There are deeper and more systematic reasons for thinking that transparency damages trust. We can only judge whether there is deception, hence reason not to place trust, when we can tell whether we have been fed deliberate falsehoods. But how can we do this when we cannot even tell who has asserted, compiled or endorsed the supposed

information? In a world in which information and misinformation are 'generated', in which good drafting is a vanishing art, in which so-called information 'products' can be transmitted, reformatted and adjusted, embroidered and elaborated, shaped and spun, repeated and respun, it can be quite hard to assess truth or falsehood.

Paradoxically then, in the new information order, those who choose to make up information or to pass it on without checking its accuracy, have rather an easy time. Positions are often maintained in the face of widely available and well-authenticated contrary evidence. Supposed sources proliferate, leaving many of us unsure where and whether there is adequate evidence for or against contested claims. In spite of ample sources we may be left uncertain about the supposed evidence that certain drugs are risky, or that fluoride in the water harms, or that standards for environmental pollutants in water or air have been set too high (or too low or at the right level), that professional training of doctors or teachers are adequate or inadequate, that waste disposal by incineration or by landfill is safer. Proponents of views on these and countless other points may not heed available evidence and can mount loud and assertive campaigns for or against one or another position whether the available evidence goes for or against their views. As the quantity of (mis)information available rises, as the number of bodies with self-conferred credentials and missions and active publicity machines increases, as the difficulty of knowing whether a well-publicised claim is a credible claim increases, it is simply harder to place trust reasonably. Milton asked rhetorically "Who ever knew truth put to the worse in a free and open encounter?". Today the very prospect of a 'free and open encounter' is drowning in the supposedly transparent world of the new information order.

Information and Informed Consent

Global transparency and complete openness are not the best ways to build or restore trust. We place and refuse trust not because we have torrents of information (more is not always better), but because we can trace specific bits of information and specific undertakings to particular sources on whose veracity and reliability we can run some checks. Well-placed trust grows out of active inquiry rather than blind acceptance. In

traditional relations of trust, active inquiry was usually extended over time by talking and asking questions, by listening and seeing how well claims to know and undertakings to act held up. That was the world in which Socrates placed his trust-and his reservations about publishing. Where we can check the information we receive, and when we can go back to those who put it into circulation, we may gain confidence about placing or refusing trust.

But where we can do nothing to check or investigate sources of information and their credentials we often, and reasonably, withhold trust and suspend both belief and disbelief in favour of cynicism and half-belief. We may end up claiming not to trust, and yet for practical purposes place trust in the very sources we claim not to trust. Where possibilities for checking and questioning supposed information are fragmented, trust too may fragment. Even if we do not end up with a crisis of trust we end up with a culture of suspicion.

So if we want a society in which placing trust is feasible we need to look for ways in which we can actively check one another's claims. Active checking has to be more than a matter of checking that many sources of information concur: reading extra copies of a newspaper or extra newspapers lends no extra credibility. Nor can active checking reduce to citing sources such as well-frequented or favourite websites and channels: arguments from authority, to use the old term, however deliciously congruent with favourite beliefs, establish nothing. In an information order in which 'sources' borrow promiscuously from one another, in which statistics are cited and regurgitated because they look striking or convenient for those pursuing some agenda, in which rumour can readily be reprocessed as news, active checking of information is pretty hard for many of us. Unqualified trust is then understandably rather scarce.

Ought we then to conclude that unqualified trust belongs only in face-to-face relationships, where information is provided directly by people we know, whom we can question and monitor? Certainly direct relationships between individuals-intimate or not-can be good for establishing trust, but they are not enough. We need to place or refuse trust far more widely.

We can place trust beyond face-to-face relationships when we can check the information and undertakings others offer. This is after all the function of informed consent requirements, where consent is given or refused in the light of information that should be checkable. Informed consent procedures have a place all the way from choosing socks to choosing university courses, from getting an inoculation to getting married, from choosing a video to choosing a career. Of course, even if all informed consent were given in the light of good and trustworthy information, those who consent can get things wrong. They may choose flimsy socks and boring videos, they may marry philanderers and embark on university courses with which they cannot cope. There are no guarantees. But informed consent can provide a basis for trust provided that those who are to consent are not offered a flood of uncheckable information, but rather information whose accuracy they can check and assess for themselves. This is demanding.

Capacities for testing others' credibility and reliability often fail and falter. Sometimes they falter because the information provided is too arcane and obscure. But sometimes they fail because those asked to consent cannot check and test the information they are offered, so can't work out whether they are being deceived, or whether they can reasonably place their trust. So Socrates' misgivings are not obsolete today. It is very easy to imagine that in a world in which information travels like quicksilver, trust can do the same. It cannot. Placing trust is, I suggest, as demanding today as ever it was in Athens.

A brief exchange of words, a few questions, a short meeting and we begin to place some trust, which we then revise, extend or reduce as we observe and check performance. But how are we to test strangers and institutions? How can we judge claims and undertakings when we can't talk with others, or observe them, let alone send them on lengthy quests? How can we tell that they are not deceiving us?

Perhaps we are in luck. We live in an age of communication technologies. It should be easier than it used to be to check out strangers and institutions, to test credentials, to authenticate sources, and to place trust with discrimination. But unfortunately many of the new

ways of communicating don't offer adequate, let alone easy, ways of doing so. The new information technologies are ideal for spreading reliable information, but they dislocate our ordinary ways of judging one another's claims and deciding where to place our trust.

Socrates worried about the written word, because it travelled beyond the possibility of question and revision, and so beyond trust. We may reasonably worry not only about the written word, but also about broadcast speech, film and television. These technologies are designed for one-way communication with minimal interaction. Those who control and use them may or may not be trustworthy. How are we to check what they tell us?

Informed Consent and Trust

Informed consent is one hallmark of trust between strangers. For example, when I understand a pension plan, a mortgage, or complex medical procedures, and am free to choose or refuse, I express my trust by giving informed consent. We give informed consent in face-to-face transactions too, though we barely notice it. We buy apples in the market, we exchange addresses with acquaintances, we sit down for a haircut. It sounds pompous to speak of these daily transactions as based on informed consent: yet in each we assume that the other party is neither deceiving nor coercing. We withdraw our trust very fast if we are sold rotten apples, or deliberately given a false address, or forcibly subjected to a Mohican haircut. So everyday trust is utterly undermined by coercion and deception.

Informed consent is supposed to guarantee individual autonomy or independence. But I think this popular thought is pretty obscure, because so many views of autonomy are in play. Some people identify individual autonomy with spontaneous choosing. Other people identify individual autonomy not with spontaneous, but with deliberate choosing. But deliberate choosing doesn't guarantee that much either. The real importance of informed consent, I think, has little to do with how we choose. Informed consent is every bit as important when we make conventional and timid choices, or thoughtless and unreflective choices, as it is when we choose deliberately and independently. Informed

consent matters simply because it shows that a transaction was not based on deception or coercion.

Informed consent is therefore always important, but it isn't the basis of trust. On the contrary, it presupposes and expresses trust, which we must already place to assess the information we're given. Should I have a proposed operation? Should I buy this car or that computer? Is this Internet bargain genuine? In each case I need to assess what is offered, but may be unable to judge the information for myself. Others' expert judgement may fill the gap: I may rely on the surgeon who explains the operation, or on a colleague who knows about cars or computers or Internet shopping. But in relying on others I already place trust in my adviser: as Francis Bacon noted, "the greatest trust between man and man is the trust of giving counsel" 1. When we draw on friendly-- or on expert-- help we ultimately have to judge for ourselves where to place our trust. To do this we need to find trustworthy information. This can be dauntingly hard in a world of one-way communication.

Trust and the Media

Today information is abundant, but it's often mixed with misinformation and a little spice of disinformation. It can be hard to check and test what we read and hear. There are easy cases: we can check weather forecasts for their accuracy by waiting for tomorrow; we can rumble supermarkets that don't sell goods at advertised prices. But there are hard cases: how can parents judge whether to have a child vaccinated or to refuse a vaccination? How can we tell whether a product or a service will live up to its billing? Yet for daily and practical purposes we need to place our trust in some strangers and some institutions, and to refuse it to others. How can we do this well?

We know what we need. We need ways of telling trustworthy from untrustworthy informants. And we have tried to make this possible by promoting a revolution in accountability and requirements for transparency in public life. I have argued in previous lectures that we need more intelligent forms of accountability, and that we need to focus less on grandiose ideals of transparency and rather more on limiting deception. Do we really gain from heavy-handed forms of accountability?

Do we really benefit from indiscriminate demands for transparency? I am unconvinced. I think we may undermine professional performance and standards in public life by excessive regulation, and that we may condone and even encourage deception in our zeal for transparency.

Meanwhile, some powerful institutions and professions have managed to avoid not only the excessive but the sensible aspects of the revolutions in accountability and transparency. Most evidently, the media, in particular the print media-while deeply preoccupied with others' untrustworthiness-have escaped demands for accountability (that is, apart from the financial disciplines set by company law and accounting practices). This is less true of the terrestrial broadcasting media, which are subject to legislation and regulation.

Newspaper editors and journalists are not held accountable in these ways. Outstanding reporting and accurate writing mingle with editing and reporting that smears, sneers and jeers, names, shames and blames. Some reporting 'covers' (or should I say 'uncovers'?) dementing amounts of trivia, some misrepresents, some denigrates, some teeters on the brink of defamation. In this curious world, commitments to trustworthy reporting are erratic: there is no shame in writing on matters beyond a reporter's competence, in coining misleading headlines, in omitting matters of public interest or importance, or in recirculating others' speculations as supposed 'news'. Above all there is no requirement to make evidence accessible to readers.

For all of us who have to place trust with care in a complex world, reporting that we cannot assess is a disaster. If we can't trust what the press report, how can we tell whether to trust those on whom they report? An erratically reliable or unassessable press might not matter for privileged people with other sources of information. They can tell which stories are near the mark and which are confused, vicious or simply false; but for most citizens it matters. How can we tell whether newspapers, web sites and publications that claim to be 'independent' are not, in fact, promoting some agenda? How can we tell whether and when we are on the receiving end of hype and spin, of misinformation and disinformation? There is plenty of more or less accurate reporting, but this is very small comfort if readers who can't tell which are the

reliable bits. What we need is reporting that we can assess and check: what we get often can't be assessed or checked by non-experts. If the media mislead, or if readers cannot assess their reporting, the wells of public discourse and public life are poisoned. The new information technologies may be anti-authoritarian, but curiously they are often used in ways that are also anti-democratic. They undermine our capacities to judge others' claims and to place our trust.

Less about trust, more about trustworthiness

We say that we want to end the supposed crisis of public trust, and we've tried to do so in part by making many professions and institutions more accountable so that they are trustworthier. In these lectures I have queried both diagnosis and remedy. We may constantly express suspicion, but it is not at all clear to me that we have stopped placing our trust in others: indeed that may be an impossible form of life. We may constantly seek to make others trustworthy, but some of the regimes of accountability and transparency developed across the last 15 years may damage rather than reinforce trustworthiness. The intrusive methods that we have taken to stem a supposed crisis of trust may even, if things go badly, lead to a genuine crisis of trust.

If we want to avoid this unfortunate spiral we need to think less about accountability through micro-management and central control, and more about good governance, less about transparency and more about limiting deception.

Talking about trust

Dr Laura James

Trust & Technology Initiative

The Trust & Technology Initiative has been awarded a Mozilla Research grant to help bridge between sectors and disciplines when exploring issues of trust around internet technologies. This project seeks to build a

shared understanding of trust and distrust and the dynamics around them, by identifying and showcasing relatable case studies and examples which highlight behaviours, attitudes and challenges around real human experiences of trust and distrust.

Why does this matter?

To build internet technologies which work well for society, we need to have more effective collaboration across different disciplines and sectors, connecting technology development, social science and humanities, policy-makers and more. Today, we lack the shared understanding and terminology to make this work around one of the most critical concepts: trust.

Trust in technology and the organisations which make it is essential for a future in which the internet is inclusive, supportive, diverse, and benefits everyone. A good understanding of trust and trustworthiness, and also distrust and the dynamics of trust, is essential. Trust is the basis on which all people, and organisations, engage and transact online — whether this is for work, play, civic and democratic duty, learning or caring. It is critical that the diverse groups working to build next generation internet systems and governance understand trust, and are able to discuss aspects of it when designing technology and internet infrastructure, and the communities and organisations which build and operate it. This enables the design of good governance structures and the creation of appropriate accountability for connected systems.

Recent internet developments include both aspects of its early promise, and challenging problems. The future internet needs to respect people and communities more than it does today, tackling inequalities from information power and surveillance, and to be built to be more accountable. Accountability means good governance, transparency in some areas, and a balance of power with means of protection and redress for those who need it. Should we build governance structures for the internet, that reflect this, for the new power centres online, sometimes not publicly owned or controlled but instead held by dominant corporations or powerful founding individuals? To even

consider this, we need to examine trust and power - and to do so across boundaries.

Trust and distrust, and the dynamics around them, are a key part of the human experience. We need an appreciation for trust, to enable greater reliance on trustworthy systems, and detection of those which are not. This is not simply a technology challenge. It is a challenge about society, organisations and people. To address it requires interdisciplinary and cross-sector collaboration.

The Trust & Technology Initiative has been learning from researchers in and around Cambridge, and in our discussions with scholars across political and social sciences, arts and humanities, and computer science and technologists, we've found that there are often real misconceptions about basic ideas of trust, which impede collaboration.

Technology developers (and politicians) often assume that trust can be built, and that more trust is necessarily good. But psychologists and political scientists see this as naive. Taking an example outside technology - democratic institutions are designed for a dynamic where the public often will not trust politicians (partly because of the power they wield). This means checks and balances are built around the changing trust/distrust landscape. Trustworthiness – that technology, systems and organisations are honest, competent and reliable – is a more valuable concept than trust.

We're exploring trust in media online, challenges with networked trust, conspiracy theories, the tensions between trust and distrust in power relations, trust making and breaking in cooperative activities, cons and scams, trust in open source communities and collectives as alternatives to corporates for technology provision, how concepts of trust and confidence translate (or don't!) across languages, and how trust issues drive people to adapt tech to their needs.

We're not seeking to create a single universal definition of trust. Such a thing is impossible; cultures and communities have such deep and varied experiences of trust that there is no single shared definition that would make sense. Instead, we're looking to illustrate the different facets

of experience and understanding around trust, distrust and mistrust, to help communities involved in building, designing, governing and evaluating internet technologies to better understand each other.

Trust, technology and truth-claims

Dr Ella McPherson

Trust & Technology Initiative; Department of Sociology

My research focuses on the production, evaluation and contestation of truth-claims in the digital age, and my path into this tangled topic is the empirical case of human rights fact-finding. Because it is so high-risk and so contested, this practice is a canary in the coalmine for wider professions and publics struggling to get to grips with the new information order. Indeed, human rights practitioners working with digital evidence were sounding alarm bells about fake news well before the problem became mainstream and are at the cutting edge of verification methodologies. A concern with trust (and with the associated concepts of trustworthiness and credibility) is at the centre of their work – and thus it is at the centre of mine.

This concern has many dimensions, but I would like to highlight two here that are particularly relevant to the launch of our new and exciting Trust and Technology Initiative. First, we should reflect on the methods we use to evaluate and establish trustworthiness and credibility. As we increasingly encounter unknown sources of information in our hyper-mediated world, we increasingly need to use these methods. Verification is, however, resource-intensive; it requires time and knowledge. Technologists have therefore been seeking and implementing ways of building credibility and trustworthiness cues into ICTs. These practices have significant implications for inequalities in our societies, a second key concern of my research – yet we are so often caught up in protecting ourselves from bad intentions and deceptions that we often overlook these implications. I often use the example of Twitter's blue verified badge to explain this; a user who has the badge has been verified by Twitter as 'authentic,' and as a result, the badge may be used as an

identity verification shortcut by fact-finders evaluating a tweet's truth-claim. But who gets the badge? Twitter says the verified user must have 'an account of public interest. Typically this includes accounts maintained by users in music, acting, fashion, government, politics, religion, journalism, media, sports, business, and other key interest areas.' So it is a pretty elite (and gendered) subset who have the privilege of this shortcut to credibility. As these verification technologies proliferate, we should be mindful of whose cultural understandings of trustworthiness and credibility are built into them, who can meet these standards, who is excluded, and what the implications are for truth-claims in the public sphere.

The second dimension of the relationship between trust and technology I wish to briefly explore is how technologies interfere with and even displace interpersonal trust, which is often built over time through demonstrations of performance and reciprocity. Though new ICTs connect human rights fact-finders to previously inaccessible information, fact-finders still state that face-to-face interviews with witnesses are the gold standard for gathering evidence. This is in part because the information exchange between human rights fact-finder and witness depends on a mutual trust supported by being in each other's presence. By mediating across time and place, ICTs can interfere with this trust-building, so much so that some fact-finders interviewed by The Whistle team said they eschew technology out of the concern that it renders information exchange into information extraction. Other technologies are deliberately developed to replace trust through decreasing the risks we use trust to overcome. As Onora O'Neill explains so well, we trust when we don't have guarantees. We used to have to trust that our children would walk home safely from school – specifically, we would have to trust not only our children but also all the people they encountered on that walk. Now, we can track them real-time on our iPhones with the Find my Friends app; we can guarantee their locations, or at least the locations of their phones. The displacement of trust with technologies is of significant consequence when trust is good for the citizens of a society (which it not always is).

Because of its interdisciplinarity and its reach, the Trust and Technology Initiative is well-poised to explore these dimensions as relates to both research and practice. I am delighted to be a part of it!

Fundamentally more secure computer systems: the CHERI approach

Prof. Simon Moore

Trust & Technology Initiative; Department of Computer Science and Technology

In collaboration with SRI International (California), members of the Computer Architecture and Security groups in the Cambridge Computer Laboratory have spent over eight year exploring fundamentally more secure ways of building computer systems. Starting with a conventional microprocessor, we have added augmented the hardware/software interface with new compartmentalisation primitives that allow security critical properties of software to be better represented. This ensures that the hardware better understands the software that it is running, so is better able to run the code as the programmer intended, not as the attacker tricked it. Compartmentalisation allows software to exploit the principle of least privilege, a fundamental idea in computer security dating back to the 1960s but is ill supported by prior computers.

Our new microprocessor (CHERI), operating system (CheriBSD based on FreeBSD) and compiler support (based on Clang/LLVM) can run existing software while allowing security enhancements to be added automatically via recompilation of the software, and through the developer adding compartmentalisation. We have demonstrated automatic exploit mitigation of security vulnerabilities like Heartbleed (that his banking) and WannaCry (that hit the NHS), as well as mitigating common place buffer overflow/underflow and many return-oriented programming (ROP) and jump-oriented programming (JOP) attacks.

This radical approach was made possible through substantial US government funding (tens of millions of dollars) for a large team over eight years. We have been able to straddle many levels of abstraction that commercially are typically different industries, resulting in market failure. In particular, the hardware and software industries are separate, limiting their ability to optimise across the hardware/software divide.

We are currently working with large industrial players and government actors to bring the technology to the masses. We believe that this new technology can make computer systems far more trustworthy than they are today.

The Political Economy of Trust

Prof. John Naughton

Trust & Technology Initiative; CRASSH

Much of the discussion of trustworthy technology is understandably focussed on the technology itself. But this ignores the fact that the kit doesn't exist in a vacuum. Digital technology is now part of the everyday lives of four billion people and in the process has raised clear questions of trust, reliability, integrity, dependability, equity and control. Some of these issues stem from technical characteristics of the equipment; others stem from the fallibility or ignorance of users; but a significant proportion come from the fact that network technology is deployed by global corporations with distinctive business models and strategic interests which are not necessarily aligned with either the public interest or the wellbeing of users.

An interesting current example is provided by VPN (Virtual Private Network) technology. This enables users to create a private network that runs on a public network, thereby enabling them to send and receive data across the public network as if their computing devices were directly connected to the private one. The benefits of VPNs include enhanced functionality, security, and privacy protection and they are a boon for Internet users who need to use 'free' public WiFi services in hotels, cafes

and public transport. In that sense VPN is a technology that enhances the trustworthiness of open WiFi networks.

Earlier this year, Facebook offered some of its users Onavo Protect, a VPN developed by an Israeli company that Facebook owns. A link to the product appeared in the feeds of some US Facebook IOS users under the banner “Protect”. Clicking through on this led to the download link for “Onavo Protect — VPN Security” on the Apple App Store.

The blurb for the App included a promise to "keep you and your data safe when you browse and share information on the web" but omitted to point out that its functionality involved tracking user activity across multiple different applications to learn insights about how Facebook customers use third-party services. Whenever a user of Onavo opened up an app or website, traffic was redirected to Facebook's servers, which logged the action in a database to allow the company to draw conclusions about internet usage from aggregated data.

Needless to say, close inspection of the Terms and Conditions associated with the app revealed that “Onavo collects your mobile data traffic. This helps us improve and operate the Onavo service by analyzing your use of websites, apps and data”. Whether non-technical users — who presumably imagined that a VPN would provide security and privacy for their browsing (rather than enabling Facebook to track their online activities outside of its ‘walled garden’) understood what this meant is an interesting question. In August 2018, Apple settled the issue — ruling that Onavo Protect violated a part of its developer agreement that prevents apps from using data in ways that go beyond what is directly relevant to the app or to provide advertising, and the app was removed from the Apple Store. (It is still available for Android users on the Google Play store.)

And the moral? In assessing trustworthiness the technical affordances of the technology are obviously important. But they may be only part of the story. The other part — the political economy of the technology — may actually turn out to be the more important one.

Compliant and Accountable Systems

Dr Jatinder Singh

Trust & Technology Initiative; Department of Computer Science and Technology

The “Compliant and Accountable Systems” research group takes an interdisciplinary (tech-legal) approach towards issues of governance, agency, accountability and trust regarding emerging technologies.

ICT continues to underpin everyday life. But what happens when it fails? How are those responsible held to account when things go wrong? How can we even determine who is responsible? How do we manage such risks? Can technologies be interrogated to ensure they are fit for purpose before deployment?

Such questions are particularly pertinent, as systems become gradually more pervasive and complex; technical environments progressively more data driven, autonomous and physical; and as visions of smart cities, smart homes and the Internet of Things become a reality.

In line with this, we see that technology is increasingly the subject of public discussion and regulatory attention - the EU General Data Protection Regulation is a prominent example. There is growing demand for improving levels of accountability regarding the technology that influences everyday life, not least as ICT/data related scandals are reported most daily.

Issues of compliance and accountability directly relate to trust, as they are key to both the adoption and public acceptance of technology, and to ensuring that the technologies deployed are, and remain, appropriate and fit for purpose, align with social norms, where those responsible can be held to account when and where necessary.

Towards this, the newly formed Compliant and Accountable Systems research group works to better aligning technology with legal concerns, and vice-versa. The team, based at the Department of Computer Science

& Technology, is multi-disciplinary, consisting of computer scientists and lawyers. Its focus is to tackle these socio-technical issues, by (a) exploring technical responses to legal problems, (b) providing legal input to guide technology development, (c) developing technical means facilitating compliance and accountability, and (d) interrogating legal frameworks for new and emerging ICT.

Core research team

Dr. Jennifer Cobbe
Dr. Heleen Janssen
Dr. Chris Norval
Dr. Jatinder Singh

AI Trust & Transparency with the Leverhulme Centre for the Future of Intelligence

Dr Adrian Weller

Trust & Technology Initiative; Department of Engineering

This project is developing processes to ensure that AI systems are transparent, reliable and trustworthy.

As AI systems are widely deployed in real-world settings, it is critical for us to understand the mechanisms by which they take decisions, when they can be trusted to perform well, and when they may fail. This project addresses these goals in three strands.

Transparency

We need to be able to understand the internal processes of AI systems. This is particularly challenging for approaches such as neural networks or genetic algorithms, which learn or evolve to carry out a task without clear mappings to chains of inference that are easy for a human to understand. This strand will study ways to make interpretable the reasons for an AI's predictions or decisions.

Reliability

Real-world AIs need to perform reliably in settings that could be very different to their training environments, with associated risks of unpredictable and unwanted behaviours. We seek to develop new approaches that can guarantee good performance for scalable probabilistic reasoning, even in unforeseen settings. This may include notions of learning and inference which can supply proofs of accuracy, as used in formal verification systems. Another approach is to explore ways for an AI to monitor its situation dynamically to detect if its environment has changed beyond its reliability zone (allowing an alert and shift to a fallback mode).

Trustworthiness

Human studies indicate that a theory of mind may be essential for empathetic trust, and for reliable initiation of acts of kindness. Equipping AIs to infer beliefs and goals of agents such as humans may improve human-machine collaborations; yet such insight may prove a double-edged sword, allowing deception and even manipulation. We shall explore these themes with researchers on the Agents and Persons, and Kinds of Intelligence projects, and with leading experts from psychology.

Why and How Email Communication Undermines Trust in Teams and Organizations

Prof. David De Cremer[1], Jack McGuire[1] and Dr Tessa Haesevoets[2]

[1] *Judge Business School, University of Cambridge*
[2] *Ghent University*

Even though the technological use of Email has been predicted to come to an end, within organizations it is still one of the most commonly used communication channels. Recent research indicates that with the use of

open work spaces – aimed at promoting more face-to-face communication – the use of email has been increasing again. Email thus remains an important communication tool because it primarily helps to distribute information among different parties involved. Transparency is key in ensuring that this communication technology is trusted by all recipients. Email, at the same time also offers several different possibilities on how to communicate. It is very easy to add people to cc and bcc, which makes that Email can also be turned in a more strategic tool of communication. Despite that these different communication options are easy to select the consequences can however be detrimental to team work. In fact, Justin Rosenstein (ex-Facebook and co-founder of Asana) commented that Email has "become a counter-productivity tool." In what way can Email as an important technology tool for organizational communication disrupt negatively team work?

We conducted a series of experimental studies revealing that using cc and bcc can significantly influence trust within teams and make people suspicious towards the use of Email. A first line of research demonstrated that when co-workers emailing colleagues include their supervisor often in cc, those colleagues felt trusted less by that co-worker and questioned his/her motives. Furthermore, including the supervisor in cc not only decreased trust between team members but also led people conclude that the organizational culture was low in trust and fostered a culture of fear in which psychological safety is lacking. In a second line of research we discovered that that cc-ing the supervisor is nevertheless a more acceptable communication strategy than the use of the bcc-option. Objectively speaking the cc-option does not violate expectations of transparency and therefore does not adhere to a "secrecy" strategy like bcc does.

Our results overall demonstrated that the default seems to be that most employees perceive others making use of the bcc-option to include the supervisor as having immoral intentions. It is these perceptions that make that once the use of bcc is found out serious trust violations emerge that are hard to repair. An alternative approach to the bcc-option – and one which is regarded as part of standard business etiquette - that many employees use concerns the strategy of forwarding emails. The strategy of forwarding emails to a supervisor after other employees had

received the email first was considered more legitimate than bcc-ing the supervisor. However, it was still evaluated as presenting a threat to one's privacy because the exact same message is forwarded and therefore trust in the one forwarding the email eventually decreased as well. In conclusion, our research underscores the need to design clear and directive protocols regarding the use of communication technology like emails in the context where team collaboration and the fostering of a trusting work culture is expected.

Digital Trust Dissonance: when you've got them by the app, their clicks and minds will follow

Richard Dent

Department of Sociology

Lisa Khan's 93 page article about Amazon's business practices suggests an antitrust legal intervention may be required. This has sustained an ongoing public debate about trust in "big tech". However, Amazon customers are not giving up their Prime accounts, including Khan's husband who is a regular user according to her. Over at Facebook, the fallout from the Cambridge Analytica scandal hasn't had a major impact on their bottom line, despite an increase in distrust. According to a recent survey, 81% of respondents reported that they 'have little no confidence Facebook will protect their data and privacy', which is line with Business Insider's Digital Trust annual survey (Business Insider, 2018). Yet Facebook report that their "daily active users" and "monthly active users" have not declined and analysts suggest advertisers are not looking elsewhere (Business Insider, 2018a). An independent study by the Pew Research Centre (2018) showed that more people are changing their privacy settings on Facebook. However, a mass exodus has not taken place. Have Amazon and Facebook users entered a state that we might call 'digital trust dissonance'?

This concept tries to explain why research shows that millions of people express distrust for major technology corporations (Google, Facebook, Microsoft, Amazon), yet continue to use these platforms with little or no restraint. Even when a company has admitted to losing users' personal data in a major hack, for example the recent British Airways hack. This behaviour seems to echo the privacy paradox, a cognitive dissonance discovered by academics like Susan B. Barnes (2006) when studying early use of social networking sites like Myspace. Barnes found that people express genuine concerns about their online privacy, yet continue to broadcast personal details in public forums and on websites that warn them that they are collecting their data. Are we seeing a similar effect with trust in digital technologies?

Distrust in technology is nothing new (E.g. the luddites). One of more vocal groups in society, older adults, will "frequently deploy the concept of distrust" (Knowles & Hanson, 2018) when talking technology as a reason for non-use. However, these are likely to be outliers, Digital trust dissonance could have several causes. A primary cause is likely that individuals are simply 'locked in' when it comes to using specific technology platforms, either by their employer or family. The time cost and compatibility issues associated with switching to alternative platform are too high. For example, try using Open Office instead of Microsoft Office when all your colleagues use the Microsoft platform. You may trust the makers of Open Office (Apache) more than Microsoft, but look out of the inevitable email from a friend or co-worker who cannot open your documents.

Another reason for digital trust dissonance could be a lack of visibility of the negative impacts of technology usage. It is hard to see how one's trust has been betrayed if one cannot observe any real-world impact or 'direct betrayal'' of their trust. This sets a dangerous precedence as individuals' may become resigned to the fact that their trust in a technology inevitably comes with some downsides, leading to a dependency that becomes hard to break.

One example of where this can go wrong is with the collection of our health data. As we channel shift our medical records to online platforms, such as Patientaccess.com, we make ourselves vulnerable. This might

not have a major impact on a person initially, but health data in digital form can more easily find its way to credit score companies, potential employers or governments. This can happen without our knowledge (or because we didn't read the terms and conditions or privacy policy). A more dystopian view is that technology providers become the agent for a political regime that seeks to target specific ethic groups or classes. This happened in Europe when smartphone meta-data was collected by EU authorities when a change public attitude prompted policy changes favouring the deportation of refugees instead of integration.
An explanation for the digital trust dissonance may also from an explanation for the privacy paradox as proposed by Hallam and Zanella (2017). They suggest "a temporally discounted balance between concerns and rewards". In other words, the more distant to the individual a privacy breach is, the more that individual will discount it. The same may be true with trust. If we had the specific details of what data we lost in a breach, and which agents received that data, our trust would break more profoundly. If we are bundled in with millions of others, with few or no details about our individual data, we may discount our distrust.

Entities like the EU are showing how regulation can place a 'check' on large multinational tech companies (E.g. Google), which might actually increase our trust in them. However, can regulations go far enough when companies lose millions of user details to hacking from criminals or hostile regimes? How can we build healthy levels of trust and distrust? Trust in technology to improve our lives with the right amount of distrust to lobby for better security, regulation and fair use of our data. It seems that the major tech companies have learned that major user data breaches, negative press or wider breaches of social trust seem to have little effect on their business. Perhaps we've become so dependent on technology in the 21st century that when you've got 'em by the app, their clicks and minds will follow.

Digital Media, Voice and Power in Africa

Dr Stephanie Diepeveen

Department of Politics and International Studies (POLIS)

Digital media appear to disrupt power in profound and polarising ways: opening up new channels for voice, and also bringing unprecedented forms of surveillance by state and private actors. In Africa, the stakes are high. The 2017 elections in Kenya bring these two sides into sharp relief. A critical and satirical political commentary of Kenyans on Twitter (#KOT) contrasted with attempts at control and surveillance as the governing coalition contracted the behavioural insights/marketing firm Cambridge Analytica, 'fake news' was reportedly rife, and the technology manager at the electoral commission was murdered a week prior to election day.

Yet, the ways that scholars have responded to the intersection of power and digital communication in Africa remains unconvincing. They are heralded as the new tools for development of the continent, feared as new weapons for coercive state control and surveillance, imbued with agency as the means for quick, decentralised political organisation and protest. Leading scholars emphasise that our understanding of what is changing in the nature and distribution of political power must get beyond impatient and crude binaries, with digital communications as either utopian or dystopian (Papacharissi, 2010), as emancipatory or coercive (Dahlgren, 2013), or as the democratic second coming or tools for domination and accumulation (Dean, 2001, 2008).

Our research agenda takes as its premise that understanding how and why digital communication technologies intersect with power cannot be answered by looking at forms of control alone, or solely new collective actions. These are only incomplete insights into what, in sum, is at stake with the politics of digital communication technologies. Missing from the picture is the way that communication technologies entangle both forms of power, simultaneously, in ways that appear in tension but are also mutually constitutive of the other. Digital communication technologies are recalibrating the basic paradox of political power: the

tense, co-constituting mix of coercion with consent, domination with agency, and rule with political action.

Working from this premise Cambridge's Centre of Governance and Human Right's (CGHR) research agenda in Digital Media, Voice and Power attempts to re-think the relationship between power and digital media through conceptual and empirical research. Over the past year, we have been doing framing research into first, how to conceptualise forms of power, and second, how alignments of power have intersected with communication technologies historically on the African continent. Going forward, we plan to conduct case studies focused on specific technical 'objects' that have been integrated into the exercise of authority and its contestation in a selection of East African countries. Our conceptual and historical framework forms the reference point for interrogating newer communication technologies and their intersection with forms of power concurrently. The first case study, looking at methodologies of data collection at the county level in Kenya, and how it relates to power, perceptions of power and public trust, will begin in October 2018.

Govtech Requires Many Relationships of Trust

Dr Tanya Filer

Bennett Institute for Public Policy

Governments around the world are beginning to support the growth of domestic Govtech, or government technology, industries. Govtech companies—typically start-ups and SMEs—seek to serve the public sector as client, maximising the efficiency of their public service provision. For governments, the development and sustainable growth of the Govtech industry holds a double allure: the promise of economic growth as the global Govtech market courts valuations of $400 billion annually; and the possibility of innovating the domestic public sector at a moment when the institution of government is at a crisis point—frequently perceived as retrograde and excessively bureaucratic. My research explores policies for building sustainable and citizen-centred

Govtech ecosystems. It finds three relationships of trust—belief in the reliability and capacity of others—to be crucial to this effort.

Institutional Trust

A crisis of trust in governance institutions has long been stirring across democracies. We are also bearing witness to increased scepticism towards the trustworthiness of digital and new technologies and the people who create them. Govtech emerges at the convergence point of this double dip in public confidence. As govtech companies gain power, they can help to recuperate citizens' trust, for example by reducing fraud. But a deficit of accountability could amplify the crisis. It may hamper delivering efficiency and innovation, which are better served by trust-based legitimacy, thus pushing citizens further towards anti-system options. To mitigate this risk, we should settle on a definition of Govtech as intrinsically a dual-purpose sector—efficiency boosting and accountability boosting—and evaluate individual govtech ventures against both those criteria

Ecosystemic Trust

The need for shared definitions feeds into a broader question of trust across the ecosystem, or community of individuals and organisations who will build, buy, use and regulate Govtech products and services. Policymakers and entrepreneurs often view each other with mutual curiosity but scant understanding. Their modi operandi diverge extremely. To work together productively, governments and Govtech companies require skilled 'translators' equipped to navigate between with the different languages, cultures, priorities and ambitions both across the technological and policymaking dimensions of government, and between tech firms and the state. These translators can create trust through enabling communication and building comprehension.

Examples of translators' work are becoming evident. Bird, a mobility company providing electric scooters, recently launched a GovTech platform for city governments. The approach is commercially shrewd, appealing to regulators who may otherwise treat them heavy-handedly.[5] But it is also empathetic, demonstrating understanding of the challenges

of urban governance by providing city administrators with the tools to oversee how vehicles occupy the public spaces that they govern.

Investor Trust

Govtech enterprises need time to build public-sector knowledge and relations. Private venture capitalists, whose financing many Govtech founders seek, often expect sizeable returns on their investments in just three to seven years. Funding deployed with a demand for unrealistically speedy profit could yield a cohort of companies that fail to acquire the depth of experience to negotiate government procurement processes or build familiarity with institutional cultures and key decisionmakers in the departments into which they seek to sell. It may also limit knowledge diffusion across the industry.

The economist Mariana Mazzucato elucidates how public funding bodies have proved adept at providing 'patient finance'—investment accepting of uncertain conditions and long and inexact schedules—for high-risk technological development, through alignment with the missions of their investments, and a willingness to bet 'long' on them.[6] This provision of time is an act of trust in the capacity of high-risk industries to flourish. Public funding bodies have extended patient funding premised on that faith (alongside an acceptance of occasional failure). As a result, they have been crucial to the development of high-risk, new technology-based industries.

Many Govtech ventures sit at the applied end of the innovation lifecycle, sidestepping the uncertainties of basic technological feasibility that underpin high-risk technology industries. Yet their risk profile—working to slow-moving and often unclear demand-side timeframes—means that they, too, require time, and the attendant patient financing, to grow. From Israel to the United Kingdom, public funders are beginning to step up to the mark. This move is encouraging.

TRVE Data: Secure and Resilient Collaborative Applications

Dr Martin Kleppmann

Department of Computer Science and Technology

Cloud-based collaboration tools such as Google Docs, Evernote, iCloud and Dropbox are very convenient for users, but problematic from a security point of view. At present, most such services are provided by companies through a centralised server infrastructure, which is vulnerable to operational mistakes by the service provider, security breaches, and cyberattacks.

The goal of the TRVE Data project (pronounced “true data”) is to build the foundation for the next generation of collaboration software, providing stronger security and resilience than the current practice. We are developing algorithms, protocols, and code that allow real-time collaboration and data synchronisation across several devices without relying on central servers. Our research is based on the following principles:

End-to-end encryption

Today's Internet services typically process data in unencrypted form on their servers, and employ encryption (e.g. TLS) only for communication between servers and end-user devices (such as laptops or smartphones). Hence, users depend on the cloud provider to prevent unauthorised access and to maintain integrity of the data. A security breach of the provider could have disastrous consequences: a hacker who gains access to the servers, or a rogue employee, can potentially read and tamper with vast amounts of sensitive data.

In contrast, we are designing systems to use end-to-end encryption, which secures data all the way from one user's device to another user's device. In this approach, servers only ever handle encrypted data that

they cannot decrypt. Thus, even if communication networks or servers are compromised, the confidentiality and integrity of sensitive data are protected, giving users better ownership and control over their data.

Making servers optional

At present, services typically transmit all data via a central server. Even if the communicating devices are in the same room, their data might be sent via a server on another continent. This approach is not only slow and wasteful, it also makes the system susceptible to disruption: if the server is blocked or subjected to a cyberattack (e.g. a DDoS attack), or if the operator goes out of business, the software stops working.

To improve the resilience of applications, we are using peer-to-peer communication where possible, sending data directly between collaborating devices, and utilising fast local networks when applicable. Servers may still be used, but the software continues working if servers are unreachable. Using local storage and local networks further improves users' control over their own data.

Open source and open standards

All software developed in this project is made freely available as open source, so that it can be easily adopted by application developers.

We have implemented this approach in Automerge, a JavaScript library for building collaborative applications. Automerge allows users to read and modify data even while their device is offline, and it performs data synchronisation and automatic conflict resolution when a network connection is available. Unlike most existing data synchronisation systems, Automerge does not require data to be sent via a centralised server, but rather allows local and peer-to-peer networks to be used, and it is compatible with end-to-end encryption protocols.

The Autonomous City

Dr Ian Lewis

Department of Computer Science and Technology

We have a broad range of research in the Department of Computer Science and Technology to tackle issues and opportunities arising from the global densification of populations into large urban centres. Our Adaptive Cities Programme is designed to exploit high-volume sensor deployments, collecting and acting upon urban data collected in real-time. The use of the word 'Adaptive' (we could have chosen 'Future') emphasises that we are collecting the data because we are likely to want to do something about it. For example traffic congestion might be improved by changing the signalling and similar considerations will apply to air quality, waste collection, power distribution and other infrastructure areas.

Our expectation is that this 'adaptation' will be most effective if based upon something the city is predicting is going to happen, rather than waiting for an issue to occur and taking action then. Returning to the example of traffic management, the predictive aspect of the real-time sensor analysis enables the city to adjust to what's coming in a timely fashion, while the large-scale deployment of both the traffic sensors and the control fabric supporting the signalling adjustments means that the optimal adjustment can intelligently be made across a large area of the city. The hypothesis is that these control enhancements in both time and space should lead to more effective management of the urban environment i.e. less congestion, better air quality, more efficient waste collection etc.

Trust, Evidence and Local Democracy; how Cambridgeshire County Council has bridged the gap

Ian Manning

Cambridgeshire County Council

Cambridgeshire County Council needs the trust of its residents; it needs to be able to back up decisions it makes with expert evidence, and show it's open to criticism - and we've done that, in a way that no one else has.

We've created a direct link between researchers and policy makers, taking evidence into policy changes for the Council.

We worked with CUSPE and CSaP and found volunteer researchers interested in tackling challenges facing the County Council; from elected Councillors we got a long list of challenges; resulting in three teams working on three challenges:

1. “What are the next generation of models to transform organisations, and how could they benefit Cambridgeshire County Council?”
2. “Why is Cambridgeshire’s educational achievement gap so bad, and what can be done about it?”
3. “What actions would have most impact in reducing deprivation inequalities in Cambridgeshire?"

Policy changes came from this: the Council invited best practice from all schools; looked at outcome based models for service reconfiguration - validating the work the transformation team was doing; created a template for how schools could report on the use of pupil premium spending; informed the schools forum;

Success in politics *is* trust. One can vote the right way, do the right thing, support your constituents; but if they don't *trust* you, you won't get re-

elected, you won't be *able* to support your constituents, do the right thing, vote the right way.

If both politicians and civil servants can work with external researchers to improve policy outcomes then we can improve that trust.

Giving Voice to Digital Democracies

Dr Marcus Tomalin

CRASSH

This exciting new project will begin on 1st October 2018, and it is one of the inaugural projects for the Centre for the Humanities and Social Change that is based at CRASSH. The Centre forms part of the Humanities and Social Change International Foundation (https://hscif.org). The project will explore the profound social changes induced by Artificial Intelligence (AI) and Information and Communications Technology (ICT) in modern digital democracies.

The importance of this subject cannot be overestimated. Currently, there are serious concerns that these technologies endanger digital, physical, and political security, induce social fragmentation, create filter bubbles and echo chambers, facilitate the spread of mis/dis/malinformation, adversely affect mental health, and therefore risk undermining the very fabric of democracy itself. However, other experts predict that the very same technologies will enable progress in health, education, transportation, energy, the environment, and welfare, promote free speech, human rights, and government accountability, while contributing up to $15.7 trillion to the global economy by 2030. Despite these starkly conflicting perspectives, China has outlined detailed plans for AI and ICT research, and there are comparable ambitions in America, Japan, South Korea, and Singapore. Clearly, this growing international consensus concerning the need for the rapid development of these technologies requires urgent disinterested scrutiny. The recent success of neural-network-based techniques have prompted the current interest in intelligent technologies, and open-source software libraries such as

‘Tensorflow’ and ‘PyTorch’ have facilitated research involving deductive reasoning, pattern recognition, and automated inferences. However, any sense of homogeneous uniformity is a chimera. AI and ICT actually constitute a sprawling set of loosely interconnected approaches that pose markedly different challenges.

Therefore, to avoid superficial generalisations, this project will focus specifically on technologies at the intersection of AI and ICT (from henceforth ‘AICT’) – namely, speech technology, natural language processing, smart telecommunications, and social media. Intelligent conversational agents such as Siri (Apple), Cortana (Microsoft), and Alexa (Amazon) exist at the very centre of this intersection, and the societal potentialities of such systems are enormous. In the next 5-10 years, they will increasingly influence all aspects of our lives, from how we turn the heating on and off, to how we encounter news stories and vote in national elections. Surprisingly, though, the specific linguistic, ethical, psychological, sociological, legal, and technical challenges posed by AICT rarely receive the focused attention they deserve.

Consequently, this project will examine the social impact of AICT in modern digital democracies. The core aim is to establish a viable research framework that enables cutting-edge interdisciplinary thinking to influence directly the development of AICT, so as to make it more trustworthy, accurate, unbiased, and transparent. Therefore, this project will provide a unique opportunity to determine how existing AICT infrastructures can be reconfigured to enable the resulting technologies to change society more beneficially in the future.