THE ROYAL SOCIETY

האקדמיה הלאומית הישראלית למדעים
المجمع الوطني الإسرائيلي للعلوم والآداب
THE ISRAEL ACADEMY OF SCIENCES AND HUMANITIES

Israel-UK privacy and technology workshop
note of discussions

Ralph Kohn memorial seminar series on science, technology and ethics

Jerusalem, 25 – 26 October 2017

Contents

Foreword

The Ralph Kohn Memorial Seminar Series on Science, Technology and Ethics has been established by The Royal Society and The Israel Academy of Sciences and Humanities to enhance collaborative activities between the two Academies, in areas relating to science and society. The first seminar in this series focused on Privacy and Technology, exploring the impact of new digital technologies on notions of privacy, and different approaches to preserving privacy, to ensure we derive maximum benefits from these technologies.

Modern information communication technologies create new opportunities and challenges for individuals and society. Data collection and generation is now near ubiquitous, carried out by new actors, and at new scales. The internet and social networks facilitate the acquisition, storage, manipulation and dissemination of information globally and in real time. In parallel, increasingly powerful analytics tools have become available, making it possible to infer valuable and sensitive knowledge about individuals, with applications ranging from personalised services to counter-terrorism. Such technological applications change the boundaries between public and private spheres, challenging traditional concepts of privacy, and creating new roles and responsibilities for the state, the private sector and the individual.

These developments raise questions about what privacy means in the information age, and how it can be protected effectively in this new era. What is privacy? What is an appropriate level of privacy? How has technology changed privacy? How do individuals and societies manage privacy? What role is there for technology in protecting privacy?

These questions were addressed by experts from Israel and the UK present at the workshop, covering a range of disciplines, including computer science, cryptography, medicine, social psychology, behavioural sciences, and the law.

The organisers are grateful to the Kohn Family Foundation for its support of enhanced collaborative activities involving the Royal Society and the Israel Academy.

Disclaimer: This is a note summarising the discussion and debate at The Israel Academy of Sciences and Humanities and Royal Society's workshop on Privacy and Technology. It is not intended to represent the views of either of the Academies, nor does it represent the views of individual attendees at the event.

Notions of privacy
What is privacy?

Privacy can broadly be defined as what facilitates the maintenance of the private, what is not shared, not open to the public. What this entails exactly is subject to much discussion.

An evolving context

Arguably privacy has always existed in human communities. However, as communities evolved, so has the notion of privacy. Some tens of thousands of years ago, people lived in small communities which were probably the size of an extended family with a few friends. People were barely aware of, and hardly interacting with, the world beyond their own circle. At that time, the confine of the public and the private was defined within that small group. Since then, the boundaries of the private and the public have been blurred. Globalisation increases the fluidity of borderlines between communities, with every internet user connected with billions of other people.

A multi-dimensional concept

Even today, scholars from different disciplines have very different concepts of privacy. For example, many cryptographers consider that privacy means maintaining anonymity. Defined as such, privacy would be an 'on or off' variable. This contrasts with the view that there is a spectrum of privacy – from full privacy to full transparency. Some consider privacy to be multi-dimensional. Privacy can indeed be related to the protection of information related to a person's identity, but it is also intimately linked with the protection of values such as freedom, security, autonomy, and dignity.

Individual good and public good

Privacy can be considered as a benefit to individuals, allowing them to choose whether to share information about them with whoever they want.

It is not as straightforward to grasp the importance of privacy for society. Indeed, full privacy, or in other words the absence of transparency, would put a break on the progress that can be achieved through information sharing. It would also make the work of national security agencies much harder, and thus could threaten security and stability.

The relative anonymity perceived by most individuals in Western societies may lead them to undervalue the dangers of loss of privacy to themselves, to other individuals and society. In addition, an emphasis on individualism, human rights and freedom may even lead some to be suspicious about the need for privacy.

Privacy, however, is instrumental to both individual and social flourishing. It underpins intimacy, autonomy and to growth. Provided they have appropriate levels of these, individuals may develop the strong core in themselves that is needed for them to potentially go out to the public and perform great deeds in the arts, sciences and politics. This in turn will benefit society, through innovation and adaptation to change.

What is an appropriate level of privacy?

Illustrating the importance of privacy as a collective good, the next section sets out examples, fictional and real, of the consequences of either a total absence of privacy or total lack of transparency. Such extreme scenarios are helpful starting points to think about what might be an acceptable balance between private and public.

Fictional depictions of privacy pushed to the limits

Fictional accounts of privacy make the picture clearer. Winston, the tragic hero of *Nineteen Eighty-Four*[1] and Mae, the heroine of *The Circle*[2], are people who experience total loss of privacy. In *Nineteen Eighty-Four*, Orwell depicts extreme state surveillance, where every act and thought is policed by "Big Brother". In *The Circle*, a high-tech company gives its workers a very attractive deal of marvellous working conditions, but the mission of the company is the abolition of privacy. Privacy is theft. People have a right to know everything.

These dystopias are very suggestive. They identify, in a chilling way, processes involving loss of privacy that are both realistic and very troubling to most readers. They illustrate very powerfully how the absence of privacy may negate freedom, undermine possible dissent or attempt to change, how the denial of privacy is an important and necessary element of totalitarianism, and how it undermines intimacy, growth and autonomy.

At the far end, where everything is anonymous, is *The Invisible Man*[3]. It is the story of someone who believes that once he gains anonymity he will be able to act freely, but, in fact, he ends up losing identity and feeling alienated. The character tries to cover his identity as he interacts with the real world, but he cannot interact because without his identity he cannot be represented in the minds of other people.

Real-world dystopias

Such extreme situations are not limited to fiction. During the Holocaust, the Nazis were able to know how many Jews there were, where they were and what their names were because of automatic tabulation of census information. Later on, during the Cold War, the East German regime applied state surveillance in a thorough way that disturbingly echoes *Nineteen Eighty-Four*. The estimate is between 2.5% and 10% of the East German population were informers for the Stasi, the 'state security service'.

Real-world benefits of information sharing

Dystopian scenarios powerfully illustrate the need to have checks in place to prevent the disproportionate collection and misuse of private information. On the other hand, this must be balanced by considerations of the benefits that an acceptable level of transparency and information sharing can bring to society.

National security services in the UK have prevented a number of terrorist attacks. This involves the surveillance of individuals who present a threat, and therefore access to their private information. The legal framework under which national security operates in the UK is one of necessity and proportionality. Before accessing private information, services need to assess what is an unjustified intrusion into privacy.

Sharing information is also instrumental to prevent the spread of diseases. During the 2014 outbreak of Ebola, the disease spread in West African countries that had no precedent epidemics and were poorly prepared. It did not spread into Nigeria, which had a solid enough information infrastructure to prevent that from happening.

1. George Orwell (1949) *Nineteen Eighty-Four*
2. Dave Eggers (2013) *The Circle*
3. H.G. Wells (1897) *The Invisible Man*

Not all data is equal

When determining where an acceptable boundary between private and public should lie, it is also important to note that not all data is equal.

Many in the scientific community would argue, for example, that there should be full transparency regarding scientific data. However, it is a different story when it comes to medical data. And within medical data, genetic information is considered particularly delicate, with professionals talking about "genetic exceptionalism". Indeed, an individual's genome is fundamentally linked with their identity, making it extremely sensitive information.

The Human Genome Project is a good example of an effort to balance increased access and use of personal information with sufficient checks. Genomic data bears the promise of incredible advances for health through diagnosis, prognosis and personalised medicine. Yet, the Head of the Human Genome Project foresaw the broader psychosocial issues that human genomic research would entail and, consequently, he set aside a portion of the Project's budget to support research related to ethical, legal and social implications (ELSI) of the newfound genetic knowledge, to guide the subsequent development of policy options for public consideration. The organisation called ELSI is still very active and important with regard to these issues.

Differential privacy

When releasing information, is it possible to reason mathematically about the extent to which the privacy of individuals will be affected? A mathematical definition of privacy, known as differential privacy, has emerged from work by computer scientists Dwork, McSherry, Nissim and Smith[4]. The differential privacy definition is that, whether you include a specific individual in the input set or not, the two different outputs are very close to each other and it is not possible to tell whether that individual was included or not. One of the strengths of differential privacy is that it is quantifiable and has composition properties, so that it is possible for example to make several queries from a database and argue that the combined queries do not violate this notion of differential privacy.

4. Dwork, McSherry, Nissim and Smith 2006 Calibrating Noise to Sensitivity in Private Data Analysis https://link.springer.com/chapter/10.1007/11787006_1

How has technology changed privacy?

More data and more insight

Digital technologies have been transforming the way we work, live and learn. More and more data is being created, uploaded and shared. Each of us generates large amounts of data through the use of our mobile phones alone.

Technology is widely deployed and is allowing organisations to get much more data, to process it in a cost effective manner, to bring sources together, and to link them in order to generate insights and produce services – such as the national security and health applications described above.

Fueled by big data, analytics have developed at an unprecedented pace; artificial intelligence and machine learning in particular. For example, London-based Google DeepMind has developed a computer game called AlphaGo, which beat the best human Go players. Google uses similar deep learning technology to improve spam filters, speech recognition and photo search, which enables anybody to search for 'sunset on the beach' and find it in their photos without any labelling. Technology can, in real time, enable you to look at a street sign and see it in different languages. Google have announced the development of earphones that will listen to somebody talking to you in one language and play it back in your own language, in real time. These are all enabled by Google's Translate, which is now arguably competitive with translation experts.

Connecting data

In the era of big data, anonymity is virtually impossible to guaranty, or at least it would require considerable effort. In addition, the increasing use of technology has been accompanied by a move from private by default to public by default – for example Facebook users have had to change their default settings in order to make their profiles more private.

Big data and analytics have increased the accuracy with which individuals can be identified and targeted. For example, data-driven technologies have prompted a move towards the mass-personalisation of services. Computer systems can learn about people's preferences and habits to suggest suitable offers. This is seen by some as an intrusion of their privacy and a limitation to their freedom of choice.

The generation of insight is facilitated by the linkage of multiple datasets. The data an individual generates through various apps on their phone can easily be linked, as the mobile phone provides a single identifier.

New technological developments can create new insight from data that was not previously deemed private. For example, face recognition technology and image search make the identification of a person from a photo much easier than before. Technology can potentially infer even more personal characteristics. A team of researchers at Stanford University declared they had developed an algorithm that, from looking at a picture, determines with a very high level of confidence anyone's sexual orientation[5].

Data and the human

As data is central to digital technologies, it risks becoming the focus or even a goal in itself. This is illustrated by a dehumanisation of the discourse. For example, one of the first legal texts about data protection, a 1981 European treaty, was called "Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data". Since then there has been a shift from "individual protection" to "data protection". Similarly the term "Internet of Things" somehow hides the humans that it connects, and some have called for it to be renamed "Internet of People". Interestingly, some health professionals report a similar dehumanisation in medicine, perhaps encouraged by a same drive for efficiency, saying that society has become better at treating diseases but not better at treating patients.

While computers can generate incredible insight, they do not have a sense of what privacy means. As a consequence, they may fail to understand context and make insensitive decisions. Some experts say that current machine intelligence is comparable with human's intuitive thinking, which is prone to all kinds of bias. This corresponds to the "fast thinking" that Daniel Kahneman opposes to "slow thinking", the more reflected form of human intelligence[6]. As humans become increasingly reliant on computer systems, these observations point to the need for individuals and societies to consciously and proactively consider questions of privacy.

5. Wang, Y., & Kosinski, M. (in press) Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology.* https://osf.io/zn79k/
6. Thinking, Fast and Slow (2011) Daniel Kahneman

Technologies to protect privacy
How do individuals and societies manage privacy?

How do individuals manage their privacy?

Most individuals consider privacy, and the information they share, as a matter of personal preference. They decide who they want to share information with and how. Generally, they might share more with relatives and close friends, less with acquaintances, and even less with strangers. This can be described as spheres of privacy. People might even have different personas depending on who they are communicating with. For example, young people across various cultures tend to readily share information with Facebook friends, but they are worried about what their parents or teachers might find out.

Individuals do make risk-based decisions about how they share data. However, they do so on the basis of perceived risk rather than objective risk. The difference between the two can be substantial. In order for people to improve their understanding of risk, they would need more information about what data organisations collect and hold about them, how they share it, and how it may be used.

A number of people might not realise the extent of the security risks that come with certain technologies. For instance, connected objects such as baby monitors or toys can be hacked. Out of convenience, people might also underappreciate the security risks associated with certain Wi-Fi networks.

Similarly, as technology has brought a lot of convenience, most people have willingly shared much of their information. However, many also feel they do not have a choice, and what people want may evolve with time.

How do companies manage privacy?

The users' trust is an asset for companies, and therefore an incentive for them to care about privacy. Indeed, technology companies have implemented the use of privacy-preserving technologies. For example, differential privacy has been adapted by Google in some of the projects relating to the web browser Chrome, and Apple encouraged all their vendors to use differential privacy.

Ultimately, privacy is managed by people within the company, directly or through technology. Whether a company's software developer or engineer will care about privacy depends on the organisation's culture.

Business models might be more or less compatible with privacy. For example, some are concerned that Google and Facebook, who heavily rely on revenue from advertisement, have more incentive towards collecting information about their users than towards the protection of their privacy. In fact, in 2010, Facebook's founder Mark Zuckerberg implied that privacy was no longer the "social norm"[7].

Currently there is more than one incentive for data accumulation. Companies have tended to do "data hoarding", i.e. collecting and storing data with no clear purpose but with the hope of possible future uses. Furthermore, the data economy is driven by business valuations, which are dependent upon the option value of data that they are currently collecting. Despite the fact that the data may not be immediately useful in some way, in another way their business value is critically dependent upon it.

How do states and public institutions manage privacy?

States and public institutions also have incentives towards the collection and use of data. Indeed, big data and analytics offer new possibilities for more efficient and smarter administration. In Israel, the Ministry of the Interior decided to create a national bank of fingerprints and photos of all the citizens in the country, for the purpose of national security. Many scholars and other citizens saw a privacy risk and were opposed to it. In the end, the biometric database has been set up. Citizens must give facial biometrics and have a choice of whether to provide fingerprints or not. If they provide fingerprints, their identification documents (national ID and passport) will last for 10 years; if they only provide facial biometrics, their documents are valid for only 5 years.

7. https://www.theguardian.com/technology/2010/jan/11/facebook-privacy

In parallel, Israel's Ministry of Health contributed funds towards the creation of a national gene bank, motivated by the prospect of personalised medicine and effective public health spending. While the benefits for research are clear, there are concerns that such a large repository of personal information could be misused, by the current or subsequent administrations. Therefore, states have an important responsibility regarding the privacy of their citizens.

States do have the power to introduce checks and balances about the protection of individuals and the use of data. For example, the Israel Judiciary Administration can decide to post court cases on its own website, but they can also add a feature that prevents Google and other search engines from indexing the decisions, so it will not come up when people conduct regular searches. Some companies could challenge that decision, but this is an example of a public administration introducing barriers to information sharing in order to protect the privacy of its citizens.

States and international institutions can also introduce legislation. For example: the European Union's 'right to be forgotten' gives EU citizens the right to require search engines like Google to de-link specific pages that show up as a result of searching their name. This limits the access to information about them. In addition, the EU's General Data Protection Regulation has introduced standards and rules that limit 'data hoarding' – the accumulation of data without specific purpose, which may give organisations detailed insights for example about consumers' profiles.

However, legislation increasingly struggles to keep up with the pace of technological development, thus prompting the need for more adapted governance solutions.

Issues with the current data ecosystem

The vast amounts of data being collected by states and companies has led some to coin terms such as "dataveillance". There are concerns about power asymmetries, with some talking about "data feudalism"[8]. One possible risk is that pervasive dataveillance may restrict our freedom and our capacity for identity formation. It could also be exploited for pervasive manipulation and threaten the social foundations of democracy. The risk of manipulation both of the individual and at a population level is real, given the power of automated behavioural profiling of individuals in real time, which can continue to update and therefore nudge an individual in the direction desired by the choice architect.

Another issue is that, as data brings new efficiencies, privacy comes at a cost. For example, public administrations can save money by automating certain services, but they will still need to offer manual processing for people who will not engage with such systems. The cost will either be to the organisation, or to individuals. Thus privacy risks becoming costly and for the privileged few.

Overall, it is unclear how much the current data ecosystem fulfils the interests of individuals and society. This has led some to propose alternative models.

Alternative models

A data trust could be created as an entity that acts in the interests of individuals in the ecosystem. There is a need to design mechanisms by which different groups could propose how they want data to be used, and people could vote so that there would be more data in the trusts that people are comfortable with. Such an organisation would then be in a position to balance large corporate and national entities.

Following another idea, The Hub of All Things in the UK is trying to create a service that manages your personal data and then selectively shares it with different services. It is a starting point showing a mechanism through which people can gather their desires and information, and manage their spheres of privacy. It takes the administrative burden away from the individual and allows use of these services, but also respects the spheres within that person's perspective.

To be most effective, privacy policies should be machine-readable. Ideally they would be encoded the same way that Creative Commons codified copyright licences. A program could help an individual to work through different privacy statements as they are interacting with services, so that the system would alert users when something requires their attention.

Another way of dealing with the power imbalance brought about by large companies accumulating data would be if, instead of entrusting their data to Facebook or others, users would keep it themselves on the cloud. Companies could issue smart requests to the users, in combination with a micro payment. The user would provide them with fine grain data and a smart contract that wraps that data. The technologies seem mature enough to make a substantial change of system architecture possible.

8. https://www.theguardian.com/media-network/2016/jun/03/data-trusts-privacy-fears-feudalism-democracy

What role for technology in managing privacy?

Technology and the Law

Traditionally, the law has been the instrument of choice for societies to formalise and impose norms. Some argue technology is another tool that could be potentially effective to achieve the same regulation, especially as it might be in a better position to keep up with the fast-developing digital systems it regulates.

However, there are concerns that the overreliance on technological means for governance could lead to a loss of normativity. Technology would, in fact, act best as an addition rather than a substitute to the law. In particular, law is considered less effective when it comes to ensuring enforcement, and this is a space where technological approaches could be especially helpful.

Early attempts at privacy management tools

Mix networks and Tor

In 1981, David Chaum, an American computer scientist and cryptographer, proposed the idea of an anonymous communication network. His proposal, called mix networks, allows a group of senders to submit an encryption of a message and its recipient to a server. Once the server has a batch of messages, it will reorder and obfuscate the messages so that only this server knows which message came from which sender. The batch is then forwarded to another server who does the same process. Eventually, the messages reach the final server where they are fully decrypted and delivered to the recipient. Mix networks are the conceptual ancestor to modern anonymous web browsing tools like Tor. Chaum has advocated that every router be made, effectively, a Tor node. In practice, nowadays only few internet users employ such tools.

Lucent Personalised Web Assistant

In the late 1990s, Bell Labs, part of Lucent, developed Lucent Personalised Web Assistant. Its question was not really how to remain anonymous on the internet, but how to control interactions and privacy with respect to whomever you talk. You may have different senses of privacy when you go to your doctor compared with when you go to your banker. Obviously you share a lot of information with each, but you might not want that information to cross. The basic idea with this web assistant was to have a proxy that manages your personas, and in practice use a universal password and ID. For every destination you interacted with, it created different credentials and you had full control over them. Lucent Personalised Web Assistant was a research project rather than a widely-available product.

Passport

In the early 2000s, two of the leading companies at the time, Microsoft and Netscape, sponsored an application called Passport. It was spearheaded by a startup that thought every user should define their privacy preference and this should somehow be taken into consideration. The Passport identification scheme kept track of Web users and could speed online shopping and other Internet transactions.

P3P

The Platform for Privacy Preferences Project (P3P) is a protocol allowing websites to declare their intended use of the information they collect about web browser users. Designed to give users more control of their personal information when browsing, whilst sparing them reading the privacy policies at every site they visit, P3P was developed by the World Wide Web Consortium (W3C) and officially recommended in 2002. Development ceased shortly thereafter and there have been very few implementations of P3P. Microsoft Internet Explorer and Edge were the only major browsers to support P3P. Microsoft has ended support from Windows 10 onwards. Some stated that P3P has not been implemented widely due to impracticality and lack of value.

Privacy Enhancing Technologies

Around the same time, the term "Privacy Enhancing Technologies" (PETs) started to be used. PETs are often developed by third parties, and offered to end users, for them to use as an additional part in their system. PETs can be as simple as a plaster stuck onto the camera in your computer – that is a privacy enhancing technology – or far more sophisticated.

Secure multi-party computation

Secure multi-party computation is about allowing mutually distrusting parties to cooperatively compute over their private data. It is, for example, the ability of parties P1, P2 and P3, each in possession of private data A, B and C, to be able to compute a function over A, B and C, such that they can get an output, but without having to reveal to each other the details of their private data. The classic example of this might be the millionaire's problem, in which you have a number of very wealthy people, all of whom wish to determine which is the wealthiest, but do not wish to reveal how much money they actually have to the other millionaires in the room. You might have a trusted entity out there, but there are potential techniques that allow you to engineer this without assuming that there are any trusted parties in that setting.

A use case is data sharing across a set of different parties, each of which has an interest in security and in protecting their data, for example national security, health and commerce. Using secure multi-party computation, none of them reveals their data to another, but they are able to compute over it.

Statistical disclosure control: *k*-anonymity, *l*-diversity and differential privacy

Statistical disclosure control is the ability to anonymise data. It can be achieved in different ways. *k*-anonymity is essentially the ability to control the size of a group within which a subject is anonymous. *l*-diversity is the ability to do that in the context of homogenous groups.

Differential privacy is essentially the ability to add noise to data in a way that still permits valid statistical inferences to be made. Differential privacy, as a definition, allows one to reason about how to maximise the level of information you can get from a dataset while minimising the risk to the privacy of the individuals included in the set.

These present a series of opportunities to be able to compute statistically interesting properties without actually revealing data that might otherwise be regarded as intrusive.

Homomorphic encryption

One of the most interesting technologies that is emerging, thanks to IBM in significant part for making some of the key intellectual contributions here, is homomorphic encryption. Homomorphic encryption allows computations to be carried out on a cipher text, which is encrypted, generating an encrypted result that, when decrypted, matches the results of the operations as if they were performed on the plain text. An example is a homomorphic concatenation. In other words, one party encrypts the word 'hello'. The other party encrypts the word 'world'. The homomorphic concatenation happens on the two encrypted results, yielding an encrypted output which, when decrypted, gives the concatenated 'hello world'. The idea of actually being able to do analytics while the data is encrypted radically changes the privacy question. There is a lot of interesting data mining issues that this can be applied to. Multi-party computing applications might be enabled by these technologies.

Functional encryption

Related to this is the idea of functional encryption, in which users can grant authorisation to perform only certain functions on the encrypted data. An example application could be for a national security organisation to hand over to a social work organisation some data that might contain sensitive national security information. Functional encryption would protect information sensitive from a national security perspective and allow an authorised organisation to find information that allow social work interventions at an early stage in the genesis of radicalisation.

Searchable encryption

Searchable encryption is the ability to outsource the storage of data to another party in a private manner, while maintaining the ability to search over it. You can encrypt on a key provider to you, then submit a token to that service and achieve a search result, without disclosing to the storing party the data that you need. This offers possibilities for cloud storage and re-engineering search services in innovative ways. Searchable encryption could be very disruptive for search engine providers and their business models.

Private information retrieval

Private information retrieval is the ability to retrieve an item from a database server without revealing which item is retrieved. Let us say that somebody wants some information about an organised criminal from their bank accounts, but does not want to reveal to the person to whom the search is submitted the nature of that search query. They do not want to say who they suspect and it is either not practical or the other party is unwilling to give all the data to be searched. How can you do this in a way that means the item being searched for is not disclosed to the party holding the data? There are some interesting emerging techniques in this area.

Zero knowledge proof

Some of the points raised in the initial conversation addressed questions of transparency. Interesting techniques here are ideas of zero knowledge proof and the ability of one party to prove to another party that a given statement is true, without conveying any information apart from the fact that the statement is indeed true. The classic intuitive example is that of the colour-blind friend. You want to prove to them that you can distinguish between red and green, without telling them which balls are actually red and green. You get the friend to take the balls, hold them behind their back, maybe swap them around and then hold them up. You tell them whether or not they have swapped the balls. You can reliably tell them that they have swapped the balls, without telling them how you do it and without saying which is red and which is green.

Smart contracts

Smart contracts are effectively computational protocols that enforce the performance of particular contracts. They are self-executing contracts as programs. They are increasingly stored in blockchains, which are distributed tamper-proof registers. Here we have a situation in which an oversight party could prescribe what actions can be performed by way of analytics on data. You can directly execute that contract and store the outcomes in a tamper-proof way. Together with an overseer, you can have a tamper-proof log of the actions that you have taken and some demonstrable proof that you have not, in fact, performed actions on the data that you have not been authorised to perform.

Privacy by design

While PETs may be seen as an external patch to an existing application, 'privacy by design' elements are embedded within it. Privacy by design is developed by the engineer who devises the application itself; hopefully they know what is going on inside the technology better than a third party developing a ready-made, ready-to-use patch. Privacy by design was initially promoted by Dr Ann Cavoukian, the former Ontario Information and Privacy Commissioner, and others in a 1995 report. The General Data Protection Regulation (GDPR) references 'privacy by design and by default', requiring organisations to put in place appropriate technical and organisational measures to implement the data protection principles and safeguard individual rights.

PETs empower the user, if they would like to use them, but require some knowledge. With privacy by design, if the engineering took into consideration privacy to begin with – perhaps using or adapting PETs – the user's privacy is protected even if they do not know what is going on in the complex system they are using.

A good example of privacy by design is the iris-based identification at Schiphol Airport. The iris information is on the card you carry and when you put your card into the machine, it scans your eye and compares it with your card. There is no central database of your iris in that system, but it is still providing the security access.

The body scanners at airport security are another illustrative example of privacy by design. The first generation of scanners used in US airports made some travellers feel uneasy, because, whilst they would stay in their clothes, an operator somewhere in the airport saw them naked in great detail. As there was growing criticism about this, the US Congress ordered the provider to fix the issue. The company, however, pointed out that they could not fix it without having to redesign everything. This technology is not in US airports anymore. A new generation of body scanners was developed by another company. Scanners now produce a generic silhouette – if the traveller carries some object on their leg or somewhere else, a dot flashes at the corresponding location on the figure.

Privacy as a multidisciplinary effort

From a professional point of view, privacy by design is something that we should be focusing on and promoting in the training of software engineers and other professionals.

In fact, this could mirror progress towards security by design. There is a move in the software development community towards embedding security in design. In the software development at the moment there are far faster cycles of development and delivery, encouraged by the move to cloud infrastructures, by continuous integration, continuous delivery, and by agile software development methods. They use the term DevSecOps meaning development, security, operations all happening in one. This implies considering security upfront in the software development lifecycle, which a number of cybersecurity experts have been recommending for years.

Privacy is a complex, multidimensional notion. Solutions will require an in-depth multidisciplinary effort, involving a range of tech-savvy and privacy-savvy people.

Annex

Acknowledgements

The workshop, held under the Chatham House rule, brought together experts from Israel and the UK, covering a range of disciplines including computer science, cryptography, medicine, social psychology, behavioural sciences, and the law. The Israel Academy of Sciences and Humanities and the Royal Society are grateful to all the workshop participants.

The Royal Society and The Israel Academy of Sciences and Humanities Staff	
Galia Finzi	Executive Director, IASH
Franck Fourniol	Policy Adviser, The Royal Society
Bob Lapidot	Director of the International Division, IASH
Jess Montgomery	Senior Policy Adviser, The Royal Society

The Royal Society is a self-governing Fellowship of many of the world's most distinguished scientists drawn from all areas of science, engineering, and medicine. The Society's fundamental purpose, as it has been since its foundation in 1660, is to recognise, promote, and support excellence in science and to encourage the development and use of science for the benefit of humanity.

The Society's strategic priorities emphasise its commitment to the highest quality science, to curiosity-driven research, and to the development and use of science for the benefit of society. These priorities are:

- Promoting excellence in science
- Supporting international collaboration
- Demonstrating the importance of science to everyone

For further information
The Royal Society
6 – 9 Carlton House Terrace
London SW1Y 5AG

T +44 20 7451 2500
W royalsociety.org

Registered Charity No 207043
March 2019 DES6113