ePRIVACY AND DATA PROTECTION
The Complex Web of Data Privacy and its Influence on AI Ethics, Competition and Tech Evolution

Citi GPS: Global Perspectives & Solutions
July 2020

Citi is one of the world’s largest financial institutions, operating in all major established and emerging markets. Across these world markets, our employees conduct an ongoing multi-disciplinary conversation – accessing information, analyzing data, developing insights, and formulating advice. As our premier thought leadership product, Citi GPS is designed to help our readers navigate the global economy’s most demanding challenges and to anticipate future themes and trends in a fast-changing and interconnected world. Citi GPS accesses the best elements of our global conversation and harvests the thought leadership of a wide range of senior professionals across our firm. This is not a research report and does not constitute advice on investments or a solicitation to buy or sell any financial instruments. For more information on Citi GPS, please visit our website at www.citi.com/citigps.
Catherine T O'Neill  
European & CEEMEA Media Analyst, Citi  
+44-20-7986-8053 | catherine.oneill@citi.com

Richard Webley  
Head of Citi Global Data Insights  
+1-212-723-4526 | richard.webley@citi.com

Expert Commentators

Vivienne Arzt  
Chief Privacy Officer, Refinitiv

Martin Ashplant  
Digital Media Consultant & Former Chief Digital Officer, Beano Studios

John Bowman  
Senior Principal, Promontory Financial Group, an IBM company

Anne Fealey  
Global Chief Privacy Officer, Citi

Emma Maconick  
Partner, Intellectual Property Transaction Group, Shearman & Sterling

Simon McDougall  
Executive Director - Technology and Innovation, Information Commissioner’s Office (ICO).

Prash Naidu  
Founder and CEO of Rezonence

Bill Philhower  
Global Head, Privacy and IT Regulatory/Audit Program Management, Citi

Todd Ruback  
Director, Promontory Financial Group, an IBM company

Dr Johnny Ryan  
Chief Policy & Industry Relations Officer, Brave

Anthony V Turiano  
Managing Director, OTRC/Privacy & Compliance Services, Citi

Justin B Weiss  
Global Head of Data Privacy, Naspers

Danny Weitzner  
Founding Director of MIT Internet Policy Research Initiative and Principal Research Scientist at the MIT Computer Science and Artificial Intelligence Lab

Andres Wolberg-Stok  
Head of Strategy, Office of the CTO, Citi
ePRIVACY & DATA PROTECTION
The Complex Web of Data Privacy & Its Influence on AI Ethics, Competition & Tech Evolution

At a conference I attended a few years ago one of the presenters posed the question “Do you own your digital identity?” Most people in the room thought if they didn’t already, they definitely should, but at the very least they should have control over how their digital identity is used. Once the conference ended most of us probably logged onto a phone or laptop to read the news, shop online, or watch a dancing kitty somewhere on social media. A thought about how much personal data we were freely supplying the Internet while we browsed probably didn’t happen.

As we’ve moved faster and faster into the digital age, the amount of data generated and collected has skyrocketed. What’s different today is a greater recognition by corporates, governments, and individuals about the value of that data. With this recognition comes increased calls for global regulation plus increased scrutiny over Big Tech and the amount of data they now control. And in the coming regional race for dominance in artificial intelligence (AI), the ethical use of those large pools of data is essential to avoid discrimination, bias, and widening societal divides.

In our first two Citi GPS report on ePrivacy we went into the details of the EU’s General Data Protection Regulation (GDPR) and talked about what companies needed to do to avoid substantial fines for non-compliance. The GDPR was intended to increase transparency, tilt the balance of power back towards consumers, and place privacy by design and default at the heart of how businesses operate.

Two years later, it has elevated the prominence of data privacy to the boardroom level but the full fining power of the regulation has yet to be tested as regulators lack sufficient funding and resource and there is evidence it may have unintentionally reinforced the strength of the Big Tech platforms. AdTech is being closely scrutinized by the regulators, prompting Big Tech to drive changes which could upend how this industry operates today with far reaching consequences for brands.

New GDPR-style regulation is cropping up in multiple markets. Privacy technology solutions are emerging to help manage regulatory requirements. Consensus is building for global cooperation around data protection and privacy and an attempt to develop global AI Ethics could trigger more innovation. True harmonization of global policy looks to be a long way off due to cultural, political, and societal differences meaning corporates will need to continue navigating the complex web of data-related regulation.

The race between countries and regions to become an AI superpower exacerbates the regional divides further, especially when data sits in the hands of a few. Europe is losing the tech battle to the U.S. and China and instead will look to regulate its way to a level playing field. Increasing recognition of the value of data has also led to a jurisdictional crossover between data protection regulation and competition authorities. As Big Tech controls more of the data, the lens through which antitrust concerns are judged has been shifting and platform-specific regulation could arise.

In the future when someone asks if you own your digital identity, innovation should drive the answer to ‘absolutely yes’. Only then will your secret be safe about watching dancing kitties.
The data privacy debate is expanding

Data is valuable and access to it is essential but consumer data currently sits in the hands of a small number of cash rich technology companies. Varying (and divergent) data privacy considerations and fostering consumer trust have a huge role to play in how the technology landscape evolves. The GDPR increased consumer privacy rights but also reinforced the strength of Big Tech platforms. Lack of regulator funding and resource has been a challenge for enforcement.

**EARLY DAYS OF REGULATORY EVOLUTION**

We are still early in data privacy regulatory evolution. The GDPR has prompted a swathe of similar regulation in other markets but divergent cultural, political, and societal backdrops mean global harmonization of data privacy rules will prove challenging. Emergence of privacy technology may provide some solutions.
CONTINUUM OF DISRUPTION FROM POSSIBLE NEW REGULATORY HURDLES

The privacy debate is bleeding into that of competition and ‘AI Ethics’. Europe is losing the battle for tech to the U.S. and China but is leading the regulatory charge. The Big Tech gatekeeper status will be hard to erode but with a tougher and more holistic regulatory landscape, if enforced properly, could lead to some rebalancing. There is a danger of AI evolving without enough consideration for ethical implications. Formal AI Ethics guidelines and principles could be a positive impetus for the development of AI, although it’s in its infancy.

Effort by Industry to Mitigate AI Risks Varies by Type of Risk (% of respondents who have adopted AI mitigating risk)

<table>
<thead>
<tr>
<th>Risk Category</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cybersecurity</td>
<td>48%</td>
</tr>
<tr>
<td>Regulatory Compliance</td>
<td>35%</td>
</tr>
<tr>
<td>Personal/Individual Privacy</td>
<td>30%</td>
</tr>
<tr>
<td>Explainability</td>
<td>19%</td>
</tr>
<tr>
<td>Organizational Reputation</td>
<td>19%</td>
</tr>
<tr>
<td>Workforce/Labor Displacement</td>
<td>17%</td>
</tr>
<tr>
<td>Equity and Fairness</td>
<td>13%</td>
</tr>
<tr>
<td>Physical Safety</td>
<td>11%</td>
</tr>
<tr>
<td>National Security</td>
<td>4%</td>
</tr>
<tr>
<td>Political Stability</td>
<td>2%</td>
</tr>
</tbody>
</table>


WHAT ARE THE EXPERTS SAYING?

“California’s new privacy law [CCPA] is definitely a game-changer...”
Anne Fealey

“The logical consequence should be that if all those countries have privacy legislation that there could be free movement of data between them. But it’s absolutely the opposite. It’s now creating barriers...”
Vivienne Artz

“I think there’s going to be some tension between, just in the European context, the approach that has been taken with the GDPR...on the one hand, but on the other hand, what gets called AI policy...”
Danny Weitzner

“We know from competition enforcement that fines are immaterial if continued wrongdoing offsets the cost of the fine. I think the regulators who matter understand that now...”
Johnny Ryan

“You don’t want the water to be over running the banks and hurting people or structures...”
Justin Weiss
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Summary</td>
<td>7</td>
</tr>
<tr>
<td><strong>Section 1</strong></td>
<td></td>
</tr>
<tr>
<td>What Impact has GDPR Had?</td>
<td>14</td>
</tr>
<tr>
<td><strong>Section 2:</strong></td>
<td></td>
</tr>
<tr>
<td>How Will Global Privacy Regulation Evolve?</td>
<td>27</td>
</tr>
<tr>
<td><strong>Section 3:</strong></td>
<td></td>
</tr>
<tr>
<td>Why has the Regulator Focused on AdTech?</td>
<td>54</td>
</tr>
<tr>
<td><strong>Section 4:</strong></td>
<td></td>
</tr>
<tr>
<td>The Intersection between Data Privacy and Competition</td>
<td>69</td>
</tr>
<tr>
<td><strong>Section 5:</strong></td>
<td></td>
</tr>
<tr>
<td>Innovation in Data Privacy and the Importance to the AI Revolution</td>
<td>94</td>
</tr>
<tr>
<td><strong>Expert Views</strong></td>
<td></td>
</tr>
<tr>
<td>A Conversation with Vivienne Artz</td>
<td>111</td>
</tr>
<tr>
<td>A Conversation with Martin Ashplant</td>
<td>120</td>
</tr>
<tr>
<td>A Conversation with John Bowman &amp; Todd Ruback</td>
<td>130</td>
</tr>
<tr>
<td>A Conversation with Anne Fealey</td>
<td>141</td>
</tr>
<tr>
<td>A Conversation with Emma Maconick</td>
<td>144</td>
</tr>
<tr>
<td>A Conversation with Simon McDougall</td>
<td>156</td>
</tr>
<tr>
<td>A Conversation with Prash Naidu</td>
<td>165</td>
</tr>
<tr>
<td>A Conversation with Bill Philhower &amp; Anthony Turiano</td>
<td>172</td>
</tr>
<tr>
<td>A Conversation with Dr Johnny Ryan</td>
<td>176</td>
</tr>
<tr>
<td>A Conversation with Justin B Weiss</td>
<td>188</td>
</tr>
<tr>
<td>A Conversation with Danny Weitzner</td>
<td>200</td>
</tr>
<tr>
<td>A Conversation with Andres Wolberg-Stok</td>
<td>207</td>
</tr>
</tbody>
</table>
Executive Summary

How would the world look if there was complete trust of the online ecosystem? It’s hard to hypothesize but it would probably be even more advanced than it is now as consumers would not feel conflicted about transacting on a new site or using voice activated systems to access their banking or healthcare information. Personal data protection and privacy is about protecting consumer data but it is also about fostering trust so we can have a more open and vibrant digital world.

Privacy concerns are one aspect of this but we can’t ignore the debates around fake news, disinformation, online fraud, and cyber-attacks as other factors that make gaining consumer trust challenging.

Figure 1. Percentage of Consumers that Agree…

Source: Edelman’s Trust Barometer (2020)

Personal data protection and privacy is seen as a rather dry topic. It’s not nearly as interesting as thinking about how the Internet-of-Things could change our everyday lives, for example, or what smart cities might look like. What underpins these more exciting areas and conversations is the ability to access, process, manipulate, and benefit from the reams of data being generated every day, both personal and non-personal. Data should be treated as a precious commodity. Used in the right way, it has the potential to drive significant change for good, but there are risks; risks that data is used in ways which widens societal divides, creates a ‘digital underclass’, and increases opacity. In recent months we have seen the COVID-19 pandemic drive an acceleration in digital consumption patterns and this has escalated the need for an increased focus on personal data protection and privacy.

Danny Weitzner

“think first of all, privacy questions are hard to separate clearly from competition issues, from AI policy issues, security issues, etc.”
This is the third in our Citi GPS report series on ePrivacy and Data Protection (see Privacy Matters: Navigating the New World of Data Protection and Who Watches the Watchers – How Regulation Could Alter the Path of Innovation). We dug into the development of data protection legislation in Europe in detail in our last two Citi GPS reports and concluded that the General Data Protection Regulation (GDPR) in Europe may well be the start of a step change in approach to data protection regulation globally. At the same time, we warned of it unintentionally pushing more power into the hands of the large tech platforms, companies which may be reluctant to embrace data privacy rules if they harm the economic viability of their business. Two years on from the May 2018 implementation date of GDPR, the market seems comfortable that not much has changed; if anything, it has decided that Big Tech has managed the changes as the share prices of these companies have continued to rise.

Figure 2. Average Share Price Performance of U.S. Big Tech vs. MSCI World (since May 2018)

Source: Datastream.

We argue that those who believe data protection and privacy is a non-issue, will do so at their own peril, which could also raise governance questions. It cannot be viewed in isolation; it permeates all sectors. It’s not. There is increased recognition of the value of data and that access to it is essential to realize the Artificial Intelligence (AI) vision many countries and regions have set. The data, however, currently sits in the hands of a small number of cash-rich technology companies who have built significant digital platforms.

In this Citi GPS report, we look more closely at the impact data privacy is having on broader regulatory and technology developments. We are seeing the debate bleed into that of ‘AI Ethics’ and the domain of the antitrust regulators and legislators, and we should expect the intersection between data protection and competition to become even more visible. Historical antitrust decisions have been instrumental in tech industry concentration. The Internet business model (inherently advertising-funded revenue) incentivizes data collection and profits over privacy which prompted the introduction of the GDPR. But the unintended consequence of the legislation is that the dominant platforms have become more dominant. This is prompting more scrutiny from a competition perspective.

How regulation evolves is influenced by many factors (e.g., politics, behavior, society) and the direction of regulation will dictate the ability to access data and the pace at which (and how) technology develops.
As nations turn their focus to the battle for AI dominance, varying (and divergent) data privacy considerations have a huge role to play in how the battle evolves. This may create tension with the bubbling nationalistic and politically-driven technology power struggle. Regulation usually has negative connotations in the investment world but in the case of AI Ethics, formal guidelines and principles could be a positive impetus for the development of AI.

We delve into these topics via in-depth interviews with a number of experts in the data protection and technology space, including senior executives at Citi, Naspers, and Refinitiv, the Executive Director of Technology and Innovation at the U.K. Regulator (ICO), consultants, specialists within the AdTech industry, an IP lawyer, and the Chief Policy Officer for Brave (a well-known privacy advocate). Throughout the report we include key commentary from the experts (in pull quotes, with the full interviews in the back of the report).

In this Citi GPS report, two years on from GDPR coming into effect, we look to answer:

- What impact has the GDPR had?
- How is global privacy policy emerging and how is this influencing the AI Ethics debate?
- Why has the AdTech business model been at the heart of the regulatory focus?
- What is the interplay between privacy and competition concerns and could it prompt any changes to the tech landscape?
- What is the innovation in data privacy and how important is it to the coming Era of AI?

**Roadmap to this Report**

We break this report into five main sections:

- In the first section we **look at the impact the GDPR has had since it was introduced in May 2018**, in particular how corporates have responded and how regulators have enforced it. The main conclusion is that enforcement has been slow to build and the full fining power has not yet been tested as regulators face informational, resource, and funding disadvantages versus the Big Tech platforms. However, while the GDPR has helped elevate the importance of data protection to the boardroom level, there is evidence it has also reinforced the dominance of the large tech platforms.
In the second section we investigate how data protection regulation has developed globally, in particular in the U.S. with the California Consumer Privacy Act (CCPA). We dig into how cultural differences influence the approach to data protection regulation and the increasing tension arising from the drive for AI superiority while also addressing AI Ethics, including data privacy. The GDPR has prompted the introduction of new privacy regulation across many markets but harmonization of global privacy policy feels some way off as the fundamental starting points (both political and cultural) differ and the race for tech supremacy creates barriers. Emerging AI platforms raise new ethical issues; there are clear tensions between what’s needed for AI to succeed and what the privacy regulation is trying to achieve in terms of increased transparency and data minimization. If governments aren’t proactive in regulating AI, we risk an unfortunate misstep leading to overregulation. We believe we could see divergent paths being taken as Europe is likely to look to regulate its way to increasing its tech presence. In the U.S., inevitably the Big Tech companies will play a role in the battle for AI superiority, which could limit the degree to which the U.S. will look to clip the wings of these companies. These challenges are less apparent for China.

In section three we provide some insight into how the U.K. data protection regulator is tackling the AdTech industry, a central player in why the GDPR was established. It is also a live case study in the pitfalls and benefits of attempting to unpick a well-established business model in an industry where power is concentrated. For the regulator it is a little like walking a tightrope given the number of stakeholders with varying motives. We conclude that Big Tech has significant influence due to their gatekeeper positions, which means they will have the ultimate say on how the AdTech industry will evolve. This could increase the data advantage that already exists with the main platforms. An upheaval of the AdTech landscape (e.g., an end to third-party cookie usage and in-app access to data) could also make it more challenging for advertisers and AdTech operators to reach audiences in the way they do now, increasing the importance of first-party data access and direct-to-consumer relationships.

In section four we dig into the interplay between data protection and competition authorities. Data is increasingly the lifeblood of the modern world and the interest of the two regulators is now converging around data. We consider the key tenets of the current competition law, implications of some pivotal decisions that have been taken by the competition authorities in the past, and what the most likely future regulatory actions will be. We conclude the future of Big Tech-driven M&A will be challenged, structural remedies could be announced, and some regions will move to create tech-specific regulation/codes of conduct and agencies to oversee it, in order to level the tech playing field and open up access to data (see the European Commission’s Digital Services Act and the U.K. CMA’s recommendations to the U.K. government). As we learned from the GDPR, any new regulations and agencies that are set up need to be well-resourced and funded for enforcement to stand a chance. The economic value of data has mostly accrued to the Big Tech platforms. If enforced properly, an accumulation of regulatory changes could lead to a more balanced picture.

In section five we explore the types of innovation occurring in the personal data arena. Some will support the Big Tech companies themselves whereas some will support the governments and regulatory bodies seeking to protect the consumer, but all will be important contributors to the data ecosystem that is beginning to emerge. In particular, we investigate: Data Privacy Technology, Data Control & Portability, Safe & Sound AI, and Open Data.
With new encryption techniques, replication and scaling techniques, and decentralized technologies providing for tokenization of identity, the requirements are achievable. They will, however, require cooperation and agreement between private and public sectors, both of which are facing many pressures, and are not all equal participants at the table as the digital platforms are now so dominant and well-funded they outweigh the others. Big Tech continues to be a step ahead of governments, having developed AI principles and data sharing frameworks. The main weapon of governments to address the balance of power is regulation.

Headline Conclusions

- The GDPR elevated the prominence of data privacy in boardrooms. This will increase as COVID-19 has prompted an acceleration in brands adopting a direct-to-consumer model in which first-party data will be critical. The GDPR has also prompted a swathe of GDPR-style regulation in other markets, which is still just the beginning of the data privacy regulatory evolution (U.S. federal law is likely). However, GDPR enforcement appears to have been slow due to the complexity and lack of funding/resource (relative to the tech platforms) and it appears to have unintentionally reinforced the strength of the Big Tech platforms. This has exacerbated other challenges around competition, media plurality, and (over the next couple of years) will lead to big changes in the AdTech landscape etc. The shifting technical landscape makes it hard for regulators to keep pace with innovation, leaving the industry to drive the narrative.

- Everyone now realizes the value of data and access to it is essential to realizing the AI vision that many regions have set out; the race is on between countries and regions to become an AI superpower. AI requires large pools of data and (outside of China) it largely sits in the hands of a small number of cash-rich platforms, which act as gatekeepers.

- The current patchwork of regional data privacy regulations make it an increasingly challenging environment for global companies to operate in. Despite several markets adopting a GDPR-style regulatory approach, divergent cultural, political, and societal backdrops mean global harmonization of data privacy rules will prove tricky as concepts such as data localization and state surveillance (in some markets) create geographical boundaries, encouraged by an increasing nationalistic stance.

- The tech world is dominated by the U.S. and China. One is heavily influenced by the sovereign state and the other by private companies. The uneven regional exposure to tech platforms means Europe is losing the battle for tech but is leading the regulatory charge in its attempt to avoid slipping further behind. The European Commission (EC) has to balance its unwavering position on protecting privacy while attempting to limit the power of the Big Tech names and create a path for the creation of national champions. The U.S. is leading on tech but behind on regulation. It has the biggest dilemma as it recognizes the concentration of power with the Big Tech platforms and that by allowing them to continue unhindered, it will likely raise bigger competition and societal-related questions further down the line. But these platforms are also the U.S.’ path to power in AI. Now is not the time to dismantle the national champions. China doesn’t face these same challenges as it is transparent the importance of the state overrides the private sector agenda and an individual’s rights to privacy.
In the race for AI there is the **danger of AI evolving without enough consideration for the ethical implications**, which could undermine the ability to protect consumer privacy and society at large from bias, inappropriate profiling, and discriminatory uses of data.

One of the questions we raise is *will there be more regulatory action taken?* We believe this could well be the case as governments and regulators take a more holistic approach. While fines may capture the headlines, it will be **structural and behavioral remedies that have the most impact**. M&A could become more challenging for Big Tech in developed markets and we see increased impetus to explore structural separation remedies. The regulatory stance in some markets may get tougher still, as regulation moves beyond competition law and towards platform specific ex ante regulation (see The Digital Services Act).¹

If regulation increases, will the competitive landscape (and balance of power) shift? In other industries, years of regulation have led to more competition and lower margins. The economic value of data has mostly accrued to the Big Tech platforms over the past 20 years or so, although consumers have arguably also benefited from the rise of these digital ecosystems. The position of Big Tech platforms is not likely to change quickly as their gatekeeper status is well-established with consumers, but an accumulation of regulatory changes, if enforced properly, could lead to some rebalancing. As we learned from the GDPR, any new regulations and agencies set up need to be well-resourced and funded for enforcement to stand a chance.

Are there a range of innovations that will address the privacy concerns with the digital platform technology ecosystem? We are seeing an emergence of privacy-focused tech which could provide consumers with alternative means to take more control over their data. Open data access and data portability are seen as solutions to take back control of where the data sits/who accesses it. At the same time, the changes large tech platforms are introducing mean access to (third-party) data in the AdTech industry will be tightened and direct-to-consumer relationships and first-party data access will be increasingly essential for brands. AI Ethics is an area where the technology industry itself is calling for the governments to provide more direction but any regulation is currently in its infancy. A global attempt at creating an AI Ethics framework could trigger an acceleration of innovation (data privacy technologies and solutions), but we are also likely to see divergence between industries and regions. The world may need a ‘Paris Climate Deal’ moment for Data Privacy and AI Ethics regulation but the consensus building for this to happen is a long process and will be many years in the making.

Section 1
What Impact has GDPR Had?

The aim of this section is to understand what impact the General Data Protection Regulation (GDPR) has had on corporates and industry, two years since it came into force in Europe (in May 2018).

The GDPR shifted data protection from regulation to legislation, bringing with it greater enforcement powers. The headline grabbing points of GDPR related to the fines as well as the focus on consent as a legal basis by which personal data can be processed. The GDPR intended to increase transparency, tilt the balance of power back towards consumers, and place privacy by design and default at the heart of how businesses operate.

We won’t dig into the details of the legislation here as we covered it in detail in our previous Citi GPS report (here). As a reminder, here is a brief summary of the key points:

The first and foremost thing we need to think about is what GDPR says, which is to think of data as a liability rather than an asset. Every time you are holding onto someone’s personal data, you are sitting on a ticking time bomb. It could be leaked, it could be hacked and you could not have the right checks in place. If there’s no enforcement action you are not sitting on a liability, you are sitting on an asset.”

Catherine T O’Neill
European & CEEMEA Media Analyst, Citi
Corporate Response

The Center for Data Innovation estimates GDPR has cost Fortune 500 and FTSE 350 companies around $9 billion to comply with the regulation, and many non-EU companies have decided to forgo operations in Europe.

There was an initial flurry of emails to European consumers ahead of the implementation date, requesting consent to remain on lists etc., but the work hasn’t stopped there. Corporations are still getting to grips with GDPR and the best way to ensure compliance while limiting the disruption to business and data flows.

And it is not just those based in Europe that have felt the impact of GDPR. The global nature of the Internet means the regulation has been far reaching.

> There’s definitely a global impact, there’s no question. Whether it comes to direct applicability of GDPR, whether it comes through your vendors imposing GDPR requirements on you through contracts, even in developing markets, and as a result developing economies professionals know about GDPR. Many of their practices are having to align to it anyway.

— Justin Weiss

> Privacy and protection personal data is on the top of senior leaders’ minds at Citi and I believe it is a top priority here.

— Anne Fealey

Our analysis of hiring trends in the data privacy field shows there was a spike in 2018 that only started to come down from mid-2019. The spike was most significant for those industries with direct customer relationships and likely to hold personal data.

Figure 4. Data Privacy Hiring by Sector

![Figure 4. Data Privacy Hiring by Sector](source)

Source: LinkUp, CGDI

Figure 5. Data Privacy Hiring by North America vs. Western Europe

![Figure 5. Data Privacy Hiring by North America vs. Western Europe](source)

Source: LinkUp, CGDI
Regulated sectors have perhaps adapted quickest.

A lot of our clients, but not all of them by any means, are highly regulated — so the banks obviously, the insurance companies, and so on. These clients get it from a regulatory point of view a bit more. Actually, another sector we do work a lot in is pharmaceuticals and life sciences. So again, they’re more likely to be risk adverse. They get the idea of strong regulation and doing the right thing and being ethical about the way they use things like data, but it does vary by sector. We’ve done some work in media, for example, and they’re not bound by the same rules, which banks are, for example, financial transactions and things like that. And so I think there has to be a bit more of a cultural shift in those kinds of sectors.”

John Bowman

Data protection and privacy considerations have gone right up to the boardroom level where the topic is a feature of the Environmental, Social, and Governance (“ESG”) policies that companies adopt, driven by concerns over reputational risk, operational risk, and financial risk (i.e., fines).

There have been very public and widespread privacy concerns emerging, particularly for the Internet and digital sector which, I think, boards recognize must be taken very seriously in a proactive way. Of course the risk of fines, which has certainly increased and features as one of the key changes of the GDPR, brought the recognition that something akin to the level of competition fines could be appropriate in the context of data protection violations, and that is, of course, a motivator for those portions of the risk community and those advising Boards of Directors on risk.

Justin Weiss

Regulatory Action

The regulator’s actions are as important as the response of the corporations. As regulators police and enforce GDPR, their behavior and approach influences that of the operators. It isn’t just the fines that necessarily prompt a change in behavior, as the large platforms can swallow the fines without flinching, but using the power to force a change in non-compliant behavior is also vital.

We know from competition/antitrust enforcement that fines are immaterial if continued wrongdoing offsets the cost of the fine. I think the regulators who matter understand that now. What matters is their power to ban processing. In other words, they have the power to force companies to change how they do business. That is the power that matters.”

Dr Johnny Ryan

There is growing frustration amongst some within the privacy community that the regulators have been slow to act. The fact that the U.S. has been the first to take action in relation to Facebook and Cambridge Analytica, with a $5 billion fine, has not gone unnoticed. Max Schrems, Chairman of privacy group noyb, issued an open letter on the two-year anniversary of the GDPR to the European Commission and the European Data Protection Board (“EDPB”) asking them to “shine light on the shortcomings of GDPR enforcement”. The grievance was directed at the Data Privacy Commissioner as the pace at which it is investigating cases is slow. He calls for the documentation to be made available to all Data Protection Authorities (“DPA”) in the European Union.²

On 24 June 2020, the European Commission published its two-year review of the GDPR. It believes the regulation has been successful in enhancing individual's rights and ensuring the free flow of data within the EU. It also acknowledges improvements are needed in several areas including better harmonization across the EU, both the legislation in place in member states and the coordination between DPAs across member states as well as ensuring DPAs have adequate resourcing and funding. We dig into the consequences of this below.

**Regulators Have Funding and Resources Challenges**

There are clearly challenges for the regulatory authorities in enforcing GDPR, not least because of the breadth and depth of the legislation (covering all sectors and operators who provide a service in Europe), but also because more precise guidelines on specific aspects continue to be published by the EDPB. Coordination and resource constraints are also a factor. Each EU member state has its own DPA that acts standalone on investigations, but other member states can intervene before a final decision is taken. Not only does this create a bottleneck as certain regulators face a high proportion of the enforcement burden (mainly Luxembourg and Ireland as these jurisdictions are where the large tech companies typically have their European HQ), but it leads to a prolonged and drawn out process. Each DPA has different approaches and views and cracks between them will show. An attempt to create one single piece of legislation is very sensible on paper but coordinated enforcement is less straightforward in practice. An adverse outcome for Big Tech is also likely to get tangled in long legal proceedings which means the regulators need to ensure the investigations are as water tight as possible when decisions are announced.

Funding and resources are a challenge. A recent study by privacy browser Brave found that half of the EU DPAs have an annual budget of €5 million or less from their national governments and only two have annual budgets of over €50 million a year (the U.K. and Germany). Compare this to the significant cash piles the big U.S. technology companies have accumulated. Not only does this mean the DPAs likely have an informational disadvantage (given the Big Tech companies sit on the data) but there is clearly a huge funding disadvantage when it comes to legal action.

---

3 https://ec.europa.eu/info/sites/info/files/1_en_act_part1_v6_1.pdf  
4 www.brave.com
Headcount levels at the DPAs do not provide much in the way of reassurance. While the Data Protection Authorities in Europe have been increasing staff numbers, it still appears insufficient based on the deluge of requests, notifications, and cases they are required to look at, and compared to the data privacy hiring trend outlined above within industry (see Figure 4 and Figure 5). This means the regulators are likely to be reactive rather than proactive. The Irish Data Protection Commission (DPC) oversees many of the large tech companies with operations in Europe and the DPC reported it had only 140 staff as at the end of 2019. This compares to the Information Commissioners Office (ICO) in the U.K., which is amongst the largest of the DPAs by headcount, with close to 700 people and Germany almost 900 employees. It is worth noting the U.K. authority is no longer part of the EDPB following Brexit, which is a blow given its resourcing levels compared to other markets. Even in the U.S., only 4% (or 46) of the Federal Trade Commissions (FTC’s) 1,141 employees work in the Division of Privacy and Identity Protection.

---

6 Stigler Committee on Digital Platforms (2019)
It isn’t just about the headcount but also the expertise within the DPAs. A Brave study found that only six of the DPAs have more than 10 tech specialist staff (Germany has the most at 101 and makes up 29% of the tech specialists in Europe’s DPAs).7

> Some of the better-funded, larger regulators in the world have been investing in technical expertise and consultant groups, in trying to really beef up their understanding of how technology works and what solutions may be out there to address the privacy impacts of technology. That trend informs enforcement as well, because it strengthens the regulator’s capacity to negotiate good solutions with companies and also to call out bad acts and to even, and not be blinded by any lack of expertise about the technology itself.1

Justin Weiss

A survey by the EDPB in 2020 found that of all the DPAs in the EU, only five believe they have enough resources.8 One area where the DPAs have struggled is with the significant number of breach notifications.

> One thing that isn’t working that well right now operationally is breach notification. I know some DPAs have said they are struggling with the volume of notifications they get. The ICO is lucky as it’s well-resourced and can process them, but I don’t think, in its current form, it generates lots of insight for us or for elsewhere. We hear from data controllers that it’s not generating better protection for users. We shouldn’t get rid of it, but is there a better way to do it? Elsewhere, it’s very early to say. We do need to make sure that all the countries who have implemented the GDPR keep talking to one another.9

Simon McDougall

---

8 https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf
The Irish DPC published its 2019 annual report in February 2020, the first full year since GDPR came into effect, and noted a 75% increase in the number of complaints (to 7,215) and 71% increase in notified data security breaches (to 6,069). As at the end of 2019, the DPC had 70 inquiries on hand and six enquiries were opened in relation to multinational technology companies’ compliance with GDPR.

The number of lawsuits filed by privacy activists has been mounting since the GDPR came into effect. Of 70 statutory inquiries ongoing by the Irish DPC, 21 of them are cross border (see below) and at least six of them were triggered by privacy advocates.

### Figure 9. Irish DPC Ongoing Statutory Inquiries

<table>
<thead>
<tr>
<th>Company</th>
<th>Inquiry Type</th>
<th>Issue</th>
<th>Year</th>
<th>Inquiry Opens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook Ireland Ltd</td>
<td>Complaint based</td>
<td>Right of access and data portability - Examining whether Facebook has discharged its GDPR obligations in respect of the right of access to personal data in Facebook 'Hive' database and portability of 'observed' personal data.</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Facebook Ireland Ltd</td>
<td>Complaint based</td>
<td>Lawful basis for procession - Examining whether Facebook has complied with GDPR regarding the lawful basis it uses to process personal data of Facebook users</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Facebook Ireland Ltd</td>
<td>Complaint based</td>
<td>Lawful basis for processing - Examining whether Facebook has complied with GDPR regarding the lawful basis it uses to process personal data of Facebook users for behavioral and targeted advertising</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Facebook Ireland Ltd</td>
<td>Own volition</td>
<td>Facebook September 2018 token breach - examining whether Facebook Ireland has discharged its GDPR obligations to implement measures to safeguard personal data of users</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Facebook Ireland Ltd</td>
<td>Own volition</td>
<td>Facebook September 2018 token breach - examining whether it has complied with breach notification</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Facebook Inc</td>
<td>Own volition</td>
<td>Facebook September 2018 token breach - examining whether it has complied with measures to safeguard and secure personal data</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Facebook Ireland Ltd</td>
<td>Own volition</td>
<td>Breaches notified to DPC since May 2018 - examining whether Facebook Ireland has discharged its GDPR obligations to implement measures to safeguard personal data of users</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Facebook Ireland Ltd</td>
<td>Own volition</td>
<td>Facebook passwords stored in plain text format in internal services - examining whether it complies with GDPR</td>
<td>2019</td>
<td></td>
</tr>
<tr>
<td>WhatsApp Ireland Ltd</td>
<td>Complaint based</td>
<td>Lawful basis for processing in relation to terms of service and privacy policy - whether WhatsApp complies with GDPR regards the lawful basis it relies on to process personal data of users</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Instagram (Facebook Ireland Ltd)</td>
<td>Complaint based</td>
<td>Lawful basis for processing in relation to terms of service and privacy policy - whether Instagram complies with GDPR regards the lawful basis it relies on to process personal data of users</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Apple Distribution Int'l</td>
<td>Complaint based</td>
<td>Lawful basis for processing in relation to terms of service and privacy policy - whether Apple complies with GDPR regards the lawful basis it relies on to process personal data of users for behavioral analysis and targeted advertising</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Apple Distribution Int'l</td>
<td>Complaint based</td>
<td>Transparency - examining the information and transparency of information regarding processing of personal data of users.</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Apple Distribution Int'l</td>
<td>Complaint based</td>
<td>Right of Access - whether Apple has complied with GDPR in relation to access request</td>
<td>2019</td>
<td></td>
</tr>
<tr>
<td>Twitter International Co</td>
<td>Complaint based</td>
<td>Right of Access - whether Twitter has complied with GDPR in relation to right of access to links access on Twitter</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Twitter International Co</td>
<td>Own volition</td>
<td>Breaches notified to DPC since May 2018 - examining whether Twitter has discharged its GDPR obligations to implement measures to safeguard personal data of users</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>Twitter International Co</td>
<td>Own volition</td>
<td>In response to a breach notification with Article 33 of GDPR</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>LinkedIn Ireland Unlimited Co</td>
<td>Complaint based</td>
<td>Lawful basis for processing in relation to terms of service and privacy policy - whether LinkedIn complies with GDPR regards the lawful basis it relies on to process personal data of users for behavioral analysis and targeted advertising</td>
<td>2019</td>
<td></td>
</tr>
<tr>
<td>Quantcast International Ltd</td>
<td>Own volition</td>
<td>In response to a submission. Examining compliance with GDPR. Principle of transparency and retention practices will be examined.</td>
<td>2019</td>
<td></td>
</tr>
<tr>
<td>Google Ireland Ltd</td>
<td>Own volition</td>
<td>In response to a submission. Examining compliance with GDPR. Principles of transparency, data minimization and retention practices will be examined.</td>
<td>2019</td>
<td></td>
</tr>
<tr>
<td>Google Ireland Ltd</td>
<td>Own volition</td>
<td>Processing of location data and whether it has a legal basis for the processing of location data of users</td>
<td>2020</td>
<td></td>
</tr>
<tr>
<td>Verizon Media/Oath</td>
<td>Own volition</td>
<td>Examining company's compliance with requirements to provide transparent information under Articles 12-14 of GDPR</td>
<td>2019</td>
<td></td>
</tr>
</tbody>
</table>

Source: dataprotection.ie

### Fining Power Has Not Been Fully Tested

The cumulative value of the fines issued since July 2018 is close to €500 million, across a total of 240 fines. The number of fines issued has picked up pace since October 2019 (although April and May were impacted by COVID-19). Spain has issued the most fines but the U.K. has issued the largest fines by value although the fines issued by the U.K. ICO have been delayed twice.
The largest fines have been issued for data breaches or weak information security measures.

Providing some context, these fines are not significant given the GDPR enables fines of up to 4% of global annual turnover and when compared to those imposed for antitrust reasons. The antitrust fines have been much larger. The European Commission has fined major tech companies > €12 billion, with the bulk of the fines coming in the past couple of years, some of which are currently being appealed.
Signs of Action on Data Protection Enforcement

While the impact of COVID-19 may cause some delays, we are starting to see some movement. There were a number of rulings from the European Union Court of Justice during 2019 in relation to data protection. One of the most high profile court cases currently ongoing is at the European Union Court of Justice; a case brought about by privacy activist Max Schrems (in 2013), questioning the validity of Standard Contractual Clauses as a means to transfer data under GDPR. The decision is due on July 16th 2020.

The DPC released its findings on its Cookie Sweep investigation in April 2020 (launched in August 2019). The DPC used a traffic light system across the 38 controllers and only 2 were given a green rating. It found several firms relying on implied consent, i.e., the continued use of site implied consent for cookies, and raised concerns that special category data is being shared with online platforms through use of explicit profiles of logged in customers or predictive profiles based on unique identifiers. Of the 38 sites, 34 had third-party plug-ins. The DPC concluded there were systemic issues in the use of cookies and other tracking and is giving controllers a six-month period to comply before following up with enforcement.

As noted above, issuance of fines does not necessarily solve the challenges. Regulators have the power to request a company to stop processing data. This is likely to be much more effective in altering non-compliant behavior. Action by the regulators can also limit, slowdown, or prevent the roll out of new products.
The regulators are also well aware the current regulation may not be perfect. The click and access privacy notices have, arguably, been a failure. Consent fatigue has kicked in and, if anything, the overload of notices has probably had a detrimental impact on consumer trust online.

Mr. McDougall points out that companies are investing in designing notices which ‘encourage’ users to remain engaged and accept terms (known as dark patterns⁹), while continuing with irresponsible data practices under the guise of having obtained consent. The Stigler Center at the University of Chicago Booth Business School, which studies the relationship between the economy and the state, recently published a report on Digital Platforms¹⁰ which highlights ‘Notice and choice’ has failed in practice as the information provided is too detailed and intricate for consumers to make a well-informed choice and companies have become sophisticated about how to ‘nudge’ consumers into accepting terms that work for corporates.

The report noted interfaces that nudge user actions can increase acceptance rates related to data protection by 228%. The regulators have to be mindful of consumers having control over their data versus increasing the burden on consumers when there is a clear information asymmetry compared to corporates.

---

⁹ https://www.darkpatterns.org/types-of-dark-pattern
¹⁰ https://research.chicagobooth.edu/stigler/media/news/committee-on-digital-platforms-final-report
The GDPR is supposed to give consumers the ability to decide what data they want to share. This has left the responsibility sitting with the consumer to trawl through extensive privacy terms and language and make a choice. If there is no credible alternative then there is no choice. In practice, do consumers really understand how data is being used now, and could be used in several years, in order to make an informed choice? Do we go with the default choice due to the need for instant gratification or because we simply don’t have viable alternatives? Do we become desensitized to data breaches as consumers have little control over rectification?

While there are clearly still GDPR-related challenges to iron out, the European data protection regulatory landscape continues to evolve. We are still watching the ePrivacy regulation and whether it comes through, as this could further tighten up requirements for communications (see our last Citi GPS report). It has been under review for over three years now and appears to have hit a stalemate position between the member states and the European Parliament.

Impact on Big Tech & Investment in European Tech

One of the conclusions we came to in our previous Citi GPS reports was that the GDPR could unintentionally reinforce the dominance of the large tech platforms. Not only can they manage the financial impact that deploying the regulatory requirements has had, but they are in an enviable position of being embedded in consumer’s digital lives, which creates an advantage when it comes to pushing for consent and laying out the value exchange. Studies conducted since GDPR came into effect appear to support our thesis.

A group of academics recently published the European Privacy Law and Global Markets for Data report, which included its findings from tracking 110,000 websites between May 2017 and November 2018 to follow the web technology being used and how this changed post GDPR implementation.11 It noted four main findings:

1. A reduction in the use of third-party cookies and an increase in first-party cookies.

---

2. The impact has been felt globally; even those sites and consumers not directly subject to the regulation.

3. Privacy policies of tech vendors provide more information.

4. Market power shift with an increase in Google’s market share post GDPR versus other web technology providers.

This is consistent with the findings of a study by whotracks.me which looked at 2,000 websites and what happened to tracking since the GDPR. They found there was a 4% reduction in the number of trackers per page in the EU from April 2018 to July 2018, compared to an 8% increase in the U.S.

Figure 15. Average Trackers per Page by Region (% change)

Figure 16. Reach of AdTech Vendors

Source: whotracks.me

However, this trend masks the divergence between operators. In Europe since April 2018, the study found there had been limited change to the reach of Google’s advertising-related trackers compared to notable declines for those of other advertisers.

While we can’t definitively conclude the GDPR has led to a reduction in investment in technology in Europe, in the 12 months following GDPR, venture capital investment in European start-ups declined 33.8% and the EU is home to only 11.5% of the world’s tech unicorns.

At the same time, tech companies have been increasingly calling for data-related regulation (there may well be an ulterior motive too) as a clear regulatory framework means companies know the boundaries they need to operate within.

---

12 https://whotracks.me/blog/gdpr-what-happened.html
If the boundaries, however, differ significantly by market or interpretation varies from operator to operator, then it is not a level playing field. This leads to the competition concerns that are now rearing their head. The U.K. Competition & Markets Authority noted as much in its recent publication of online platforms and digital advertising:

‘Our concern is that such platforms have an incentive to interpret data protection regulation in a way that entrenches their own competitive advantage.’

Implications

To come back to our question of what impact GDPR has had, there are signs the number of trackers online have fallen and privacy notices have increased in prominence. More importantly, it has helped to elevate the importance of personal data protection globally at a government, regulatory, and board room level. There are however many challenges remaining:

– Enforcement is building momentum but it has been slow. Regulators are under resourced and underfunded for the task at hand and enforcement is likely to get tied up in legal wrangles for many years. The regulators also face an informational (and financial) disadvantage versus Big Tech.

– There are signs the spirit of the law is not always being followed, evidenced by practices such as dark patterns.

– Consent notices have led to consent fatigue and consumer understanding of shared data remains patchy. Instead of fostering trust it potentially increases lack of trust.

– The regulators have not tested their full fining power yet but it won’t be fines that change behavior. They may grab the headlines but limitations on processing, for example, will address non-compliant behavior.

– The regulatory burden has been more painful for smaller entities. There is evidence the GDPR has reinforced the already strong position of the Big Tech companies. We will turn to this later as this is where we see the interplay with competition authorities increasing.

14

https://assets.publishing.service.gov.uk/media/5efc57ed3a6f4023d242ed56/Final_report_-_1_July_2020_.pdf
Section 2:
How Will Global Privacy Regulation Evolve?

New Data Protection Laws

In this section we look at how global data protection regulation has developed, in particular in the U.S. with the California Consumer Privacy Act (“CCPA”). We dig into how cultural differences influence the approach to data protection regulation. It will become increasingly complex as the battle for AI superiority is heating up but there are tensions with privacy regulation requirements. Addressing data privacy within a framework for 'Ethical AI' is likely to be a significant distinguishing factor in how the uses of AI unfolds.

While some may be frustrated with the pace of GDPR enforcement, the legislation has raised the importance of data privacy and brought the conversation to the fore. We believe it has been a factor behind how the global data protection regulatory landscape is changing with many markets introducing/updating data protection regulation and legislation.

The number of data protection standards have been rising since the 1970s, but we are still at the beginning of the evolution. Regulation has a lot of catching up to do with the pace of technological change.

GDPR may not be future-proofed enough even though it tends to dominate discussions about privacy and data protection regulation as the detail, depth, and breadth of what the GDPR covers is much further reaching than any other piece of data protection regulation. We will come back to this later when we dig into AI.

Most major economies have either individual pieces of data protection regulation or have proposals and bills to bring forward data protection legislation. The GDPR is absolutely a benchmark and a standard for markets to consider. When we look at Brazil, when we look at India, when we look recently in Japan, when we look at the United States, these economies are certainly noticing that data protection regulation is important in this age — from a trade perspective, from a governance perspective, from a citizens trust perspective. It's becoming a cost of doing business in the sector. There are really not very many places to hide from good data protection practice, for data intensive companies."

We are at a very early stage. If you zoom out, the fact that so much of the focus has been on GDPR and CCPA masks the breadth of the spectrum of cultural environments in which very different approaches to consumer data control and privacy are going to end up emerging — or not emerging. I would put us on guard against looking for convergence and coincidences between GDPR and CCPA, and thinking this is where the world is going. There's going to be a lot more and it might be very different."

I think the GDPR does represent the global state on the articulation of individual privacy rights. It's a thoughtful, comprehensive framework, but it really doesn't provide the specific guidance needed by more data intensive businesses. The revolution, if you will, of the GDPR was to say that you can't process data without some legal basis. The one that I think is really important, but underexplored, is this idea of legitimate interest. I think it's actually a very powerful and potentially agile tool for companies that have a lot of personal data and want to make use of it for advanced analytic purposes in a machine learning context. “

© 2020 Citigroup
The global data protection landscape is by no means static. There are a number of amendments to existing or new data protection laws recently introduced and coming down the pipe, some of which we have listed below. Common among many of them is increased transparency and consumer control over data, specifications on data localization, and increased fining power. The COVID-19 pandemic, however, could delay implementation in some regions — Thailand and Brazil are considering delaying new laws to 2021.

### Figure 17. New or Amended Data Protection Laws

<table>
<thead>
<tr>
<th>Country</th>
<th>Date</th>
<th>Law/Amendment</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>Current</td>
<td>Amendments to Data Protection bill</td>
<td>Bring it in line with GDPR (right to be forgotten, right to data portability)</td>
</tr>
<tr>
<td>Brazil</td>
<td>January</td>
<td>General Data Protection Law (LGPD)</td>
<td>Replaces 40 regulations with one that is aligned with GDPR plus creation of a DPA</td>
</tr>
<tr>
<td>Canada</td>
<td>Current</td>
<td>Personal Information Protection and Electronic Documents Act (PIPEDA) 2000, amended in November 2018</td>
<td>To include mandatory data breach notification and record-keeping laws</td>
</tr>
<tr>
<td>Chile</td>
<td>Current</td>
<td>Considering net bill</td>
<td>Additional rights for data subjects, obligations for data controllers</td>
</tr>
<tr>
<td>China</td>
<td>2020</td>
<td>Several</td>
<td>Legal Committee of the National People's Congress Standing Committee announced in late December 2019 that the enactment of legislation on personal data protection and security is a priority in the next legislative year</td>
</tr>
<tr>
<td>Colombia</td>
<td>Current</td>
<td>Bill in Congress</td>
<td>Supplement to law No 1581. Allow DPA to impose fines</td>
</tr>
<tr>
<td>India</td>
<td>Current</td>
<td>Indian Personal Data Protection Bill 2019</td>
<td>Similar to GDPR, establishes a data protection authority, increases consumer rights and requires additional IT security, data localization (in India), consent management</td>
</tr>
<tr>
<td>Japan</td>
<td>Mar-20</td>
<td>APPI Amendment 2017</td>
<td>The APPI Amendment restricts data transfers to a third country without obtaining data subjects' consent where the level of data protection is insufficient. Data breach notification requirements and increase in penalty for breach. The amendment has helped to get Japan on the EU’s “white list” of countries with adequate data protection legislation</td>
</tr>
<tr>
<td>Malaysia</td>
<td>Current</td>
<td>Personal Data Protection Act</td>
<td>Mirrors many aspects of GDPR</td>
</tr>
<tr>
<td>Russia</td>
<td>Dec-19</td>
<td>Data localization</td>
<td>Data localization rules have been in place since 2015 (to maintain Russian personal data in Russia). The update brings in financial penalties for non-compliance</td>
</tr>
<tr>
<td>Singapore</td>
<td>Current</td>
<td>Changes to Singapore’s Personal Data Protection Act</td>
<td>Draft Personal Data Protection (Amendment) Bill 2020: (1) Mandatory breach notification; (2) Expansion of deemed consent; (3) Individuals can port data between services; (4) Fines increased up to 10% of annual turnover or $1 million (whichever is higher)</td>
</tr>
<tr>
<td>Thailand</td>
<td>May-21</td>
<td>Personal Data Protection Act</td>
<td>Mirrors many aspects of GDPR</td>
</tr>
<tr>
<td>United States</td>
<td>2020</td>
<td>California Consumer Privacy Act (CCPA)</td>
<td>Gives residents of California four rights giving them more power over their personal data: right to notice, right to access, right to opt in (or out), and right to equal services. Any organization collecting the personal data of California residents, not just businesses located in the state, must comply with CCPA</td>
</tr>
</tbody>
</table>

Source: Citi Research

The situation in Europe also remains fluid. In the U.K., there are question marks over whether GDPR will remain in place post Brexit. The U.K. Prime Minister, Boris Johnson, said in February 2020 that the U.K. will develop ‘separate and independent policies’ in a range of areas as part of the withdrawal agreement, but the timing is tight and it likely depends on the U.K.’s willingness to maintain standards similar to the GDPR. The area of surveillance is likely to create tension and the commentary from the EU negotiator Michel Barnier suggests that an agreement on the free flow of data between the U.K. and Europe could be challenging. He claimed the U.K. is ‘lowering current standards and deviating from agreed mechanisms of data protection — to the point that it is asking the EU to ignore its own law and the jurisprudence of the European Court of Justice (ECJ) on passenger data (e.g., PNE rules, the rules for presentation of European Standards)’.

The California Consumer Privacy Act

Before we explore the global landscape further, we want to spend some time on the California Consumer Privacy Act (CCPA). This law really came about as a reaction to the Cambridge Analytica/Facebook events which unfolded in early 2018. The GDPR was looking unlikely until the Snowden scandal, which provided impetus to those pushing for its approval within Europe. These are illustrations of how unexpected events in the real world can trigger a long lasting change in the direction of regulation. Maybe we are cynics, but it seems highly unlikely these will be the last ‘unexpected’ data related events.

The U.S. has typically followed a sector-based approach to data protection. The introduction of the CCPA from 1 January 2020, with enforcement from 1 July 2020, has created waves. The CCPA came about much quicker than the GDPR. It has similarities to GDPR, although more narrow and arguably more prescriptive, and the penalties have potential to be more significant.

Background to CCPA

Our discussion with IP lawyer Emma Maconick (Sherman & Sterling) detailed the background to CCPA, what changes it brings, and how it will be enforced.

CCPA Requirements

We summarize below the main requirements of the CCPA. There have been several tweaks and the California State Attorney General (“AG”) continues to update and issue guidelines.
Figure 18. Main Requirements of the CCPA

<table>
<thead>
<tr>
<th>Personal Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CCPA Applies to</strong></td>
</tr>
<tr>
<td>• For profit businesses with gross revenue &gt; $25 billion</td>
</tr>
<tr>
<td>• Businesses that receive or disclose information of more than 50,000 Californian residents</td>
</tr>
<tr>
<td>• Derive more than 50% of annual revenue from selling California resident’s data</td>
</tr>
<tr>
<td>• Applies to California residents (37m people, ~12% of U.S. population)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Definition of Personal Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Information that identifies, relates to, describes, is capable of being associated with, or could reasonably be linked directly or indirectly, with a particular consumer of household, with many specific examples including geolocation data and biometric information</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CCPA Main Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Businesses inform consumers prior to collection of personal data the purposes for use of the data</td>
</tr>
<tr>
<td>• Provide option to opt out of third-party data sharing and sale of personal information. Do Not Sell link on homepage for those businesses that sell data so consumers can opt out</td>
</tr>
<tr>
<td>• Online privacy policies to describe consumer’s CCPA rights and detail all of the categories of information being collected</td>
</tr>
<tr>
<td>• Consumer rights enhanced - right to request information on the data and right to request data deletion</td>
</tr>
<tr>
<td>• Have equal service and price regardless of whether they exercise privacy rights</td>
</tr>
<tr>
<td>• Exemptions for personal information covered by the Health Insurance Portability and Accountability Act (HIPAA), Gramm-Leach-Bliley Act and other federal and related California privacy laws</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Enforcement</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Responsibility of the California Attorney General. Fines of up to $7,500 for each intentional violation and $2,500 for unintentional violations</td>
</tr>
<tr>
<td>• Individual class action possible in relation to data breaches (up to $750 per consumer per incident)</td>
</tr>
</tbody>
</table>

Source: Citi Research

There are aspects of the CCPA that are elegantly crafted. Other aspects are less so, with unintended consequences, or where legislators had not fully addressed how these laws would be operationalized. The law was amended multiple times, and the California State Attorney General also published draft regulations in October 2019, which were intended to help clarify the ‘how are you going to do it’ question.

As for the key requirements of the CCPA, the first thing you have to figure out is whether or not a business is subject to the CCPA. A "business" defined under the CCPA, is essentially a for-profit legal entity that collects consumers’ personal information, determines how that information is processed, and does business in California. A business also needs to meet at least one of three thresholds: it has to have annual gross revenues of more than $25 million, or it collects more than 50,000 consumers, households, or devices worth of data annually or it derives 50% or more of its annual revenue from selling consumer’s personal information. That third threshold is directed to data brokers and data aggregators.

After you determine whether a business is covered, then you need to take a look at whether a person is covered. A “consumer” under the CCPA essentially means a natural person who is a resident of California — meaning somebody who lives and intends to stay in California.

And the activities regulated under the CCPA are the collection and sale of personal information. This is one area where people are especially exercised because that definition of sale is extremely broad, capturing any making available, transferring, renting, releasing, disclosing, etc., of personal information. A lot of analysis has gone into determining whether an activity is a “sale” under the CCPA.
What CCPA Changes

If you think about the changes from the previous California regime, there is much more required to comply with the law under the CCPA than was required previously. More attention needs to be paid to the collection, processing, and storage of personal data, and you need to be very granular about your business operations to understand how the data flows — where you receive data from, the purposes for which you use it, who you share it with, how you transmit it. All of this requires a greater allocation of resources and time. You should also understand the security around your business’s personal information, even though the CCPA is not a data security regime.

Another entirely new aspect for California law is PIARs, or personal information access requests. Under the CCPA, consumers have rights they can exercise by a verifiable consumer request — the right to know, the right to delete, the right to information about a business’s data practices, the right to opt out, etc. These are similar to rights granted under GDPR, but there are differences between the laws. So companies are operating under both CCPA and GDPR will need to keep differences between the data regimes in mind.

Consumers also have a right to notice at or before the time of collection under the CCPA. And they have a right to non-discrimination, meaning that consumers can exercise their privacy rights under the CCPA and not be discriminated against by the business, including in pricing.

Critically, consumers also gain a private right of action against businesses that suffer a security breach resulting from the business’s failure to do certain things. That’s quite new under U.S. law.

Ten years ago, personal data in California was a list of nine or 10 things right? Name, address, Social Security number, bank account, that sort of thing. Now with CCPA, it includes anything that can reasonably be used to identify someone, including things such as IP addresses, cookies, device identifier — so, boom, all cookies by definition collect at a minimum IP addresses and browser information. So right there the CCPA is saying third-party online data collection is collecting personal information.

With the narrow focus on the sale of data, some organizations (especially in the online advertising space) are assuming not all requirements apply.

A lot of organizations saw the service provider exemption as a kind of residual catch-all if you will, as a way to get around the sale analysis. The digital advertising industry is actually taking a position at large that many of its players in that ecosystem could constitute as service providers, meaning there is not a sale occurring, so your legal obligations are not triggered with regard to a sale.
Enforcement is by the California State AG and the Fines Could be Significant.

Only the AG can enforce the CCPA requirements. There is also some process around this. First, the AG is prohibited from bringing any enforcement actions until 1 July 2020. Second, for any enforcement action the AG takes, there is a bit of breathing space. The AG’s office must first provide the entity with notice of the alleged violation or nonconformity with the CCPA, and the entity then has 30 days to cure it. If you can cure an issue raised by the AG, there should be no further action from the AG. If you can’t cure it, then the AG can bring a civil action with civil penalties of up to $2,500 per violation, or $7,500 per intentional violation. I’m not sure how much enforcement activity we’re going to have, but that’s to be determined.

Separately, the CCPA also provides a private right of action for individuals, for California consumers… The personal right of action only comes into play if there’s been a security breach that exposes Joe and Jane’s personal information, and to recover damages, they also need to show the breach results from the business’s violation of its duty to implement reasonable and appropriate security procedures. Of course, what the duty to implement security procedures requires will be debatable in individual cases.

If a California consumer succeeds in a private action, the CCPA provides statutory damages of at least $100 per consumer per incident, up to $750 per consumer per incident. Actual damages are also available, but it’s been historically difficult to show actual harm in data breach cases. And we can see how the statutory damages can scale up very quickly.

A number of companies and bodies wrote to the California AG requesting enforcement be delayed from 1 July 2020 to 2 January 2021 in light of COVID-19 which has led to a slowdown in creating processes to ensure compliance. In addition, the interpretation and finalization of the law itself is still unclear. The California AG’s office said it remains committed to enforcing the law from 1 July although the state Office of Administration Law is yet to review the proposed regulation, which is required before it takes effect, so this could lead to enforcement being pushed back to October.

The CCPA is not the end for consumer privacy law in California. The CCPA 2.0 or California Privacy Rights Act (CPRA) could appear on California’s ballot in the November election and it could have more teeth than CCPA for the AdTech industry. If adopted, amongst other things, it would come into force from January 2023 and it looks to tighten up definitions and the use of sensitive data, expand breach liability, introduce a new enforcement agency (California Privacy Protection Agency), and enhance consumer data rights further.

Likelihood of U.S. Federal Data Protection Law

The challenge in the U.S. is that privacy regulation has not yet been tackled at a federal level. There are sector level laws and new privacy laws are evolving at a state level, starting with CCPA, and a similar approach is being considered in several other states. This could make it extremely challenging for businesses operating across the U.S. as it means a growing number of privacy laws to navigate. Many tech companies and corporates are pushing for a federal framework for privacy.
The composition of a federal privacy framework may vary depending on the political party in power, but we see it as highly likely a federal privacy law of some kind is introduced, although not in the near future.

There have been a number of draft U.S. federal privacy bills proposed but data protection (and any) regulation is not created in a vacuum. Political influence is unavoidable. The difference in views between the two political sides in the U.S. can be neatly illustrated by the proposals on data protection during the COVID-19 pandemic. Senate Republicans proposed the COVID-19 Consumer Data Protection Act on May 10th. Democrats in the House and Senate then proposed the Public Health Emergency Privacy Act. Many areas of each bill are consistent, but many are not. Both bills require express consent from consumers to collect or process COVID-19 data — a clear privacy policy — a requirement for recurring deletion, and a data minimization requirement that data should not be collected beyond what is necessary and proportionate to public health needs (such as for advertising, offers of finance, credit etc.). In terms of major differences between the two the Democrats proposed bill is much broader in the scope of who it covers (private and public entities), the data it covers, the privacy-related elements, and enforcement.

**Challenges Created by Data Protection Regulation**

While we are seeing an increase in global data protection regulation, there are two main challenges: (1) technology infrastructure and (2) lack of a consistent global approach by governments and regulatory authorities.

---

Regardless of the specifics around the regulation, one of the biggest challenges in the realms of data protection is the current technological infrastructure is not capable of dealing with all that the regulation requires. This will require investment and innovation (see more details on privacy tech in the final chapter).

"Which ever legal framework you’re talking about, you have this trajectory where the regulators are expecting companies that use personal data be able to exercise more control, have more predictability, and give users more control over their personal data than has been the case in the past. In fact, we do not have the technical infrastructure today, in any organization to actually comply with those rules in a serious way. I think that’s the fault of computer science, to be really honest. We have not built systems which allow the kind of accountability and control over personal data privacy laws are actually now starting to demand.

I think there’s actually a lot of really interesting innovation opportunities here, because I think there are new capabilities that enterprises all around the world, and governments all around the world, are really going to need and we just don’t have today. Number two, I think they’re critical to having a curious, but agile privacy regulatory environment. People don’t want to completely shut down the use of personal data, they want to know that surprises won’t happen. That’s what I pick to be a really core privacy principle: just don’t surprise people, but right now we don’t have systems that can handle that.”

The biggest technology challenge we had with GDPR was manual. Some of the processes are manual. We collect the requests manually, the letter processes are manual, etc. With the right to be forgotten and need to delete and not maintain data, we’ve had some significant challenges. Over the last year plus, we’ve had a reasonably large program overseeing the storage limitation aspect of GDPR to make our technology applications compliant with all those requirements. Frankly we need to change the way we architect applications, so we build that into all of them from the beginning: No one wants to continue doing this as a fire drill. It’s expensive, as you can imagine."

The lack of a global approach to data protection is also a challenge. This creates complexity for corporates who operate in multiple markets as they have to adhere with a number of national policies, many of which come with specific requirements about data localization, creating geographical boundaries around the physical location of data.

Figure 19. Number of Privacy Laws Privacy Professionals Are Complying With (by region)

Source: IAPP and TrustArc (Nov 2019)
The approach to implementation of data protection regulations can have implications for competitiveness.

To Mr Weiss’ point below, the challenge may be greatest for those businesses where privacy concerns are explicitly tied to the structure underpinning the business, regardless of their size. One could certainly argue that is the case for online advertising-funded models which are based on amassing more and more granular data.

We have to be mindful of the impact an increasing regulatory burden can have on a sector over a long period. As we have seen from the telecom and bank industries, regulation has attempted to alter barriers to entry, and has led to reduced pricing power and impacted margins. It is not necessarily one single piece of regulation driving this but an accumulation of regulation over time. In the case of telecoms, the product has become more desirable as consumers live their lives on their devices but the benefits have not accrued to the operators.
Cultural and Geopolitical Tensions Create Challenges

There is clearly a movement for more data protection regulation but it is worth exploring why a lack of consistent global approach has evolved, and what the implications are because this will be a key factor behind how technology can evolve within each region.

We see two interlinked drivers:
1. Corporations and governments have come to realize the potential power data and technology can provide. If data is an asset, you don’t want anyone else to get their hands on that asset. In the Western world, corporates are ahead of the curve. The governments realize this means the power is concentrated in very few hands and they need to level this out.

Vivienne Artz

“If data can flow freely, data can be copied, data can be derivied, and there’s a lot more you can do with data. The potential is extraordinary. It is interesting to see the growing awareness and realization as to who has access to the most (and diverse) data in the world, and it is not governments who know everything about us, it is actually the global tech companies. … The ability to monetize that data is huge.”

2. The political and cultural environment is instrumental in how regulation evolves. What makes this particularly challenging when it comes to data protection is that our digital lives are borderless (consumers use sites, apps, and information from across the world) but the regulation is far from borderless as the fundamental beliefs on consumer rights underpin the approach to data privacy.

John Bowman

“The important thing about GDPR is it is rights-based legislation. It circles back to the Charter of Fundamental Rights as well. Culturally in Europe, which might differ from the U.S., there’s a kind of historical legacy building upon the horrors of Nazism and communism and the sort of surveillance society. Particularly in places like Germany, which has experienced both those things, there is an idea that people should be free of state surveillance because some of the worst abuses of power have arguably been perpetrated historically by the state operators and government. In some ways, people are shifting their thinking about whether it’s the state who are the bad people, or other parties now, like the big corporations? There is a cultural background which I think plays out less in the U.K., actually, because we weren’t under those sorts of totalitarian regimes, but I think in the U.S. they don’t really think it from that perspective.”

Emma Maconick

“It’s more helpful to think of the U.S. as coming from a more capitalist view. Its perspective is that data is a kind of asset, something of value, and personal data can be an individual asset or an asset on the balance sheet.”

Both of these factors have manifested themselves in the concept of data localization, incorporated into data protection regulation in many markets. The idea of data localization stems from both a lack of trust of how personal data will be used in other markets (that may not meet data protection standards in the citizen’s market) as well as the growing understanding of both the power and value of data (data retention within the country it was generated). In some markets it goes further than data localization with carve outs in the data protection regulation for the state vs. private companies.

- Under the GDPR, countries must have ‘adequacy’ status (there are only 13 countries deemed adequate) to enable the transfer of data. The European Commission acknowledges the value of data being able to flow between markets, and is updating Standard Contractual Clauses, streamlining the Binding Corporate Rules approval process, talking of promoting data sharing with trusted partners, and is currently considering providing ‘adequacy’ status to more countries in Asia and Latin America.
However it is against access of foreign public authorities to data, so will stop well short of allowing data to freely flow globally. 17

– The Indian data protection bill was introduced in Parliament in December 2019 identifies privacy as a fundamental right and has many similarities with the GDPR but there are exemptions for government access to consumer data for ‘national security’ purposes. The law also brings in a requirement for social media companies to provide consumers with a voluntary means to verify identity.

– The law in Russia requires data on Russian consumers is retained within Russia and created a regulatory framework promoting a ‘sovereign Internet’, which will enable state-based centralized control. A law was also recently introduced requiring Russian software to be pre-installed on hardware.

Data localization is challenging for companies to manage in practice.

We are seeing data localization in a number of countries. African countries have become a challenge in that they want data maintained locally. In some case we’ve been able to convince the government that having a copy that isn’t used is sufficient, but we’ll see where that goes over time. If we have to start running data in local countries, it’s going to impact our ability to serve global clients."

Bill Philhower

"We can say, “Oh, isn’t this great, more and more countries are adopting privacy legislation”, and the logical consequence should be if all those countries have privacy legislation there could be free movement of data between them. That should be the natural and logical consequence. But it’s absolutely the opposite. It’s now creating barriers. Look at India with the proposal for data localization; look at China, also data localization. Indonesia, data localization. The list goes on.

India, unashamedly, in their legislation, has focused on protecting its citizen’s data by proposing to localize data. China is localizing on the basis that the national interest takes priority. They are pretty open about it. India’s approach is to protect its citizen’s data against use and misuse by what are seen as large unregulated companies hoovering up the data. The value transaction between what individuals get, i.e., a free service, and what the entity hoovering up the data gets, are seen to be oceans apart. ... That value transaction had worked for a long time until it was realized these big companies have more money than many economies in the world. That’s how profitable the data opportunity is." 17

Vivienne Artz

State surveillance is a contentious topic and one that epitomizes cultural differences between regions. The idea of state surveillance in Europe (and the U.S.) has often been seen as sinister and goes against human rights, hence the backlash when the Snowden revelations were reported. As Vivienne Artz notes, the drawing of the line varies by market. In many instances this has really highlighted where the power sits.

17 https://ec.europa.eu/info/sites/info/files/1_en_act_part1_v6_1.pdf
In China, the Internet is highly censored and foreign platforms aren’t generally allowed to operate but it doesn’t mean there are no data protection requirements in place. They just don’t look the same as the rules in Europe, for example, when it comes to the state. The 2017 Cybersecurity Law included the concept of consent and limits network operators to collecting only the necessary data. China is set to introduce a new Personal Data Protection Law in 2020. While data privacy requirements are tightening in China for private companies, the same rules don’t apply to the state as the state’s interest ranks above that of an individual.

There have also been media reports of a Social Credit System incorporating data from across various platforms, transactions, movements etc. to assign a score for each citizen. Some academics and journalists have questioned whether the way consumers are profiled in the Western world is actually that different to China. Casey Newton at The Verge wrote:

“"The more I look around, the more it seems like an American social credit system is springing up around us—and it doesn’t look all that different from China’s.""^19

Edward Snowden highlighted the level of state surveillance in the U.S., which provided impetus for the introduction of the GDPR in Europe. Google and Facebook both publish requests for information they receive from government agencies, and both have seen a consistent increase in requests over time.

Figure 23. Google: Global Government Agency Request for Information

Source: Google

---

18 https://en.wikipedia.org/wiki/Social_Credit_System
19 https://www.wired.com/story/china-social-credit-score-system/
While state surveillance is off limits in Europe, has this opened the door to commercial surveillance in capitalist societies?

Shosana Zuboff coined the term “Surveillance Capitalism” in her recent book “The Age of Surveillance Capitalism”, and she sees the development of free services in exchange for the ability to monitor behavior of users as the latest phase in the evolution of capitalism.

“Surveillance capitalism unilaterally claims human experience as free raw material for translation into behavioral data. Although some of these data are applied to service improvement, the rest are declared as a proprietary behavioral surplus, fed into advanced manufacturing processes known as ‘machine intelligence’, and fabricated into prediction products that anticipate what you will do now, soon, and later.

Finally, these prediction products are traded in a new kind of marketplace that I call behavioral futures markets. Surveillance capitalists have grown immensely wealthy from these trading operations, for many companies are willing to lay bets on our future behavior.”

“Demanding privacy from surveillance capitalists or lobbying for an end to commercial surveillance on the Internet is like asking old Henry Ford to make each Model T by hand. It’s like asking a giraffe to shorten its neck, or a cow to give up chewing. These demands are existential threats that violate the basic mechanisms of the entity’s survival.”

The author provides 16 reasons as to how surveillance capitalism has been able to thrive. One of them relates to the timing being fortuitous. She believes surveillance capitalism was invented in 2001 benefitting from: (1) A view at the time that self-regulation was possible as imposed regulation was an impediment to the free market and (2) In 2001 Congress was considering new privacy related legislation, which could have limited some of the practices, but 9/11 quickly shifted the focus away from privacy to a need for information.

Against this, there is the argument that consumers don’t care all that much about data privacy. Consumers are provided with terms and conditions and consumers accept the services are free in exchange for data and, in fact, place a significant value on those services. Whichever view you are swayed by, the reality is that what society accepts and trusts will also determine the path of innovation. This is where differences in cultural and societal norms play a role.

20 https://www.theguardian.com/technology/2019/jan/20/shoshana-zuboff-age-of-surveillance-capitalism-google-facebook
Deployment of facial recognition in China is the norm, for example, whereas in many Western markets the intrusiveness and privacy-related challenges have limited the roll out. Facial recognition has more recently come to the fore as an example of AI exacerbating biases, which we discuss in more detail below.

According to Edelman’s Trust Barometer the average level of trust of non-governmental organizations (NGOs), business, government, and media is the highest in China, followed by India, and the lowest in the U.K. and Russia. None of the European countries scored particularly well. An interesting outcome when we take into account that European regulation is built around trust, transparency, and protecting consumer rights.

**Figure 24. Edelman’s Trust Index Barometer**

![Edelman’s Trust Index Barometer](image)

Source: Edelman. Average % trust in NGOs, government, business and media on a 9 point scale (% scoring 1-4)

**COVID-19: A Study in Cultural Differences**

The COVID-19 pandemic has really highlighted these cultural data privacy variations by country. South Korea and China have been able to access a wide range of data for the purpose of contact tracing (credit card and geo location information and use of creation of a color coded system in China based on QR codes). Gaining trust and consent from citizens has not been a necessary, or restrictive, requirement.

Ironically, many developed market governments are relying on Big Tech operators for the development of tools and services for health authorities to build contact tracing apps on. Big Tech are prioritizing privacy with a decentralized system (information is retained on the device) despite some governments pushing for a centralized system with location tracking to identify hotspot areas and trends.
The EC’s guidelines specify national health authorities should be responsible for the app and usage should be voluntary with consent required for each separate function of the app. The data should be stored on the user’s device and data minimization should be adhered to. Margarthe Vestager, the European Commissioner for Competition said in an interview with the New Yorker: ‘I think it is essential that we show that we really mean it when we say that you should be able to trust technology when you use it, that this is not a start of a new era of surveillance. This is for virus tracking, and this can help us open our societies.’

In the Western world, COVID-19 tracking has certainly created a dilemma as the integral and essential role of Big Tech in helping to provide a data-driven means to control/manage the virus has become apparent. Privacy concerns at a consumer level may reduce the effectiveness of using apps to track the virus in the Western world. Where contract tracing apps have been deployed countries have seen about a 20% take up by the general public, according to Forrester.

One question coming out of the pandemic in some markets is whether COVID-19 is the foot in the door that opens up for mission creep and mass surveillance. Reuters reports the city of Huangzhou in China is proposing to permanently assign each resident with a colored health badge based on medical records and lifestyle habits e.g., exercise, eating, and drinking habits, which faced backlash on social media platforms.

**AI Power Struggle**

Clearly there is a power struggle brewing (or already occurring) between countries and between governments and private operators when it comes to data. While GDPR is being held up as the international standard of data protection regulation and aspects of it are being mirrored in many markets, the complete harmonization of policy globally is highly unlikely due to cultural and political differences (even if what society accepts differs more by market than the actual behavior of the states).

This will have direct implications for the evolution and uses of AI across markets. AI is a battleground that is heating up as world leaders recognize the development of AI being an important force in driving economic growth and social good. At the same time, emerging AI platforms raise a host of new ethical issues.
Commentary on AI highlights the tensions and possible power struggle in this area, which could undermine the ability to drive a global approach to protecting consumer privacy and personal data and, more broadly, society at large from bias, inappropriate profiling, and discriminatory uses of data. This is not to say technology developments always come at the expense of data protection but the two are not always going to develop in unison, especially when amassing huge quantities of data underpins the technology.

The focus tends to be on national development, as opposed to global collaboration, and data privacy and ethical considerations will clearly need to be addressed. There is a risk the race to establish a strong position within the development of AI leads to privacy and ethical considerations not being properly addressed and developed; either almost entirely overlooked or a too heavy-handed approach. Or, varying approaches to AI ethics and privacy by region will play a role in determining which regions are well placed to lead the AI battle. It certainly makes for an uneven playing field.

In 2017 Russia’s president Vladimir Putin said ‘Artificial intelligence is the future, not only for Russia, but for all humankind. It comes with colossal opportunities, but also threats that are difficult to predict. Whoever becomes the leader in this sphere will become ruler of the world.’\(^{21}\)

In 2017 China announced its ambition to be the world’s leader in AI by 2030.

In December 2019 the European Commission added another string to the bow of Margarethe Vestager’s role, appointing her as Executive VP for a “Europe fit for the Digital Age” alongside her position as the Competition Commissioner. This means she will be responsible for steering and coordinating Europe’s approach and strategy to propel its position as well as ensuring the competition policy is fit for purpose. This dual role appears to be unique to the EU. In its announcement of her new role, the EC acknowledged ‘Europe is lacking in global digital champions; so far these have come mainly from the U.S. and Asia.’\(^ {22}\) Europe sees AI and digitization of industry (focused on traditional sectors and SMEs) as its opportunity to level the playing field with other markets.


In February 2020, the EC released its “European Strategy for Data”. The commentary included in the report highlights Europe’s ‘fear of missing out’ (FOMO) as well as its concerns about reliance on non-EU entities for areas such as cloud provision, which it sees as creating external data threats, and concerns about Chinese laws related to national intelligence and cybersecurity.

The EU has the potential to be successful in the data-agile economy. It has the technology, the know-how, and a highly skilled workforce. However, competitors such as China and the U.S. are already innovating quickly and projecting their concepts of data access and use across the globe. In the U.S., the organization of the data space is left to the private sector, with considerable concentration effects. China has a combination of government surveillance with a strong control of Big Tech companies over massive amounts of data without sufficient safeguards for individuals.23

Europe’s approach continues to focus on individual rights but its longer-term vision is the creation of a single European “data space”, encouraging data sharing and driving data-driven innovation. It has identified nine data spaces across areas such as Industrial, Health, Mobility, Financial and Agriculture. Europe faces the problem of lack of availability of data. The data advantage sits with a small number of large cash-rich tech platforms (which are mostly not European). In order to realize their vision, the EU appears to be looking to address it through additional legislation (which we dig into in the section on ‘Intersection between Data Privacy and Competition’).

The EU should create an attractive policy environment so that, by 2030, the EU's share of the data economy — data stored, processed, and put to valuable use in Europe — at least corresponds to its economic weight, not by fiat but by choice. The aim is to create a single European data space — a genuine single market for data, open to data from across the world…..It should be a space where EU law can be enforced effectively, and where all data-driven products and services comply with the relevant norms of the EU’s single market. To this end, the EU should combine fit-for-purpose legislation and governance to ensure availability of data, with investments in standards, tools, and infrastructures as well as competences for handling data. This favorable context, promoting incentives and choice, will lead to more data being stored and processed in the EU.24

It certainly feels like the societal and cultural backdrop in Europe (consumer privacy as a fundamental human right) creates bigger hurdles to achieving the vision of a single data space in Europe than the societal-related hurdles that may be faced in China, for example, and potentially creates a moral dilemma. There are questions over how Europe achieves its aim of data spaces incorporating personal data, while protecting privacy, limiting the power of the Big Tech names, and creating a path for the creation of national champions. The risk is Europe attempts to (over) regulate its way to AI power relative to other regions.

The Challenge of AI Ethics

Data privacy is core to the development of ethical AI. Ethical concerns relate to challenges such as the amplification of embedded bias within datasets, the risk that it creates wider societal divides and inequalities, increases potential for misuse, and there is tension between data privacy and the vast volumes of data required for AI to work.

Figure 25. Definitions of AI Ethical Challenges

A recent survey by Deloitte of 1,400 U.S. executives found ethical issues are one of the top three risks of AI (32% respondents).

A number of examples already exist of the ethical implications the use of AI could have.

The recent withdrawal by Amazon, Microsoft, and IBM of the use of their facial recognition technologies by the police in the U.S. is a very interesting live debate around AI Ethics. These announcements followed the recent killing in the U.S. of George Floyd by a police officer. Until this point Amazon had been selling its facial recognition technology to U.S. police departments. There had been questions previously raised over the justification of its use by police as well as whether it exacerbates biases and racial profiling due to issues over misidentification.26 It took this event to prompt companies to take action and increase pressure for legislation on the use of facial recognition technology.

Despite this, AI Ethics is not being considered often enough. A recent Global AI survey suggests where companies are embedding AI into their business functions, less than half are putting in place means to mitigate the risks. The proportion is much lower in relation to ethical factors.

---

Data privacy is one of the elements of AI Ethics. There are clear tensions between what’s needed for AI to succeed and what privacy regulation is trying to achieve in terms of increased transparency and data minimization.

I think there’s going to be some tension between, just in the European context, the approach taken with the GDPR, which does tend to be a maximalist approach to privacy on the one hand, and on the other hand, what gets called AI policy.

I think it’s going to be really interesting when the AI policy debate inevitably merges with the GDPR, as a lot of the key industries affected by European AI policy are really European champions, whether it’s in financial services or insurance or automotive or pharmaceuticals and life sciences. These are all markets where you have very substantial European competitive positions, unlike the Internet platforms, and where there’s obviously a real stated interest from European policymakers to maintain and extend that economic frame. In simple terms, if you’re a European policy maker, making policy about German auto or insurance companies or any of the European based pharma companies, all of whom are making major steps in the AI arena, could be a very different dynamic than the GDPR dynamic.

### Figure 27. Differences between GDPR and AI Requirements

<table>
<thead>
<tr>
<th>Scope</th>
<th>GDPR Requirements</th>
<th>AI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>Data minimization</td>
<td>AI needs large volume of data for research, analysis, operation, training, and to avoid bias</td>
</tr>
<tr>
<td>Purpose</td>
<td>Purpose limitation (to purpose original intended)</td>
<td>Uses of data for new and unforeseen purposes beyond the original scope</td>
</tr>
<tr>
<td>Legal Basis</td>
<td>Legal bases for processing</td>
<td>Insufficient/limited variety of legal bases may undermine the potential of AI</td>
</tr>
<tr>
<td>Data Retention</td>
<td>Data retention (for period needed for the original purpose)</td>
<td>Need to retain data for AI training, deployment, and oversight</td>
</tr>
<tr>
<td>Transparency</td>
<td>Transparency</td>
<td>Operates in black box and may produce low, no explainability, and unanticipated outcomes</td>
</tr>
<tr>
<td>Rights</td>
<td>Individual rights e.g. right to erasure</td>
<td>Cannot always facilitate access, correction or explanation of the logic</td>
</tr>
<tr>
<td>Automation</td>
<td>Rules on automated decision making</td>
<td>Human input slows down the result and can generate less accurate results</td>
</tr>
</tbody>
</table>

Source: Vivienne Artz

Exploring some of these a bit further:

AI is about identifying patterns and this requires data at scale and a process to understand correlations in data. Under GDPR personal data should only be used for the purpose originally intended. One solution could be pseudonymization and/or anonymization of data.

Article 22 of GDPR states that a person can object to decisions based on automated processing, including profiling, unless explicit consent has been obtained, the EU country allows automated decision making for a specific purpose, or it is used for performance of a contract e.g., job application screening, mortgages, and insurance screening.

People have a right to know how data is being used and the logic behind a decision. This makes decisions generated by the ‘black box’ algorithms used in AI challenging and means human reviews are required.

Fairness and transparency pervades GDPR guidance on data processing, which brings into question whether decisions made about individuals should be based on typical behavior of similar people based on AI, e.g., assessment on credit limits. This links back to the challenges of explainable AI and bias in datasets.

The right to erasure could disrupt AI systems, as there has to be a system that enables deletion of data. This could also make it less accurate.

Regulating AI

As noted above, there is existing regulation AI has to adhere to, such as GDPR, including unfair and deceptive trade acts and practices. We address the AI principles and approaches being adopted by key industry participants in the final section of this report, but it is clear governments and regulators need to step up to take control of the narrative. Decisions made on the regulation of AI could dictate development of future technology and the technology industry itself is calling for regulation in this area. Sundar Pichai, CEO of Google, wrote an opinion piece in the FT early this year: "Now there is no question in my mind that artificial intelligence needs to be regulated. It is too important not to. Companies such as ours cannot just build promising new technology and let market forces decide how it will be used." He also calls for large-scale regulation to avoid a patchwork of approaches across the globe.
Establishing the right balance between the role of government in regulation and that of self-regulation will be critical but is also likely to diverge between industries as well as regions. While many see the need for allies in the development of AI there is a realization that national interests and intellectual property require protection and cracks between regions are already showing. There are a number of published guidelines and principles but full blown AI regulation is yet to appear:

- **OECD AI principles (which the U.S. has aligned with).**
  
  The EC published a whitepaper on the European approach to AI in February 2020. Developing ethical and trustworthy AI is the core focus, opening a consultation on policy and a future regulatory framework for the EU to avoid fragmentation within the EU. The EU also has a ‘High Level Expert Group’ made up of 52 experts, which has identified requirements for trustworthy AI and policy recommendations (and has driven the view on creation of data spaces), advocating for a risk-based and principles approach (with a focus on ‘high-risk’ AI applications such as in healthcare, transport, energy and public sector). The whitepaper sees international cooperation on AI as essential but believes it has to be based on protecting fundamental rights including ‘human dignity, pluralism, inclusion, non-discrimination, and protection of privacy and personal data and it will strive to export its values across the world.’

- **The G7 established an international study group for AI (launched by France and Canada in 2018) based on principles of human rights, inclusion, diversity, innovation, and economic growth. The U.S. finally signed up in May 2020. The suggestion is the U.S. joined as it believes there needs to be a counter to AI developments from Chinese companies with the White House CTO stating: ‘Although the U.S. and our allies have diverse approaches to AI policy and regulation, the democratic principles that unite Global Partnership for AI (GPAI) nations matter far more than our differences. This is especially true as authoritarian regimes abuse AI and try to bend existing international organizations to serve their interests. Chinese technology companies are attempting to shape international standards around facial recognition and surveillance at the United Nations International Telecommunication Union, while the Chinese government’s misuse of such technologies has been well documented.’**

---

28 https://oecd.ai/ai-principles
30 https://www.wsj.com/articles/artificial-intelligence-can-serve-democracy-11590618319?shareToken=st31e9696eb9954d70a66eb5cbe6740a92
In January 2020 the U.S. proposed ten binding principles to help agencies manage AI development. This was based on an Executive Order issued in 2019.\(^{31}\) The U.S. is conscious of regulation not being too prescriptive.

China issued guidance on AI in 2017, identifying AI as the ‘new focus for international competition.’\(^{32}\) In May 2019, the Beijing AI principles were published by a coalition of academics and industry.\(^{33}\)

Conclusion:

GDPR is the high-water mark for data protection regulation and has prompted the introduction of new privacy regulation across the world, some of which incorporate similar requirements to GDPR.

The common thread running throughout new and amended regulation is increased transparency and consumer control over data, data localization, and increased fining power, which means risks around use of personal data will rise (and some practices may even come to an end). There are two challenges as it stands:

1. The technological infrastructure to manage the regulatory requirements are not where they need to be. This creates opportunities and growth for privacy-based technology. At the same time, regulation will need to try keep pace with the evolving new technologies (such as AI) that are evolving and creating privacy related conundrums;

---


2. The lack of a global approach to data protection creates complexity for those operating in multiple markets. How global corporates incorporate data protection regulation may have broader repercussions for competitiveness within markets versus local operators. Some business models have been built on utilizing data (and personal data) — online advertising springs to mind — which makes adapting to a regulated world more challenging.

Both the California Consumer Privacy Act (CCPA) and the GDPR were given legitimacy by data-related scandals (Cambridge Analytica, Snowden). These are illustrations of how unexpected events in the real world can trigger a long lasting change in the direction of regulation. These are unlikely to be the last of this kind.

Following the rapid introduction of the CCPA in the U.S. (with signs of other states looking to follow suit) which significantly tightens compliance requirements and raises fining levels for corporates, there are industry-wide calls for a data protection law at the federal level. We see it as highly likely, although not imminent, and there may well be many political hurdles to clear.

More broadly, over the longer term, we have seen industries where onerous regulation has led to lower margins (i.e., telecoms and banks). We believe we are still at the start of the data privacy regulatory evolution. We might see an increase in bilateral agreements between countries in relation to data but the increase of data localization and the varying degrees of acceptance of state control over, and access to, personal data suggests complete global harmonization of policy is highly unlikely (at least for now). Political, cultural, and societal expectations have a huge role to play in why and how regulation evolves — Europe’s history means that personal freedom (free of state surveillance) underpins the GDPR while in some markets the data protection laws have carve outs for the state (e.g., India’s draft bill). Other markets, like the U.S., probably sit somewhere in the middle.

Corporates and governments have realized the potential power of data and technology. The corporates appear to have a head start; the vast bulk of consumer data across the world sits with a handful of cash-rich companies. It has become a political football and the free flow of data is becoming increasingly challenging as nationalistic stances are on the rise. Cultural and societal differences means it may not be a level playing field here — the approach to COVID-19 contact tracing technology has perfectly highlighted this. In China, it is very transparent that national interests come before those of individuals, which means technology developments benefiting national interests can occur at speed.

Technology evolution will not always develop in harmony with data protection considerations, which could create a moral dilemma for some and slow down the pace of innovation. In its aim for Europe to command its fair share of the data economy by 2030, driven by single data spaces as the core vision, the EC has to balance its unwavering position on protecting privacy while attempting to limit the power of the Big Tech names and create a path for the creation of national champions.

We see increasing evidence of technology and data-related power struggles, with the battle heating up in the field of AI — corporates vs. governments and regions vs. regions. As one of our experts noted, AI is a great tool as much as it is a weapon and international cooperation is paramount. Emerging AI platforms raise new ethical issues and corporates are not always taking these considerations into account. Principles have been developed but regulation will come and, much like we have seen in the area of data protection, there could well be divergent paths.
The EC is focused on developing trustworthy AI which means ironing out the conflicts with data protection requirements. The U.S. joined the G7 AI group recently but likely driven by political motivations to counter AI developments out of China.

Regulators need to be proactive with AI rule-making rather than reactive otherwise we risk an unfortunate misstep leading to overregulation. Europe is likely to look to regulate its way towards increasing its tech presence. In the U.S., the Big Tech companies will inevitably play a role in the battle for AI superiority, which could limit the degree to which the U.S. will look to clip the wings of these companies. These challenges are less apparent for China.

“Any decisions on that front are going to carry huge political connotations and are not going to be just about technical competition. Regardless of the outcome, there is going to be a political reading, and that might create a pendulum, which swings every four years in one direction then the other, depending on who feels most aggrieved or benefited by the way the data ended up being used. It becomes a political football, it already is. You have the left-right political football. You have the "European distrust of American Big Tech" political football.”

Andres Wolberg-Stok
Why has the Regulator Focused on AdTech?

Turning back to Europe, we provide some insight into how the data protection regulator is tackling the AdTech industry, an industry which was central to why the GDPR was established. It is also a live case study in the pitfalls and benefits of attempting to unpick a well-established model in an industry where power is concentrated. For the regulator it is a little like walking a tightrope given the number of stakeholders with varying motives.

Advertising revenue underpins the Internet and it has enabled the Internet to thrive giving consumers access to content and services for ‘free’. Central to this business model is the ability to utilize consumer data, which has fueled the growth of the online advertising market. The lure of online advertising vs. traditional media channels has been driven by the increasingly sophisticated means to target and measure audiences. However, this model has created and incentivized an industry to prioritize commercial gains over data protection.

Ethan Zuckerman (Director, Center for Civic Media at MIT): I have come to believe that advertising is the original sin of the web. The fallen state of our Internet is a direct, if unintentional, consequence of choosing advertising as the default model to support online content and services. Through successive rounds of innovation and investor storytime, we’ve trained Internet users to expect that everything they say and do online will be aggregated into profiles (which they cannot review, challenge, or change) that shape both what ads and what content they see.  

The prize has been significant. In 20 years online advertising has gone from nothing to a $300 billion industry, half of the global advertising market. Share gain is set to continue.
There are broader ethical and societal considerations prompted by the evolution of online advertising. Here we highlight three that have become a source of debate:

1. It has created an ethical debate for brands and advertisers as they have funded this opaque model. The World Federation of Advertisers published its first report on data ethics in advertising (Data Ethics – The Rise of Morality in Technology35 ) on 1 June 2020, following a year’s work by the WFA Data Ethics Board (made up of 19 senior experts from global companies). It sees it as the responsibility of brands to improve trust in the use of data.

We are witnessing a data arms race. From our shopping habits to our sleeping patterns, our personal lives are becoming more intensely and more intimately tracked by technology. Despite global adoption of some of the strictest privacy laws in human history, regulation continues to lose ground to the eye-watering pace of data-driven innovation. Consequently, the gap between what we can do with data and what we should do with data is getting ever wider. The health of our digital society depends on businesses going beyond their legal obligations and living by a new code for the responsible and ethical use of data.36

Consumers also believe brands should take action with 72% of the view that companies should stop advertising with any media platform that fail to prevent the spread of fake news and false information.37

---

35 https://wfanet.org/knowledge/item/2020/06/01/WFA-launches-worlds-first-guide-on-data-ethics-for-brands
36 WFA Data Ethics – Rise of Morality in Tech (June 2020)
37 Edelman’s Trust Barometer (2020)
2. The other is the impact on the traditional news publishers. Online advertising has eaten into print ad revenue and loosened direct relations between publishers and consumers (less than 50% of traffic to publishers are direct). The complexity of the online ad industry means that even when publishers can monetize inventory on their sites, they appear to take a disproportionately small share of the ad spend at the end of the chain due to the number of intermediaries. Plum Consulting concluded that the publisher retains between 43% and 72%.

Figure 30. Share of Advertiser Investment: Programmatic Indirect (Open Exchange) Sales of Display Ads, U.K.

The impact on the industry also risks the loss of media plurality and choice, which has a knock on effect for the knowledge and information accessible to society.
We are starting to see signs that governments are stepping in to support the traditional publishing industry, for example the announcement in Australia that a mandatory code between digital platforms and media companies is established, which includes the publishers getting a cut of online platform’s advertising revenue.

3. The complex nature of the online advertising industry has led to the rise of digital ad fraud which, according to Statista, is estimated to have cost $19 billion globally in 2018 and expected to rise to $44 billion by 2022.

**Murky AdTech Industry**

AdTech is a crowded industry with over 7,000 ad tech vendors just in the U.K. As we noted in the previous ePrivacy Citi GPS reports we believe the extent of consumer tracking and the opaque nature of the AdTech industry was a major factor behind the evolution of data protection regulation in Europe. Therefore starting with the online AdTech industry post GDPR made sense for the regulator.

**The Regulator’s Concerns**

The U.K.’s data protection regulator, the Information Commissioner’s Office (ICO), launched its review of online advertising, in particular the practice of Real Time Bidding (“RTB”), in February 2019:

> “When we are looking at AdTech, in particular real time bidding, what we see is personal data being processed on a massive scale. What we see in many cases is that in the drive to personalize and target advertising there are economic incentives to get as much data about the individual and link as much data to the individual as possible, in order to generate more ad revenue. You have an industry and a model that incentivizes greater personalization. It has notions of scale and speed that concern us and also notions of transparency and personalization that concern us.”

Simon McDougall

The ICO’s research found that 42% of consumers feel they have no control over which advertisements they are shown, while only 18% felt they have full control.
Once informed about the workings of the AdTech industry, the ICO found consumers are much less accepting of display adverts being shown in return for freely using a website compared to the acceptance level before being given information on AdTech. This goes back to our original point about the privacy paradox. Consumers may prioritize convenience over privacy but consumers may not be making as informed a choice as they think they are.

**Focus on Real Time Bidding**

The focus of the regulatory review was narrowed to the RTB system because the system is opaque and complicated with many intermediaries, it lacks transparency and makes it possible for unsuccessful bidders to retain information: ‘The process — known as real time bidding — relies on the potential advertiser seeing information about you. That information can be as basic as the device you’re using to view the webpage, or where in the country you are. But it can have a more detailed picture, including the websites you’ve visited, what your perceived interests are, even what health condition you’ve been searching for information about.’


Figure 33. The RTB Market is Large and Growing

We include a simplified diagram of the RTB ecosystem (sat within it is thousands of AdTech companies) and an explanation below of the process, as explained by one of our expert interviewees who is familiar with the AdTech ecosystem.
In a June 2019 report[39] the ICO stated there are systemic problems and it honed in on two priorities: (1) Processing of special category data which require explicit consent and (2) Issues caused by relying solely on contracts for data sharing across the supply chain. RTB has also been a focus for privacy groups and alternative technology providers.

---

The ICO gave the industry six months to work on the points it raised in June 2019 and was keen to work with stakeholders on a solution. In January 2020, the ICO provided an update stating there had been some progress from the Interactive Advertising Bureau (IAB) around guidance and principles and Google around commitment to phase out third-party cookies within two years. However, it noted that many organizations had not taken action and were using ‘legitimate interest’ as a justification for processing personal data in RTB, and the ICO saw that as insufficient.

The ICO is working with the industry and maintaining active dialogue to try to coerce the industry into a proactive approach that is consistent with what the ICO is looking for and acknowledges that regulatory action may be needed but further down the line.

---

For now the pandemic has stalled the process with the ICO releasing a statement on May 7th noting the investigation has been put on hold as it does not want to ‘put undue pressure on any industry’ and will restart the working in the coming months. Even prior to this, some industry operators did not think the ICO had moved fast enough.

The ICO says it wants to move the industry and operate in a cooperative fashion. Knowing what I know about how the industry operates, it’s hard to see how this will work. I’ve spoken to Chief Marketing Officers (CMOs), who have said they won’t stop what they are doing because there’s too much return on investment (ROI) to be had in using data in a non-compliant way and they can’t afford to stop doing it unless there’s a very real danger that the cost of being compliant significantly outweighs the cost of not. There’s a lot of bigger companies not being compliant so why should they. If one of them gets a big fine then they’ll think about becoming compliant too. There are AdTech companies who have built their entire business on using personal data in a non-compliant manner and taken millions of pounds of investment. If they were thinking about trying to become compliant then it would either cost them a lot or they would go bankrupt. They will probably just carry on being non-compliant because the cost of becoming compliant is too great and they’ll just hope the ICO isn’t enforcing anyway. The ICO might believe it’s better to get the whole industry to move as enforcement action could get tied up in lawsuits.

All the while, the gap between Big Tech and others continues to widen.

The way the ICO should tackle it, in my view, because that’s what the law says, is that to use someone’s personal data in any form you need their consent, at least within Europe right now. For Google and Facebook to serve targeted advertising, they need that person’s explicit consent... Their ability to serve targeted advertising while the rest of the industry doesn’t have the ability, that disparity needs to go away.

What Action the Industry Is Taking & Implications

A combination of pressure from the regulator and proactive changes within the industry will lead to a change in the AdTech ecosystem.

I think it’s a combination of consumer desire for things to change, more awareness and better visibility of what’s going on, plus the technology partners like Google, Apple and the rest of them, actually making changes. Then the regulatory bit coming over the top of that. It definitely feels like there’s a change in the narrative this time as opposed to a few years ago when we were talking about cookies in a much broader context.

The ICO is comfortable that the tech companies drive the change.

I think Google and Chrome has made far more difference than we have and I’m quite happy with that. We are not saying we’re driving this market. As a bit of sledgehammer to tackle some aspects of this issue, that’s going to make a huge difference.
If the solution is going to be led by the industry then online platforms will have the ultimate say over what that looks like. According to Statcounter, Chrome has 64% share of the browser market. According to the Competition Markets Authority (CMA) in the U.K., Chrome has significant market share across several parts of the AdTech chain.

Figure 35. Global Browser Market Share (Dec 2019) Browsers vs corporate

![Chrome 64%](image)

Source: Statcounter

Figure 36. Google’s Role in Advertising Technology Chain - CMA

![Google's Role in Advertising Technology Chain - CMA](image)

Source: CMA

The decisions the online platforms take could have significant repercussions for the industry, while acting to protect (and potentially increase) their own position within the online ad market.

I can definitely sympathize with the argument that this [data protection regulations] seem to make Google’s position even more powerful from that perspective [regulations didn’t break their dominance], because there isn’t a place for others to play if Google are essentially holding all the data in a pen, albeit in a collaborative fashion, but ultimately they’re the ones who are determining where that [data] then goes. Then you can see why there are people in that industry who are nervous that all this [data protection regulation] is doing is accentuating that control over it.

Martin Ashplant
This outcome could still be better for the industry than if the regulator dictated a more aggressive and abrupt change.

Prior to the ICO review the industry had started to move away from reliance on third-party cookies as these are seen as the worst offending trackers following consumers across sites. Prior to the ICO review, Safari and Firefox had already taken the decision to disallow the use of third-party cookies. This was mildly disruptive to the AdTech industry (pricing for ads targeting Safari users fell 60% over the following couple of years) but was manageable due to the relatively low market share of Safari and Firefox.

Industry Initiatives

Google announced in January they would phase out third-party cookies by 2022 and announced a product called the Privacy Sandbox (focused on privacy by default) as the alternative to third-party cookies, requesting industry feedback. According to the CMA report on Online Digital Platforms Google will phase out third-party cookies dependent on the other solutions being successful (which is based on decentralized processing on device). The CMA notes that the changes being proposed could increase the gatekeeper status of the browser and device operators for AdTech and both the CMA and ICO are working together to consider the risks and challenges that could arise.

There are concerns within the AdTech industry that Google's internal advertising teams will get access to more data than it provides to the industry\(^4\) giving Google will have even more control over the ecosystem.

---

\(^4\) [https://digiday.com/marketing/wtf-googles-privacy-sandbox/]
At the same time, the only practical means for the industry to adapt is if Google leads as the online advertising industry is so heavily reliant on them.

Apple announced in June 2020 a tightening of consent requirements for its Identifier for Advertisers (which is the unique ID on an iPhone used to track app usage and often used for targeted advertising). With the roll out of iOS14 in autumn 2020, app developers will need to obtain consent from iOS device users for third parties to access user data (such as AdTech companies and advertisers).

The IAB Europe published a guide in May 2020 to the Post Third-Party Cookie Era and identified the changes which will make certain practices challenging such as retargeting, frequency capping (limiting the number of times an ad is seen to someone), cross device tracking with identity linkages, last or multi touch attribution, and audience targeting using third-party data.

There are also alliances being formed such as The Ozone Project in the U.K. which is a programmatic advertising platform created by a consortium of publishers which provides a single access point for audiences across a number of sites based on publisher first-party data and context. There are a number of industry initiatives including the Internet Advertising Bureau and W3C ’Project Rearc’ also looking at the evolution of online advertising.

**Industry Implications**

For valued publishers and strong brands, the outlook may not be as challenging as they increase the focus on first-party data (logged in users) and, in the case of publishers, subscriber revenue.

For the long tail of publishers with no established, regular user base, the tech platforms are likely to be the main source of traffic and advertising the main source of revenue. A world without third-party cookies could be painful as the value of their traffic falls. Google said it believes the removal of third-party cookies could reduce publishers’ ad revenue by 52%.
For brands/advertisers, the ability to target audiences could get harder. Unless consumers are convinced of the value exchange, getting them to part with their data may be challenging. This is increasingly relevant in a world where brands are more focused on direct-to-consumer relationships; to execute on this first-party data access will be essential.

We have started to see announcements from large publishers about how models are being adapted. The New York Times announced in May 2020 it will no longer use third-party data for targeted advertising. From July it will offer 45 first-party audience segments, and will increase that number over time. Allison Murphy, SVP of Ad Innovation, said “This can only work because we have 6 million subscribers and millions more registered users that we can identify and because we have a breadth of content. While a differentiator — and I’m thrilled about it — this isn’t a path available for every publisher, especially not local who don’t have the scale of resources for building from scratch.”

For consumers, it means we could see more paywalls and less freely available content and services. For some who can’t afford numerous subscriptions, this could lead to a reduction in choice. On the other hand, it may mean a reduction in funding to some of the less-desirable, long-tail publishers squeezes them out of the market, helping reduce fake news sites.

The business model most under threat is the retargeting model. That whole side is severely under threat but there’s a lot of pivots those companies can make. They can go to the context space. The desire for an advertiser to put an ad on a page next to relevant content won’t go away. If IDs are removed and retargeting is removed then it hurts the ability of the vendors to spread the money across the long tail. If an ad agency ran a report of where they delivered ads for their clients, they’d get a list that’s over one hundred thousand long. You’d only recognize a handful of the sites and only a fraction of the budget would go on them. If you can’t retarget in that same way, then what brands will say is that they want to be safe and serve it on titles they know. Agencies will be forced to deliver the ad spend to publishers the brand knows. The spend will get concentrated with the large publishers. The downside is the funding of that small blog site, which has acquired some traffic, will almost entirely disappear. It could be good for journalism. Fake news sites get a lot of advertising because of programmatic advertising, which puts the audience first and context second.”

Money may stay within the digital ad ecosystem but flow to different parts of it, for example connected TV and video games.
We could see an increase in contextual advertising as well as the use of machine learning models based on larger cohort groups.

It feels unlikely we will do a complete reversal and go back to contextual advertising. Although if publishers can build up strong first-party data, matched with contextual advertising and show it works, I think there will be opportunities there. But I remain unconvinced we will suddenly shift right back as a result of regulation. I think it will be more likely we will see this evolve over time and the industry will come up with different ways of trying to achieve the same thing.

There will certainly be less reliance on third parties (and more scrutiny of third parties) as the focus on first-party data increases. This could lead to a more transparent AdTech chain and less leakage, squeezing the vendors in the middle.

The feeling amongst some is the Big Tech companies will be the main beneficiaries of any change due to the data advantage they have over the rest of the industry.

The online advertising industry is now worth almost $300 billion (half of the global ad market). We have a lot to thank the industry for; if it wasn’t for the advertising-funded model, consumer access to ‘free’ content and services online would not have developed as it has. However, an unintended consequence is that it has encouraged the gathering, amassing, and concentration of consumer data and the development of a complex and opaque AdTech ecosystem. Consumers’ understanding of the data used within the online advertising ecosystem is far more optimistic than in reality and consumers feel they lack control over their data. The GDPR was put together with the online advertising industry in mind, so it is not surprising the regulators (the ICO in the U.K.) have started by reviewing this industry (in particular, RTB) due to concerns of non-compliance.
A year into its review, the ICO clearly has concerns about the industry but is keen for the industry to develop a solution. Attempts to alter or unpick the very model driving the success of online advertising, in order to better protect consumer privacy, will be easier said than done unless it is tackled from within. Big online platforms have significant influence due to their position as gatekeepers, which means they will have the ultimate say on how the AdTech industry will evolve. This also gives them the ability to limit the pain inflicted upon their own business. Google’s announcement of the Privacy Sandbox (an end to third-party cookie tracking within two years) and Apple’s tightening of consent on requirements for app data access may be enough to hold the regulator off from taking a more prescriptive approach. At the same time, these changes could dramatically change the AdTech landscape.

- It could increase the transparency of the AdTech chain and lead to a shake out of vendors in the middle with business models entirely reliant on third-party cookie tracking. This could be positive for publishers if it reduces the current level of revenue leakage in the value chain but it could put downwards pressure on advertising pricing.

- Publishers will need to become more creative with diversifying their revenue streams. Brands will need to rethink how they target audiences and customers. First-party data will become more important for the large, well established publishers and brands and those with a loyal following. Publisher paywalls are likely to increase.

- It will be challenging for the long tail publishers lacking a loyal user base but, on a positive note, it could help to reduce funding to fake news sites.

- It could lead to a reallocation of spend to contextual advertising and, ironically, social media.

What seems clear at the moment is the direction will be driven by the large tech platforms and everyone else will have to fall into line, thereby increasing the already existing data advantage with the main platforms. As the EC is pushing for more interoperability and data access, this could have the opposite effect.

Other ethical and societal concerns driven by the online advertising industry could also start to be tackled:

- For brands (that have funded the online advertising model), data ethics considerations are now coming to the fore as there is a realization that data protection regulation alone is not enough. Brands also have a responsibility to use data in a responsible way in order to regain the trust of consumers. Better oversight of the third parties they work with is a good starting point.

- Online advertising gain has come at the expense of traditional media, eating into advertising revenues and altering direct relationships with readers. The complexity of the online ad ecosystem means a publisher’s online presence has not been enough to offset the offline pressures as it has led to significant leakage of spend to intermediaries before any it reaches the publisher. The sustained pressure this has placed on the traditional publishing industry over a number of years now has implications for media plurality, choice and distribution of information. We are yet to see if any changes to the online AdTech landscape will alleviate or exacerbate the publishing industry pressures.
Section 4:
The Intersection between Data Privacy and Competition

This leads us to the interplay between data protection and competition authorities. The remit of a data protection regulator is to ensure compliance with a data protection regulation. The remit of an antitrust regulator is to ensure a fair, open, and competitive market. Data is increasingly the lifeblood of the modern world and the interest of the two regulators is now converging around data. The changing data protection regulation landscape has clearly had implications for the online ecosystem and, as outlined in previous chapters, has unintentionally reinforced the strength of the big platforms as the control of consumer data is concentrated in hands of only a few. Regulators cannot operate in isolation and have to take a more holistic approach.

The Competition and Markets Authority (CMA) in the U.K. acknowledged this in its recent review of online platforms and digital advertising:

*Privacy concerns and the application of GDPR are likely to have a significant impact on the market, reinforcing the trend towards vertical integration and potentially increasing the data advantage of the large platforms that have their own sources of first-party user data, making it harder for third parties to compete. These trends could exacerbate the existing competition concerns.*

Industry experts see the intersection between data protection and competition being one that comes to the fore. One of the outcomes of the UK CMA’s review of online platforms is that it will work with the ICO on the interaction between data protection, consumer control and competition to ensure that consumers have more control of their data than platforms. The question is if a coordinated approach will alter the landscape.

In this section we look at the key tenets of the current competition law, implications of some pivotal decisions taken by the competition authorities, the potential outcome of the numerous reviews of the digital platforms, and what the most likely future actions are.
Competition Law to Address Big Tech

The concept of limiting harm (consumer harm) tends to be at the heart of antitrust regulation. In the U.S., based on the Chicago School, the idea that markets self-correct is core; the general thesis is monopolistic positions tend to lead to price increases which means a focus on pricing is central to the approach. Monopolies in themselves are not seen as illegal unless they have achieved the status through anti-competitive behavior. In relation to mergers, the focus is also on the creation of monopolies and the risk of eliminating competition (horizontal mergers), which means the approach is potentially narrow and ignores the network effect and ability for platforms to leverage their position and data in one market to create an advantage in another (thereby building their ecosystems). Government and antitrust concerns about Big Tech are not new but, with hindsight, the lens through which they were looking may not have been the right one. Or, it may well be that competition law (in its current form) is not the best means to address data asymmetry.

Last year we saw the first attempt by a competition regulator to combine data protection restrictions with competition concerns. The German competition regulator (Bundeskartellamt) announced in February 2019 it was prohibiting Facebook from bundling and merging data from across its services and third-party sites without user consent. The regulator argued that combining the data enabled it to build unique profiles and gain market power (with no clear social network alternatives). In cooperation with the German DPA (data protection authority), it concluded the terms of service and uses of data by Facebook were a violation of data protection rules and, from a competition law perspective, deemed the terms of service to constitute exploitative abuse. The President of the Bundeskartellamt said:

*Today data are a decisive factor in competition. In the case of Facebook they are the essential factor for establishing the company's dominant position. On the one hand there is a service provided to users free of charge. On the other hand, the attractiveness and value of the advertising spaces increase with the amount and detail of user data. It is therefore precisely in the area of data collection and data use where Facebook, as a dominant company, must comply with the rules and laws applicable in Germany and Europe.*

Facebook appealed the ruling in the Higher Regional Court of Dusseldorf and won the appeal in August 2019. The Court did not think there was a clear causal link between Facebook’s abuse of its position due to its alleged non-compliance with GDPR. The court also stated it is not an essential service therefore consumers are making a rational choice when they use Facebook, aware the service is free due to the use of personal data to finance it with advertising. The court also observed the data Facebook gathers could be available to other parties and as the competition authority didn’t show that Facebook’s data processing is disproportionate due to its market position, it is not clear that it causes consumer harm.

---

42 https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemeldungen/2019/07_02_2019_Facebook.html
This ruling is now being appealed so it is not yet over, but whatever case is made by competition authorities on the basis of data suggests it needs to be watertight with clear evidence of consumer harm (which is definitely more tricky to prove with zero priced services) and market abuse.

**Pivotal Antitrust Decisions**

Before we delve into the recent reviews of tech platforms that have taken place, it is worth looking at some of the pivotal historical antitrust related decisions playing a significant role in how the landscape looks today:

- The FTC launched an investigation into whether Microsoft was creating a monopoly in the early 1990s, which was subsequently closed. The Department of Justice (and several states) then filed antitrust charges against Microsoft in 1998 on the basis of it bundling programs into its operating system and giving away browser software for free, as well as making it difficult for consumers to install competitor software on Windows operated computers. The judge ruled Microsoft violated the Sherman Antitrust Act and ordered it to break up into two (operating system in one entity and the software business in the other). Microsoft appealed the decision and won, thereby avoiding a break up and instead agreed to a settlement and some behavioral remedies, such as limiting the tying/bundling of software. Ironically, the restrictions on Microsoft paved the way for Google’s rise.

- The clearance of the 2007 Google/Doubleclick merger as search advertising was viewed as a separate market to display advertising.

- Facebook’s $1 billion acquisition of Instagram was cleared in 2012 by the FTC. In the U.K., the Office of Fair Trading did not refer the deal to the Competition Commission as it took Instagram’s growth as a sign the barriers to entry in social networks were low and it was not deemed as a threat to the display advertising landscape (at that point Instagram was not monetizing the platform).

- Facebook’s acquisition of WhatsApp (for $19 billion) in 2014, a privacy focused, user paid messaging app, was cleared by the FTC and the European Commission. The Commission identified the relevant market for WhatsApp as the communication apps for smartphones and while it acknowledged lines between communications apps and social networks were blurring, it identified important differences between the two. It saw sufficient communications alternatives for consumers and took the view that switching costs were low. It also considered data concentration in the context of online advertising and concluded the lack of advertising on WhatsApp, and an adequate number of competitors in the online ad market, meant it did not raise concerns.
It also stated: ‘Any privacy-related concerns flowing from the increased concentration of data within the control of Facebook as a result of the Transaction do not fall within the scope of the EU competition law rules but within the scope of the EU data protection rules.’

More recently, in 2013, the FTC in the U.S. decided against taking any enforcement action following its investigation into Google’s online search dominance.

**Market Share is Concentrated**

Now let’s jump forward a few years and look at market share positions of the Big Tech platforms.

Google and Facebook have around 50% share of the global online advertising market. This rises to over 55% in the U.S., with Amazon now at 7% (and increasing). Google’s share of the global search engine market on mobile devices (i.e., smartphones) is >90%. Facebook’s share of the display advertising market in the U.K. is almost 50% (including Instagram). Amazon’s share of the ecommerce market in the U.S. is around 45%, according to Statista. Consumer time online is also highly concentrated. According to the CMA, consumers spend 86% of their time online on 1,000 properties and just over a third of time on sites owned by Google and Facebook.

---

**Figure 37. Market Share: U.K.**

**Figure 38. Global Online Advertising Market Share**

Source: tatcounter, Comscore, eMarketer and Company Reports

Source: eMarketer

---

https://ec.europa.eu/competition/mergers/cases/decisions/m7217_20141003_20310_3962132_EN.pdf
The technology world has changed significantly in 20 years while the approach to antitrust hasn’t. There is not one Western technology super power but a small concentrated number of owners of the digital ecosystem we all engage with. Each dominate a different sphere of the online world. The major advantage each has is access to user data that is superior to any other and a network effect that is nigh on impossible to replicate from scratch. When services are ‘free’ to consumers (potentially subsidized by other profitable parts of the business) it makes for a tougher case when antitrust laws are narrowly focused. As the Stigler Committee report highlighted, it is possible that a digital market has a negative equilibrium price because of the value extracted from the consumer data for targeted advertising. In other words, consumers are paying with data which has a price/worth.

The other challenge is whether the regulators have sufficient resources and expertise. In a world where the regulator is at an informational (and resource) disadvantage to the Big Tech companies, it makes it very challenging for the antitrust investigations to really establish the mechanics of the inner workings of the digital markets. There is a risk that antitrust enforcement may be relatively ineffective (the tech companies can swallow large fines and tweaks around the edges to how they operate) as the pace of change is too rapid for regulators to keep up.

**Scrutiny from Regulators on the Increase**

There are several investigations on Big Tech that have been launched in the U.S. over the past year or so:

- In February 2019, the U.S. Federal Trade Commission launched a task force to monitor technology markets. The Technology Task Force will examine industry practices in technology markets, conduct law enforcement investigations, and review completed mergers in technology markets.

- The Department of Justice launched an investigation in June 2019 into Google’s position in the AdTech market, in particular its search practices and could reportedly bring a case in the summer 2020.

---

44 https://www.axios.com/tech-summer-antitrust-24cfc68a-41f8-9b91-5be6d8269202.html
The House Judiciary’s investigation into competition in the digital markets is expected to complete soon and could lead to legislation. As part of the review, it has held hearings on the impact of digital platforms on the news industry, handling of user privacy, impact on smaller businesses and innovation and entrepreneurship. This includes claims that Amazon accesses data from third-party marketplace sellers to inform decisions about its private label brands, both Amazon and Google use ‘predatory pricing’ for the sale of voice-controlled speakers and charging unreasonable rents due to market positions (Apple’s 30% app store charge to app makers) etc.45

A number of State AGs are conducting a multi-state investigation into Google and Facebook.

The FTC announced in February 2020 that it is investigating acquisitions made by Amazon, Apple, Facebook, Microsoft and Google since 2010 that had not been reported to the antitrust authorities due to their size.

In June 2020 the EC opened an antitrust investigation into Apple’s rules for app developers on distribution of apps via the App Store and its requirement that they use Apple’s own in-app purchase system (which comes with a 30% commission rate on subscription fees), restricting use of alternative cheaper options. It is particularly focused on areas where Apple also competes with other apps, such as music and books.

The Wall Street Journal reported the EC is set to file antitrust charges against Amazon relating to its use of third-party seller data on its marketplace to compete against them.

There are also signs of regulation forming that will specifically target Big Tech platforms, moving beyond competition law. As part of its European Data Strategy vision (announced in February 2019) the EC is planning two Acts. One is a Data Act (in 2021) that will encourage and incentivize data sharing and it may enhance consumer control over data portability (under GDPR) with interfaces for real-time access in machine readable formats. The other it is planning is a Digital Services Act (in 4Q20). Part of this will focus on the need to level the playing field in the European digital markets and ensure there is choice, innovation, and competition, which could entail additional ex ante rules ‘for all platforms of a certain scale’ in areas such as self-preferencing, non-personal data access obligations, and personal data portability or interoperability. The other part of the Act relates to clearer rules around online intermediaries (in relation to content). Acknowledging the regulatory framework needs modernizing, the EC launched a consultation in June 2020 for the Digital Services Act.46 At the same time it also launched a consultation on a possible new competition tool to deal with ‘structural competition problems’ (preventing markets tipping) not addressed by current competition rules. This tool would enable the Commission to impose behavioral and structural remedies as opposed to fines.47

Providing us with some additional insight, two of the European Parliament Committees have put forward recommendations to the Commission on the Digital Services Act:

- **Committee on Internal Market and Consumer Protection:** It believes ex ante regulatory remedies imposed on the large platforms could open up the market to new entrants and increase customer choice and innovation, beyond what competition law can achieve. It cites practices that should be prohibited include self preferencing, discrimination in intermediary services, making use of data for making market entry difficult for third parties, and practices aimed at locking in consumers. The burden of proof to justify behavior would be with the platform. High levels of interoperability should be imposed so tools, data, and expertise are shared to limit consumer lock in. Technologies should be explored to enable this. It also believes a central regulatory authority should be established to oversee compliance with the Digital Services Act.48

- **Committee on Legal Affairs:** It recommends the establishment of a European Agency to monitor and enforce compliance on content management and moderation and for content hosting platforms to publish decisions on removing user-generated content. It does not believe platforms should automate control of content. It also believes targeted advertising should face stricter regulation, limitations on data collection, and users should have a choice to consent to targeted advertising based on their interactions and have more control over which content is curated for them. It is also supportive of data portability and interoperability as a means to address market power imbalances.49

### Findings of Competition Related Reviews

During the course of 2019 there were several market studies conducted on the power of online platforms, in particular in the area of digital advertising. These were conducted by competition authorities, expert panels, and consultants (sometimes at the request of governments). We focus on five that were conducted in the U.S., Australia, and the U.K. The overarching takeaway is that power is in the hands of few, which means they control what data is shared, how it is used, what content is displayed, and how regulation is interpreted. These could form the basis of changes to the landscape, not necessarily reliant on competition law to achieve them. We include below a summary of each and identify the themes consistently appearing across the reviews:

- **Winner takes all or most:** Digital platforms benefit from strong network effects, economies of scale, marginal costs close to zero, high and increasing returns from use of data, low distribution costs, and switching limitations. They act as gateways to consumers.

- **Unequal Access to Data:** GDPR is identified as having raised the barriers to entry as data is the single biggest barrier. Access to data leaves the current large tech platforms with an advantage (an ACCC review reported Google has trackers on 70% of sites and Facebook on more than 20%) and leaves them well-positioned for the coming artificial intelligence (AI) and machine learning (ML) technology wave. Lack of transparency (no independent verification of inventory, changes to algorithms, opaque open display ad market) and lack of interoperability were also highlighted, as was the ability to leverage positions into other related services.

---

- **Lack of Competition Means Consumers Give Up More Data:** Consumers are paying with data. The burden has been placed on consumers (click and access notices) which means companies can manipulate choice. In a more competitive market, data protection may have been higher and consumers could give up less (or even get paid for data).

- **Consumers Stick with Default:** Few alternatives means consumers stick with the default option. Limited incentive to move due to loss of data and history, lack of interoperability, bundling of services/hardware, inertia, lack of financial incentive. One study highlighted Google is paying $12 billion a year to Apple to be the default search engine on the iPhone (which is interesting given Apple positions itself as more privacy focused than the other large tech companies).

- **Power over their Market:** Strong bargaining power over the long tail and ability to block competition, e.g., Facebook reportedly banned cryptocurrency ads a year before announcing Libra. The network effects and scale of the platforms means they can ‘de-emphasize’ smaller competitors on their platforms while staying on the right side of the antitrust laws, e.g., copy innovations, contracts that require single-homing.

- **Stifles Investment and Innovation:** Lack of competition reduces the incentive to innovate. Venture capital is reluctant to invest in businesses directly competing with the large platforms. Creation of ‘kill zones’ whereby the platforms acquire to protect them from the risk of rising competition and use acquisitions to distort innovation. There has been a long run effect from mergers with none of the Big Tech acquisitions blocked in the past 10 years (many under the size threshold).

- **Media Plurality:** Lower revenue for publishers impacts investment in news. Digital platforms control news distribution and curate/personalize content shown to maximize consumer time on the platform which can lead to fragmentation of society. No legal liability for online platforms (S230 of the Communication Decency Act).

- **Emerging Tech:** Increased use of smart speakers amplifies some risks such as increased data collection (impacting privacy) and a further reduction in consumer choice of news (risk of echo chambers).

- **Lobbying and Information Asymmetry:** The U.S. tech platforms are amongst the biggest spenders on lobbying. Regulators are at an informational disadvantage. Engagement with users means they can influence choices and views. Risks of online harm such as fake news and political manipulation.

- **Conflicts of Interest:** Vertical integration creates conflicts of interest, e.g., operating across digital advertising chain.

The head of the CMA, speaking at a conference in March 2020, said the digital revolution has led to ‘amplification of fake news, the erosion of the tax base, the huge increase in the risk of fraud and identity theft, illegal content” and “thriving ‘dark markets’ for drugs and weapons’ and said ‘the unprecedented challenge of digital will indeed necessitate a strengthening of both competition and consumer protection powers.”
**Figure 41. Summary of Digital Market Reviews**

### Competition & Markets Authority (U.K)  
**Purpose:**  
Market study into online platforms and digital advertising to look at: market power of digital platforms; lack of consumer control over data; competition in supply of digital advertising; Furman proposals.

<table>
<thead>
<tr>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Over a third of U.K. users’ Internet time is spent on Google and Facebook sites. Google has &gt;90% share of the search advertising market and Facebook almost 50% share of the display advertising market.</td>
</tr>
<tr>
<td>Costs of digital advertising ($13bn in U.K. in 2018) are reflected in prices of goods and services.</td>
</tr>
<tr>
<td>Lower share for publishers could impact investment in news and online content.</td>
</tr>
<tr>
<td>Lack of competition could mean people have less control over their data.</td>
</tr>
<tr>
<td>Consumers may receive inadequate compensation for their data.</td>
</tr>
<tr>
<td>Other online harm: Fake news, political manipulation.</td>
</tr>
<tr>
<td>Search and social media have significant network effects and economies of scale: Lack of scale in click and query data would limit ability to compete with Google search.</td>
</tr>
<tr>
<td>Consumer decision making and power of defaults: Default choice of platform and default settings influence the ability of the platforms to collect user data. Google paid $1.3bn to be the default search engine on Safari which inhibits competition.</td>
</tr>
<tr>
<td>Unequal access to user data: Google claims U.K. publishers earn 50-65% less revenue if advertising isn’t personalized. Google and Facebook have data from across their ecosystems as well as tags on other sites. Interpretation of GDPR has reinforced this advantage, entrenching vertically integrated platforms.</td>
</tr>
<tr>
<td>Lack of transparency and asymmetric information: Google and Facebook provide just under 40% of traffic to large publishers and changes to algorithms can have a significant impact on traffic levels. Neither party reportedly allows independent verification of their own inventory, which risks overstatement and increased prices.</td>
</tr>
<tr>
<td>Vertical integration and conflicts of interest: This can arise from a strong position across the digital advertising chain.</td>
</tr>
<tr>
<td>Evidence of harm: Auctions to set price means there is less influence over price but platforms can influence factors that drive price, e.g., reserve prices, number of ads per search query. Profitability of the platforms far exceeds the cost of capital. Lack of transparency in the open display market can lead to excessive/hidden fees to intermediaries. Google and Facebook can also leverage their existing positions into other related services, making it difficult for competition.</td>
</tr>
</tbody>
</table>

### Chicago Booth - Stigler Committee on Digital Platforms  
**Purpose:**  
George J Stigler Center formed an independent committee on digital platforms to assess the impact on economy, privacy, news media, and functioning of democracy.

<table>
<thead>
<tr>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winner takes all due to: (1) Strong network effects; (2) Economies of scale; (3) Marginal costs close to zero; (4) High and increasing returns to the use of data; and (5) Low distribution costs.</td>
</tr>
<tr>
<td>Perception is that it’s free but consumers pay with data: Market power can mean lower data protection (e.g., Facebook became more aggressive with data collection once MySpace disappeared), less investment, and lower quality.</td>
</tr>
<tr>
<td>Power over advertising: Can charge higher prices and block competition, e.g., Facebook banned cryptocurrency ads a year before announcing Libra.</td>
</tr>
<tr>
<td>Consumers stick with default options: Harm is greatest when there are few alternatives and a lot is known about the potential customers, e.g., Google pays $12bn a year to Apple to be the default search engine on iPhone.</td>
</tr>
<tr>
<td>High margins are a sign of barriers to entry: Venture capital is reluctant to fund business directly competing with largest digital platforms (know as ‘kill zones’).</td>
</tr>
<tr>
<td>News media: Media plurality is essential but digital platforms control news distribution. Editing is aimed to maximize consumer time on platform and news is personalized which fragments society. In the U.S. there is no legal liability due to Section 230 of the Communications Decency Act.</td>
</tr>
<tr>
<td>Privacy: Burden placed on consumers with click and access notices which means companies can manipulate choice. Lack of data protection.</td>
</tr>
<tr>
<td>Dark patterns: User interfaces that make it difficult for users to express actual preferences or manipulate users into taking actions not aligned with preferences. This can increase acceptance rates of a data protection plan by 228%. Combined with limited alternatives means there’s limited consumer backlash.</td>
</tr>
<tr>
<td>No accountability in relation to political governance: Digital platforms spend a lot on lobbying (in the U.S. Google, Amazon, and Facebook were 2nd, 6th and 9th biggest spenders on lobbying in 2018). They are media outlets (with immunity due to Section 230). Information asymmetry means they can ‘sponge’ regulation. Connectivity with consumers means they can engage users to challenge anything that disadvantages them. Seen as national champions (clear in the Trump tariff tit-for-tat).</td>
</tr>
<tr>
<td>Platforms are opaque and deny access to data.</td>
</tr>
</tbody>
</table>
Digital Competition Expert Panel (Furman Review) - Unlocking Digital Competition

Commissioned by U.K. Government

March 2019

Purpose:
Review to make recommendations on changes to competition and pro-competition policy to unlock opportunities of the digital economy.

Findings
- Acknowledges benefits of digital economy but it's winner takes most. More competition may have meant consumers give up less in terms of privacy or even get paid for data.
- Many reasons for concentration: economies of scale, network effects, limitations to switching, access to finance. Incumbent access to data seen as single biggest barrier to entry. This diminishes the incentive to innovate.
- Consumers don’t move due to loss of data, history, lack of interoperability, services bundled with hardware, inertia, and lack of financial incentive to switch. Zero price could be too high given the value extracted from consumer data.
- Current dominant tech companies are well positioned for the next tech wave being AI and ML due to their access to data. Regulators at an informational disadvantage.
- Creation of a 'kill zone' so dominant tech companies acquire innovation to protect from competition and use acquisition to distort or kill innovation.
- Long-run effects from mergers: Over the past 10 years the five largest tech firms have made over 500 acquisitions globally and none have been blocked.
- GDPR created higher barriers for new tech companies. Misuse of consumer data and harm to privacy is an indicator of low quality caused by lack of competition. A means to cement market power.
- Unfair terms: Platforms have strong bargaining power over the long tail which could increase fees. Findings that publishers receive <70% of digital ad revenue. Commission investigation into Amazon's user of data from sellers on site to boost its retail sales.

Australian Competition Commission (ACCC) Review of the Digital Market

June 2019

Purpose:
Review of impact of digital platforms on consumers, business, news media.

Findings
- Google and Facebook act as gateways to reaching Australian consumers. Opaque nature means it’s hard to determine standard of behavior.
- Profound impact on media markets: News and journalism is crucial to hold government and decision makers to account. Where digital platforms perform comparable functions to media they should be regulated in the same way.
- Where Google and Facebook’s clients are also competitors, the playing field may not be level. Online advertising is opaque and hard to ascertain true value of the inventory.
- Consumer and privacy law as important as competition law: T&Cs of several large digital platforms requires user grants broad license to store, display, and use uploaded content.
- Breadth and depth of user data is a strong advantage: Google has trackers on 70% of sites and Facebook on more than 20%. 88% of apps on Google Play send data back to Google and 43% to Facebook.
- Emerging tech: Risks resulting from widespread use of smart speakers include: (1) potential foreclosure of news media businesses either through refusal to supply their content or charging exorbitant fees for supplying their content on the device; (2) reduced consumer choice of news supplier; (3) where devices supply news based on consumer preferences or previous searches, this could create an ‘echo chamber’ or ‘filter bubble’ effect; and (4) increased threats to consumer privacy due to increased data collection.

European Commission - Competition Policy for the Digital Era

Comissioned by European Commissioner for Competition - Margarethe Vestager

March 2019

Authors: Jacques Crémer, Yves-Alexandre de Montjoye, Heike Schweitzer

Purpose:
Explore how competition policy should evolve to continue to promote pro-consumer innovation in the digital age.

Findings
- Extreme returns to scale: Digital production is lower so scale is a competitive advantage.
- Network externalities: Network effects can prevent a superior platform gaining traction. The incumbent has an advantage.
- Role of data: Ability to use data is a competitive advantage which will increase. It’s the key to AI and online services.
- Most favored nation or best price clauses. Multi-homing and switching. Platforms act as regulators which can lead to self-preferencing.
- Acquisitions are often under the radar due to size (don’t meet turnover thresholds).

Source: Citi Research
**Free Market Solution or Intervention?**

Facebook’s response to the CMA’s interim report notes that holding data is not a competitive advantage by itself, as consumers often share the same data with other platforms and the cost of distribution is near zero so restrictions on a consumer’s ability to multi-home is time available to them. It also notes there is a large marketplace for data accessible at a relatively low cost. Facebook believes its success (and other platforms’) has been due to innovation and differentiation rather than access to data (as this was not the case upon launch).

The political implications of increasing the regulatory burden on the ‘national tech champions’ is highly likely to be a consideration in the U.S. and an argument which will be used by the industry. In an interview with Recode 2018, when Facebook’s CEO was asked about the calls to break up the Big Tech companies, his response was:

> “I think that the alternative, frankly, is going to be the Chinese companies. If we adopt a stance which is that, “Okay, we’re gonna, as a country, decide that we wanna clip the wings of these companies and make it so that it’s harder for them to operate in different places, where they have to be smaller,” then there are plenty of other companies out that are willing and able to take the place of the work that we’re doing.”

It is not that Big Tech companies are opposed to any form of regulation, in fact, they have actually called for regulation and more specific guidelines in certain areas, or at least they realize the regulatory environment will tighten and want to have a seat at the table in formulating how it will look.

Facebook’s CEO Mark Zuckerberg published an op-ed in the FT in February 2020 called ‘Big Tech Needs More Regulation’ and cited four areas where regulation is needed: elections, harmful content, privacy, and data portability. There are plenty of examples where the Big Tech platforms use tightened privacy rules as a hurdle to the industry achieving openness with data. He states ‘Without clear rules on portability, strict privacy laws encourage companies to lock down data, refusing to share with others, to minimize regulatory risks.’ A subtle way of tying the EC in knots with regards to its approach to data privacy and its drive for data access and openness.

From a competitive standpoint, both Google and Facebook responded to the CMA report arguing the competitor set is broad and well-funded in digital advertising and provided examples of digital business models where multiple operators exist, e.g., ride hailing, online travel agencies, and price comparison.

To be fair, there is evidence of competition opening up in some areas where the Big Tech platforms operate, e.g., Facebook’s entry into ecommerce to link up with SMEs (Amazon’s domain), AI-driven video app TikTok’s entry into online video/social media (Facebook has singled out the rapid growth of TikTok), and Amazon’s entry into online advertising and search (Google stated in its response that 51% of product searches are conducted on Amazon in the U.K.).

---


In 2020 TikTok has seen unique web visitors globally rise (helped by COVID-19 related lockdowns), and the app hit 1 billion monthly active users in just three years, faster than other social platforms.

In the U.S., app downloads have exceeded Facebook this year and daily time per active user on the app has surpassed all of the other social media platforms, and is closing the gap with YouTube.

The rising competitive threats tend to come from one of the other big platforms.
The smart speaker/home hub market are a good example of this. There is a race to own the home as this positions Big Tech as the gatekeeper or gateway between suppliers and consumers in the home. It’s no coincidence, therefore, the market share of home hub devices are the same Big Tech names the regulators have been so focused on (according to eMarketer, over 70% of smart speaker homes in the U.S. have the lead smart speaker). If we are moving to a world where we conduct searches by voice using in-home devices, then the days of lists with links could be long gone. This shift alone could upend the advertising market but it could amplify regulatory concerns around market power, data usage, privacy etc. Data protection regulation will still dictate how the operators use the data, but that alone won’t change the fact that personal data is in the hands of few.

The smart speaker market is estimated to be worth $30 billion by 2024 according to Global Market Insights. The U.S. had the highest penetration of all markets, followed by the Nordic markets and the U.K. not far behind. Penetration in China was only about 10% in 2019.
Figure 48. Summary of Solutions & Interventions: Based on the Recommendations from the Digital Market Reviews

**Commissioned by U.K. CMA**

**Competition & Markets Authority (U.K)**

**July 2020**

**Purpose:**
Market study into online platforms and digital advertising to look at: market power of digital platforms; lack of consumer control over data; competition in supply of digital advertising; Furman proposals.

**Potential Interventions/Solutions**

- **Government legislation** to introduce new regulatory regime for digital platforms for both a code of conduct and pro-competitive interventions. Establish a Digital Markets Unit (DMU) to enforce it.
- **Behavioral rules for firms with market power**: An enforceable code of conduct may help address a number of concerns. The code should be principles-based rather than prescriptive, focused on fair trading, open choices, and trust and transparency. They would apply to platforms funded by digital advertising with strategic market status (SMS).
- **Pro-competitive interventions**: (1) Increase consumer control over data (especially in area of personalized advertising); (2) Mandate interoperability; (3) Mandate third-party access to data where it is valuable; (4) Mandate data separation and data silos (within businesses). The DMU should have the power to introduce consumer choice and default interventions and power to introduce separation intervention (operational or full ownership).
- **Potential interventions to address sources of market power and promote competition**.
- **Search**: (1) Google provides click and query data to rival search engines; (2) Restriction on the ability of Google to enter into arrangements to be the default search engine on devices and browsers so consumers are prompted with a choice (beyond Android).

**Chicago Booth - Stigler Committee on Digital Platforms**

**September 2019**

**Purpose:**
George J Stigler Center formed an independent committee on digital platforms to assess the impact on economy, privacy, news media, and functioning of democracy.

**Potential Interventions/Solutions**

- **Interoperability**: Lack of network externalities. Interoperability of sites (like U.S. forced with phone companies) should be considered. Open Banking and Brazilian Good Payers Credit Act are examples of this.
- **Increase antitrust hurdles**: Many mergers have passed without scrutiny. Criteria needs to change to force the digital platforms to show how it won't harm competition.
- **Reduce power of data**: Open up access to data. FTC should have power to access internal databases and perform research on platforms. To protect privacy anonymization techniques should be applied such as differential privacy and secure multiparty computation.
- **Single regulator for digital platforms**: Concerns are all linked to the power of data. The U.K. is considering a Digital Markets Unit. In the U.S. it could be a sub-division of the FTC. The focus of the unit would be on ‘bottleneck power’ and create regulation to address it, e.g., merger review authority even over small transactions, oversee mandate for interoperability.
- **Political disclosure**: Digital platforms should increase disclosure such as which content is being demoted, the politicians being supported and funding to academia.
- **Privacy**: Notice and consent does not protect privacy and top down regulation can be rigid. Could adopt ‘consumertarian default rules’ based on preference of majority of U.S. consumers based on scientific studies. Stringent constraints on waiving default to avoid dark patterns. Top down regulation should limit data collected and shared.
- **S230 Communications Decency Act**: This means internet companies are not liable for content as they don’t edit it. They do monetize content and are responsible for the algorithms so should be subject to the same rules as media. Additionally, media plurality should be considered as a key merger metric.
Potential Interventions/Solutions

- Digital markets will only work well if supported with strong pro-competition policies that counter forces leading to a single winner. Reliance on merger and antitrust enforcement won’t do this. Need a Digital Markets Unit with new powers.
- Merger assessments: CMA should consider harm to innovation and potential competition. Strategic market status digital companies should notify CMA of planned mergers. Update merger guidelines to address under enforcement and improve effectiveness.
- Antitrust: Quicker use of interim measures and sufficient information gathering powers. Prioritize consumer enforcement work. Retrospective evaluation of historical cases not brought and assess how market evolved since.
- Government should encourage cross-border cooperation between competition authorities to develop a common approach.

- New Digital Markets Unit should: (1) Establish a code of conduct; (2) Promote personal mobility and systems with open standard (like Open Banking); (3) Use data openness as a tool to promote competition (access to others); (4) Cooperate with stakeholders and powers to impose solutions, investigate and penalize non-compliance; (5) Focus on companies with market power and avoid burden on smaller companies; and (6) Sufficiently resourced and with expertise.
- Data mobility to move data between businesses at request of the consumer or user of a single app to manage a profile, e.g., friends networks across platforms. Data Transfer Project is a collaboration to create a common framework. Data openness with anonymized or non personal data.

- Potential market failure, require data handed over from the platforms and report issues of concern below the threshold of the competition regulator.
- Cross border coordination is essential to address competition and consumer concerns.
- Competition law may not be effective due to the lack of transparency of the market, reliance on data evidence (that may not be available) and slow pace of investigation.
- A separate unit focused on the digital market should be set up to investigate potential market failure, require data handed over from the platforms and report issues of concern below the threshold of the competition regulator.
- Creation of harmonized media regulatory framework that is platform neutral.
- Advance notice of acquisitions and changes to merger law.
- Strength privacy regulation: Advises adopting some of the GDPR rules such as definition of personal information, strengthens notice and consent requirements, enable erasure of personal information, introduce direct rights of action for consumers and higher penalties for breaches.

- Merger assessments: CMA should consider harm to innovation and potential competition. Strategic market status digital companies should notify CMA of planned mergers. Update merger guidelines to address under enforcement and improve effectiveness.
- Antitrust: Quicker use of interim measures and sufficient information gathering powers. Prioritize consumer enforcement work. Retrospective evaluation of historical cases not brought and assess how market evolved since.
- Government should encourage cross-border cooperation between competition authorities to develop a common approach.

- Data pooling and sharing is pro-competitive and already exists in banking and insurance. There are risks if competitors are denied access, it discourages competition from differentiating and improving its own data collection, and it includes competitively sensitive information. Regulation may be needed.
- Some states have introduced transaction value-based thresholds for assessing M&A. Need to consider the ability for competition to enter if large platforms are acquiring small businesses with high market potential. Need to include horizontal elements into conglomerate theories of harm.

- Current competition law framework is insufficient but needs refining: (1) Clear consumer welfare gains should be required for strategies used by dominant platforms (even if consumer harms can’t be measured); (2) Less emphasis on market definition and more emphasis on harm and identification of anti-competitive strategies; (3) Take into account consumer bias towards default option and assess access to data not available to competition; (4) Place burden of proof on the incumbent for showing pro-competitive conduct; and (5) Competition law and regulation can reinforce each other.
- GDPR allows data portability of personal data which can overcome pronounced lock in effects. Data interoperability could address market concentration. It can reduce multi-homing and barriers to switching. It may require sector specific regulatory regime or (under Article 102 TFEU) could be confined to dominant firms. Where competitors request access to data from a dominant firm, analysis is needed to ascertain if access is indispensable. Legitimate interests of both parties to be considered.

- Some states have introduced transaction value-based thresholds for assessing M&A. Need to consider the ability for competition to enter if large platforms are acquiring small businesses with high market potential. Need to include horizontal elements into conglomerate theories of harm.

Source: Citi Research
Time to Break Up?

The most extreme approach would be to break up the platforms (which would likely get caught up in court for a number of years). This outcome may not lead to a more competitive and fair market, which is acknowledged by most of the reviews/regulators, and a decision would likely need to be unilateral.

The Australian Competition and Consumer Commission (ACCC) does not see a break up or divestiture of subsidiaries as a solution as it believes a break up could mean less investment and lower quality of service and economies of scale, which could impact on the consumer. It also may not reduce market power as there are other barriers to entry which insulate the platforms (brand, economies of scale).

A breakup is not a route that the EC’s Competition Commissioner believes in either as it might not resolve the challenge that the regulator is looking to tackle:

‘Well, first from a very plain, practical point of view, how do you do that? How to make sure that you get companies split up? Second, you know the story about the hydra. If you cut off one head, I can’t remember if it’s two or five or seven, but at least you have more of what you wanted to get rid of. It’s not a given that this is an idea that would reintroduce competition, because we are talking about markets with very strong network effects and with marginal costs that approach zero.’

Update to Competition Laws

The Stigler Committee report believes new competition legislation doesn’t need to move away from the idea of bad conduct and harm, but may need to change its tolerance of antitrust for risk of over-enforcement and under-enforcement, reducing high proof requirements imposed on antitrust plaintiffs, and placing a more rigorous burden of proof on defendants. There should be greater intervention around how platforms deal with actual or potential competitors and modification to the idea of predatory pricing (the idea dominant platforms can price below cost).

The review for the EC also suggested a new competition law isn’t needed but may require changes to specific aspects. The EC is consulting on a new competition tool and what that should look like. The review last year concluded the current competition law framework is sufficient but needs refining as follows: (1) clear consumer welfare gains should be required for strategies used by dominant platforms (even if consumer harms can’t be measured); (2) less emphasis on market definition and more emphasis on harm and identification of anti-competitive strategies; (3) taking into account consumer bias towards default option and assess access to data not available to competition; (4) place burden of proof on the incumbent for showing pro-competitive conduct and (5) competition law and regulation can reinforce each other.

Restrictions and Higher Hurdles on M&A?

Acquisitions by the digital platforms are typically on a smaller scale. These deals are often under the radar of antitrust agencies and, even when they aren’t, it is difficult to hypothesize about the path of the target if a deal didn’t occur. Many deals have enabled vertical integration in the industry, creating conflicts of interest (note CMA’s comments on Google/DoubleClick).

Scrutiny of mergers (even smaller scale) could increase, especially for large platforms where acquisitions could lead to further concentration of data, potentially increasing the advantage over competition. Increased scrutiny could come via a tightening of the merger competition laws or the creation of a framework specifically focused on the large platforms.
The EC review notes some markets have brought in value-based thresholds to assess M&A and consideration of the ability for competition to enter markets where large platforms are acquiring small businesses with high market potential.

We are already seeing evidence of increased scrutiny of M&A. For example, the FTC is reviewing acquisitions by the Big Tech companies back to 2010 that were under the radar due to size. The CEO of the CMA in the U.K. described the merger of Facebook and Instagram and Google and DoubleClick as examples of ‘merger control gone wrong’ and said it is even more vital to protect competition in concentrated digital markets, arguing that structural remedies or prohibition are more likely solutions than behavioral remedies.

As we noted earlier, there is also a significant political focus across markets to cultivate national tech champions and to avoid them landing in the hands of foreign entities. COVID-19 has triggered concern that foreign investors will take advantage of recent volatility and depressed valuations of many tech companies. The EC issued guidance to EU member states on protecting assets from foreign investment during the pandemic. There are several recent instances of more restrictive criteria for foreign acquisitions, especially in the areas of technology innovation, and security. On March 31, 2020 Spain introduced a regulation stating any acquisition of a stake of 10% or more in a Spanish company by a non EU/EEA based company in strategic sectors (which includes critical technologies such as AI, cybersecurity etc., and those with access to personal data and media) requires prior approval from the government. In April 2020 Germany’s government announced plans to tighten the rules on transactions with implications for national security as well as screening of stakes of 10% or more in the area of AI, robotics, semi-conductors, biotech, and quantum technology. The U.K. is also reportedly planning tougher laws for foreign takeover of firms which could increase national security risks. In June 2020 the French Finance Ministry announced an initial €150 million fund to invest in local tech start-ups if they are approached by unsolicited foreign investors. The funding could increase to €500 million from 2021.

**Structural Separation**

The move away from focusing on market definition, as proposed by the report for the EC, would be one means of forcing regulators to look more holistically at the range of operations of a large technology platform. For example, should the leading ecommerce player be allowed to operate the marketplace as well as compete on the marketplace? Should a smartphone manufacturer be able to provide the app store as well as offer apps (which are often pre-loaded onto the hardware)? In fact, both of these are now subject of antitrust reviews.

The CMA, for example, suggests a possible solution in the open display market is a range of separation remedies such as separation of the ad server from other the Demand side platform (DSP) where there is market power.

---

Separation or limitations could come in many ways; not just asset separation, but also limitations on internal data sharing. We have seen the competition regulator in Germany (unsuccessfully) attempt to limit data sharing across the Facebook assets, which could be viewed as some level of separation.

The U.K. CMA has proposed that the new Digital Markets Unit (which is proposed to act as regulator and enforcer of tech) should have the power to enforce both ownership and operational separation and that data separation should be implemented within businesses where it impacts competition.

**New (Tech Specific) Regulation/Code of Conduct**

A general consensus from the reviews of digital markets and competition is there should be additional regulation of online platforms, especially those deemed to have a significant market position, as competition rules can’t necessarily deal with some of the main concerns raised and it could take many years going down that route. This could mean a new piece of regulation or legislation or some kind of enforceable code of conduct to promote competition.

The CMA, Furman Review, and ACCC reviews concluded a principles-based code of conduct should be introduced for platforms with ‘strategic market status.’ The idea of a code of conduct could cover various areas from ensuring fair trading, open choice and transparency to governing relationships with news media around transparency and data and revenue sharing.
Alongside the proposal for an enforceable code of conduct, the CMA proposes a range of pro-competitive interventions: (1) increased consumer control over data (option not to share data for personalized advertising); (2) mandating interoperability (especially in social media); (3) mandating third-party access to valuable data (in search, for example, this relates to Google providing click and query data to third-party search engines) and; (4) Mandating data separation and silos where it impacts competition. The CMA has recommended the government legislates to introduce a pro-competitive regulatory regime for platforms.

Both regulation and codes of conduct would have an ex ante approach to try to prevent unintended outcomes and, in theory, could also mean action could be taken more swiftly than it can under competition law.

There is already evidence of this occurring beyond the U.K. As noted earlier, the EC is planning to implement new legislation in the form of the Digital Services Act, which will include rules to address market imbalances.

In Australia, following the ACCC’s review, the treasurer instructed the ACCC to develop a mandatory code of conduct for digital platforms to negotiate how to pay news media for their content, to provide more transparency on algorithm changes, and to share data with the media companies. More details are expected by the end of July 2020. In the meantime, Google and Facebook are pushing back, arguing the platforms generate much less revenue from news related queries than is claimed and there is a net benefit to the news outlets.53

In Germany, there is draft legislation proposing that the national competition authority can designate platforms as dominant and prohibit them from a range of conduct.

In Australia, following the ACCC’s review, the treasurer instructed the ACCC to develop a mandatory code of conduct for digital platforms to negotiate how to pay news media for their content, to provide more transparency on algorithm changes, and to share data with the media companies. More details are expected by the end of July 2020. In the meantime, Google and Facebook are pushing back, arguing the platforms generate much less revenue from news related queries than is claimed and there is a net benefit to the news outlets.53

In Germany, there is draft legislation proposing that the national competition authority can designate platforms as dominant and prohibit them from a range of conduct.

Japan released an interim report in June 2020 that proposed ex ante regulation applied to digital platforms.

We could also see changes to existing regulation. In the U.S. currently there are debates about S230 of the Communications Decency Act,54 which excludes the platforms from being liable for the content uploaded as they are not the editors of the content (like the media outlets). The Stigler Center Committee argued in its report the platforms monetize content and curate content due to their algorithms so should be subject to the same rules as media. How this is handled is also a matter for debate in Europe as part of the Digital Services Act. The Committee for Legal Affairs believes platforms should not automate content and there should be more transparency and scrutiny of how content moderation and selection occurs on the platforms. Facebook also published a white paper on the topic.55

---

53 https://australia.googleblog.com/
**Tech-Specific Regulatory Unit**

Leading on from the idea of establishing a code of conduct, most recent reviews of the digital market make the argument for establishment of an expert agency/body to regulate and oversee the platforms, either as a standalone group or as part of an existing agency. The idea being it would be able to act faster than the competition regulator, promote competition, focus on those with market power, and address potential market failure.

In December 2019, the Australian government announced it would invest $26.9 million over four years into a new unit (under the ACCC) to monitor, report on and act on competition and consumer protection concerns related to digital platform markets. It also requested an investigation of the online advertising market.

The U.K. Government has put together a Digital Taskforce under the CMA, comprising experts from the CMA, Office of Communications, and the ICO (data protection regulator) to advise the government on how to implement a pro-competitive regime (this runs until September 2020). The Taskforce will help the government to design the Digital Markets Unit that could act as the regulator for digital platforms and, according to the CMA, will be able to act more quickly and work to improve the functioning of the digital market, which is not possible when approaching it from a competition law standpoint only.

Cross-border cooperation is also likely to become increasingly important (although tougher to achieve) as markets narrow their focus on the approach they want to take. The CMA believes that ex ante regulation through an agency such as the OECD could be a sensible approach.

**Data Portability, Interoperability, and Opening up Data Access**

Each of the studies advocates the idea of opening up access to data to other operators and/or regulators. The CMA, for example, suggests Google provides click and query data to competitors.

Most reviews concluded data portability and interoperability are also means to level the playing field. The consumers right to data portability was introduced as part of the GDPR, so this is really where the intersection of antitrust and data protection could complement each other. But the GDPR lacks details on how it would work practically given lack of a standard technology, security risks etc. If these ideas are to be further developed, additional regulation or guidance would be needed. The European Commission review believes the encouragement of data portability by the GDPR is beneficial as it increases consumer control over data and helps avoid data driven lock ins, although it argues for stricter rules for dominant platforms. This means users could transfer data from one service to another, e.g., music playlists. Most of the reviews refer to Open Banking as a framework that could be adopted for data portability across other sectors.

Data mobility or data interoperability goes further and isn’t covered by the GDPR. There are different types of interoperability:

- **Protocol interoperability**: Ability of two services to interconnect with each other.
- **Data interoperability**: Like data portability but enables continuous access to user data.
**Full protocol interoperability:** Standards allowing substitute services to interoperate, e.g., messaging systems. The CMA highlights the potential for Facebook to interoperate with other social networks (so users can upload posts simultaneously and view posts across platforms).

The challenges with the latter are a need for standardization across competing platforms which could limit innovation and even lead to collusion.

Any system that enables interoperability would also need to take data privacy and security into account. We include more details on data mobility and privacy tech initiatives in the next section.

The CMA also identifies data mobility and privacy-enhancing technology as potential solutions but notes neither have taken off at scale yet. One of the challenges for data mobility is incentivizing consumers to participate. The ACCC was the only report not to recommend data portability at this stage because it sees the short term impact as limited due to lack of competing platforms to port the data to as well as lack of financial incentive for consumers to move platforms (unlike in banking services).

There have already been some industry initiatives in response to the GDPR data portability requirement. Several Big Tech companies are involved in the Data Transfer Project to build an open source framework that can connect online services. The industry has also raised concerns over the concept of data portability, such as privacy risks related to sharing data (due to the ease at which data users can be re-identified).

Facebook published a white paper on the topic in 2019 and argues in it that data portability raises challenges and questions about who is responsible for protecting the information when it moves between services and the tension between increased privacy versus increased data sharing would need to be navigated carefully.56

**Conclusion:**

The Big Tech platforms each dominate particular areas of the online world but they are also much broader ecosystems with a significant data (and cash) advantage over others, which gives them a powerful starting position when expanding into new areas. There are signs of them competing more with each other but limited signs of real threats from other operators. The antitrust market definitions are narrow, focused on pricing and consumer harm, ignoring the network effect and data asymmetry these platforms benefit from. Antitrust regulation, which is supposed to ensure a fair, transparent, and competitive market, has contributed to the landscape evolving as it has (due to some pivotal competition related decisions).

The race is already on to win the smart speaker/home hub markets, potentially the next means to control the gateway to consumers in the home.

Ironically, the GDPR was seen has having reinforced the dominance of the large tech platforms, prompting more scrutiny from governments and competition regulators. Data is the battleground for the data protection regulators and the competition authorities now. While the two come from different standpoints, the desired outcome, i.e., a more transparent and level playing field, is much more likely if a more coordinated approach is taken, leading to a more holistic view. The data protection authorities dictate how data is used but not where the control sits. Additional regulation or increased antitrust enforcement will be needed to tackle the concentration of power.

Momentum appears to be building with several investigations in the works. Remedies as opposed to fines are needed if the authorities really want to level the playing field. Competition fines have been headline grabbing but are easily dealt with by the cash-rich Big Tech platforms.

Our conclusion from the numerous reviews conducted over the past year or so of the market power of the large platforms is that the regulatory environment will get tougher. The platforms are clearly aware of this and are calling for regulation and clarity in certain areas but presumably would like a collaborative approach when any new regulation is formulated. We include a continuum below with the most disruptive move being a break up and the least disruptive being status quo or allowing a free market solution to organically emerge. We mark those in red that are highly unlikely, those in amber that are possible and those in green that we see as highly likely (if not already occurring).

We see both ends of the continuum as highly unlikely. We could see calls for structural remedies, i.e., separation of certain business operations/units within a corporation, as it has become clear that vertical integration has created conflicts of interest and potentially encourages anticompetitive behavior. The U.S. is likely to focus on areas that comport to anti-trust laws but not dismantle the national champions.

We could see refinement of competition laws, especially when it comes to M&A as there is little appetite to allow the Big Tech firms to get bigger. Scrutiny here will increase from both a competition and political standpoint; and deals by the large platforms may be harder to get through (regardless of size).

Moving beyond reliance on competition law (which often entails years of legal wrangling) to achieve change is looking likely. Platforms could face a pincer movement from competition law being refined on one side to more specific regulation to address the Big Tech platforms on the other side. The EC is planning a Digital Services Act, which could include specific ex ante rules for the large platforms as well as new competition tools. This is likely to include requirements to increase open access to (non-personal) data (fitting with the EC’s idea of a single data space) and implement data portability and/or interoperability. In the next section we look at the data portability/mobility initiatives that are being developed in the market.
All changes come with the risk of unintended consequences: structural separation could lead to less efficiency; increasing data openness could lead to a lack of differentiation and increased security risks; making platforms liable for the content posted means they are potentially playing god on what is right and wrong. Data portability and mobility face challenges related to technology harmonization, security, and incentivizing consumers to move beyond the existing walled gardens of the large platforms.

More global coordination and agreement will increase the likelihood of the changes being more effective, given the platforms operate cross-border, but political considerations will also play a role. The EC is likely to come down hardest as it looks to create a nurturing environment for local tech businesses. The U.S. is likely to remain conscious of the Big Tech platforms also being their ‘national champions’ and it may well be a case of ‘better the devil you know’.

The position of Big Tech platforms is not likely to change quickly as their position as gatekeepers is well established with consumers and dramatic changes would have repercussions for the U.S. as a tech super power. If the governments and regulators want to address the concentration in the digital landscape then they need to take more control of the narrative. This is likely to mean a step up in coordinated regulatory action, as weak enforcement (from both data protection and competition regulators) could have economic and political implications longer term. As we learned from the GDPR, the set-up of any new regulations and agencies needs to be well resourced and funded for enforcement to stand a chance.

The economic value of data has mostly accrued to the Big Tech platforms. If enforced effectively, an accumulation of regulatory changes could lead to a more balanced picture over time.
Section 5:
Innovation in Data Privacy and the Importance to the AI Revolution

As noted earlier, technological solutions are needed to manage the regulatory requirements; these are still in the early stages of development. As we enter a new decade in which data privacy regulation is likely to develop substantially around the world, it is important to explore the types of innovation occurring in the personal data arena. Some will support the Big Tech companies themselves whereas some will support the governments and regulatory bodies seeking to protect the consumer, but all will be important contributors to the data ecosystem that is beginning to emerge.

In this final section of the report we break these innovations and topics down into four broad categories, each of which we will explore and cite examples of companies leading in their field of expertise.

- **Data Privacy Technology**: Technologies such as privacy browsers, encryption software, and virtualization of personal information offering to improve an individual’s data privacy.

- **Data Control & Portability**: Technologies and business models exploring ways to make personal data interoperable between digital platforms and services, putting control of this interchange in the hands of the user.

- **Safe & Sound AI**: Technologies and industry groups focused on how to use huge volumes of personal data in a way that does not create inequality and unfair bias in AI implementation.

- **Open Data**: The provision and enabling of data access for all users to facilitate social and medical advancements, e.g., COVID 19 response.

Each of these categories overlap and are complex fields of expertise. We will try to explore some of the leading voices at a high level to set the stage for the coming era of innovation and digital platform development.

**Data Privacy Technology**

**Privacy Browsers**

If the Internet browser has become the source of the most personal data leakage then the first place to begin to solve for that risk is in the browser itself. Privacy browsers have emerged over the last decade as a different type of browser that does not change the experience for any single individual user profile and seeks to keep individual identity secret and secure. Privacy browsers do not store your search history and therefore do not build a profile of an individual’s interests over time, in addition to rejecting third-party cookies seeking to track usage.

The company DuckDuckGo ([www.duckduckgo.com](http://www.duckduckgo.com)) launched in 2008 and has seen steady growth over the past decade, but with around 65 million searches each day in recent data, still only represents 1% of the global search performed on Google.
Privacy browser Brave (www.brave.com) launched in 2016 and has fast become a competitor to DuckDuckGo, offering to block ads and website trackers, and in a newer version offering users compensation if they do choose to let advertisers use their profile to target advertisements. This is done using a cryptocurrency token called the ‘Basic Attention Token’ (BAT) which seeks to return ~70% of ad revenue to the end user who agrees to see the ad in the browser. Brave recently announced it has 15.4 million active monthly users, so again a small fraction of the overall browser market but growing rapidly (2.25x monthly active user growth YoY).57

57 https://brave.com/15-million/
Google’s own ‘Privacy Sandbox’ proposal will see cookies replaced by five application programming interfaces (API’s) and advertisers will use each API to receive aggregated data about issues such as conversion and attribution. The user data will be stored and processed in the browser so the data stays on the users’ device. The industry believes a universal identifier will be the outcome, and the concept of the Privacy Sandbox is still new and attracting significant debate. Whether Google can transition to a new cookie-free model without increasing its own AdTech dominance is an open question.

**Homomorphic Encryption**

When we browse, and when we record actions on websites we use, our activity is directly recorded as plain text, or ‘input data’ that is encoded and related to us directly. Homomorphic encryption is an encryption technique that transforms that plain text into a scrambled version referred to as ‘cipher text’ that represents the original text but protects the confidentiality of the digital data.

The unique characteristic of homomorphic encryption, as opposed to more traditional techniques, is that mathematical analysis can be performed on the transformed data without having to go through a decryption process, which is often the most vulnerable part of the encryption process. Data can remain confidential while being processed and analyzed.

---

58 [https://coinmarketcap.com/currencies/basic-attention-token/](https://coinmarketcap.com/currencies/basic-attention-token/)
59 [https://oko.uk/blog/privacy-sandbox](https://oko.uk/blog/privacy-sandbox)
60 [https://www.chromium.org/Home/chromium-privacy/privacy-sandbox](https://www.chromium.org/Home/chromium-privacy/privacy-sandbox)
Homomorphic encryption allows data to be encrypted and outsourced to commercial cloud environments for research and data-sharing purposes — all while protecting user data privacy. It can be used for businesses and organizations across a variety of industries including financial services, retail, information technology, and healthcare to allow people to use data without seeing its unencrypted values. For example, homomorphic encryption allows predictive analytics service providers to safely analyze medical data without putting data privacy at risk.

Many of the technology giants have launched products focused on evolving the use of homomorphic encryption to safely analyze personal data. Microsoft created SEAL (Simple Encrypted Arithmetic Library) in 2016, which is a set of open-source encryption libraries allowing computations to be performed on encrypted data.61 Their aim is “to put our library in the hands of every developer, so we can work together for more secure, private, and trustworthy computing.”

Last year Google also announced their support for homomorphic encryption by rolling out their own set of open-source libraries called ‘Private Join and Compute’, which allows for secure multi-party computing to “help organizations work together with confidential data sets while raising the bar for privacy”62.

The concept of homomorphic encryption has many applications across industries which generate and rely on processing large volumes of personal data. Banking and healthcare are obvious candidates, but as we move into a more connected digital world it will extend to almost every business model and privacy requirement.

**Synthetic Data**

Another approach to protecting personal data as it passes around the digital ecosystem is to generate artificial data replicating the statistical components of real-world data but not including the source personal identifiable information (PII). The demand for this type of data generation is being driven by the increasing needs of data science teams and machine learning algorithms that require vast amounts of sample data to develop and test models.

Synthetic data improves data privacy because it constructs the new data from scratch and does not link back to the original record or transaction. If you swipe a credit card and your purchase is recorded that transaction is replicated, with the same data elements and characteristics of the original transaction, in essence creating new people who are synthetic but represent the activity. Synthetic data is being explored by many financial organizations who need to model their customers at scale but have challenges with the handling and movement of PII around an organization.

Privacy browsers, homomorphic encryption, and synthetic data are all different technologies and approaches seeking to anonymize our digital interactions with companies we choose to engage with, and companies seeking to engage with us through advertising and targeted services. All three innovations have been developing gradually over the past decade and have seen increased momentum since the events related to Facebook and Cambridge Analytica uncovered in 2018 and then the implementation of GDPR that summer.

---

Data Control & Portability

As our world has moved into the digital age and companies and services have moved online, we are increasingly asked to connect different digital platforms using common log-on and access methods. Good examples of this are the Google and Facebook plug-ins now prevalent on many websites allowing for a seamless subscription to a new service or product. This has served to deepen our reliance on a small number of Big Tech platforms as gatekeepers to other services and products, but it does not enable us to move our data around from company to company easily.

There is going to be increased regulatory focus on the ability to move our digital identify and data around, referred to as ‘data portability’, at our request. The concept is a key tenet of both GDPR and the move to ‘open banking’ which in Europe has been incorporated into the Payment Services Directive 2 (PSD2) regulation which was enacted in September 2019.63 Open banking encapsulates the right of an individual to move their own data between financial organizations and the law requires that banks provide Application Programming Interfaces (APIs) to allow third-party financial technology firms to interact and provide services around the financial organization. Both topics have been explained in detail in previous Citi GPS reports (insert references), but some recent innovations are important to discuss in the context of the data privacy debate.

Tokenization of Personal Data

Tokenization has become an important topic in relation to data privacy and it refers to the substitution of specific personal information with a key or unique identifier, that can then be ‘exchanged’ with a recipient as a proxy for the original users identify and PII.64 If you replace the sensitive individual data with randomly generated numeric keys and then only share those keys in a secure network you can avoid moving the sensitive information outside of primary organization, but at the same time provide the seamless connection between services and different platforms. Figure 52 below illustrated how the concept of tokenization works for personal data being passed between devices.

63 https://www.openbankproject.com/psd2/
64 https://en.wikipedia.org/wiki/Tokenization_(data_security)
One of the main drivers for tokenization of identity has been the growth of the financial technology (FinTech) market since the Global Financial Crisis and the increasing demand for access to financial data at the individual level. Many web-scraping and data aggregator vendors have facilitated this growth by sourcing details from user activity on websites and the sharing of PII over unsecure networks. In response to this, the banking industry is now pursuing a number of utility solutions that will use tokenization as the core principle for exchanging data.

In February this year Fidelity announced the launch of a new FinTech company called Akoya (www.akoya.com) to address the need for a more secure network for passing personal data around the financial markets ecosystem. Akoya is a joint venture between the largest U.S. banks, Fidelity and The Clearing House, which is the largest U.S. payments network also itself owned by the major banks.

Akoya is built on a principle that consumers should have the right to access and control who sees and uses their personal data and that consumers should provide affirmative consent to a third-party using their data to provide a service. The platform uses tokenization techniques based on a secure API network built on the Financial Data Exchange (FDX) protocol and aims to become the central utility for financial data transfer. Figure 53 below shows an overview of the data access network envisioned using tokenization of identity and sharing of unique keys with participating firms.

---

66 https://www.theclearinghouse.org/
67 https://financialdataexchange.org/pages/fdx-api
Another recent development promoting better personal data control is the launch of Solid (https://solid.mit.edu/) by the original creator of the modern Internet, Professor Sir Tim Berners-Lee, and a team at MIT. Solid, which derives its name from ‘Social Linked Data’, is developing an open-source protocol that allows a user to control their data across the Internet. Solid is working on a combination of standards and protocols that developers will be able to use to build applications where the data itself can be maintained independent of the application, and with a high degree of user control. This decentralization approach promises an interesting new phase of Internet and application development, with data privacy at the core of its vision and mission.

Data portability, in its different forms and with different approaches, will be a central theme within further developments of the personal data ecosystem. It is already supported within the GDPR construct and will likely be core to any national regulatory framework in the U.S. As Simon notes, there are a number of hurdles that need to be overcome such as standardization of technology, definition of data liability as the data is passed around, trustworthy authentication, and a means to incentivize consumers to understand the value of data sharing.
Several government-backed efforts at identity portability are also underway across the world. These are also highlighted in a The Stigler Committee report and include:

- Estonia’s e-Estonia initiative to give citizens a unique digital identifier.
- India’s Aadhaar, a verifiable 12-digit identity number issued for each citizen which serves as an identifier and authenticator for a variety of offline and online services.
- Sweden’s and Norway’s Bankid, which allows companies, banks, and governmental agencies to identify and conclude agreements with individuals over the Internet.

The report also discusses ‘self-sovereign’ identity solutions that would use blockchain technology to allow individuals to own their identity credentials and control who can access their data in online services. If an individual could then port their identity and data to the platforms and providers they wish to use, this would again promote entry of new services and erode the switching costs of established platforms.

**Edge Computing & Federated Learning**

The second important innovation relating to data control is the development of ‘edge computing’ and the ability to store data and run complex processing on the physical device in the users hand. The topic of edge computing, which looks to localize storage and calculation on a device, has developed as a counterweight to the growing dominance of centralized cloud infrastructures.

The local computing and storage resources of edge computing are referred to as ‘cloudlets’, micro datacenters, and ‘fog nodes’ and are placed on the Internet at the far reaches of the network. This is typically an individual’s phone, or a device associated with the IoT (Internet of Things) like a smart speaker. By serving as the first point of contact for a user to join the network, the cloudlet can enforce the privacy policies of its owner before releasing them to the broader cloud network. This in essence localizes control of personal preferences and can protect personal data attributes at the choice of the individual.
All of the major telecom companies have initiatives relating to edge computing and many are collaborating on development of the technology. Numerous conferences have sprung up to bring industry resources together with academia as the topic of edge computing becomes critical in the interconnected world of IoT and with the rollout of 5G technology across networks. The European Data Strategy cited Gartner in its report noting that 80% of data is processed and analyzed in data centers and centralized computing facilities and 20% in smart connected objects and edge computing and this will be inverted by 2025.

One of the challenges and complicated debates relating to edge computing is whether the device ID itself just becomes the proxy for an individual, given that geolocation tracking is now so advanced, is difficult to separate the individual from his or her device. This implies that both Google (Android) and Apple (iPhone) have much to gain if the third-party cookie model dies out, edge computing protects PII, but the device ID itself becomes the proxy for the person.

Federated learning is a machine learning technique that uses data from a distributed network of edge computing devices to train algorithms on local data samples, not requiring the data to leave the host device. This approach to machine learning allows many participants to use the data elements while protecting the underlying attributes of the data and personal information, and may become another important component of the data privacy dialogue.

**Safe & Sound AI**

The third area of innovation with broad implications for the future of society relates to the development of safe and sound Machine Learning and AI algorithms, which rely so heavily on the vast amounts of data we generate as users of the digital platforms.

As discussed in Section 2, the regulatory landscape for AI is beginning to unfold in many jurisdictions as part of an innovation agenda for governments who realize the enormous potential of being leaders in this area. As this agenda develops the topic of AI Ethics, how to deploy and leverage the technology for societal good, and how to avoid bias and unintended inequality, is becoming central to this debate.

---

68 https://tmt.knect365.com/the-edge-event/
Innovation in and around AI Ethics can be explored by understanding which groups are being formed to tackle the debate, and then which technological developments are beginning proposed to provide additional products and services that make the AI industry fair and transparent. There are many collaborations and forums focused on addressing the topic and we have identified a few notable groups for discussion in this report. This list is by no means exhaustive and we welcome additional referrals to understand other initiatives.

**The Future Society**

The Future Society[^69] is a non-profit organization, founded out of the Harvard Kennedy School in 2014, which in 2015 launched the ‘AI Initiative’ to foster debate with policymakers and the private sector over the development of safe and sound AI. This initiative has been working across the public and private sector to develop a common playbook for the development of AI seeking to set standards, agree norms and best practices, and provide guidelines to businesses and governments.

The group is also exploring the development of a **Global Data Access Framework (‘GDAF’)**[^70], collaborating with the United Nations, OECD and many private sector groups, which would make data broadly available if it could support the advancement of the UN’s Sustainable Development Goals (SDGs), which aspire to address a range of global issues by 2030. The GDAF aims to democratize access to critical data needed for AI to be used to benefit society, and will likely gather momentum with the COVID-19 crisis and global responses by governments and NGO’s.

[^69]: https://thefuturesociety.org/
The Future Society were also integral in the establishment of the IEEE ‘Global Initiative on Ethics of Autonomous and Intelligent Systems’, which as the leading trade organization for the technology industry carries significant importance for software engineers in both the private and public sectors. This steering committee, founded in 2016, has sought to develop a set of general principles under the banner of ‘Ethically-Aligned Design’ (EAD), which include a key pillar on Data Agency, stating that artificial intelligence systems should empower individuals with “the ability to access and securely share their data, to maintain people’s capacity to have control over their own identity”. One of their recommendations for this pillar is that organizations explore, test, and implement technologies and policies that let individuals specify their online agent for case-by-case authorization decisions on whom can process an individual’s data and for what purpose.

Al Commons

Also formed in 2016, Al Commons is another non-profit organization seeking to foster the development of AI for the good of society. With members from many of the world’s leading universities, technology companies, and AI industry players the group is promoting a range of initiatives including a proposal to develop a Global Data Commons (GDC) to provide mechanisms for ‘rules-based access to data’ to help leverage AI to support progress against the UN SDGs.

Google AI Principles

Given their prominence as one of the leading digital platforms, it has been necessary for Google to take a proactive role in the discussion on AI Ethics, and in June 2018 they launched their own set of AI Principles. These include a pillar on the incorporation of privacy design principles, stating Google will include these principles in the development and use of their AI technologies. They cite they will give an individual “the opportunity for notice and consent, encourage architectures with privacy safeguards, and provide appropriate transparency and control over the use of data”. These statements were a pre-cursor to the announcements on the Privacy Sandbox.

Facebook AI Research

Facebook has also sought to re-establish trust following the Cambridge Analytica scandal in 2018 and has continued to invest in the Facebook AI Research (FAIR) division of the company. In early 2019 the company announced a partnership with the Technical University of Munich to support the incubation of an independent AI Ethics center. The Institute for Ethics in Artificial Intelligence (IEAI) recently launched a new Global AI Ethics Consortium (GAIEC) to accelerate the use of data and AI to assist with the global COVID-19 crisis. This group unites many of the world’s leading academic AI Labs in an effort to collaborate on the use of big data to help governments deal with the ongoing pandemic.

---

71 https://standards.ieee.org/industry-connections/ec/autonomous-systems.html
72 https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e_general_principles.pdf
73 https://ai-commons.org/about-us/
74 https://ai.google/principles/
75 https://ieai.mcts.tum.de/launch-of-the-global-ai-ethics-consortium-gaiec/
Facebook was also a founding partner of a group called the Partnership on AI in late 2016, which is a collaboration between other peers such as Google, Apple and Microsoft, and a number of academic and social justice organizations such as the American Civil Liberties Union (ACLU). The organization now has ~100 partners across the private sector, non-profit sector and academia, representing 13 countries. One of their principle tenets is to maximize the benefits of AI while managing the challenges, specifically working to protect the privacy and security of individuals.

**Apple AI Principles**

Apple CEO Tim Cook laid out their views on data privacy in a keynote address at the International Conference on Data Privacy and Privacy Commissioners (ICDDPC) in Brussels in 2018, where he described the massive data collection of online platforms as having ‘exploded into a data industrial complex’. He acknowledged in the speech that what the digital platforms do ‘is surveillance’ and proposed the U.S. government develop a comprehensive federal privacy law addressing four essential human rights.

1. **The right to have personal data minimized**: Companies obligation to de-identify customers and limit the amount of PII collected.

2. **The right to knowledge**: Users ability to always know what is being collected and what it is being collected for.

3. **The right to access**: Companies recognizing that the data belongs to the users who should be able to edit and maintain any personal records.

4. **The right to security**: A foundational principle that user data should be protected in order for users to trust the companies they entrust their data to.

As the company behind a popular smartphone, one of the leading operating systems, and the developer of many apps which are used daily by their customers, Apple is setting a high bar for all developments across their product suite. The question is whether these provisions themselves lead to the development of ethical AI when all of the data is combined to feed algorithms driving and enhancing service offerings.

Apple differs from Google and Facebook in that it does not derive advertising revenue from its user base, but it does have a unique capability to understand what users need and want because of their integrated hardware, software, and operating system design. AI is therefore developed and deployed within a closed loop system with a very broad access to user data so the development of specific AI principles has not been at the forefront of Apples’ public commentary.

The company does of course make money from the apps developed on the iOS (the operating system of the Apple products) and developers of apps pay annual fees to list their products on the App Store, which proliferates across Apple hardware. This gives them a unique ability to act as a gatekeeper to access the millions of customers they have worldwide, and a unique edge in developing their own products as they see all development activities.

---

76 https://www.partnershiponai.org/
Data Privacy within the AI Ethics debate (see Figure 54 below) is a central concept, and the dominant digital ecosystems of Facebook, Google, and Apple are very conscious of the importance and the contradiction that it creates.

Figure 54. Appearance of AI Ethics Challenges in Principle Documents

To make data protected and anonymous is to throttle back the power of AI, and to scale data for machine learning requires key features which are indelibly human and individual in nature. The private sector wants to lead the debate and be seen to act in the best interests of society, but the concentration of just a few companies risks a collective consciousness and consensus view which is narrow, and likely conflicted.

Innovation within AI Ethics will come from the application of techniques and protocols that allow for data control, portability, and transparency, and it is likely this small and elite community will agree on common approaches to addressing individual data privacy rights. Whether those protections also increase the dominance and power of these platforms will play out in the anti-trust debates outlined in Section 4.

‘Open Data’

Since the onset of the COVID-19 pandemic in early March we have seen the acceleration of a new concept, which also underpins the Ethical AI debates, which is how much of the world’s data should be ‘open’ and available to all participants. The crisis has revealed an inconsistency in how governments, health organizations, academics, and other public sector actors can access and share important individual data specific to mobility, health, and economic standing. It has also prompted many of the technology giants to initiate ‘open data’ initiatives to aid response and recovery efforts.

Microsoft in April 2020 launched the Open Data Campaign, to ‘close the looming data divide’ and help all organizations leverage data for the common good. The initiative highlights that over 50% of all data created each day feeds to only 100 companies globally, and that the U.S. and China dominate initiatives around AI.

78 https://news.microsoft.com/opendata/
The open data proposal has a number of principles, shared with the Open Data Institute, which include security and privacy for data exchange. They plan to work with governments and NGO’s to address social and environmental issues, using shared data to drive learning and AI that enhances economic growth. The creation of a collaborative lab called ‘Open.Data.Policy.Lab’\(^79\) is already beginning to share data between these constituencies and COVID-19 has highlighted the need for state and local organizations to have broad access to useful data.

Both Apple and Google have made their mobility data (i.e., their tracking of individuals) available as open data sets that can be accessed by anyone directly from their sites. This data is aggregated and abstracted so individual movements cannot be seen, but it is very granular to locations and modes of transport. Google mobility reports show global trends derived from the use of Android smartphone devices\(^80\) and use geolocation to identify destinations and therefore activities. See Figure 55 below for an example excerpt.

---

**Figure 55. Google Mobility Report**

Apple generate their mobility reports from the tracking of Apple iPhone’s and by linking to the type of transport used, derived from pattern recognition in the device movement itself. This allows them to separate out driving, walking, and public transport to study movement patterns, which are useful in understanding economic recovery and social attitudes towards resuming the use of public transport. Figure 56 below is an example from London, U.K.

---


80 [https://www.google.com/covid19/mobility/](https://www.google.com/covid19/mobility/)
It is worthwhile noting that neither Apple nor Google Mobility data sets include any date from China or North Korea, but do include data for all other countries including Russia.

Conclusion

There will likely always be a divide between publically available and private data, but it is interesting to consider a world where data was more democratized and therefore less valuable to the small number of large digital platforms. It is likely however that the COVID-19 crisis increases the dialogue around Open Data and this topic becomes part of the Data Privacy debates and regulations, which will play out in the coming decade. The idea of data portability and interoperability and sector-based data spaces are all solutions put forward by governments and regulators. There are a number of initiatives in these areas, but there are also a number of hurdles to overcome for these solutions to be scalable, which will take time.

A nirvana future state for personal data would have to include a nirvana future state for AI, and that itself is a hugely complex societal and ethical debate. Explaining the decisions machines make, based on our data that they use to make recommendations, track whereabouts, and promote services, has to also be part of the design of AI for this to be a realistic vision of the future. The large industry operators have also developed their own AI frameworks but we come back to the idea this is an area where governments need to provide more steer.
Organizations like the Alan Turing Institute have embarked on efforts to benchmark and baseline ‘explainable AI’, but this work is only in its infancy and research and the resulting framework will require many iterations.

Being able to audit the decisions made with personal data, and being able to audit the lineage of the data itself, presents enormous technological challenges for data engineers and innovators around the world. With new encryption techniques, replication and scaling techniques, and decentralized technologies providing for tokenization of identity, the requirements are achievable. They will however require cooperation and agreement between private and public sectors, both facing many pressures and are not all equal participants at the table.

The coming decade will witness the next chapter in the use and control of our data, and will experience dramatic changes as the result of an explosion of AI applications in every aspect of our lives. How we participate in this future will be fought out between a handful of companies headquartered on the west coast of the United States, a select few in Asia, and the major regulators of the EU, U.K., and the U.S. They will all want to establish new rules and strike a balance between over-regulation of innovation and fair competition, and they will all be aware of the Chinese landscape as an illustration of how surveillance capitalism can develop if allowed to be owned by only a few actors. The world may need a ‘Paris Climate Deal’ moment for Data Privacy and AI Ethics regulation but the consensus building for this to happen is a long process and will be many years in the making.

Ultimately individuals need to decide how much they are willing to learn and pushback, and how they want to hold governments accountable to govern the Internet and Big Tech in a balanced way. The next decade will be a defining one for all of us, and we are unlikely to desire less free services and less online interaction so the bargain is ours to decide on.

81 https://www.turing.ac.uk/news/project-explain
Expert Views
A Conversation with Vivienne Artz
Chief Privacy Officer, Refinitiv

Where are we post GDPR? Have you seen much change in how corporates are using data?

Yes, yes. You can see it just on websites — we continue to be inundated with cookie opt in notices. That's driven by two things. One of which is the ePrivacy Directive — it is going to get revamped to the ePrivacy Regulation. Given the years that the Regulation has been debated, no one is willing to guess if/when it will be finally agreed and then implemented. One of the key issues they've got stuck on is marrying the realities with the regulations. Each time they're, ‘oh, and it needs to work like this’, industry comes back and says, ‘Yes, but if you don't allow us to screen without consent, how do we protect children from slavery or exploitation? If you don’t allow us to screen without consent, then how do you prevent malware viruses?’ So, how do you craft wording that creates an exception for legitimate purposes, including law enforcement, but still keeps a reign on cookies and the like? This is proving to be a long and challenging journey for the draft Regulation.

If the Regulation was limited to specific relevant sectors, then it would be easier to agree and implement, but one of the challenges is the aim that the Regulation is horizontal in approach. As a result, it is taking a huge amount of time and effort to work through the intended and unintended consequences, whether legal or practical, and to address them.

Another example of the challenges posed by cross sector legislation is the CCPA in California. That's driving a lot of the consent-based approach as well. You will recall that the CCPA it came into effect in June 2018, so a month after GDPR. The California Legislature hastily wrote it as an alternative to the stricter privacy legislation instigated by Alastair Mactaggart. The California Consumer Protection Act, is what it says on the tin. It's about enabling California consumers to know what personal data is being collected, with whom it is sold or shared, and the option to object to the sale of their data or request its deletion. It is interesting that it originates in California, where breach notification started and where the tech sector is based. It’s quite clear the Californian public are finding their voice and wanting to exercise more control over their data, and are challenging the current perceived commoditisation of their data via various means including social platforms.

The CCPA came into effect in January 2020, and it has created real challenges for existing or more traditional business models which sit outside of the tech sector targets. The carve-out for lawful purposes, the interests of national security, public policy reasons, or if it's necessary to perform a contract, is limited. That's been a real headache, and existing and legitimate business has found the requirements are somewhat tied, because once they start broadening the exceptions, this raises concerns that the CCPA is being watered down. The CCPA continues to evolve as does the guidance, so it is far from the end of the story on this one.

Going back to your question on post GDPR, I think what we could say is that it has had the impact that the EU probably desired it to have, as a global benchmark. The Brazil legislation feels a lot like a version of GDPR. If you look at Africa, and all the legislation popping up there, you can see GDPR concepts in the Kenyan Data Protection Act. Look at South Africa — it borrows terminology from GDPR. Look at Nigeria — it borrows terminology from GDPR.
They talk about things like portability, right to be forgotten, but don’t expand on what it means. What we’re seeing is GDPR is being held up as a — I wouldn’t want to say gold standard — but as an international standard for privacy.

What I think many countries are not doing, is reflecting on the fact that GDPR evolved over many years out of a unique cultural and political and historical reality that has shaped Europe, i.e., two World Wars. That does not apply in the same way to California, or Africa, or to Indonesia, or to India, or to Brazil, as the case may be. Some of the concepts are very helpful, but you need to be quite careful about how you use them, because they’re going to need to work for you in your context, as opposed to the European context. The European context is embedded in privacy as a fundamental human right, whereas in the U.S., there is no parallel, and the U.S. privacy laws are sector specific, not horizontal like the GDPR. Importing terminology based on a concept comes with a lot of assumptions. If you take those assumptions away, then you end up with something like a Gouda cheese — it’s got lots of holes in it, and doesn’t really work.

How do you think corporates manage the regulation that is popping up in various places, especially if you’re operating across several markets?

This is the irony of it, isn’t it? There is a sense that if GDPR were to be the global standard, and if everyone’s using the same terminology, that there would be consistency. But oh, my gosh, it all means a different thing. What we’re seeing, particularly in Africa, is a huge reliance on consent, not on legitimate interests, as a basis for processing. That creates all kinds of challenges one way or the other.

What it does mean is that two-thirds of the world now has privacy legislation. What we don't have is any significant level of cooperation, consistency, collaboration, equivalence, or adequacy between them. We can say, “Oh, isn’t this great, more and more countries are adopting privacy legislation”, and the logical consequence should be that if all those countries have privacy legislation that there could be free movement of data between them. That should be the natural and logical consequence. But it’s absolutely the opposite. It’s now creating barriers. Look at India with the proposal for data localization, look at China, also data localization, Indonesia, data localization. The list goes on.

Is that because you think regulators and governments have become aware that data is just a huge asset? Maybe it’s not tangible but it’s a huge asset and therefore, you don’t want it leaving your borders? Is that how you think about the control factor?

Yes. I think its two things. It’s very much about concerns about the monetization of data and control over data.

So just on data localization, is it implicit or explicit in the regulations?

It’s both. So in India, it’s absolutely explicit, as it is in China and Russia and other countries. Even the EU is regarded as a data localization jurisdiction because if you read chapter five of the GDPR, it’s actually really hard to export data. The legislation creates barriers to the sharing of data globally.

I think a lot of this has been driven by the opportunity. If data can flow freely, data can be copied, data can be derived, and there’s a lot more you can do with data. Base data can create more data. You get this proliferation of data. Big data can lead to more data. Having data lakes, i.e., a whole bunch of data together, means lots of different organisations can play with it and create different opportunities and insights. It’s a bit like Lego really, except it’s self-replicating all the time.
It’s the same base pieces of data available to users and they can do different things with it and you end with so many different forms of it. The potential is extraordinary.

The challenge is that those jurisdictions where they are being most creative with data and monetizing it to the greatest extent, are the U.S. and China. There is a growing realization that the restrictions that apply in the EU are adversely impacting innovation. The EU has started focusing on ethics and innovation. I think it’s great, and they need to talk about ethics, but it’s something that they need to get the rest of the world to adopt, not just the EU, otherwise they are going to be innovating slowly and carefully, but find that they are also being left behind. So actually, it’s a dialogue that needs to happen globally, and get the big players like the U.S. and China involved.

It is interesting to see the growing awareness and realization as to who has access to the most (and diverse) data in the world, and it is not governments who know everything about us, it is actually the global tech companies. Some of these companies could write my life story for me in a way that no amount of government data I’ve provided through my tax returns and through CCTV ever could. They know when I move, when I sleep, where I work, what my schedule is, who I communicate with, what I say, what I listen to, what I buy — all via apps, location data, and so on via our phones. It’s everything. The ability to monetize that is huge.

India, unashamedly, in their legislation, has focused on protecting its citizens data by proposing to localise data. China is localising on the basis that the national interest takes priority. They are pretty open about it. India’s approach is to protect its citizens data against use and misuse by what are seen as large unregulated companies that are hoovering up the data and the value transaction between what individuals get, i.e., a free service, and what the entity that hoovers up the data gets, are seen to be oceans apart. Of course the providers of those free services want to say ‘Well, you want free services online, you want to post your photos, you want to chat to your friends, hey, all of this requires servers and infrastructure and people. It doesn’t just magically happen. You want free, I’ll give you free, but the trade-off is that you give me your data.’ That’s the trade-off. That value transaction has worked for a long time until the growing realization that these big companies have more money than many economies in the world. That’s how profitable the data opportunity is.

If a lot of these things that come off the back of looking at, like you say, how basically the biggest companies in the world have built their businesses on data and they have been receiving fines for various reasons but it makes very little difference. What’s the next thing that governments and regulators might do if they think they want to curb that benefit if fines make very little difference?

I think it’s the ethical debate and the ethical discussion that’s getting people to vote with their feet. So increasingly, we’re seeing alternatives, such as Tor. Unfortunately, it’s also used for shady activity, but increasingly legitimate users are using browsers like Tor, just because they don’t want to be followed all the time, or tracked, or traced as the case may be. It’s this fine line between privacy and national security. I think governments are really struggling with where they draw the line. Particularly in the U.K. and the U.S.
An issue which raised some interesting issues was the tracking of the Boston Marathon bomber a few years ago. The authorities had the phone and they asked Apple for the key and they said ‘sorry, can’t help you there’. They managed to crack the security in the end, but it raised the interesting issue of when “can you require privacy/security can be put to one side in order to track down a threat actor?” That is one of the big ethical debates that we are now engaging in. Where do you draw the line between the two, particularly in a world where we’re becoming increasingly aware of both the national security perspective as well as the importance of privacy? The EU knows where it wants to draw the line, which may mean sacrificing the national interest to preserve privacy. In the U.S., the line is not in the same place. In the U.K., we’re somewhere in the middle of the U.S. and the EU, particularly post Brexit.

There’s a cultural dimension and legal dimension to consider as well. For example, WeChat is much bigger than WhatsApp in Asia, but WeChat is operated by a Chinese company, making it subject to Chinese law, jurisdiction and enabling access by the Chinese authorities; what you’re saying, to whom, what you are buying — none of this is “private” in the Western understanding of the concept. Culturally, this reality is just accepted, and WeChat is hugely popular.

*Post GDPR, in terms of regulating and enforcing, have you seen much change in behaviour there?*

Yes, the U.K. is a case in point with the Marriott (£99 million) fine and another proposed fines, which are staggeringly large compared to previous fines. The Equifax fine in the U.S. at $700 million is also eye watering, and creates a new precedent.

However, there are many who feel that these enormous fines, as a percentage of turnover, are not enough. The question is what would be a meaningful and impactful fine for some of the largest companies who are in the sights of the EU? Or do we need to look beyond fines to other enforcement measures?

*How does it evolve? Can it evolve beyond just the fines getting bigger and bigger? Can it become more restrictive in other ways?*

This is it, there are a lot of tools in the toolkit, but the one they keep focusing on is the fines. The fines have an impact, but it varies depending on the company and the sector. What can have a far greater impact is the impact on reputation and share price. We are increasingly seeing companies being publicly questioned about their privacy practices, whether at the government level or amongst privacy professionals or by the public.

In the regulatory toolkit, they can do a number of things. They’ve been focusing on the fines, they have the ability to stop firms from processing data. If they were to actually use some of the other tools, it could be much more effective. To give you an example in the U.K., the ‘84 Act had to be amended in recent years to allow the ICO to issue fines of up to £500,000. This then changed from May 2018, with the new Data Protection Act providing for fines of £20 million or 4% turnover, whichever is the greater. But a powerful example under the ‘84 Act was when the a U.K. bank CEO was required to give an undertaking, as CEO, that he would change the practices of the bank and improve and invest in the data protection teams and infrastructure. That drove more change in the firm than any fine because the CEO was being held accountable, and had to make a public commitment to improve.
Pressure on the CEOs is increasing from ESG and sustainability factors. How much do you think data privacy is really considered a health factor?

I think when you're talking about social responsibility, part of that has to be around trust and how you use data. So even if it's not explicit, it is absolutely part of that broader engagement piece. You know, who are you as an organisation? What is your culture? Are you transparent, fair and honest about your data practices? It's very much a part of the culture of the organisation. From my perspective, I think that those sorts of tools can be much more effective than a fine.

Is there anything stopping the regulators using that tool set, rather keeping going with fines?

They're learning. On the GDPR side, they're learning. One of the criticisms post GDPR is that after all the big talk about big fines, we are not seeing much happening on the EU side. The CNIL fined Google €50 million for its marketing practices, but the other fines have been a fraction of this. However, the privacy regulators are all saying that it takes time, just watch the space, we've got a lot in the pipeline. I suspect they do, and they're going to come down hard on some big names.

I think one of the challenges that they're facing is that when you're going to issue a significant penalty to a significant entity, you need to make sure that it is watertight before proceeding to an announcement, and there may be other issues to take into consideration, such as disclosures to regulators, shareholders and so on, which may complicate the process.

The ICO only appointed their first general counsel a couple of years ago. Prior to that they had no internal legal function, which meant that they were practical, rather than legalistic. They were policy driven. However, as the stakes have increased, all the DPA's are going have to ramp up their legal teams as the GDPR is a complex and detailed piece of legislation.

I was going to ask you about the ICO's review of ad tech and programmatic trading they're currently going through? Could they change practices?

What they're doing is they're signposting. It is interesting to see that the ICO, generally regarded as the most business friendly of the GDPR regulators, is the first one to deal with this really complex issue, and call out the issues around GDPR compliance. What the U.K. has done is they've come out and they said ‘Look, we've consulted with you and we can't see how you can marry these practices with those rules. I suggest you go back to the drawing board and come back to us with how you think you are going to comply. If we don’t see this happening, then we're going to have to take action against you.’ So they've sent a shot across the bows. AdTech now has to come back and explain how they will comply. It is creative – rather than leaping to enforcement, they encouraging firms to comply.

And in terms of sort of Europe, specifically, do you think there's any more regulation coming down the pipe? You mentioned the ePrivacy regulation, which we've looked at in the previous GPS as potentially imminent and it seems to have gone very quiet so I wonder if that’s then died a death?

No, e-privacy hasn’t died a death. The file continues to be picked up by the rotating Presidency and the next Presidency, Germany, will pick up where the last one left off. The challenge is the technical and practical issues, trying to cater for every eventuality, and the drafts are subject to continued challenge and negotiation.
Obviously, a big issue for Europe is around international transfers, and that's brought to the fore with Brexit. The EU has said that by 2021, they would review all 13 countries where they have adequacy provisions currently in place. So what we are seeing in terms of trends is that those adequate countries are revisiting their privacy laws, because they know this revisit is going to take place, and they're amending them to address any material gaps with the changes that GDPR has introduced. Nobody expects those changes to be a mirroring of GDPR, which highlights that the U.K. has privacy legislation most closely aligned to the GDPR. So, the question is, how will adequacy be assessed given how very different the privacy regimes are across the current adequate countries – will this be a legal, practical or political assessment or all three?

*In terms of the California Privacy Act, as a corporate of what are you seeing from firms and the discussions you have within the industry around how corporates are dealing with that? If they're already GDPR compliant, if they've already dealt with GDPR?*

If you dealt with GDPR it's less of a heavy lift because you're familiar with transparency measures such as notices; you're comfortable with the process around individuals asking you for details of the information you process on them. But what CCPA asks for is also different from what GDPR asks for. The result is that firms have had to repaper policies and notices and processes, as the information you have to provide is different and the time scales are different. There are still a number of practical issues to be ironed out, and questions around whether existing concessions such as the processing of HR data will be carried over on a permanent basis, so the landscape in California continues to evolve at pace.

*Other signs of the other state laws?*

Thirteen other states are developing mirror type legislation. I think Washington State's gotten as close as it can to passing their legislation, but nothing's gone live as yet. What it has done is to reignite the debate around U.S. federal legislation. I have lost count of how many drafts there are out there, and it is a fascinating process. With the U.S. elections looming at the end of the year, I don't think there is a realistic expectation that we will see federal privacy legislation in the U.S. this year, but maybe one day in the not too distant future.

*So you think there will be a federal rule?*

Under the current Trump administration highly unlikely, but whatever the next administration is, I suspect that this will then become a political football. Most of the rest of the world now has privacy laws, and now California has as well. Who was the first to introduce breach notification laws? California. Now every U.S. state has a breach notification law and it's been enshrined in the GDPR. The question is whether California will continue to set the agenda in this space.

There really is a very good argument on the federal side to get ahead of this issue, as 50 versions of CCPA will be challenging for both businesses and consumers to deal with. I suspect, as I said, it'll become a bit of a political football for the elections. A big question is what to do with the existing sectoral privacy legislation in the U.S., and whether and to what extent U.S. federal privacy legislation will pre-empt state level legislation. This isn’t just a civil society issue — business would benefit from increased certainty and consistency in the privacy landscape in America, rather than having to address something similar to the current 50 flavours of breach notification, which is hugely complex and challenging to deal with.
Compliance becomes expensive, bureaucratic, and takes a lot of time, and it is not clear why different versions are better for the consumer. One of the things that has been learned from the breach notification law, is that people have notification fatigue. The impact and awareness is lost when lengthy breach notification letters become a regular feature, and there is not much that an actual individual can do to make or see a difference.

Any particular sectors where you think genuinely they could become good custodians or seen as good custodians of data from a consumer perspective?

The big cloud providers generally have a good reputation as they have invested heavily in the infrastructure and understanding of the privacy requirements. On the other hand, we have seen a huge number of breaches reported in the public sector, and this has caused a significant erosion of trust amongst consumers. This may be due in part to the fact that the public sector has limited budgets to invest in top of the line security infrastructure and talent, although with the new focus on cyber, this looks like it is changing. It is more difficult to assess the private sector, across the broad spectrum of SME’s and large organisations, and because we have just two years of mandatory breach reporting to work from, the picture is still evolving. However, feedback suggests that despite the financial crisis, people are more likely to trust their bank with their data than any other type of organisation. Whether it is physical security or digital security, both confidentiality and security have been essential to the operating model with customers, and one with which individuals have a long standing relationship.

Are you hearing of any innovations that have been driven by data privacy?

There is a whole industry coming to maturity, with tools to automate data privacy compliance, tools to help demonstrate accountability, tools to manage breach and subject access processes, tools to manage cookies, and of course tools to report on privacy compliance metrics. So if you’re in the world of privacy, there is plenty of innovation taking place. And in addition to the innovation and new players, we are also seeing some consolidation taking place with One Trust acquiring Data Guidance and Nymity being acquired by TrustArc — all very big names in privacy. I think what people are realising is that the reporting and accountability obligations for privacy are so detailed and broad and complex, that you can’t do this in a spreadsheet and in Word anymore. Tools for reporting and creating visualisations of where your data is, particularly around data mapping, are becoming essential tools of the trade.

A whole industry around privacy is evolving, well beyond consultancy and advisory. Privacy qualifications are becoming increasingly important, and for many roles, are a pre-requisite. And it isn’t just the IAPP that provides qualifications anymore. The CNIL, French privacy regulators, are looking to provide their own qualification. There are a couple of universities that are now providing a privacy qualification. There’s no independent standard as such, but the IAPP has been around the longest and is probably the most well recognised.

You mentioned the ICO are looking at AI. Is this an area that will be focused on?

When it comes to AI, the future is today. AI is a significant theme for the U.S., China, the U.K. and also the EU. The EU has already explored ethics and AI and continues to consult on whether and how to regulate AI. But you can’t talk about AI without having the ethics discussion. The ethical dimension to AI is key, as is the bias perspective.
You’re obviously very well connected in the privacy world. Are there any topics of conversation within your sort of world, an industry that it’s been focused on at the moment that we’ve not touched on or that you think is interesting to explore?

How the DPA’s approach enforcement is proving to be an interesting area of focus. The U.K. ICO has announced some eye watering numbers in its intentions to fine two companies, and the CNIL in France issued a record breaking fine of €50 million against Google in 2019, but, two years after GDPR, there is interest to see how regulators are using the full range of tools in the toolbox, and not just fines.

If they put in something much more stringent, like, stop processing, would that kick off disagreements of other countries?

Regulatory cooperation is becoming an increasingly important issue given the ramifications that decisions inevitably have across borders. Sometimes it is not feasible to ring-fence measures to just one jurisdiction, as people can just access a search engine in another jurisdiction to obtain access or services. Given the widely differing privacy rules, whether they exist or not, whether they are general or industry specific, whether they apply to private sector or also include the public sector, etc., means that it will take considerable time and effort to see an effective and joined up approach between regulators, although we are already seeing more informal approaches evolving.

Do you think they get to that point where they’re all on one page across countries?

It’s a long journey. It would be helpful to see increased cooperation and collaboration, but usually based on your own jurisdictional perspective! One of the challenges is that the fundamental basis for processing differs between jurisdictions, which means that it is difficult to have a level playing field for compliance. While the GDPR is likened to a ‘gold standard’, it is not necessarily one that fits all legal regimes, cultures and societies, and so if data transfers and collaboration are to be assured, legislators and regulators need to be flexible in their criteria for adequacy or mutual recognition, as models will differ across the globe. This certainly provides the U.K. with the opportunity to be a leader in this space with their practical and outcome focussed approach to regulation which does not seek to impose U.K. standards, but instead seeks synergies.

An interesting example of a recent EU adequacy approach was that with Japan, where an adequacy assessment was made, but only subject to an additional eight EU principles being adopted in addition to the Japanese requirements when processing EU personal data – a rather odd approach adequacy.

Is there anything else within the industry that’s a big discussion topic?

Don’t lose sight of ethics and AI. This is a huge and dominating theme, heavily linked to innovation. What is the future without digital innovation? How do we harness the power of data to grow innovation, but still meet legal and ethical criteria that are often perceived to hold back innovation? Do strong privacy controls create trust, security and good quality data which in turn fosters transparency and fabulous innovation, or does it hold back the pace and creativity around digital innovation? With the U.S. and China leading on digital innovation, the European model has a lot of catching up to do. Will the European gold standard for high quality innovation lead the way, or will the more flexible data regimes in the U.S. and China continue to lead? Is it an either/or, or is there a middle ground to be achieved?
I think time will tell, and those countries that are open to innovation and willing to be flexible, while supporting trustworthy data principles and standards, will reap the benefits.
A Conversation with Martin Ashplant
Digital Media Consultant & Former Chief Digital Officer, Beano Studios

*Martin, please tell us a little bit about your background?*

I've been in digital media for 20 years now. My background is as a journalist; I started in the usual journalist fashion working for a local newspaper and then working for a news agency, but very quickly ended up in digital, around 2000/2001, when digital was becoming a thing. It was clear to me quite early that the way digital was evolving meant it wasn't just going to be about content in future. You could have the best possible content, but if you didn't have the whole piece, then it wouldn't work. So that's why I think my career evolved more around the whole digital product side of things.

My background is working for places like Metro, City AM; I also worked as a digital media consultant for the likes of Johnston Press, ITV News, and various others as well. All of that has been around growing digital audiences, leveraging the relationships with the consumers and the audience in order to generate revenue streams, whether that is looking at a traditional revenue stream around digital advertising or something different. Of course, back in early 2000s, digital advertising was still quite a nascent industry in itself; there was a sense that trying to reach scale would offset any losses in the print advertising world, which I think has been shown to be naïve in its thinking, but that was the world in which we inhabited back in the 2000s.

More recently, for two and a bit years, I've worked for Beano Studios. What is really interesting for me, in reference to this conversation, is that Beano Studios operates in a kid's world. Our target demographic is broadly seven to twelve year olds, so very much under that age 13 level, which is determined as being the difference between child and adult online. Since inception in 2016, Beano Studios has been operating within very strict regulatory conditions and we talk about this a lot. When we were analysing what GDPR meant, what was reassuring from our perspective, was that a lot of the approaches that we were already taking were very much in alignment with what GDPR was suggesting. I think that's on the basis of our approach being very much in line with COPPA, the Children's Online Privacy Protection Act, and working within that - and also sometimes working even harder from a privacy perspective. I guess I've seen what a really regulated data privacy world would look like, based on the experience in a child directed proposition. It's been really fascinating for me to see how more and more of the things that we have as requirements in the kids' world are filtering out into the mainstream digital media world as well.

*Given you’ve seen what the regulated landscape looks like because of the demographic you’re focused on, what could that look like for the broader digital advertising or publishing industry?*

I think it has the potential to genuinely change the industry as it is at the moment. There has been a tendency within the industry to almost turn a blind eye to some of the regulatory discussions that are going on; often because it was such a complicated, such a sophisticated landscape and ecosystem. You've probably seen the very famous chart which shows all the middle parties between the publisher and advertiser. If you start looking at that, you might think, okay, I don't know what to start with to try and make this right.
There has definitely been a tendency within the industry to say ’I hope someone else will work this out.’ I think that is starting to change and the penny is starting to drop that this isn’t going to just go away — particularly because I think consumers are demanding a little bit more. Everything that happened with Cambridge Analytica, and the very high profile, political related data-gathering there, has changed the viewpoint of consumers.

One of the things that everyone says they do is privacy by design, but we genuinely have to do this at every stage. It feels like that will be the biggest change to the way that digital publishers operate; rather than looking at it and hoping someone else will sort out the bits in the middle, you have to start from that very inception of what it is you’re doing and ask how can we ensure that privacy is central to that.

**What does that mean in practical terms? How do you go about implementing a privacy by design approach versus where publishers are currently?**

From a product development perspective, the very first discussion around what it is you are trying to do and what your objectives are should include privacy. When you do your design thinking, it’s bringing privacy into the room straightaway. So one of the things we do is also consider what’s the worst case scenario from a data privacy perspective, what could go wrong, and then work back from that. In a product development world, you are always talking about doing a risk assessment around how this is going to work and do we have to do this. It’s about bringing data privacy in at that stage, rather than doing what I think some people or some organisations have done in the past, which is see privacy as a bolt-on at the end.

It's really hard to ensure privacy if you've already created your feature or released your product and then you're trying to retrospectively paper over the cracks. Also the relationship with a third-party is something that is really challenging. Knowing exactly what is going on with all the third parties is hard and there's been all sorts of research pieces done which have suggested how many different third-party cookies there are on every different website. Particularly in media, it is in the hundreds quite often because there are there are so many advertising technology integrations that you almost lose track of what cookies are there and what they're doing and who put them on in the first place.

There was a stage when it was a little bit fast and loose with third-party cookies and those integrations. I think there has to be a much more considered approach to why you would use a particular third-party. It probably won't be based on cookies within the next couple years, but whatever the new solution might be, whenever an implementation with a third-party is discussed, it's important to ask: what does this do, what could go wrong, and how do we ensure that the privacy is protected within that implementation?

**Who regulates the third-party system? Is that the ICO which is looking at those third parties?**

Yes, I think in the U.K., it is the ICO and although I can't speak for the ICO I would imagine they would acknowledge that it hasn't always been the most regulated of worlds. It is difficult because of the kind of cross-territory basis of it in that so many of these ad-tech businesses are based out of the U.K. jurisdiction and it becomes quite challenging to police that. I think one thing that I certainly took from recent ICO guidance is they are really keen to ensure that those middle people in that process are aligned with what they are trying to do from a regulatory perspective. I think that definitely will start to happen.
What do you think is the main intention of the ICO right now? Is it to try and clean up the middle man that sits within the landscape or do you think we can see a complete upending of how the digital advertising and ad tech landscape works?

For me it's been interesting to see how the narrative has modified slightly in the last six or seven months. When the ICO first put out their report and their guidelines into ad-tech and real time bidding in July of last year, it was a very all-encompassing overview. It was talking at quite a high level of granularity around things like analytics cookies and needing to get consent for analytics cookies. It was seen as being really wide reaching and questions were asked about whether they had fully considered what that means and the fact that there aren't really any alternatives for this. It feels like after consulting with the industry, they have almost refined their focus to be much more around that ad tech ecosystem and working with the people who are dealing in those data points, and there's been quite a lot talk about the more sensitive data points, so political stance, religion, sexual orientation. It feels like it has gone from a very broad overview to a bit more of a targeted approach; it would seem like that's going to be the focus for the next few months or so.

Given a more targeted approach, do you think the impact will be manageable? Or do you think it could still lead to quite a big change in how the industry operates?

I think a lot will come down to what their appetite for enforcement is. It depends on the ICO but there's also a growing desire to effect change. We've seen it with Apple and Firefox making breaking changes at a browser level, which at a stroke stopped third-party cookies being collected, and then the more incremental approach that Google Chrome is taking over the next couple of years. I think it's a combination of consumer desire for things to change, and more awareness and better visibility of what's going on, plus the technology partners like Google, Apple, and the rest of them, actually making changes. Then the regulatory bit coming over the top of that. It feels like all the factors are in play to actually make things happen, whereas there has been a tendency in the past for it being purely regulatory, to wait and see what happens from a regulation perspective. It definitely feels like there's a change in the narrative this time as opposed to a few years ago when we're were talking about cookies in a much broader context.

GDPR came in May 2018, do you think that's given the light to the ICO/the regulator to look into it further because it's now law and there is potentially greater fining power?

Yes, and I don't think it was ever really clarified what GDPR actually meant. There was a lot of talk at the time when GDPR was implemented around PECR, the e-privacy directive, and almost waiting for that to help define in more detail what this actually meant for things like cookie collecting. It's been quite a long process, but GDPR has certainly given that initial legal framework for interpretations of it to evolve over time and to be used as the rationale for why we will enforce these regulations because 'if you are not doing what our interpretation of that means, then you are in breach of that GDPR law.'

As you said, there are lots of different players and operators of different sizes across parts of the value chain, do you have a sense of what the view is across those different parts of the value chain, especially publishers, what the ICO is looking for and then what Google's proposing around third-party cookies?
So I've had a number of discussions with various peers in the publishing industry. One of the things that is really clear is there is a sense that the opportunity around programmatic advertising has changed. I find it really interesting. The Reuters Institute for Digital Journalism, which works with Oxford University to do various reports, do a prediction piece at the beginning of every year where they ask about 250 Digital execs across 30-odd different countries, what their expectations are for the year ahead. This year 50% of them said their main revenue stream is expected to be reader revenue, as opposed to 14% expecting it to be advertising. Now, I don't know the exact figures of that two, three years ago, but I can be very confident that that has shifted almost on its head because there is a real acceptance now when you speak to publishers that digital advertising is being buffeted by so many winds right now that you need to diversify those revenue streams. But more than diversify those revenue streams, you actually need to really create some very strong robust revenue streams that are reliant on your direct relationship with the consumers in the audience that you've built up.

I think the regulatory discussion that's happening at the moment is adding to that because it's putting much more emphasis on the need to generate first-party data that you can use to control the value exchange and make sure that your audience is giving you that data willingly, and then you can use that to improve their experience. A few years ago, that third-party data world was almost adjacent to what you were doing as a publisher. Whereas now, first-party data is so important that you're seeing publishers take a lot more interest and starting to bring in people who are data scientists and able to understand data and optimise the flow around reader revenues, subscriptions; much more than what was previously focused around providing advertising.

**How much reliance is there still on third-party data by the broader publishing industries?**

Well, I guess the big one is the third-party data that's provided by Google, as I would be hard pushed to name one big media owner who doesn't have some reliance on Google, whether that's simply Google Analytics, or Google Ad manager, alternatively known as DoubleClick. Google powers so much of that ecosystem and provides so much of the data back to publishers, which is almost the lifeblood of that publishing industry, which anyone that was to say that you're not relying on third parties, is almost ignoring Google; Google is the biggest player here and it is the organisation that is determining the future landscape.

So what Google's proposing around the change to third-party cookies in February and then the longer-term two year transition around third-party cookies on Chrome? How does that look in terms of the access to the Google third-party data and broader third-party data as these changes come in?

I think it remains to be seen. At the moment, there are very much two schools of thought. One is that what Google is doing around its Privacy Sandbox is trying to ensure heightened privacy while still providing that lifeblood data for the industry, whether that's around conversion metrics for advertising, or analytics, or whatever it might be. If you look at it from that perspective, it's brilliant what Google is doing, they're trying to come up with a solution to ensure that both requirements are met. Now, of course, the other way of looking at it, which I think is probably the way that other businesses within the ad tech industry are looking at it, is that actually all that Google have done here is protect their own interests by building up the system, where they are the arbiters of what data is allowed in and what data is allowed out.
Given that they are, by far, the biggest online advertiser in the world, you can see how some organisations, some people, might look at that and think, well, that's protecting their own interests. But they're in a difficult place. I think what they have done is they have sought to engage as many different people and organisations and industry bodies as they can throughout this process. I think the consensus among the publishing industry is that it is far better to work with someone taking that approach, than to find out that Apple have just blocked third-party cookies with their latest release. So, it's a tricky one because Google are obviously not an independent party in this, but they are also probably one of the better places to try and come up with genuine solutions that will work.

And from the publisher perspective, when Apple and Firefox blocked, what impact did the publishers feel? Either revenue or pricing around the digital ad inventory? What was the most notable impact?

I guess it will depend on what the makeup of your audience is but broadly, it probably represents somewhere between 10 and 20% of your audience, which is navigable. So it has an impact, but it's not going to drastically change everything overnight. Whereas, because Google Chrome accounts for something like 70% of all Internet activity, you can see that the impact of that will be much, much bigger.

So, it's almost useful from a publisher perspective to be able to look at what happened with Apple, so you can then segment that audience and extrapolate that to forecast what the impact of what Google has said it is going to do will be. So what does that mean for us? What's the impact? There is definitely a sense within certain publishers that I'm aware of that the revenue opportunity from programmatic advertising is being impacted by these kind of changes. Now, it might be a short term thing; as you transition into new ways of doing things that comes back up. It might be that it presents opportunities for publishers as well, as we were touching on before, the fact that first-party data is going to become so important does give publishers an opportunity to build up their own first-party datasets, build up that audience and then fill that gap with third-party data, maybe not so freely available. Well, does that mean that contextual advertising across a big media owner suddenly becomes more valuable? That'll be the glass half full way of looking.

What's your view on how regulation continues to evolve? Specifically in Europe, but then if you have an idea around other markets as well?

I think it still feels to me that in Europe, it is about really defining what GDPR actually means; it still feels like we're in that place at the moment. You often hear discussion about the spirit of GDPR, or the essence of GDPR, not being met with certain solutions. With any regulation that's what's going to happen until we find an understanding what it means and what's the precedent around what's okay and what's not okay, what boundaries are acceptable and what's not. It feels like we're navigating to that point.

What will be interesting is whether other territories come in and make that happen quicker. So what's going on in California at the moment is interesting, because is that the framework of what will become the North American approach or North American equivalent to GDPR or is that not really going to work or not going to be enforceable? I think that's really interesting because you could definitely see a scenario where regulation in the state supersedes what's happening for GDPR or clarifies that.
That's one of the things that we found to happen with the children's element of GDPR. It was as if GDPR took COPPA as the framework and then where there was ambiguity, or things didn't quite seem to be relevant to what is going on in real life, there were some amendments or some changes, which made it easier to interpret and to understand what does this actually mean.

**Have you seen much in the way of regulatory enforcement, or do you have any idea around how regulators are going to look to enforce GDPR longer term?**

It will be interesting to see what the response is to privacy groups launching legal action. There's been quite a bit from the Irish equivalent of the ICO, the DPC, because Ireland is where Google has its European base. It'll be interesting to see what impact that sort of thing has. One of those privacy organisations has actually recently threatened the ICO with legal action for not doing enough to enforce GDPR, so that's something that could happen quite a lot in the space; it's certainly something that we see happen a lot with children's privacy. You actually see it happening from the perspective that the privacy group makes a big noise about something that they are passionate about and then the regulator responds to that, and more awareness grows around this stuff and more publicity is garnered by this sort of thing. It will be interesting to see what individual responses emerge relying on their own interpretation of what the regulators should be doing and whether they try to take legal action against specific organisations.

**As a publisher, have you seen much change in consumers either awareness or behaviour on online? Do you think there's a real push from consumers for any change there?**

So one thing I have noticed is from the large amount of user testing we do at Beano. We put things in front of both kids and their parents quite often. We've done that with our privacy notice recently to make sure that there's a bit that makes sense to kids but there's also a bit that makes sense to parents; one thing that has come across quite clearly is that kids are more naturally talking about data privacy as something that they're interested in. I think the younger generations are growing up being taught about that and are being given guidance on making sure you know what data you are providing and know what personal data is. It feels like there is a generation coming through that is much more savvy about data, the value of their own data and making sure that they are not dishing it out, unknowingly.

In previous roles I've had, when we've done UX testing around things like cookie messaging buttons with adults, the sense is that most interaction with it is just to remove the message out of the way because it's stopping them from doing what they are trying to do. You would have a small proportion of people who care deeply and passionately about that data privacy element, but the vast majority were either unaware or didn't care enough to do anything about it. It was just an impediment to what they wanted to get to, and I think that has seemingly started to shift a little bit more and more people are at least a little bit aware of what this is all about. I think I'll put that down to, again, the publicity around things like Cambridge Analytica and all the data breaches that have been talked about. I think people are more aware but it doesn't necessarily mean more of them care enough to really dig deep into it.

It is imperative for publishers who have that direct relationship to be as transparent as possible. Now, the danger, or the challenge, is that transparency can sometimes come across as being overwhelming. So a privacy policy that details every single cookie that has been collected and what might or might not be going on there — I'm not sure that always provides the clarity and the summary of what is going on here.
Sometimes it feels like things are hidden behind various bits of terminology that you have to use from a legal perspective. And one of the things that's been really interesting for me from working with kids is you have to strip all that back because once you realise that it doesn't mean anything, it almost acts as a curtain across what's actually going on. So even though you are genuinely trying to increase transparency, what you end up sometimes doing is the opposite because it becomes so unwieldy that people don't even look at it, so they're not even learning anything. I think there'll be a challenge as to how publishers ensure that they are providing that clarity and transparency and being open and honest with what's going on without getting so wrapped up in concerns about what you should be doing from a regulatory perspective that you're almost not putting the end user first in that in that discussion.

*Going back to the point you made about the survey where 50% expect main income to be basically subscription type revenue from the consumer. If there is a change to how digital advertising works in terms of the ability to use data in the way it was used in the past, are we going to be looking at a landscape where everything's behind paywalls? And what does that mean for the publishing industry?*

It's the big question at the moment in the publishing industry, and I think you are increasingly looking at a need to not be reliant on having an open, scaled proposition that generates revenue through eyeballs. I think that is something that has become quite clear over the last few years. What that means will depend on each individual publisher; for some it makes absolute sense to create a subscriber model. There's some really good examples of businesses that are successful in that; New York Times is one that's just come out with some really, really impressive numbers. The New York Times is massive and it has huge content power and digital product power to get behind that sort of thing. So they're kind of at the top and you have The Economist, the FT, which are able to generate subscriptions, because they seem to be important for people's careers. Then at the other end of the scale, you've got an increasing number of smaller organisations that are really niche. So the communities that matter will be prepared to pay. So in my world, I'd be willing to pay a subscription for someone that was getting really detailed information around the digital media industry. If you were in the energy sector, you'd be willing to pay for something of high value in that sector because it's meaningful and that can aid your career but also across really, really specific interests.

The middle is where it gets interesting. The experience that News UK had with The Times being behind a paywall, and then trying to put The Sun behind a paywall, and then deciding that wasn't going to work and taking that paywall away is notable. It speaks volumes for the potential of a more commoditized mass-market publisher to actually generate the right amount of revenue from subscriptions. I think that's what we will see over the next few years. Then there are other ways to make money and that's something that I have seen at every media and publishing conference I've been to recently; all about that diversified revenue. You've got people talking about successful ecommerce strategies or successful data strategies or successful agency strategies which lean on the expertise you've built up in a particular area because you have a genuine relationship and understanding of an audience. In Beano's situation it's kids; there are a whole bunch of organisations that don't understand kids but want to and need to understand how they can better resonate with kids. Therefore, we as a publisher are able to show that relationship, and have been able to better help them. I think you'll see increasing numbers of publishers leveraging that in order to make money. I think we will see more paywalls, membership schemes, premium tiers, community bundles, whatever the flavour of it is, but that reader revenue element is going to be really important.
Do you think a better business model option here is where the consumer themselves can be compensated for the use of their data? So either Google or the publishers basically offer financial rewards for sharing data?

Possibly, it is something that has been talked about. Actually, it will very much depend on the perceived value of that data. If you can make a case that the ROI is worth it; so if an individual's data is sufficiently worth you buying it, then I think there could be. I think an example where it might work is if that individual is a either a really high net worth or a very difficult demographic to reach; you could definitely see some elements there. There are some examples where people are being paid to give their data or their opinions or their views on things because they are seen to be valuable. Whether that works at commoditized level, I'm not sure; you're then adding more costs to your business. You need to be really sure that that generated sufficiently more revenue to make it worthwhile but it is definitely an interesting concept.

One of the things that I remember when I worked at City AM is when we were the first U.K. newspaper to put an ad blocker solution on our site, which basically said: if you use an ad blocker, then you won’t be able to see the content. That was very much from the perspective that this is an ad funded business which has an audience who are real business people, so understand the need to make money and understand the economics of it. If there was ever a proposition that should be able to have that conversation with its audience, then you would think it would be one based in the City. What struck me when we did that was the feedback we got when we asked people what they thought about this; there was a lot of misinformation about how that ecosystem worked in the first place. For instance, some people said, 'well, you should be paying me to see ads on the site.' It struck me that there was a very deep misunderstanding around how publishing works and how you've made money out of the back of it, which I think would be an interesting consumer thing to understand as well; what's a consumer's perspective on the value of their data? You could definitely see a scenario where consumers say, 'well, you pay me for that data.' It's just whether that stacks up economically.

There must be some point in the value chain where the consumer themselves get to switch and say, 'you know, my subscription costs halve, or pay me $50 a month or whatever it is.'

I guess that that'd be a really interesting situation to be in because it would come back to the dominance that Google and Facebook have in that world. The digital advertising market is massively dominated by Facebook and Google; the latest figures suggest about 69% market share in the U.K., and I would imagine they would be able to claim the most valuable data from their users as well. So what would individual media owners be able to leverage from a data perspective that was even more valuable than what you could get from the likes of Google and Facebook? And then who would be the buyers of that? I think, again, it might come down to those more niche levels. If you are, for sake of argument, an energy publisher, and you are able to supply really informed, insightful data about people in the energy sector to energy businesses, or those who interact with energy businesses, then I could see that potentially having enough value for the energy publisher to say, 'well, actually, it makes sense for us to incentivize people to give us that data because we know it's more valuable than what we we’re paying for it.' I don't know whether that would be possible in the middle of the market, but everyone is looking at increasingly inventive ways to try and monetize digital content and digital products and digital media per se at the moment.
If Google and Facebook take roughly 69% of the ad market in the UK, how do you see that evolving over the next five years?

The latest figures I saw on the UK, from the eMarketer consultancy, suggested that this market share was growing over time and would be above 70% next year. That's when it starts to feel quite, quite dominant. You then think well, potentially, the only way that's going to really change is if the likes of Amazon come in and shift that. There doesn't seem to be an awful lot of the pie left for other organisations to claim much of that digital ad market. It might be the same in the data transaction market; it feels like it might be a similar situation. That's why there has been so much focus on direct to subscriber, it cuts that out, because you are then not trying to take away money that's already been hoovered up by Facebook and Google.

And so you don't think GDPR, CCPA or any of the sort of data protection regulations actually break that dominance? Or maybe it goes the other way and creates an environment that's even more favourable? What's your view on that?

I can definitely sympathise with the argument that this seems to make Google's position even more powerful from that perspective, because there isn't a place for others to play if Google are essentially holding all the data in a pen, albeit in a collaborative fashion, but ultimately they're the ones who are determining where that then goes. Then you can see why there are people in that industry who are nervous that all this is doing is accentuating that control over it.

Do you think it will go back to the days of contextual based advertising, subject based rather than specific tracking of data and individuals?

That is something that we've talked about for a number of years now; there's always been this sense of 'maybe we'll just go back to what it used to be like in print when you bought based on the perceived audience and you contextualise it based on the content that they were reading in that given time.' There may be some of that but I think the technological advances have been such that it feels more likely that there will be technology solutions that will allow a similar approach to be taken, such as Google's Privacy Sandbox. It feels unlikely that we will do a complete reverse and go back to contextual advertising although if publishers can build up strong first-party data, matched with contextual advertising, and show that it works, I think there will be opportunities there. But I remain unconvinced that we will suddenly shift right back as a result of regulation. I think it will be more likely that we will see this evolve over time and the industry will come up with different ways of trying to achieve the same thing.

Are you seeing any innovations? Any innovation around other means of still getting access to data that's going to enable them to sell advertising inventory in a very targeted way?

It feels like most of that innovation is expected to come from Google because they are so dominant. In that world, the vast majority of ads online will go through Google Ad Manager one way or another and the same with analytics. As an example, there's a real possibility that every website will soon have to ask for consent to drop a Google Analytics cookie. The vast majority of the industry is using Google Analytics to tell them what's going on their websites. And if that were to suddenly rely on consent, the estimates are that you would overnight lose about 70 to 80% of your visible audience.
So anything that relies on audience data to determine what's the best experience for our users — what they like, what they do not like doing, A/B tests to determine is this better or is this better — overnight that would just be a much harder game to play. I think there is a huge reliance on the likes of Google Analytics, Google Ad manager and the rest of that ecosystem that has been created by Google to come up with solutions to allow publishers to still get that. You could argue, well, why don't publishers try and solve that themselves, but that's a really difficult challenge. It's a bit of a vicious circle; there is a tendency to complain that Google dominates but when there is a need for a solution, there is a tendency to wait for Google to come up with a solution. I think that's just a natural state of where we are with that in the moment.

**Google Analytics, is there anything around that, that's at risk from a GDPR perspective? I don't know how much of that falls into personal data, whether it's this or legitimate interest?**

Up until last June, when the ICO came out with their new guidelines, the approach had been that Google Analytics was not part of the same story because the data collected there was to help make the experience better to make users have a more fulfilling time on the site. But the ICO were quite clear in June that analytics cookies do count and therefore do require consent. So that is one of those things that if that actually ends up being enforced, it will have a huge impact. The prevailing thought is that won't universally happen until there is an alternative. Then you again look to Google to come up with an alternative, because the impact of that would be so wide reaching that it would potentially, at a stroke, change every media owner's ability to understand their audience, to make fairly standard changes and understand how many people are coming to the site. I suspect that there is a sense that until there's an alternative, it doesn't feel the right course of action to enforce that because of the low risk of privacy breaches through that.

**Anything else as a publisher, that you're keeping an eye on from a data protection and privacy perspective, that could be coming down the pipe?**

I think I touched on it before but what's going to happen in the U.S.? Is it going to be a situation where you have multiple states doing different things on data privacy which would then become really quite difficult if you are trying to have different approaches to different places in the world; it is a real operational headache as well.

We saw it when GDPR came out, when some quite big U.S. publishers just said, ‘we're not even going to let people in Europe come to our websites because it feels more effort than it's worth.’ I think there's probably a little bit of politics involved in that as well, but you could definitely see situations where, if you start to get different interpretations in different areas, that you might see a publisher saying ‘it's just not worth us being over there, so let's lock down the site in whatever territory this regulation is coming out of.’ That said, I find it hard to imagine a situation in the U.S. where you would have 50 different interpretations on a state by state basis, it just doesn't seem enforceable, or operationally doable.
A Conversation with John Bowman & Todd Ruback

John Bowman – Senior Principal, Promontory Financial Group, an IBM Company

Todd Ruback – Director, Promontory Financial Group, an IBM Company

John and Todd, what are your backgrounds?

John Bowman: I am a senior principal here at Promontory and have been for the last five years. Previously, I spent 25 years at the Ministry of Justice (MoJ), joined there in 1989. My interest in the data area came from about 2011 when I was a policy leader, MoJ, and subsequently lead negotiator on GDPR. Five years subsequently, I have been at Promontory, and have done lots of different client engagements, across different sectors, everything from GDPR prep to digital ethics, ad tech, and tracking and all sorts of other things. So a huge variety of work.

Todd Ruback: I am a director on the U.S. data protection and privacy team. I’ve been here about eight months. Prior to joining Promontory I’ve been in the privacy and data protection space for more than 10 years. Prior to joining Promontory, I was at JPMorgan Chase and was the global privacy officer where I headed up strategic privacy initiatives worldwide. Prior to Chase, I was the CTO Chief Privacy Officer and General Counsel at a privacy technology company here in New York City. I have spent many a trip in the U.K. thinking about and talking to thought leaders about privacy issues, especially around digital privacy and self-tracking technologies.

Since GDPR, what are the notable impacts you’ve seen in terms of corporates you deal with around how they’re using data and their ability to access data? Has there been any notable change since the implementation?

John Bowman: Well, I think that certainly the organisations we engage with do a lot of data processing and GDPR was a major priority for them, in terms of being able to map their data processes in particular. So the article 30 requirement to have a record of data processing activities is, in some ways, seen as an accountability requirement, but actually understanding what your data flows are, could potentially open up the possibilities as well. The idea is that GDPR doesn’t necessarily put a brake on data processing or doing interesting things with analytics and so on. I think companies are just becoming a bit more thoughtful about privacy issues. At the same time, they don’t want to throw out the baby with the bathwater. They do want to keep and tap into the potential of the big stores of data as they say data is a really key asset – a valuable asset they may hold in different configurations, whether it’s in a data lake or dispersed across systems and so on. It’s something which they want to make use of because they might lose that competitive advantage. But at the same time, they don’t want the regulators all over them either. So it’s finding the right balance.

Up until May 2018, everyone was working to this deadline, which was a kind of a false deadline, in some ways, because there wasn’t going to be enforcement straight after the event. Organisations have sort of taken a bit of a step back and are looking at privacy maturity, but also having the right controls in place and the right governance in place.
They are saying, well, we can apply a risk-based approach, and we can do interesting things with data once we have the right controls in place.

Promontory does privacy impact assessments, or another way of putting it, we build in a notion of privacy by design, either as a process or as a cultural change. On that basis, you start to see companies thinking about doing digital strategies or this idea of being digital organisations or data-driven organisations.

I think the DPO in these organisations now has a more say and influence. They can say let's do all these great things but please bear in mind that it's not a free for all. So I think GDPR to some extent has brought about a cultural change. The real question is, do we think there is actually a big upside to this as it happened at a time when we knew we were going to go through the digital transformation anyhow, the privacy by design and digitalisation as well? Our clients are going to be self-selecting to the extent that they've thought about this and may think do we need to think about privacy by design? Or how we manage this data?

Since GDPR have you seen any notable differences in terms of how different sectors have approached it and maybe even firms by different sizes? And in particular, do you think there's been any impact on the Big Tech firms, either from a monetary perspective or in terms of their behavior within the market?

John Bowman: A lot of our clients, but not all of them by all means, are highly regulated. So the banks obviously, the insurance companies and so on. So they get it from a regulatory point of view a bit more. Actually, another sector we do work a lot in is pharmaceuticals and life sciences. So again, they're more likely to be risk averse. They get the idea of strong regulation and doing the right thing and being ethical about the way that they use things like data, but it does vary by sector. We've done some work in media, for example, and they're not bound by the same rules, which banks are, for example, financial transactions and things like that. And so I think there has to be a bit more of a cultural shift in those kinds of sectors.

Todd Ruback: If I could jump in just on two points, supporting what John said, I was heavily involved in GDPR throughout the whole process, and I think it's a brilliant piece of legislation for a number of reasons. Relevant to this conversation, it forced organisations to understand what data they collect, how they use it, and really to take a step back and 'moral-it', if you will, know where it is and what the flow is throughout the organisation globally or domestically. Even though there was a lot of moaning and groaning about the cost to do this at the front end, what organisations are seeing now because they have this information, this knowledge about data, is a whole new vista of opportunities opening up that we wouldn't have even conceived of otherwise. So there's new revenue streams that are popping up and emerging technologies from it. So I think it's really brilliant.

When you talk about new revenue streams and technologies? Do you have any examples of what kinds of things have come about?

Todd Ruback: Now the possibility of having large pools of data, whether it's data lakes or however you want to term it, those are accessible to organisations to analyse and determine if there's any way that they can extract value out of these new data lakes. So, we're seeing an explosion in artificial intelligence and algorithmic or programmatic sorts of analyses of these large pools of data.

My assumption had been that when you look at the regulation, this is going to completely stifle innovation and the idea of machine learning and AI, if you end up with much less access to data, but actually, it's almost had the reverse effect?
John Bowman: I think one of the interesting things is certainly the whole debate around the sort of trustworthy use and ethical use of AI and machine learning and big data processing. So the European Commission has started, they issued some guidelines on trustworthy AI. I think the potential is there to provide consumer and societal good in lots of different ways from new insights. One of the areas though, which has been a bit controversial and beneath the radar until recently, and this is a big data processing phenomenon, is whole ad tech driven side, particularly around real-time bidding.

The system can seem opaque to the general user, and they may have no idea that effectively the real estate on their phone or on the web browser is auctioned off in in microseconds via many parties and ad exchanges and intermediaries and so on. Effectively, that's how they get their targeted advertising now, based on potentially hundreds of different data points, which are profiled against behaviour.

If you use the Cambridge Analytica analogy, where they say in the US they had 5000 data points per individual, that's being used in real time by advertising companies to create those?

John Bowman: I think there has to be some sort of debate about what the consumer wants, whether they're happy to get free search engines or free news websites, and so on or social media or whether that stops and they have to pay for a service, which they take for granted for being free. There's also transparency issues. Well, do people actually know what's going on, or do they know that this information might be shared without their permission?

Todd Ruback: Just circling back to an earlier question – have there been any industries negatively affected by GDPR? We're starting to see data, we know the full picture hasn't come in yet, but it looks like digital marketing in the EU specifically, what John was alluding to regarding interest-based advertising or online behavioural advertising, we're starting to see some data points that it actually has dropped off in the EU.

Where would you see that data point?

Todd Ruback: Just to be clear, in the US digital advertising as a sector is about $58 billion in revenue in 2018. That's according to an IDC survey. The EU was close to that, about €55 billion euro, so big money, and it accounts for millions of jobs on both continents. What we're starting to see is even though interest-based advertising might have been negatively impacted, the money is just flowing to different types of digital advertising instead of interest-based advertising. So they're just reallocating the advertising dollars to other types of digital advertising, like look-alike campaigns, contextual advertising, addressable TV advertising. The irony is now a lot of the money may actually be flowing to social media.

If we move more to voice search and the algorithms can tell us the one thing that we need to know, the consumer has a narrower choice, right? So I have a narrower choice, but I'm still not getting compensated for that choice. Is this convenient for the consumer, therefore, it's valuable and we don't care that these goods are being sold in for machines?

John Bowman: I think the convenience is part of it. There might be some sort of value exchange in terms of sort of subsidised hardware. So, if you had a connected fridge, would you allow advertising on the display on that connected fridge shop for food that they know that you're interested in? Having the ads would subsidise the cost of hardware.
But this might present an ethical dilemma in that perhaps a “digital underclass” may need to rely on advertising to get the hardware they want whilst those who can afford the hardware upfront will not need the advertising.

**But you may not even know if you’re in that digital underclass, you may not even know that's an option, right?**

John Bowman: Well that might seem like a dystopian vision and might put us off now; it may become acceptable in 10 years’ time. We tend to accept all these things after a while once you get used to it, like the idea of CCTV and then the idea of mobile phone tracking, or even voice activated devices.

**Do you think the GPS or the phone tracking stuff slowed down after GDPR?**

John Bowman: Well, I think they do rely more on consent or a form of consent at least, whether it's effective or not I don't know. But there is usually a permission available to access location data, so you'll get a pop up saying this is tracking your location. And there seems to be a bit more choice and a bit more control over that now. Although consent isn't the only form of legal basis that you can use for data processing, in some ways it can work because you can say, I've got the consent and that's it. I think the other thing, of course, is that the EU has been trying to legislate around this area with the updated ePrivacy regulation.

**What's happened to the ePrivacy regulation?**

John Bowman: The negotiations have been going on for three years now. It stalled at the Council of Ministers Meeting at the end of 2019. The problem is replacing the 2002 directive, which originally was just dealing with telcos and the integrity and confidentiality of communications and so on. It got updated in 2009 with the with the famous cookie consent requirements. Then the European Commission said that they needed a complementary law to the GDPR effectively for things like Internet of Things, machine to machine, communications, network communications – so a level playing field for over-the-top communications providers in terms of their obligations as well compared to the traditional landline and mobile phone networks. And then it also tried to bring in some choices around consent and opting in and opting out of marketing, so ‘no legitimate interest’. For electronic communications, it's all consent based. But actually, the reason why this whole thing stalled is because the member states negotiating this were lobbied intensely by the media and ad technology companies because they're saying that effectively, this privacy regulation would be so restrictive, it's just going to undermine the business models that keep the online world running. So the ad tech side of things, whether you need to give consent to the cookies tracking. They were proposing that browser suppliers put in controls as well. Interestingly, in October 2018, companies started going a bit more public on this and, for example, Apple’s Tim Cook did a speech in Brussels at the international conference to the data protection commissioners, and he said something along the lines of, ‘Well, we do build these privacy controls into our software and into our devices and so on, and we think that's a good thing.’ Effectively, there was some sort of raising of the stakes in terms of saying, 'Well let's adopt privacy, let's say it's a good thing and, let's see what the competition says as well.'
So, where does it go from here in terms of the ePrivacy regulation?

John Bowman: Well, the European Commission has got to take it away and rethink it because the draft was just getting nowhere. And it reached the end of the road, basically. So I expect they'll come back with another proposal at some point, but it could take a while to design it. These texts are hideously complex. I think that's the problem with that particular draft. I think it was rushed out because they tried to get it in in time with the GDPR, and they missed the boat on that.

Do you know what prompted the ICO investigation into the ad tech industry and what the possible outcomes could be? What impact this could have more broadly on the industry?

John Bowman: I think there's been some grassroots advocacy on this. There's activists in the privacy community who have been remarkably successful like Max Schrems, for example, he was instrumental in ending the U.S.-EU safe harbor agreement, and is challenging the standard contractual clauses as an instrument for international data transfers. There's other organisations that have basically been concerned about the activities of advertising platforms, tech platforms, ad exchanges, and so on because they say that there's an issue with transparency, there's an issue with the lawfulness of processing, particularly data that might be sensitive, relating to health issues.

It seems quite negative commentary from the ICO: you guys get your act together otherwise, we will do it for you. What are their abilities to enforce?

Todd Ruback: Yeah, look, when we talk about the California Consumer Privacy Act, there’s a concern around this opaqueness in the ad tech industry, and that is why the CCPA really came into being. So, there is concern and there's growing scrutiny, and it's a vexing issue for a lot of organisations, how to think through the digital advertising revenue that they derive because of the ad tech ecosystem, through the lens of the CCPA. So, it's a challenge, no doubt about it.

John Bowman: The other thing to mention is that in the European context, these kinds of discussions have been linked with the market dominance discussions as well, so antitrust. Margrethe Vestager, the new Commissioner, actually covers competition and digital. In some ways, they're arguing that the Big Tech companies are using their market dominance because if they're data driven organizations, they've got more data than anyone else, which allows them to maintain their market dominance.

That is a question I had more on the U.S. side, but it'd be interesting to hear actually on the European side from what you know, in terms of Europe, how they think they can deal with corporates who have data at scale, is there anything they can do or is it shutting the door after the horse has bolted?

John Bowman: Under European competition law, the European Commission can issue fines of up to 10% of global revenue.

But that doesn’t necessarily mean that suddenly you're going to change your model?

John Bowman: I guess the threat that is always there is about whether they could break up the individual companies because one of the issues with the tech giants is that they have lots of different subsidiaries.
Do you think the European Commission believes GDPR has been successful in achieving what they hope to achieve and how are they approaching enforcement and implementation?

John Bowman: Generally, it's been considered to be a success, because I think so many companies took notice of it, and so many organisations did, and it did raise awareness of people's rights. The important thing about GDPR is that it is rights-based legislation. So effectively it derives from the Treaty on the European Union. It circles back to the Charter of Fundamental Rights as well. So there's the right to data protection, which is article eight and article seven, which is the right to privacy and a family life; I think the ePrivacy regulation was looking to address this as well. So, culturally in Europe, which is where it might differ from the US, is there's a kind of historical legacy which it builds upon the horrors of Nazism and communism and the sort of surveillance society and so on. Particularly in places like Germany, which has experienced both those things, and the idea that people should be free of state surveillance because some of the worst abuses of power have arguably been perpetrated historically by the state operators and government. In some ways, people shift their thinking about whether it's the state who are the bad people, or is it other parties now, like the big companies? So, there is a cultural background which I think plays out less in the U.K., actually, because we weren’t under those sorts of totalitarian regimes, but I think in the U.S., they don't really think of it from that perspective.

Todd Ruback: The EU and the U.S. have historically different starting points when it comes to the concept of privacy. In the U.S., it’s looked at as a contractual relationship between the organisation that wants to collect data and the individual. We have moved much closer to the EU perception of privacy than then where we started, and that's evidenced even by the expanding definitions of things like personal information or personal data. So, I personally think that privacy of our data is a fundamental right, and I think more and more people are sharing that over here.

It’s probably a good time to talk about the California Privacy Act. At a high level, could you talk about what this act will really change within the state around the approach to data privacy?

Todd Ruback: The reality is because the U.S., for whatever reason, can't get its act together and pass a federal privacy law, we've left it to the states. California, as usual is taking the lead. We have CCPA. It's similar in some respects and very different in other respects to the GDPR, but like the GDPR, it is forcing organisations to rethink their whole data strategies. It is forcing companies to take a step back and look at the data they collect in California and the states and know how it’s governed and make sure they know where it is and can access it to respond and comply with the CCPA. Because there are 50 states, one of the challenges is how do you do segment or fragment your digital business in the U.S. and only comply with and only make this relevant to the California consumers or do you standardise and make this your national standard. So a lot of companies are thinking through that.

What is the general approach?

Todd Ruback: It depends on if it is a regulated industry like financial services. They're looking at it as a standalone situation because there's 13 or 14 other states lined up with their own version of CCPA. So time will tell, but what I anticipate is similar to the data breach laws that emerged state by state, it started with California in about 2002. Fast forward to today, all 50 states have their own flavour of data breach notification law.
They're all similar with some nuances, like time to respond to a data breach or time to notify regulators. I anticipate at some point, the majority of states will have their own version of a consumer privacy law. I think it’ll happen rather quickly.

**So we could end up with a patchwork of regulation that is somewhat consistent with a few tweaks?**

**Todd Ruback:** That's right. And I think companies will say, we're going to adhere to the most stringent of the 50 states laws, and that way, we're complying with 80% of them.

**I feel like for the financial services firms, the default position is you have to be the gold standard because otherwise, you've lost business in one state?**

**John Bowman:** I guess many large companies have already adopted GDPR as a type of ‘gold standard’ so it might be more transferable over to CCPA. Briefly coming back to Europe again — it has this binding corporate rules system that basically allows for the free flow of data across the whole group, but only within the actual group itself. There are intra-group data transfers, but when you have hundreds of different offices all around the world and different entities, it saves them having to put in place lots of different contractual agreements. Now they're considered to be a gold standard as well, because they have to be signed off individually by the European regulators. But also, they contractually bind the participating entities in the company to achieve that same standard right across the board. So the BCR (binding corporate rules) is a way of having a kind of global standard, which you volunteer upon yourself, but it's monitored by the European regulators. Sometimes that can be applied not just in the U.S., but across the board as well.

**And you mentioned there’s a lot of similarities between CCPA and GDPR and some differences. Could you talk about what the main differences or the key ones?**

**Todd Ruback:** At a high level, the similarities are important to underscore as well. And I just want to circle back a little bit to the consequence of so many U.S.-based companies having to go through the GDPR compliance processes that CCPA was not nearly as painful as it might otherwise have been. A lot of the work and the workstreams and the effort for GDPR have been used for CCPA as well. The similarities are that there’s individual rights in California, including the right to access your data, know what it is, and have the right to request that it be deleted, whatever that might mean. But you know, the individual has the right to know what’s going on and the right to control it. And the other similarity besides the right to know and control is the obligation for organisations to be transparent about what they do.

Regarding differences, there's not an express or explicit accountability obligation. And there's no explicit requirement to do things like Privacy Impact Assessments. So the law itself was not well thought through or considered because it was conceived in and legislated so quickly, as opposed to the GDPR, which was heavily negotiated over three or four years. So in the GDPR, it was thought through, every aspect of it, and the CCPA, it just got pushed through very, very quickly without much negotiation or ability for organisations to comment on it. So there's lacking in the CCPA a lot of the express obligations or requirements that the GDPR has, but I think all of that will be fleshed out as we go through versions of the CCPA. We've got CCPA 2.0 that is being proposed already.
You mentioned the opaqueness of the digital advertising industry was one of the main triggers, are there other factors that lead to CCPA occurring and so quickly?

**Todd Ruback:** The main thing was what are companies doing with your digital data? Is it being sold? Are they taking your data and sharing it with third parties without your knowledge or ability to (a) know it and (b) control it? That was the main concern and that was derived from what was happening in the digital advertising industry. So the CCPA is really focused around the concept of sale of data.

*If we hadn’t had Cambridge Analytica, would we still have ended up with CCPA?*

**Todd Ruback:** My instinct tells me it had an effect on the timing of the CCPA, but I also feel if it wasn’t Cambridge Analytica, it would have been another story breaking

**John Bowman:** Take a step back to 2013. Another key accelerator is Snowden. The GDPR negotiations, which I was involved in at the time, were taking a long time, because it was almost like the e-privacy in that everyone was getting lost in the weeds and it was becoming a rather torturous negotiation. Snowden had a tremendous effect in galvanising the European Parliament and the European Commission to try and get this through. It did take a couple more years afterwards to actually get it through, but it gave it the needed boost to build up momentum again. And so, you do get these occasional game changing, catalyst events.

There’ll probably be something else which comes along, which will be the point when people start to think ‘Well, I didn’t know this was happening to my data.’

*There’s a lot of groups who are trying to get people together to think about like global ethical standards, but some of it’s very academic maybe. Do you think there’s another level?*

**John Bowman:** Europe has shown itself as influential with GDPR. Margrethe Vestager has been tasked to come up with some sort of AI strategy within the first hundred days of her commission tenure. Is there going to be a Greta Thunberg type figure emerging from all this? There may be some who are more successful than others in promulgating a message about privacy rights — you could say Snowden or Max Schrems, or a few others as well. But is there going to be one person who really changes everything?

*Will there be a backlash eventually?*

**John Bowman:** I thought of this term the other day — about people actually giving up these things — and you could have a sort of digital veganism I suppose. Someone else had already come up with that term about seven years ago now. The idea is that people actually go to rehabilitation places to detox their digital lives

*Back to CCPA. How have companies been preparing for that? And have you seen any notable variations by sector? I know, we’ve mentioned that some of the global companies were probably already quite well prepared anyway, given the GDPR approach but are there any sectors or areas of this particular legislation, which has been particularly challenging?*
Todd Ruback: The companies that already went through the pain of GDPR were best positioned, from what I saw for the CCPA, as so many of the workflows and workstreams from GDPR were transferable. So that was good news. And again, it depends on the industry – the regulated entities like financial services, they’re just so used to these large, complex compliance projects, and they just took this in stride. It was the smaller- to mid-sized companies that had to comply with CCPA that were not necessarily concerned about GDPR, they found it a bit of an earthquake.

To answer your question, companies were really good at getting the project plans and teams in place and identifying the workflows and getting budget and allocating the money and hiring their consultants and law firms, and they were good at thinking through the processes that needed to be put in place. What was hard to solve was digital advertising. Procrastinating is the wrong word, it was just a vexing issue, how to comply with the CCPA with respect to your digital advertising. Specifically the online third parties that are on your website that are collecting, sharing, slicing, and dicing anonymous website behavioural information in order to do X, serve an ad, or to support the electronic, programmatic auction system that digital advertising uses.

The reality is, companies haven’t been able to properly address that issue because of a couple reasons. Number one, after the CCPA was enacted, there were a whole bunch of proposed regulations that were put into play. And the attorney general, who is in charge of enforcement, proposed a number of regulations, and went on a road show to discuss these proposed regulations. And the regulations will not be finalised and published until after January 1, the effective date of the CCPA.

**What are the regulations specifically related to digital advertising?**

Todd Ruback: There were amendments and regulations across the board on all aspects of the CCPA; the regulations were meant to clarify ambiguity in the CCPA. It was enacted so quickly, within 30 or 45 days, so a lot of companies are saying I just don’t know how to comply with this particular aspect of it, we do not sell a product with online third-party tracking, and I’m waiting for the AG to give guidance so I can have clarity.

**Is the focus on selling data under CCPA going to be used like legitimate interest under GDPR? It’s not a loophole per se. It’s not supposed to be, but it’s open to different interpretations, or how do you think about that approach to it, versus how it’s supposed to be interpreted?**

Todd Ruback: A lot of organisations saw the service provider exemption as a kind of residual catch all if you will, as a way to get around the sale analysis. The digital advertising industry is actually taking a position at large that many of its players in that ecosystem could constitute as service providers, which would mean that there is not a sale occurring, so your legal obligations are not triggered with regard to sale. I guess the AG will provide clarity in due course and answer that question, but a lot of lot of organisations are trying to maximise the statutory exemptions as far as they can go, and even beyond.

**The AG might come out with the clarification or is there another regulatory body who might say that this interpretation is or isn’t as we would have expected?**
Todd Ruback: It would be the Attorney General. I saw some quotes on Twitter and LinkedIn, where the AG, and I'm paraphrasing here, was saying, I'm going to enforce vigorously and I'm going after low-hanging fruit, and don't think for a second that this is not top of mind; we're gearing up — so he gave fair warning. But again, he has not given guidance on this issue. So yes, we're still in the same place, and I think the lack of guidance is actually guidance and that is a statement I've been sharing with clients. If he was going to come out with a position, they would have two weeks ago. So take that for what it's worth. The other thing, just to answer your question, is the plaintiffs' bar in California. They're gearing up. This is a huge opportunity for class-action litigation in California. We view online third-party tracking as an opportunity for plaintiffs' class-action litigation.

What are the mechanisms to enforce CCPA? Financially, what kind of impact could this have for a corporate owned entity that's not being compliant or not yet complying?

Todd Ruback: So there's statutory damages of I think up to $7,500 per violation that the AG can impose. What does that mean? It sounds pretty limited. Remember, the statute, I believe, and I'm paraphrasing, is up to $7,500 per violation. So in the context of online third-party, we don't know what 'per violation' means — that could mean each time data is shared from one third-party to another. We don't know, but what we can deduce is a violation will not be a one-time occurrence, it will be a matter of process. If the AG is looking at online third-party activity, then processing will be in the millions. So conceptually, it could be $7,500 times 1,000 instances of sharing.

The second area of potential damages, besides what we have just mentioned, is when it comes to data breaches. That's when there can be private litigation -- the plaintiffs' bar is saying this opaque online activity could be looked at as a potential data breach. A lot of people have been talking about this for years. So when a consumer says, I'm opting out, don't sell my data, and don't send it down the daisy chain, that opaque ecosystem that I don't know about. If a company doesn't listen, then the legal construct is that activity is considered sharing unauthorised access to an individual's personal data. And the data breach law in California says a breach of data is unauthorised access of someone's personal information.

In terms of the definition of personal information, is it similar to GDPR?

Todd Ruback: Going back to our earlier conversation, this is what I think this wonderful, I think it is why we are moving closer to an EU set of nomenclature. Under conceptual understanding, the definition of personal information or personal data has greatly expanded from a finite set of enumerated elements. Ten years ago, personal data in California was a list of nine or 10 things right? Name, address, Social Security number, bank account, that sort of thing. Now with CCPA, it includes anything that can reasonably be used to identify someone, including things such as IP addresses, cookies, or device identifier — so, boom, all cookies by definition collect at a minimum IP addresses and browser information. So right there the CCPA is saying third-party online data collection is collecting personal information.
The telco and tech industries have been pushing for a federal law. Do you know what the rationale behind that is?

Todd Ruback: So telcos and the large financial institutions were pushing for a couple reasons—the same reasons that actually a lot of large multinationals wanted the GDPR. It's easier to comply with one law than 50 laws; it is easier to comply with GDPR than 28 different entry-level directives. So ease of compliance, they can centralise their compliance dollars, and it's easier for monitoring, testing, and auditing and documentation, but also, this way they can control the narrative with regard to exemptions. It's certainly better for a telco or financial institution to have a blanket country exemption, than a patchwork of state laws where they have exemptions in some of the states but not in others.

What appetite do you think there is for a federal law? And do you have any sense of what that might look like relative to say GDPR or CCPA?

Todd Ruback: About a year ago, I thought we had a pretty good chance of finally getting federal legislation. I thought the CCPA was going to be a catalyst for Congress to actually get something together and over the finish line. There were probably six or seven different really viable federal initiatives going on, one of them was led by the Department of Commerce. It kind of died away with everything else happening. Recently, one of the senators from Washington proposed a new federal bill, which was met pretty favourably. I can't handicap it, but I'm in favour of it and hope it gets traction.

When you think about sectors and companies and how they're positioned, do you think any of them could end up being custodians of data? From your perspective, do you think there's any sectors that could end up being stewards of data and consumers would be fairly comfortable with that?

Todd Ruback: I think to a certain degree, financial services already plays the role of data stewards. And that may continue to grow as online and digital banking becomes a bigger part of their business. And John mentioned the cloud environment and cloud providers. Is there anything else that is really pressing for the whole data privacy industry at the moment?

John Bowman: The one thing I'd say is that you've got to expect the unexpected because something will happen. And it might be in a year's time, might be in five years' time, but there'll be another game changer, somewhere along the way, which takes us in a different direction. What other things are emerging that are not necessarily regulated or which are difficult to regulate? So, once you start getting artificial intelligence-based decision making that has some sort of adverse effect on groups of people, then that could be the next big issue.
A Conversation with Anne Fealey
Global Chief Privacy Officer, Citi

Talk a little bit about your experience having been at Citi for a year.

Let me start by saying that privacy and protecting personal data is on the top of senior leaders’ minds at Citi and I believe is a top priority here, which is good.

As you’ve come into a huge franchise like Citi and you’ve inherited a massive personal data protection program, what are some of the challenges, what are some of the things groups within Citi have coalesced around?

The company and people at the company are working towards the same goal to ensure that we keep the personal data of our clients, customers and employees secure and confidential. The challenges for me have been the size and learning the Citi businesses and structure while trying to drive changes in the data privacy program.

And with that, obviously we have 200,000 employees and 100,000 contractors in the firm and millions of customers. As you’ve thought about that data architecture, have we had to do a lot of data governance work to get that architecture right? Have we had to upgrade a lot of our procedures for managing PII? How have you seen that play out since you’ve been here?

The review and updating of our policies, procedures and processes are part of an ongoing build and refinement within the data privacy program. When you have changes in laws or changes in technology that increase the type of personal data being created, a regular review of and tweaks to the data privacy program are needed. If you look back 20 years, and compare where we were then with respect to the use of technology and the Internet, the datasets we were looking at then are very different than the datasets we are looking at today. So we have to be pretty nimble and aligned across the company in managing the program and data privacy protections. It’s not just me or my team, but a partnership with stakeholders across the company to help ensure that we have the right controls and the right governance over personal data.

The roles of the Chief Privacy Officer (CPO) and the Chief Data Officer (CDO) — are they evolving as rapidly as the industry of data is evolving? How do you continually define and redefine your role and how you operate? Do you think we have a best practice at the corporate senior executive level now with these roles?

I believe the CDO role is broader in a large corporation like Citi because it covers much more data than personal data. My role (the CPO) focuses on personal data. I also believe that both roles necessarily have to evolve all the time because of the new technology and the new data being created every day.

Let’s look at the U.S. data privacy law landscape and where we are. California’s Consumer Privacy Act (CCPA) is live as of about 8 weeks ago; as you look at the U.S. legislative landscape, obviously it’s an interesting year here with an election coming up, what states do you see with proactive legislatures, which state are you thinking about next, are you thinking more nationally? How do you view the U.S. privacy legislation landscape at the moment?
California’s new privacy law (CCPA) is definitely a game-changer, and I think across the U.S. we’ll start seeing more and more state privacy laws like the one in California. And there are likely to be changes that bring the California law closer to the privacy law in Europe (the GDPR). It’s not necessarily ideal for the U.S. to have a patchwork of state laws addressing the same issues, because they could be addressing those issues differently with different requirements. It’s always tough to abide by all those laws if they’re different. But that’s likely what we’ll see. The way that Citi approached CCPA compliance ensured we built solutions that covered adherence with the CCPA, but could be leveraged across the U.S. and globally, since we’ve had legislation passed in Brazil, there’s new legislation in Thailand. We’re seeing a lot of the same type of privacy laws being passed across the globe, so we need to be able to respond as a company and not necessarily respond one way in California and another way in Brazil and a third way in Thailand.

_I was just going to ask about, in terms of the privacy laws and regulations, particularly California, the GDPR, now Brazil, Thailand, etc., do you think these are actually fit for purpose? Do you think the legislators and regulators got it right?_

I think we sometimes see certain aspects of these laws that conflict with other laws. With the GDPR, for instance, the idea around the ‘right of erasure’ or the ‘right to be forgotten’, goes at the heart of what privacy means for a lot of European regulators. But if you compare that right to the right of a free press to report on things, you have to decide which of those rights trumps the other when an individual wants a newsworthy story erased or forgotten. So that’s always challenging, when you have lawmakers passing laws that conflict with other laws.

_What do you think in terms of possible tension with respect to data localization laws?_

Data localization is probably the one type of law that I think most companies would object to. In this global economy, data are stored on the cloud in servers located in various countries and accessible by people in different countries. To require that certain personal data be kept and accessed locally in a single country, doesn’t make good business sense. And I’m not sure it achieves the goals of the lawmakers to provide greater data privacy protection.

_Do you think that culture drives a country’s view on localization?_

We’ve seen it in South Korea and Russia and I think the proposed changes to India privacy law also have an aspect around localization, so I’m not really sure. You can imagine the challenges across the globe if every country did that. It could be that these lawmakers believe it will keep businesses in their countries; I’m not sure that’s true, but it’s definitely something most companies would not support.

_The technical environment probably now is better than it ever has been to anonymize and protect data. When you look at some of these anonymization technologies, and people talk about homomorphic encryption, encryption techniques, etc., do you think that suite of technologies or techniques are helpful in this overall mission? Are we exploring some of them, and how do you think about that in terms of using the technical tools to solve some of the problems?_
Generally technology is always going to help, not just with de-identifying or anonymizing personal data, but it will also help with finding personal data and protecting it. As a company, our technology team is always looking at technologies that we can build and use to help our data privacy program and help protect people's personal data.

How involved are you in some of the discussions around ethical AI and how individual data can potentially be used to drive certain decisions in machine learning? Is that something you get involved in, or is it something you’re aware of? Any thoughts there?

It is something that I’m aware of, though not heavily involved in yet. When you talk about true artificial intelligence, where decisions are being made automatically by computer algorithms without human intervention, basic questions come up around what kind of information are we using, what decisions are being made, are these decisions correct, etc. We’re seeing across the U.S. and globally, lawmakers and legislators looking at facial recognition technology and asking whether it should be banned. Artificial intelligence provokes these types of questions and so necessitates an ethical framework. I think we’re just starting to see the reaction to some of this type of technology from lawmakers.

Anything else we should be thinking about?

We are in an age where new technology creates data all the time and we’re in a situation where it’s hard to catch up with that technology. I think taking a step back and having that kind of ethical mindset already mentioned as well as an accountability framework around the data privacy program is going to be key.

Citi launched a data privacy impact assessment tool this year — is that something that Citi will be using to assess vendors? Will this lead to a reduced number of third-party service providers?

I think everyone would love to reduce the number of third parties as much as possible, but the use of the privacy impact assessment in the third-party review is really just tightening up the review of controls relating to data privacy risks with third parties. We continue to work on linking the third-party reviews and the PIAs so that the process is aligned and not duplicative.

What do you think the biggest challenges are across the bank in terms of managing and utilizing data and changes that have to be made?

That's always a tough question to answer. I believe that people’s expectations from companies, including banks, have increased with respect to their personal information. One area in which we see that is in our consumer business, where our individual customers expect us to be able to tell them what personal data we hold about them, when and with whom we share that data, and how they can access and correct the data. Whenever you’re in a consumer-facing business, I think your challenges are always going to be greater than in a business-to-business (B2B) environment, just because of the amount of personal data involved when working with consumers. The other challenging area is with Human Resources (HR). If you think about all the personal data that HR needs for employees, and the third parties that HR works with for healthcare and payroll services, then you realize that would be another big challenge area.
A Conversation with Emma Maconick  

Partner, Intellectual Property Transactions Group, Shearman & Sterling

How does the existing landscape look in the U.S. from a data protection and privacy standpoint?

It depends where you’re looking from, or the perspective that you bring to this area. If you’re looking at the U.S. from a European vantage point, or from the outside looking in, it can look fragmented. Books and articles often start by describing the U.S. privacy field as a “patchwork” of laws and regulations. I think that actually does it a disservice, because it sounds like U.S. laws have been cobbled together without much thought. It’s more helpful to think of the U.S. as coming from a more capitalist view. Its perspective is that data is a kind of asset, something of value, and personal data can be an individual asset or an asset on the balance sheets. Increasingly, it’s represented in the form of goodwill or something similar. Europe’s perspective, largely because of its history and how it views personal freedoms, is that privacy is a fundamental human right. So Europe has a much longer history of legislation and a much more comprehensive approach to privacy than the U.S., but it’s both dismissive and overly simplistic to think of the U.S. field as slapdash or completely unorganized.

The U.S. also seems to view data as free, where people can generally do what they like with their information, and regulation should focus on those sectors of the economy that where information is particularly sensitive — information about children, about health, financial information and to some degree educational records (which are often also about children). That’s one way of making sense of the historically sectoral approach to data privacy in U.S.

But now you are seeing in the U.S. some tendency and attempts to try to harmonise this. I’m not sure if it’s necessarily going to work. In California, for example, the big news has been the California Consumer Privacy Act, which is quite different from what's happened in the U.S. before. First, it has a potential extraterritorial effect, which is necessary to capture activities in other states in the U.S. But if a company processing or collecting the information of California consumers, and you meet the definition of a business as defined under the CCPA, then that will be subject to the CCPA, whether it’s a Swiss bank or sells hardware from Colorado.

Fair information protection principles have also be gaining traction in the U.S., which may sound more like a Europe approach. But it still looks like any grand unified law of privacy is far off in the U.S.

In terms of the main focus of the regulators, or I guess what, what triggered CCPA being introduced?

The ostensible direct trigger was Cambridge Analytica; that's sited in legislative materials. Not long before that was the Equifax data breach, and those were both in a long line of significant data breaches.

But this is also a feature of California, where anybody can introduce laws through a ballot measure. If you get around 600,000 signatures, you can propose new law through a ballot measure. When Californians vote in elections, they also vote for ballot measures, and if a ballot measure passes, it becomes California law.

Although for the CCPA, it wasn’t such a straightforward path.
Alistair Mactaggart is the lead proponent of the CCPA. He’s a real estate developer based in Northern California and he put out a ballot measure who put forward ballot measure partly in the wake of Cambridge Analytica. But apparently it was also spurred by a discussion that he had had at a cocktail party, where he was talking to an engineer of a large technology company who told him, ‘If you really knew how much data we have on you and everything we know about you, you’d be freaked out.’

That got Alistair Mactaggart thinking, and he proposed a ballot initiative to give consumers rights over the use of their personal information. When his ballot initiative got the 600,000 or so required signatures, California voters would have the chance to vote on this new privacy law. The California legislature was rather alarmed by this, and there was a bit of a mad panic where the legislature tried to draft the substantive provisions of the CCPA over several months – rather than have the law adopted through a ballot measure.

There are aspects of the CCPA that are elegantly crafted. There are other aspects that were less so, with unintended consequences, or where legislators had not fully addressed how these laws would be operationalized. So there was a fairly protracted process of getting public commentary and input from interested parties, from industry, from the public, from commentators, from other government agencies, about how to improve the CCPA. The law was amended multiple times, and the California State Attorney General also published draft regulations in October 2019, which are intended to help clarify ‘how are you going to do it’.

You mentioned that some of its well-crafted, some of it has unintended consequences. Could you talk about what the key requirements of CCPA are and how that differs from current legislation? I know you said it’s fragmented so it’s probably hard to say how it differs from current regulation, but what are the real changes that have been brought to the fore?

There are a few things. One of the thoughtful points is that it recognises existing data security laws, like the Gramm-Leach-Bliley Act, which addresses privacy of certain non-public financial information, and HIPAA, which addresses medical privacy. There are specific exceptions in the CCPA that say that certain kinds of data isn’t covered, or certain provisions don’t apply, recognising the fact that there are already laws and regulations in the United States that are intended to cover those kinds of personal data. That is fairly thoughtful, where businesses that are already subject to U.S. privacy laws can continue their current practices.

As for the key requirements of the CCPA, the first thing you have to figure out is whether or not a business is subject to the CCPA. A “business”, defined under the CCPA, is essentially a for profit legal entity that collects consumers’ personal information, and determines how that information is processed, and that does business in California. A business also needs to meet at least one of three thresholds: it has to have annual gross revenues of more than $25 million, or it collects more than 50,000 consumers, households or devices worth of data annually, or it derives 50% or more of its annual revenue from selling consumer’s personal information. That third threshold is directed to data brokers and data aggregators.

There are analytical details in all of those steps. For example, there are discussions about what constitutes a household and what if people have multiple devices. My house probably has 30 devices that collect my data. Is that one household? Am I still one consumer, etc.?
In borderline cases, these are probably not places to get too cute or too creative in your arguments. I think you have to take a reasoned approach here.

After you determine whether a business is covered, then you need to take a look at whether a person is covered. “Consumer” under the CCPA essentially means a natural person who is a resident of California — meaning somebody who lives intends to stay in California.

And the activities regulated under the CCPA are the collection and sale of personal information. This is one area where people are especially exorcised because that definition of sale is extremely broad, capturing any making available, transferring, renting, releasing, disclosing, etc., of personal information. A lot of analysis has gone into determining whether an activity is a “sale” under the CCPA.

I saw an article last week about the definition of selling data and who it applies to. Is it a bit like legitimate interest under GDPR, an area that can be seen to be open to interpretation?

I think that it is. There is a lot of very careful thought that goes into whether or not you’re selling data. There are entities whose entire business is selling data, so we're going to take them out of the scope of this discussion. But think about a platform that connects users with providers of goods and services and collects data from consumers in that context. If that business makes the data available to or transfers it from the purchaser to the provider, that’s not necessarily deriving economic value from the transfer, but it could still fall within “sales” under the CCPA. The CCPA’s definition of sale is that personal information is provided to another business or third party, for monetary or other valuable consideration. That’s an extremely broad range. If there’s direct payment for the information, then it makes sense that that will most likely be caught as a sale. But how should we understand “other valuable consideration”? There has been quite a bit of debate on that point, and different people can take different perspectives.

It is also important to understand is that selling does not include certain disclosures; there are specific exceptions from the definition of “sales.” If the disclosure is at the direction of the consumer, where the consumer actually instructs a business to do it, that’s not a sale. It's also not a sale when a business shares personal information to give effect to an opt-out, and again, that's partly directed by the consumer. There’s also an exception for transferring personal information to a service provider, which is used quite often, and another exception when personal information is transferred as an asset in a corporate transaction.

When it comes to whether a transfer is or is not a sale, the devil will be in the details, and in how the business is structured.

In terms of how it differs from the current regulatory landscape, and that whole concept of being about the sale of data or providing data for monetary or valuable consideration, is that a completely new concept within the area of data protection?

Yes. If you think about the changes from the previous California regime, there is much more required to comply with law under the CCPA than was required previously. More attention needs to be paid to the collection, processing and storage of personal data, and you need to be very granular about your business operations to understand how the data flows — where you receive data from, the purposes for which you use it, who you show it with, how you transmit it. All of this requires a greater allocation of resources and time.
You should also understand the security around your business’s personal information, even though the CCPA is not a data security regime.

Another entirely new aspect for California law is PIARs, which are personal information access requests. Under the CCPA, consumers have rights they can exercise by a verifiable consumer request — the right to know, the right to delete, the right to information about a business’s data practices, the right to opt out, etc. These are similar to rights granted under GDPR, but there are differences between the laws. So companies that are operating under both CCPA and GDPR will need to keep differences between the data regimes in mind.

Consumers also have a right to notice at or before the time of collection under the CCPA. And they have a right to non-discrimination, meaning that consumers can exercise their privacy rights under the CCPA and not be discriminated against by the business, including in pricing.

Critically, consumers also gain a private right of action against businesses that suffer a security breach that results from the business’s failure to do certain things. That’s quite new under U.S. law.

*Is there a specified fine level or penalties?*

Enforcement can take one of two separate forms.

On one hand, the California State Attorney General can enforce noncompliance with the CCPA themselves — and only the AG can enforce the CCPA requirements. There is also some process around this. First, the AG is prohibited from bringing any enforcement actions until 1st July 2020. Second, for any enforcement action that the AG takes, there is a bit of breathing space. The AG’s office must first provide the entity with notice of the alleged violation or nonconformity with the CCPA, and the entity then has 30 days to cure it. And if you can cure issue raised by the AG, there should be no further action from the AG. If you can't cure, then the AG can bring a civil action with civil penalties of up to $2500 per violation, or $7500 per intentional violation. Not sure how much enforcement activity we’re going to have, but that’s to be TBD. So that's the State AG enforcement powers for the obligations under the CCPA.

Separately, the CCPA also provides a private right of action for individuals, for California consumers — but only if the consumer’s personal information was exposed in a data security breach. So there’s no general right for Joe and Jane Smith from California to tell a company: ‘We don’t think you've done enough in terms of your data mapping exercises, and you’re not securing data adequately, and you haven't responded to my request to delete my data, and also your website privacy policy was not displayed to me at or prior to the point of collection,’ and then file suit against that company. The personal right of action only comes into play if there's been a security breach that exposes Joe and Jane’s personal information, and to recover damages, they also need to show that the breach results from the business’s violation of its duty to implement reasonable and appropriate security procedures. Of course, what the duty to implement security procedures requires will be debatable in individual cases. (Though, as an aside, in FTC enforcement actions we do see companies held to account for failures to implement basic security measures, ignored warnings of security vulnerabilities and other obvious lapses. So there may be cases where violation of that duty is less controversial.)
If a California consumer succeeds in a private action, then under the CCPA provides statutory damages of at least $100 per consumer per incident, up to $750 per consumer per incident. Actual damages are also available, but it’s been historically difficult to show actual harm in data breach cases. And we can see how the statutory damages can scale up very quickly.

So, it’s limitless to some degree, depending on how many consumer incidents?

It could be; there’s no limit in the statute. And I think this freaks people out a bit. Somebody did the math on CCPA fines available for the Cambridge Analytica incident, if it had been in force at the time, and it would have been orders of magnitude greater than that the fine the FTC levied against Facebook. You see the same thing in other data security breaches, which are quite common in the U.S., and, frankly, throughout Europe. Data security incidents happen all the time, but historically one of the difficult hurdles for plaintiffs in US courts was showing harm from a data security incident. The statutory damages available under the CCPA appear to solve that problem for California plaintiffs, so that's kind of a big deal.

What are you seeing in terms of how companies are going about preparing for this new regulatory landscape in the U.S.? Or how they, how you think they should be approaching it?

The short answer would be: As best as they can. There’s a broad range of responses to the requirements. It depends on the maturity of the company meaning how much previous exposure they’ve had to working to operationalize data privacy laws. Is it a large global or multinational company that's doing business in lots of different jurisdictions and already subject to data privacy laws? Or did a local consumer-facing business just barely meet one of the thresholds, and it’s never thought of itself as having that much data, even though it actually does. There’s an enormous spectrum of compliance capability here.

That said, I've seen tremendous efforts made by many of our clients; they're quite aware of and focused on this, and there’s a real effort to try to get it right. In many cases, companies are working on retrofitting existing compliance processes to meet CCPA-specific requirements, and even that can be a big lift, especially in a big organisation with many different sources and uses for data.

What do you think the main challenge has been for companies trying to get ready for this and ensure they’re compliant?

I think the main challenge is getting a handle on their data, understanding where personal information comes from, what’s done with it, where it goes. If a company doesn’t already have that, it’s very difficult to build, especially for mid-market through large public organisations. Interestingly, smaller organizations have an easier time with this, where there’s less volume of data and less process to understand in order to see what the data sets and data profile looks like.

But this has always been step zero – and one of the biggest challenges – in any data protection compliance process. It doesn’t start with legal processes, but just figuring out what data is there, so that you can understand what legal processes are needed. When you do the data mapping exercise, be thoughtful about getting complete and accurate information, because that will put you in the best position to understand your compliance requirements and how to prioritize resources.
The next challenge is interpretive. The CCPA is a recent body of law with limited guidance. We have draft regulations from the AG, but it’s nothing like the Article 29 Working Party or the European Data Protection Board, or the 50 years of law that preceded GDPR in Europe. The public discussion and consultation period has also been quite limited in California. This means that there are a lot of questions facing the industry without clear answers. Is this in scope or is it out of scope? Is this actually a service provider? Is this a sale? Or how do we implement a verification processes without creating additional risk from phishing exercises? Operationalizing these details requires careful and detailed thought and an enormous amount of consensus and participation by multiple stakeholders across the business.

Finally, the pervasive nature of CCPA compliance, and more generally of data privacy or data governance laws requires a significant shift in mindset for many U.S. companies. Clearly, some U.S. companies are way ahead of the curve, and many of those are actively involved in crafting legislation, both at state and federal levels. But a lot of other companies are grappling with these issues for the first time, and need to think through how to actually implement this. A lot of people are probably thinking, “Hang on a minute, I have to apply this to my business. Nobody was thinking about my business when this law was being written, and now it doesn’t fit.’ With data privacy laws, the challenge is often figuring out how the law applies to your business and your data.

You mentioned earlier about maybe there’s some unintended consequences in terms of how this law is being drafted. What are those?

There were a few unintended consequences in the CCPA’s original drafting. One of those — which has now been fixed — came from the definition of consumers as essentially California residents. Initially, the law defined “consumers” as California residents. It didn’t distinguish between people acting as consumers, in the sense of buying things or using online services, and people in their capacity as employees. So, California businesses had a pretty allergic reaction to having the CCPA apply to employee data, where there’s already a very different set of rights and obligations. The legislation makes much more sense as consumer protection measures. But in many ways, it does not make sense when applied to how we use employee data. As an employer myself in California, it caused quite a bit of heartburn. The state legislature was open to hearing public commentary on that and feedback from industry. And toward the end of 2019, the law was amended with an exemption that applies to employees, contractors, agents, etc. from key requirements under the CCPA. The exemption is in effect until the first of January 2021, with a widespread understanding that something will be more permanently put in place to address how employee information will be handled. Though even with the exemption, California employees are still entitled to certain notices about what data being collected by the employer and the purposes of collection, but you don’t have most other CCPA obligations. That’s a one of the relatively simple example of unintended consequences of the CCPA.

Another example was addressed by the business-to-business, or B2B, exemption that added to the CCPA by amendment around the same time as the employee exemption. Here again, the CCPA was drafted in a way that it would regulate completely legitimate business to business communications, in a way that I don’t think was the intent of the law. If collecting information about the director of a company in the course of doing due diligence, for example because you were considering an investment in that company or because you want it to be a joint venture partner, or you just wanted to get some products and services from that company.
Under the CCPA as originally drafted, you would have had to give notice to that director. That director would have had rights to request deletion of the information, which doesn’t make much sense from a business operations perspective. So the one of the amendments to the CCPA added a B2B exemption, excluding personal information about individuals like directors, officers, employees, contractors of a company, or an entity, when the information is used in the context of either doing due diligence or buying or selling goods and services. I think that’s logical.

There were a handful of other unintended effects, like whether personal information is publicly available or whether vehicle information was included when in information about individuals. Another amendment made it very clear that de-identified aggregate data is not personal information for purposes of the CCPA.

Thinking about maybe the next steps, there’s been lots of news flow about possible federal laws, data privacy laws being proposed. What do you think the chances of there being a federal data privacy law in the U.S. and do you have any sense of what it might look like? Is there any sort of consistent theme across all these various proposals that have been suggested?

I’ll go out on a limb and say that I think the chances of having a federal law are good. But the chance of it happening anytime soon are much less good. There’s quite a bit going on right now in Washington, even more than usual. Plus we’re coming into an election year so attention is stretched thin. I think that, in many ways, the CCPA coming into effect and compliance efforts around it raised the profile of data security and privacy compliance in the United States. This isn’t just California, too. Another half dozen or so states that have some kind of data privacy law either proposed or pending before the legislature. Some appear likely to come into effect.

This has become a concern — I think rightly so — for both the government and businesses. We spoke earlier about the U.S. patchwork of laws, where we have sectoral or industry-based federal level legislation on top of state level legislation. If more states pass laws on par with the CCPA, we’ll have a very complex additional layer of state level legislation that where laws can have extraterritorial effects, that impact entities not in the state or not necessarily having a nexus with the state. Companies could feel fairly overwhelmed by that: Which of these regimes do I comply with? Or if one requires some measure that another law prohibits, how can I reconcile differences between the two?

So some sort of federal privacy law will be needed, but trying to predict the form of federal data privacy law is extremely difficult. We’ve had an enormous range of federal proposals. Senator Marco Rubio’s proposal would have essentially pushed a rulemaking responsibility to the FTC. Senator Brian Schatz put forward a proposal that would establish a fiduciary duty for online providers. At the same time, there are still more targeted legislative proposals, restricting law enforcement from using facial recognition technology or establishing security standards for Internet of Things devices. All of these things relate to data, privacy and security, so it’s quite difficult to make one law to bind them all.

I think the solution is to take an approach based more on principles and less on prescriptive requirements. This kind of solution applies, not just across the 50 states plus territories of the United States, but it addresses a global issue. Although there’s a great deal of variance in specifics, most data privacy laws have similar underlying principles that define good privacy practices.
For example: choice, notice to individuals, limited use (meaning only using data for the limited purposes of its collection), data minimization (meaning you’re not collecting, using, or retaining more data than needed). Also, the data that you hold should be accurate, with reasonable and adequate protections, like using de-identification or aggregation wherever practicable, and there also needs to be some form of enforcement, with liability for entities’ misuse of personal data.

At that level of abstraction, I think it’s possible to have a unified federal privacy law and my sense is that will probably end up being the focus on Capitol Hill. Of course, the devil will still be in the details — look at long it took to put the GDPR in place. This is not the year to do that in the U.S.

Do you think the outcome of the election will have a big impact on either the likelihood of a federal law happening, but also what it looks like? Do you think there will definitely be one at some point, it just may not be soon?

I think it’s likely that there will be a federal law, because of the potential for costly fragmentation, though the timing of that, as I said, is uncertain. I don’t see anything happening in 2020. I could be wrong, because there’s certainly a lot of activity on the Hill about this, in terms of panels and legislative hearings. The composition of the government may affect the timing. But the other influence is likely to be what happens in California as CCPA develops, which could provide the impetus and may influence the scope and content of federal privacy law. Senators and others on Capitol Hill are taking what happens in California very seriously.

How do you think governments, more broadly, moving beyond data protection, given it’s all about data at scale is increasingly an asset for corporates, how do you think governments might look to regulate that in some way? Either increase transparency or limit the advantages versus disadvantage you get from access and utilisation of data?

It’s a really complex issue. Different people have different perspectives on it, and I don’t necessarily advocate for one way or another. Some kind of direct regulation of personal information that’s held by corporations is likely. The question is, what form will that regulation take, and will it make sense? Antitrust is the current favourite medium of regulation, and there’s also the approach of viewing data through rights in property and applying that jurisprudential logic. It remains to be seen how effective either of those will be at regulating what companies can and can’t do with data. It’s not clear whether those bodies of law make sense for what regulators appear to be trying to achieve. These areas of law may not have the logical elasticity to achieve effective data regulation. It’s not necessarily what they were drafted to do. I understand the arguments that are put forth in both cases, but they start to fall apart at the seams in some cases.

It seems to me that consumer choice and transparency is probably the better way to regulate, where people can make informed decisions about what happens with their data. A lot of what the CCPA is trying to achieve is to raise people’s awareness about what happens to their data — to make people understand that if they’re not paying for a service, they are likely the product, and that they often purchase convenience with personal information. That’s not necessarily a bad thing. People should to make an informed choice. And then the CCPA creates consequences if that information or transparency is not provided or if the information provided is not followed. That may be all it takes.
This sounds simple, but it’s not. If you’re really trying to protect the consumer from privacy infractions, then that’s a good approach. But that’s a very different analysis from the understanding that data has economic value, and concentrations of data may provide an economic advantage. Those are very different problems to solve, and there’s a tendency to conflate the two, which I think need some work.

How much awareness is there amongst consumers around the change in regulation that’s coming in and therefore their increased rights to limit the usage of data, or at least have a bit more control over that? Do you think the consumers have been well educated or is it still under the radar once you move beyond the specific industries and practitioners?

Silicon Valley is the little l bubble where I currently live, and I think people here are very aware here. That’s partly because it’s generally a tech literate society, either part of the technology edifice or working with it or servicing it. So there’s probably a higher degree of awareness of these things here. I think there’s certainly a generational segmentation as well, about the extent to which people feel activated by privacy issues and the societal implications to protection of privacy. The idea that privacy is freedom, I think that resonates more in Europe than it does here, with the caveat that for the younger U.S. generation, and likely those sectors of U.S. society that have been historically disenfranchised, that message certainly resonates much more strongly.

Outside of data intensive industries, the CCPA is probably more of a shrug — as in large industrials, chemicals, hardware, farming etc. Those are actually still data intensive industries, and extremely sophisticated industries, but those are also areas where privacy and individual consumer rights are less of a direct concern or less capable of commercialization. It’s not so salient in those fields.

I also think, however, that privacy issues in general are getting more national and international attention. Whether or not this is something that you do for a living or something that’s a necessarily part of your industry, I think the average human being is much more aware of privacy issues now than they were even five years ago, even if they’re not aware of the CCPA specifically. While people may, choose, on an individual level, to provide their personal data because it serves a purpose; like it’s convenient to have products and services provided to you. But on a larger, collective level, there’s more awareness — coming back to data breaches and Cambridge Analytica — that something needs to be done about how that data is used. Even though they’re not quite sure what that something is.

In terms of corporate managing, varying requirements globally, given those different approaches, how do they go about that? Have most of the global corporates who operate in Europe just applied a GDPR type approach across the board? Or are they implementing different approaches across different markets depending on the requirements in those markets?

This depends on the entity. I’ve worked with U.S.-based entities that saw GDPR coming in and decided, ‘Look, we’re going have to geo-fence because we’re not in a position to comply with GDPR.’ They have made a rational business decision based on the compliance lift and the change in their operations that would be required to comply with GDPR and decided that it’s not worth the revenue that they generate from Europe. There’s not a direct correlation between revenue and risk, but that’s the analysis performed. That said, a lot of our clients are global, whether they’re product companies or financial institutions, and many of them have large legal and compliance departments who focus on this daily and consider the minutiae, the high level requirements, and also look at compliance issues at a practical level.
I think the larger, more well-established companies that have more sophisticated data governance programmes, tend to take a principles-based approach. Those companies try to implement good data use principles and then use those principles to develop policies and procedures — and values — in their approach to privacy compliance.

You can't realistically have perfect compliance with 47 different countries and 50 states worth of laws. So instead you can take a principled approach that says you won't collect more data than you need, and you'll have a thoughtful data deletion policy. Then you can be very transparent about your data practices and stay at the high level. I think that's the best you can do as a global entity, to establish global data governance principles and devolve the application of those to the business units, or sometimes to the product owners. And I do see that in practice.

Anything else you think is of importance that’s going on in this field that we haven’t covered?

Watch this space. It's only just beginning.

I don’t think we’ve even started to solve — at least not at a legislative level – the really complex issues around artificial intelligence, facial recognition and deep fakes, or similar types of issues. There will be a lot more complexity there, and these technologies will probably develop faster than anybody currently recognises.

Also, we spent the better part of the last hour talking about privacy, but I think security is as much if not a bigger issue. If you look at data as an asset, or data as a product, you can't really develop that data-driven product without mature privacy practices. But once you develop it, all of the value from that data driven-product is completely stripped away if you fail to secure, not just the intellectual property around it, but also the information contained in it. That should also be at the forefront of people's minds. Also, although these are separate issues, they can complement each other. Better data privacy practices often translate to a lower security threat profile.

It will also be interesting to see how industry reacts to the obvious tension between the increasing desire for customization and personalization in everything – from the outfits that people buy, to medicine and prescriptions, to the cars they drive. In all of these things, there’s a desire for the unique, and that needs to balance against the safety that comes with anonymity. If I were writing a thesis I'd probably want to explore that, but tragically I don't get to go back into academia anymore.

There’s a reason why animals on the savannah travel in herds. There’s safety in numbers, in looking like every other gazelle or zebra. It can be dangerous to stand out; there's a reason evolution had that result. It’s something to think about that the next time you personalise anything, On the other hand, I get it, there’s a basic human need to differentiate ourselves from others. I’m sure you could add a lot of societal psychobabble to this, but I think it's an interesting tension. There’s something in that tension that we as a society should look at, and companies could also look at.

I know in Europe, they’re starting to look at AI. How do you think legislation could evolve around things like AI and facial recognition and machine learning and all of these things that the corporates are putting in place now they have these huge data sets?
This is something I've been interested in for a decade, and it's been really interesting to watch the evolution of it. People sprinkle the term AI around and are often not clear as to what it consists of. But a chatbot is a form of artificial intelligence, natural language processing is a form of artificial intelligence, and facial recognition potentially, any kind of learning algorithms, etc. So, how are we going to legislate this entirely new body of autonomous agents and better, faster, more complex technology? I think it's going to have to be purpose, process, and results based.

We'll need to have to have a fair practice and fair process that addresses issues around bias and unconscious bias, where the ghost in the shell of the algorithm can tend to favour a particular result as opposed to another. You need to make sure that your engineers and your data sets are as representative of the society and of the products that you're trying to address. That's some of the fair process part, but you also need to ensure that the algorithm or the AI instrument also has a fair result. It's possible to put in place a perfectly fair process, but still have a discriminatory result, which society deems unacceptable. Even if it isn't algorithmically or mathematically unacceptable, it can be socially unacceptable.

We'll also need to have strict purpose-based restrictions, such as “thou shalt not build weaponized drones,” “you will not use facial recognition for discrimination, or for undermining people's rights. There's a start toward that in the GDPR, which in effect prohibits automated decision-making without keeping a human in the loop. Keeping that human in the loop component will be critical because applying AI to a bad process means just doing it faster, and you augment the bad, it doesn’t necessarily make things good.

This is probably why what gets lost in this narrative is the enormous potential for Artifical intelligence to be used for good purposes. You can use facial recognition for anti-human trafficking efforts. That’s a great purpose, which is very different from using it to profile people based on neighbourhoods or socioeconomic class, which is poor process and poor purpose. Or think of being able to recognize patterns in social media posts to identify at-risk youth who are more likely to commit suicide, or automatically identify those at risk based on interactions with their counsellors. Developing an algorithm to predict that, and then to intervene by giving them a positive social nudge of some kind, that can be an enormously beneficial purpose and use of artificial intelligence. We have to understand AI, as with any kind of technology, can be a source of both redemption and damnation. It is a great tool as much as it is a weapon. We have to use it appropriately, and appropriate use will boil down to these fair processes, fair purposes, fair outcomes. It's not easy.

The law already has half a dozen or so constructs for dealing with autonomous agents, so that part is hard doesn’t seem that difficult. For example, we have the idea of vicarious liability. And corporations are regulated as entities; we could regulate artificial intelligence the same way that we regulate corporations. Employees’ actions can be regulated based on whether they’re in the scope or out of the scope of their employment, which implicates liability either for their employer or for the individual. Or as a parent, I have responsibility for my kids, including for certain of their actions. There’s also responsibility for owners of animals, both domestic animals and wild animals. There are bodies of law around what happens if those animals get loose. Or there’s another extreme where we treat AI as human and only put it in ‘digital prison’ if it behaves badly — though behaving badly is obviously needs to have subjective and objective measures. As you can see, there are a number of legal principles that can address autonomous agency. I actually think that's the simple part.
And we don’t necessarily need a standalone new law for AI when we can use existing laws and extrapolate ideas of autonomous agents. And this is helpful. I think that AI is going to happen much faster than anybody thinks it will, and it will be different from what people expect. The pace of change will actually be more difficult to cope with. Even so, there’s an enormous potential for good if we get it right.
A Conversation with Simon McDougall
Executive Director of Technology and Innovation, ICO

Could you go through your background and how you have ended up at the ICO in your current role?

A very long time ago I was a chartered accountant. I moved into privacy in about 2000, setting up privacy consultancies for Deloitte UK and then for Promontory, which is a global outfit with teams in London, New York, San Francisco, DC, and a small team in Singapore when I finished. I then moved to join the executive of the Information Commissioner’s Office in September 2018. I report to the Commissioner and the full title is Executive Director of Technology and Innovation, which means I have responsibility for our policy engagement and thinking around everything to do with new and emerging technologies including areas such as AI, data ethics, cyber security, anonymization, advertising technology. Everything that is technology led and all of our innovation work, the work we are doing to learn from innovation and promote innovation, which involves a range of groups. Along the way, in terms of how we manage things, I lead our relationships with the CMA and Ofcom so I lead the work we do on privacy and competition and the work on online harms.

If we start with the work on the online advertising industry, which seems to be topical at the moment, why did the ICO decide to start with the online advertising industry as place to investigate, in particular real time bidding?

Are there other areas that you think that could be interesting to consider when we think about the regulation in place now?

The question of why the ICO chooses to focus on things is an interesting one because the nature of the modern world means the scope of the ICO is incredibly broad. The example I often use is that on the one hand with AdTech, with real time bidding, you are talking about an industry that processes billions of transactions a day, often using quite detailed personal data profiles. However, in the large majority of circumstances it just serves fairly benign ads to people who don’t really care what ads they’re seeing anyway. We are also dealing with the challenges around rape victims being asked to disclose the entire contents of their mobile phone in order for a case to go ahead. For one regulator to be trying to balance up how they allocate resources for those two issues, and many issues in between, is a real challenge. It reflects quite how much data about people is processed in so many different ways now compared to where we were even ten years ago, definitely twenty years ago. Then you were focused on big centralised databases and a small number of companies that would say they were data led. That’s our challenge in deciding what we look at any period in time. There’s a lot of comparing apples and pears and oranges to try and work out what to do.

When we are looking at AdTech, in particular real time bidding, what we see there is personal data being processed on a massive scale. What we see in many cases is that in the drive to personalise and target advertising there are economic incentives to get as much data about the individual and link as much data to the individual as possible, in order to generate more ad revenue. You have an industry and a model that incentivises greater personalisation. It has notions of scale and speed that concern us and also notions of transparency and personalisation that concern us.
It was a focus for the ICO going quite a while back. If you look at our technology strategy, that covered cross-device tracking as an area of interest, for instance, which is very much part of this world. When I came in it was one of the very first conversations I had with the Commissioner in early October. As the ICO came out of the GDPR phase, how do you get out messages about what this big, new regulation means? Where does the regulator go and where’s the real risk? Even before we received complaints around RTB, AdTech was already something that was being discussed within the ICO as an area of focus.

Where are we in the process of the review? How should we think about the potential outcomes and enforcement?

The best source material for understanding our concerns is still the June 2019 Update Report. It has been really pleasing that throughout the last 8 months or so, the factual statements we made and observations we made about concerns, haven’t actually been challenged. People have said that it really summarises the issues well and it’s an accurate depiction of the industry. This is great as we’re the first ones to say it’s all really complicated and we’re still learning. I would have been relaxed about saying that this is an update report and we now realise it doesn’t work this way, but we haven’t had to say this.

The other key report is the blog I put out in mid-January, on 17th January, which spoke of where we are now at the end of this six month period where we asked the industry to look at how it addresses our concerns. That’s where we’ve had to choose the words very carefully because as a regulator that has to follow due process and make sure that all of the work we do is considered, data driven, proportionate, risk based. We are never going to be able to deliver the headlines that various members of civil society and Twitterati would love us to do. We will not be rushed by anyone who wants us to move more quickly. We have to move in a very measured way because we want to make sure we get it right and, quite appropriately, will very often be challenged on what we do.

What we say in that blog is that we recognise that lots of good progress has been made, in particular with the IAB group of associations and with Google. In both cases, we are in a very different place to where we were a year ago. At the same time, we will carry on working with both of those groups and there’s more work to be done. Those are active dialogues. Again, what we say in the blog is that we recognise that this is such a broad industry and it’s so immature that even if we do get the bulk of industry to move with us, which is what our aim is for industry change, not just individual organisations, the ecosystem has to change. If we get that to happen it’s such an immature and diverse industry that not everybody will follow the norm, the protocols that drive them. The ethical companies and big companies will. We think it’s very possible that more action will be required after that.

I would go back to the blog for the exact wording as it was very well worded:

We are using the intelligence gathered throughout last year to develop an appropriate regulatory response. We will continue to investigate RTB. While it is too soon to speculate on the outcome of that investigation, given our understanding of the lack of maturity in some parts of this industry we anticipate it may be necessary to take formal regulatory action and will continue to progress our work on that basis.
Enforcement is something that happens way down the line once you’ve undertaken several other steps. We have a regulatory action policy that lists all the steps we can take. Internally we are working very hard on what we next do and what is appropriate to do and if that includes taking regulatory action.

**What is the industry doing? Google seems to have created waves with its announcement to limit the use of third-party cookies over the next two years. Are there any types of anonymization techniques within the industry that you think are being discussed and deployed that could be compliant with regulation? Are some of the possibilities genuinely good alternatives or is it a means to continue targeting?**

There is definitely an ever decreasing chunk of the market, but it’s still there, whose reaction to all of this change is ‘how do we do the same stuff but in a slightly different way.’ Device fingerprinting is the obvious example. It’s something that is more opaque, less ethical - and they cross their fingers and hope they can do the same things they’ve always done. It’s really weird because over the decades AdTech has been innovative and fast moving so it’s weird that there’s people who are trying to cling onto one business model when the business model has changed over the years anyway. That’s still there and there are people who are working very hard to just keep the status quo.

What we have seen is over the last year in particular, the conversations around alternative models get more serious. There’s always been a few things out there. Even in the first fact finding forum we had in March 2019, there was serious discussion around on-device processing and the idea that phones are now getting to a level of sophistication that you can run an RTB process, but the profile you are running it against doesn’t actually leave the device or control of the individual, so you’re not having all this personal data being sprayed across all these counter parties. It is within the realm of the user. There’s also interesting start-ups out there. There’s definitely now a thriving conversation going around different ways which tech solutions may help things and whether we need to reweight the business model. It’s not our job to say what business model works and doesn’t work.

I was really struck at the last fact finding forum we had as one of the very large players in this market went on stage and said: ‘We know that right now massively personalised adverts generate more revenue per advert than contextual advertising. Our guess is that even the best contextual ads are not going to be as good as a targeted advert. However, the reason the gap is so big is because for the last two years everyone has been pouring millions, well billions of dollars, into better targeting adverts and contextual adverts have just been ignored. If society and the regulators are saying that these models don’t work and the industry actually puts more effort into contextual advertising, the delta would be much smaller.’ I thought that was really interesting because if you get to a stage of saying that through different ways of reaching markets and individuals, innovative ways, different ways to work with publishers, we can find pretty good ways of joining up markets and people, which is what we’re all about, and the delta between that and a huge amount of use of person data is fairly small, well maybe that’s a cost worth paying. It doesn’t have to be that there is a barren wasteland and there’s no innovation going on. Overall, we are seeing lots of stuff coming into the market now. I think Google and Chrome has made far more difference than we have and I’m quite happy with that. We are not saying we’re driving this market. As a bit of sledgehammer to tackle some aspects of this issue, that’s going to make a huge difference.
How are you working with the CMA on their review of the online advertising industry? How do you ensure outcomes don’t put more control in the hands of a few or is that not part of your remit?

It is a challenge because on the one hand we are the privacy regulator, and our focus has to be on the GDPR, the Data Protection Act 2018, PECR and as the ICO overall there is different legislation that we also regulate on. We have our statutory remit and we have to pursue that obviously. On the other hand, data permeates everything and most of that data is personal. We don’t want to be blinkered and not see the big picture. For me, it’s really great timing, and I would say it’s reflective of the concerns that society and the market has, but it is great timing that the CMA has its market study going on at the same time as we are focusing on AdTech and technology.

Internally we have a great relationship with the CMA. We are both really engaged and recognise that privacy and competition used to be distant relations and there’s now a focus on how these two things come together. We are interested in how large organisations are using data profiles to personalise services and make decisions about people, whether it’s to serve an advert or give them a loan or grant asylum. When we are looking at these big corporations, they use these personal datasets to do that. The competition regulatory authority is looking at the same corporations and the same datasets, but how they exercise dominance in the markets. We are looking at the same activities with different perspectives. We’ve had a lot of input into the study. We’ve sat down and spoken with the CMA about the work we are doing and our concerns and we are very pleased that the interim report recognised that. I see the work we are doing and the work the CMA is doing are two sides of the same coin. It is a good thing that we are trying to look at both of these challenges at the same time.

Will they still run separately or will there be a joint recommendation?

The CMA has its own remit and that’s their report. The work we’re doing is our work. The focus is on coordination and alignment. We do a lot of trilateral coordination work with the CMA and Ofcom and we are forming the Digital Regulation and Cooperation Forum. It is a working group committee/body, not with any legal status or making any overriding decisions, but for those three regulators to work on digital regulatory issues, recognising there’s a huge amount of overlap. As a group we have been meeting for a long time and this is a way to formalise it. It includes resource sharing, coordinating policy issues.

The concept of data portability exists in GDPR but the idea of interoperability and open banking often comes up as a good template. Is this something that you have discussed? From a data protection perspective, how would the idea of an open banking type approach work under GDPR? Where do you think there might be challenges?

I joined the Open Banking Implementation Entity Steering Group. The ICO has observer status. Likewise I sat as an observer on the FCA’s Open Finance Advisory Group that’s looking to see if the concept of Open Banking could be taken elsewhere so we are really interested in this as an area. I think there are two different buckets of interests and concerns. The first one is old school privacy and that’s mainly what we’ve focused on with open banking so far. If an organisation is looking to share personal data elsewhere, it has to worry about notice and transparency, fairness, security. There’s stuff there that has always been there and is good solid data protection and privacy concerns.
Very often with open data and competition remedies in general is that the advantage has been gained from holding onto data and not making the data itself or insights available elsewhere. So, there is a drive to make this happen.

Our first bucket of concern says ok, if you are going to make this data available somehow, how do you do it in a way that is privacy compliant and privacy respectful. That's what I call old school privacy. We've known this stuff for a long time.

Then you have the right to data portability, which is part of the GDPR. It's a right which has different facets to it. On the one hand it's a fantastic right for personal autonomy and the ability of individuals to take their data and move it elsewhere and have a level of control over it. It is also a really interesting aspect of a competition remedy. We are very keen to promote it as we want all aspects of the GDPR to flourish. At the same time suddenly we find ourselves, again, close to the competition discussion.

One thing I would say about data portability is we have not been deluged with data portability complaints. The world is not waking up on a Monday morning and talking about data portability. There was a good paper by Ctrl-Shift and DCMS on data mobility. There was a good list in there of different things that need to be in place to make data mobility happen. Data portability as a right is one part of the jigsaw but if you don’t have good APIs, confidence of the public, robust security, and a lot of other things in place, it’s not going to happen. We are very pro data portability. We think it has a place that may be supportive and complementary to other competition remedies. At the same time, we would be the first to say that it hasn’t caught fire yet and there needs to be other things in place for it to flourish.

Are there particular areas post GDPR where you have seen complaints, requests or interest in from various groups?

The ICO has to be driven to an extent by what people complain about, what the public is concerned about through research and what we know is coming up around the corner. Companies are pouring millions and billions of dollars into designing experiences so they nudge and maneuver people into continuing to use their app and be engaged. There is a massive imbalance of power there vs. consumers. In terms of invisible processing, dark patterns, if all we did was only respond to things that people are immediately worried about that just incentivizes these Big Tech firms to be even more sly around how they nudge and manipulate people so they don’t know what’s going on. We have to be a bit smarter than that and get under the hood as well as on the surface. That’s the challenge in my department is to balance all of those things.

How do you think about where the responsibility sits? Brands have poured lots of money into digital advertising, the tech companies pour money into platforms to increase engagement so where does responsibility sit?

It’s really hard in this area because as a data protection authority we have to ask who is the data controller and who is the data processor and are they fulfilling their obligations under GDPR. If you look at how RTB works then the bulk of the data processing happens when publishers are interfacing with users, through to different intermediaries etc. The other thing is we have to look at who the data controllers are in the big chain and if they are doing the right thing. The challenge is that the advertisers are often data-light. On the whole the brands are normally instructing agencies to run campaigns and sharing some first-party data if they have it.
Yet, that’s where the money is coming from. One thing I have said in public a few times is we are focused on publishers and AdTech intermediaries as that’s where the bulk of the data processing is happening. Then you get to the exchanges and agencies and the data stops there.

We have not zeroed in yet on the advertisers but if you are a big brand there is a big focus on ethics, they have to start asking themselves if the regulator and others continue to say that this practice is wrong, it’s illegal and unethical, then at what point does it become an ethics issue for a brand to continue putting money into those practices? When we have bilateral meetings with the brands, there is no push back on this. As a data protection authority, we are having to go through the mechanics of who is deciding the purposes and means of processing the data. In the real world, we know that at some point we will have to look at the business models and advertisers have to step up.

*The U.K. has been proactive so far. How are you working with other data protection authorities in other markets in Europe and further afield?*

The way it has worked so far has been a lot of informal and bilateral conversations between different regulators. When we have held our industry events and fact finding forums we have had Irish, French, Dutch and some other regulators who are interested in this, and we continue to talk with them. The thing that has struck me about coordination with different regulators around the world in AdTech is that there is very little divergence in policy positions. It is fair to say different regulators have different ways of acting on those positions because they have been dealt different cards. The Irish have investigations going on into Google and Quantcast because they are lead regulators for those. We have focused on an industry wide push because we thought that’s the best way to have impact here. We and the Irish see these approaches as complementary. We are learning from one another but not stepping on each other’s toes.

It’s been a lot of informal dialogue and making sure that we try to give notice of what we are doing. It would be nice to get to a stage where there is consistent global approach on this but in the meantime the best we can do is keep talking with other regulators and make sure we are as aligned as possible.

*If post Brexit, the U.K. moves away from GDPR, is that realistic possibility and what do you think that would mean for the approach to data protection?*

We don’t have much say on this. It’s in the hands of the government. Look at Boris Johnson’s speech from a few weeks ago where he stated that there were a range of policy areas where the U.K. would form its own policy from the 1st January and data protection was on that list.

As the regulator we are here to regulate the law. We are not here to advocate new positions or the status quo. We have to look at the regulation and interpret it as best we can and enforce it. If government, through policy and Parliament decide to change things then we go with that as that’s the will of Parliament.

*In your role focusing on emerging tech, what areas are you interested in? AI and ethics has come up quite a bit but what are your thoughts on the topics of conversation?*

Let’s talk about AI and anonymization.
With AI, the first thing to recognise is it’s a really broad term. We are not here to define what it means for people but what we do recognise is that machine learning is increasingly being used to both process personal data and then make decisions about people. There are some aspects of machine learning, in terms of the kinds of decisions it can make, the way it processes data, its potential complexity and opacity that increase privacy and data protection risks. That’s why we’re engaged with this because it’s already a big issue. The FCA/Bank of England research on use of AI in financial services is a good articulation of how in certain areas there’s a lot of machine learning going on and in other areas it’s really nascent. This is not a theoretical conversation, it’s a real conversation.

We are looking to make sure we are engaged on the data protection risks. We have two big things we are doing on that right now. One is Project Explain, which is the work with the Alan Turing Institute on explainability which the government asked us to do through the AI sector deal. That is a really thorough and world leading piece of work on explainability.

The other thing is our AI audit framework. We did blogs last year about how we will go about reviewing and auditing AI in the future and what we think good looks like. We are consulting on our initial piece of material on that and we will produce more material around how organisations should manage risk around this area and audit it.

Those are two quite practical things. We will continue to stay engaged with AI as an area. It’s a fast moving area that it quickly becoming commoditised. One last comment on this, one area that is now being recognised in the market as an issue, is procurement around AI. The Committee on Standards in Public Life has produced a paper on AI in the public sector that focuses a lot on procurement. Organisations that are buying AI solutions don’t really understand what it is. It’s entirely possible to have black box AI so people are blindly buying these things because they think it will give them better decisions or make things cheaper, and they are not really thinking through the fact that, as with any third party, they should be assessing the solution they are buying. That’s a really mundane but practical risk.

The other thing I would mention is anonymisation. We are going to be renewing our old anonymization guidance because it was last done in 2012. We also want to expand it. I was on the judging panel for the FCA’s AML TechSprint in autumn and it really struck me that there’s lots of solutions around how to share data in banking around homomorphic encryption, federated learning, secure multi party computing. Some aspects of the Techsprint used homomorphic encryption, federated learning, synthetic data set. These are all privacy enhancing technologies, different ways that organisations can get insights from personal data sets without necessarily seeing or accessing the personal data. A lot of these are still nascent and evolving and lots of these don’t work at scale or how you want them to work yet, but they were all lab based technologies a few years ago that have come on leaps and bounds. I think some of the challenges we currently face around privacy and data protection can be solved by these technologies. They have to get better first, but if you can get to a stage where you can say we can’t share this data with you but we can work with you so you can get all the insights you need from this, and understand what decisions you have to make around it, you will never see a jot of the personal data and it will be cryptographically secure, then brilliant. Anonymisation is also privacy enhancing. If you can anonymise the data then in terms of the GDPR it is out of scope; it’s not personal data any more if it’s truly anonymised. The problem is that it’s hard to do and often the data you have at the end is just sludge.
Once we’ve done the anonymization work we will carry on working on all the other areas of privacy enhancing technologies and try and make sure that U.K. data controllers understand how these techniques work and which ones are good for different situations and how they should go about developing, procuring and engaging with these techniques.

**GDPR is seen as one of the tightest regulatory frameworks for data protection. You don’t think this creates a disadvantage from an AI perspective? Europe has put out its AI white paper; are there some tensions with GDPR and the idea of data minimisation? How do the two work together? Will the anonymization techniques be the answer? Is it harder work if you are in Europe?**

It may create harder work but it’s worthwhile work. If you think about why London has flourished as a financial services centre. One of the reasons is clear, consistent regulation and a clear rule of law. It has not been because there is no regulation. When you are dealing with people’s information, getting to a stage where organisations have clarity and we get to stable, good, simple regulation in these areas, is actually helpful to innovation.

When we talk to start ups and Big Tech firms, anyone involved in innovation, they don’t say ‘can you get rid of all these boring rules’ (I’m sure a couple of them think it). They say we just want to know what to do. The problem with AI is it’s moving so quickly that we are all trying hard to keep up with a moving picture. That’s what we have to aspire to, not to think that this makes us non-competitive, because the rest of the world is copying these rules. Over one hundred countries have privacy regulation and we do need to get to a point where people can say ‘if we follow these rules we’re going to be ok. That gives us a space to play in and we can do it knowing there won’t be public backlash against what we’re doing, people will be comfortable sharing their data and regulators won’t enforce against us.’ That’s the nirvana we need to move to, an environment of trust.

**Are there any other aspects of GDPR that have created particular challenges for corporates?**

There will always be room for improvement. One thing that isn’t working that well right now operationally is breach notification. I know some DPAs have said they are struggling with the volume of notifications they get. The ICO is lucky as it’s well-resourced and can process them, but I don’t think, in its current form, it generates lots of insight for us or for elsewhere. We hear from data controllers that it’s not generating better protection for users. We shouldn’t get rid of it, but is there a better way to do it? Elsewhere, it’s very early to say. We do need to make sure that all the countries that have implemented the GDPR keep talking to one another.

**Lots of notice and click and maybe it has created a privacy paradox. Can consumers ever really make an informed decision? Has it been the right approach to give consumers that level of choice or decision making?**

I think it is challenging. We have a long way to go to understand the best way to get a really large, wide, diverse population of people to understand what is happening with their data. We don’t expect everybody to be GDPR experts.

The analogy I would use is food regulation. If I go to the shop and buy a bag of mince, I don’t sit and try to understand the full supply chain of how that mince got into Tesco or worry about whether it’s been stored at the right temperature during that process. I effectively buy that product with a level of confidence and through a seamless transaction, and I don’t lose sleep about that.
When that process breaks down and there’s a scandal, it’s headline news but it’s very rare. People realise these things happen but we aren’t worrying about that on a day to day basis. Compare that to when people are going online and they click accept because they don’t have time. They don’t feel good about it. It increases their feeling of mistrust with the process but they click. The amount of time people leave things in their online shopping basket, or decide not to transact or don’t try a new product because they are scared of doing something new online. All these things are dragging us down. The journey from getting to the point where people are online and behaving as seamlessly and without friction as they are when they are buying their bag of mince. That’s a hell of a journey. That’s got to be where we get to. It’s crazy that our lives are becoming lives online but our every click and move is being done with a heavy heart. That’s just bonkers.

Anything else that’s important?

The whole PECR/ePrivacy thing. With ePrivacy regulation, there were some headlines last week about some new drafts coming out that are a bit more flexible on legitimate interest etc. It’s gone round in a few circles and it still goes on and there’s still clear blue water between what the Commission and Parliament might want to do and what the Council wants to do. When they get to a trialogue process and the commentary starts to say there’s something coming out of this, then it’s worth paying it full attention. It could be some way off. The fact the process is going on is recognition that PECR itself should change. Until then we have PECR.
A Conversation with Prash Naidu

Founder & CEO, Rezonence

What’s your background in the AdTech industry?

I was actually in banking myself; I was a quant. That helped me look into AdTech because high frequency trading is not that dissimilar to what happens in the AdTech world where you’ve got an auction for every ad served, and the winning bidder gets to serve their ad. I spotted a hole in the AdTech ecosystem with regards to effective ads, to deliver proven human engagement for brands and good levels of monetisation for publishers. At Rezonence, our product FreeWall® looks to solve this challenge for advertisers and publishers. We still operate in the display ad space because it is a banner but it’s interactive. It also operates within the real time bidding space so an advertiser can bid and, if they win, deliver their FreeWall to the end user. Rezonence is very deep into the RTB ecosystem, which has come under criticism. We have taken a forward looking stance, which is much more GDPR compliant and consent based. The changes that the ICO are looking for should help our business.

Could you explain how RTB works, at a high level? What are the challenges in relation to third-party cookies and the data flowing through the ecosystem?

One of the primary purposes of RTB is retargeting. If you went to visit the Sony web page, for example, what would likely happen is Sony would have a segment pixel that will load on that page. It’s just a small pixel so as a user you don’t notice it but as it loads, a call is made to a server and that server drops a third-party cookie on your browser. It tells the browser ‘here is an ID, can you store it please.’

Now, then when you go off to, let’s say, a publishing website; it could be the Guardian, Mirror, Telegraph, Times etc. When you load that page, lots of ads are loaded; all loaded in real time programmatically. With every ad you see, a real time auction is happening. The publisher works with Supply Side Platforms (SSPs), providing the media inventory supply. That ad slot generates a bid request. The SSP will be plugged into a data management platform, which control the IDs from third-party cookies. The bid request will have this third-party cookie attached to it and therefore the ID that was dropped on the browser from visiting the Sony website, and there could be a lot of other IDs from visiting other sites.

When the bid request goes into the ad exchange a lot of advertisers using DSPs (demand side platforms) will bid to serve an ad. This is all happening in milliseconds. At this point the DSPs will go ‘do I have an interest in this person.’ The DSP working on Sony’s behalf will look at its huge list of IDs and see if the ID attached to the bid request matches with the table of IDs that it has. If there is a match then the DSP will bid quite a lot more than if it didn’t have a match. Let’s say, if you found a match you bid £5 per thousand, and if no match you might bid 50p. As you are willing to bid so high, the chances are that you are successful and win the auction and your ad gets delivered. That’s why it looks like ads follow us around as that DSP is consistently bidding higher for you so they keep winning and keep delivering the ads. Eventually they might hit a frequency cap and stop or another product site might start bidding higher then they might start winning.

If you load up a web page and there’s 5 ads on the page, then there were 5 auctions that happened in about 50 milliseconds, sometimes a bit longer. That’s why it’s called real time bidding.
How should we think about where the control of the ecosystem sits?

The privacy breach is not whether you sit across the ecosystem or not, it’s about what’s being attached to the bid request. Let’s say I am a company that wants to acquire data about people, I can listen to the bid stream. I can participate in the auctions but never bid enough to win, so I am getting access to all the data. If you think about what’s attached to the bid request, there’s a user ID but also the URL the person was visiting, together with other data the publisher might have attached to upsell; for example, if the person is high earner, male or female. You can’t control who can listen to the bid stream; it’s a free for all and there’s lots of companies who listen to the bid stream.

Because real time bidding is happening on every website, every time you load a website, a unique user ID, the browser and the article you are reading is being transmitted to someone; they can build up a very detailed picture of who you are. Imagine your entire browser history being shared with pretty much everyone in the world with a user ID. There are lots of companies offering cross device graphs which can take the cookie ID, which is attached to the bid request that can be translated into a real person’s name, email address and address. Most people don’t know this but essentially everyone’s entire browsing history is being transmitted all the time and, if someone wanted to do, it can be linked back to a person.

Some people may say that they don’t mind people seeing what they read but there are lots of sensitive things. Johnny Ryan, a privacy advocate, talks about ads targeting people who are vulnerable and that was because, for example, they knew that these women had been reading about infertility and this data was being broadcast freely. You may be gay and not ready to come out, but if you are browsing certain sites, it’s being broadcast and someone can stitch it together. He talks about the wholesale data breach happening all the time. He is right.

The industry says there’s no personally identifiable information being transmitted, it’s just IDs. That’s why GDPR goes so far as to say it doesn’t matter if you don’t know anything about this person and all you have is an ID and a bunch of behavioural data and can deduce the person is in the market to buy a car. If you can treat that person as an individual then it doesn’t matter if you don’t know them, you are treating them as an individual. It’s very easy to work out who they are.

There was a researcher in the U.S. who was able to demonstrate that with six data points there’s a 90% chance of being able to figure out who you are. For most people there’s probably in the hundreds of data points out there that are accessible.

Going forward, how do you think the AdTech industry will evolve?

Apple has already got rid of third-party cookies. Google is proposing the same, but in two years’ time. They are saying ‘we want the industry to be able to transition otherwise business models will collapse.’ If you look at what’s happened in the ecosystem with just Apple users, the yields that advertisers are willing to pay are much lower, which means publishers are earning less due to Apple’s move.

Is that across the board for publishers?

I’ve heard different things. I’ve heard the yield has dropped from 10% up to 50%. The median number of 20-30% range. If we take the Sony example, if they didn’t find a match then they would bid low and if they find a match they bid high. If you are on an Apple browser, that ID would be deleted almost instantly so Sony’s DSP wouldn’t find a match. When Chrome does the same thing, the chance of finding a match will be almost non-existent.
You might argue that they will have to bid high on things like context. This is one of the things that’s coming back. That’s how it used to work. If you were advertising cameras and knew a publisher was publishing a piece on photo journalism, then you’d advertise on that. Context could potentially make a comeback.

You can still advertise without IDs as long as you can get consent from individuals about things they are interested in. A news aggregation app could ask that you tick the relevant things you are interested in. A publisher, to build a relationship with reader, could say ‘to help us show more relevant ads, tell us what you are interested in’ and you select your hobbies and what you do for work, maybe not all in one go, but over time. They will then serve relevant content and ads. Just because you kill of the ID it doesn’t mean you kill of programmatic. It certainly changes the way you can operate.

What GDPR forces you to do is that the end goal is the same. I can’t see how you can be compliant by maintaining these IDs. What Apple has done, what Google will do and what the ICO say needs to happen, are all the same. It’s the elimination of IDs. As long as you have an ID attached to a person, and as long as it’s constantly being used as they are going around online, then you can stitch it back together anyway.

Some companies seem to suggest that the shift away from third-party cookies is fine as they are using Device IDs and mobile IDs sat on the device and it’s anonymised. Does that also fall foul of regulation?

Yes, because device IDs are nothing more than IDs attached to a device rather than a session or vendor. There’s already rumours that Apple will kill off the device ID or rotate it on a more regular basis. It’s happening in the web world. It’s only a matter of time. Apple won’t be able to have a consistent stance if they allow device IDs while they have killed off cookie IDs. They could be accused of anti-competitive behaviour because they’ve killed off other people’s abilities to have unique IDs while maintaining their own ID. If Google kill off third-party cookies and maintain their device ID, which is Android, then again they could come under criticism for killing off the ability of others to target people except their own. Both Apple and Google will have to tread carefully if they choose to maintain device IDs. GDPR and the ICO will force the elimination of the device ID. Every time you load an app and see an ad, the device ID is being propagated. One can quickly work out all the apps a user has installed and that tells me a lot about them. With some apps, it’s possible to glean what they are doing within the app.

Digital advertising spend continues grow. If the yields have been falling, where is the money being reallocated to?

This is the worry. Advertisers will still want to serve ads in a targeted manner so where can you do that; it will be within the walled gardens. Third-party cookies allow you to transfer IDs across sites. If all you did was stay on one publisher they could use a first-party cookie to learn about the behaviour while they are on the website. Unfortunately for most publishers, people don’t spend enough time on that one site to deliver meaningful insights. However, if you are Google or Facebook, they do spend enough time within their ecosystem that they can glean more information about people and serve targeted advertising to them. The worry is that, as it is, Google and Facebook take 80% of any new ad dollar that goes into the online space and now they will take even more. This is potentially one of the reasons the ICO is moving more slowly because if they told everyone to stop using third-party cookies, it would mostly kill off the system outside of Google and Facebook, but Google and Facebook would be fine.
If an advertiser wants to target a person, they won’t be able to find them as IDs are gone, but if someone searches within the Google and Facebook then they know everything about them.

**How does the ICO tackle this? There is also the CMA investigation happening at the same time, are they working together?**

The way the ICO should tackle it, in my view, because that what the law says, is that to use someone’s personal data in any form you need their consent, at least within Europe right now. For Google and Facebook to serve targeted advertising, they need that person’s explicit consent. This isn’t just saying ‘can we show you ads’. When you’re profiling someone they will add you to segments and, according to GDPR, this itself requires consent. What they will need to do is constantly get consent to add users to various buckets and pools. It has to be freely given, so the user should have the ability to say no. Facebook are trying to be open and you can look at choices but the default is ‘yes, do serve me personalised ads’ and most people don’t turn it off.

GDPR says explicitly that the default should be turned off. Then Google and Facebook end up in the same position as the rest of ecosystem where you can’t target people based on their behaviour unless they have explicitly said they are happy for you to do that. Even within Facebook they will have to keep showing messages to ask if happy with targeted advertising. If not, then you will get untargeted advertising and they will get less yield like everyone else.

Their ability to serve targeted advertising while the rest of the industry doesn’t have the ability, that disparity needs to go away. If everyone is on the same level playing field that you can’t do targeted advertising, you are going back to the old days of print and outdoor. Broadly untargeted but you build brands.

**If we end up with a world where untargeted advertising is default and explicit consent is required, would that have a deflationary impact on the ad spend pool?**

I think it would follow Mary Meeker’s model of time spent on a medium is proportional to the ad dollars spent on that medium. If people spend more time on their phones then the ad dollars will go there. The fact you can target better and attribute better and assess ROI better is speeding up the follow of money, but it doesn’t mean money won’t keep growing. If the average person in 10 years is spending 90% of time on phones, that’s where the money will be.

**Do you think we will ever get to a point where consumers will be rewarded or paid by advertisers for using our data?**

A couple of companies have tried it. I know one that tried it and went bust. You could create your own data directory and sell it on. The problem is that the amount of value that you personally would get is quite small. Standard programmatic spend is £1 per thousand. If you are willing sell your data, and data is sold on at fraction of that, it could be 20p or 50p. If you were to sell your data then you would earn 50p from a thousand ads. You’ll have to see a lot of ads for it to be worth signing up to a data vault and broadcasting it to lots of people. There is not sufficient value in it for individuals to sell their data. It works for a data company where you have 20 million data profiles you have gathered for free and can sell it on.
If there is no disparity and it becomes more transparent, how do you see it all shaking out? Are there some types of models that won’t be able to operate and some that will benefit?

The business model that is most under threat is the retargeting model. That whole side is severely under threat but there’s lots of pivots those companies can make. They can go to the context space. The desire for an advertiser to put an ad on a page next to relevant content won’t go away.

You might argue nothing will change. The general record is 2% of GDP is spent on marketing. The channels will broadly stay the same. Every marketing department will be given a cheque and they would like to put it down on targeted advertising where they get a better return. If there is no targeted advertising, online is still as good as TV or print. If you look at how agencies operate in their pitches, it’s all about pitching relatives, guaranteeing a better price than their competitors. If you have two big brands fighting it out, all they need to do is make sure they are getting their fair share of voice. I don’t think digital advertising collapses if you take away targeting. Whatever ends up in digital will probably end up in digital in 5 years’ time but maybe on different vendors.

From a publisher perspective, do you think there’ll be a distinction between the different publishers and how the money is reallocated?

Premium publishers will do well out of this in the long run. There are two reasons retargeting is working well for advertisers. One is being able to find the person. The other is being able to find that person cheaply. In the past if you were doing print advertising, there’s no way Sony would place their ad in some random publication they’d never heard of. The premium publications knew this and would charge accordingly. AdTech came along and said ‘if you want a Times reader, we can give you a Times reader when they are on some other random website because we can track them from site to site. We can serve your ad to someone who has visited your product page and has visited the Times but we will deliver the ad when they are on site XYZ.com but at the tenth of the price the Times would charge.’ That’s why AdTech grew. The advertisers were very happy.

If IDs are removed and retargeting is removed then it hurts the ability of the vendors to spread the money across the long tail. If an ad agency ran a report of where they delivered the ads for their clients, they’d get a list that’s over one hundred thousand long. You’d only recognise a handful of the sites and only a fraction of the budget would go on them. If you can’t retarget in that same way, then what brands will say is that they want to be safe and serve it on titles they know. Agencies will be forced to deliver the ad spend to publishers that the brand knows. The spend will get concentrated with the large publishers. The downside is that the funding of that small blog site that has acquired some traffic, will almost entirely disappear. It could be good for journalism. Fake news sites get a lot of advertising because of programmatic advertising, which puts the audience first and context second.

The idea that for every £1 that is spent the publisher ends up with 30p as there’s several intermediaries taking a cut, what happens to these intermediaries? Will that change?

Hard to say. A lot of the vendors will remain and there will be new tech that has to be built and they will build it. They will have to recuperate cost and will do to by taking a margin out of the flow. A publisher with their own sales teams and programmatic desk might be able to short circuit the chain and the advertisers might spend more directly with the publishers.
If that were to happen they would take a greater share of every dollar the advertiser spends. Right now it’s hard to take a view with any level certainty. I still think the structure of the data system will stay the same but how they operate will have to change. I can’t see a whole class of vendors dropping out entirely. You will still need DSPs, SSPs, anti-fraud vendors, all those people.

**Do you have a sense of the timing of the ICO review? What’s the next step?**

I don’t know and I’m in the camp of people saying ‘come on ICO, you’ve got to move on.’ When GDPR went live, almost two years ago, there was a narrative of they will give everyone time to change. It’s not like it happened over night. GDPR was years in the making and they should have expected everyone to be compliant.

The ICO published the RTB report which said they aren’t happy with what’s going on, and rightly so, because a lot of what’s going on is basically non-compliant. They said everyone had six months and then if you aren’t compliant they will take enforcement action. Then six months has come and gone and they seem to be saying they’ll give it more time. The analogy I used recently was, imagine the law says you’re not meant to be speeding or drinking and driving but everyone knew the police wouldn’t stop you and prosecute you, people will just carry on. It’s only when there’s enforcement action that people will stop.

The ICO says it wants to move the industry and operate in a cooperative fashion. Knowing what I know about how the industry operates, it’s hard to see how this will work. I’ve spoken to CMOs who have said they won’t stop what they are doing because there’s too much ROI to be had in using data in a non-compliant way and they can’t afford to stop doing it unless there’s a very real danger that the cost of being compliant significantly outweighs the cost of not. There’s a lot of bigger companies not being compliant so why should they. If one of them gets a big fine then they’ll think about becoming compliant too.

There are AdTech companies who have built their entire business on using personal data in a non-compliant manner and taken millions of pounds of investment. If they were thinking about trying to become compliant then it would either cost them a lot or they would go bankrupt. They will probably just carry on being non-compliant because the cost of becoming compliant is too great and hope the ICO isn’t enforcing anyway.

**What is causing the ICO to keep delaying and pushing it out?**

I’ve heard rumours that they’re not geared up to begin enforcement. They’ve tried and they realised that those who look to enforce will take them to court and question the interpretation of the law. They could get locked up for years. Even if they take enforcement action and levy a hefty fine, the company might try to fight it as they might get the fine reduced. The ICO might believe it’s better to get the whole industry to move as enforcement action could get tied up in lawsuits.

**If the ICO doesn’t act, how does it all play out?**

If the ICO doesn’t do anything, the industry will do what suits it which means it will continue to operate on the basis of what was happening pre GDPR. My company was built for a post GDPR world so it’s annoying the ICO isn’t enforcing. I want the future to be here today. If my company was built for a pre GDPR era, I would be trying to delay the onset of the new world. You will end up with broadly the same happening. It may be dressed up as privacy sandbox but you won’t get the privacy standard the GDPR calls for.
The European Commission is talking about a single data lake and focus on data portability. How does that work with GDPR that’s about data minimisation?

The first and foremost thing we need to think about is what GDPR says is to think of data is a liability rather than an asset. Every time you are holding onto someone’s personal data, you are sitting on a ticking time bomb; it could be leaked, it could be hacked and you could not have the right checks in place. That’s only the case if there’s a threat of enforcement. If there’s no enforcement action you are not sitting on liability, you are sitting on an asset. It’s all well and good to talk about a way to create a unified system that can securely transmit data without encroaching on privacy. There’s lots of ways to encroach on privacy. We have all these smart speakers now, which are listening all the time. If there’s no regulation or threat of a fine, then it doesn’t matter whether you suggest a nice way. If holding people’s personal data and acting on it carries danger to an organisation if not done properly, only then will they take care to do it. In the past VCs would attribute a value to a company based on how many data points and user profiles they had. It was part of a company’s pitch and VCs would say it’s worth money because data was seen as an asset. Now they have to say the access to data is compliant. If there is no real risk attached then it will go back to being about having tons of data and making money from it.

The main question is, are the ICO going to do anything and, if so, when? There’s one camp led by the IAB and wanting the ICO to work with the industry to get it to change. I know there a lot of players in that who have no intention of moving because it’s not worth their while.
A Conversation with Bill Philhower & Anthony Turiano

Bill Philhower - Global Head, Privacy and IT Regulatory/Audit Program Management, Citi

Anthony Turiano - Managing Director, OTRC / Privacy & Compliance Services, Citi

Bill Philhower: Anthony is the program director for the California Consumer Privacy Act (CCPA) implementation, so a lot of lessons learned there. Stepping back and looking at it from a practical perspective, the most significant challenge we have right now is just the way these regulations and requirements are quickly coming online. From the US perspective, we have seen them take a state driven approach vs. federal, which is causing complexity and lack of clarity on definitions. We are trying to figure out what’s the right way to build our solutions.

Anthony Turiano: I think that’s where we’re evolving in the short term because of the requirements being in place for CCPA. We made sure we could solve for that. We wanted to build a solution that is going to be flexible. In the case of CCPA, there was also some consideration of the extra territoriality, but we have not solved that issue from a systemic perspective, and we have a manual capability around that. What I would describe as a good win for Citi was that we were able to implement consistent interpretation of the regulation. Most importantly, the solutions that we developed were using a common entry point into the organization for our Internet portal, as an example, so whether it be a prospect, client, we’re asking them all to go to the same place. We have one phone number, and we are guiding client-facing folks to respond to that. Then we route it internally to a single place for the U.S. at this point, and that is pretty impressive.

It must be complex thinking about how different individuals interact with Citi. What were some of the structural challenges you had in thinking about the differences between GDPR and CCPA frameworks? Does it require a different technical implementation?

Anthony Turiano: I think that’s where we’re evolving in the short term because of the requirements being in place for CCPA. We made sure we could solve for that. We wanted to build a solution that is going to be flexible. In the case of CCPA, there was also some consideration of the extra territoriality, but we have not solved that issue from a systemic perspective, and we have a manual capability around that. What I would describe as a good win for Citi was that we were able to implement consistent interpretation of the regulation. Most importantly, the solutions that we developed were using a common entry point into the organization for our Internet portal, as an example, so whether it be a prospect, client, we’re asking them all to go to the same place. We have one phone number, and we are guiding client-facing folks to respond to that. Then we route it internally to a single place for the U.S. at this point, and that is pretty impressive.

Do we take an approach that we build to the highest standard so we have a gold-plated version at the top?

Anthony Turiano: Bill and I got involved in GDPR post the roll out. GDPR is a law impacting our European clients, for Citi mainly the institutional bank. The consumer bank was spared from GDPR. It’s really about understanding data, responding to the requests and the volume. We couldn’t glean a whole lot from it, but we knew we didn’t want to leave it to the individual app developers to determine compliance with GDPR; we needed to create a company-wide solution.

The institutional division and consumer bank — independently — were looking at CCPA a year or so ago, and then they asked Bill to take on that broader remit to look at it at the enterprise level. That meant common solutions, common privacy hub where people interact with us. The institutional and consumer bank colleagues were busy getting varying opinions from external counsel as the focus was on that business and product line. We needed to come up Citi’s interpretation and then, with intel, we started to develop a privacy posture: where do you want to be as a company? A lot of peer banks said “we do not sell your data.” Citi was thoughtful and said when we exchange data with partners like Home Depot, there’s some level of reciprocity.

William (Bill) Philhower, Managing Director, is the Global Head of Operations & Technology (O&T) Risk and Control. In this position, Bill is responsible for overseeing the identification, assessment, and mitigation of operational risk globally across O&T. His management responsibilities include Citi’s global Business Continuity Management and Cross-Border Data Clearance Programs; as well as O&T’s In-Business Operational Risk & Control; O&T Risk & Control Analysis; and Regulatory & Audit Management functions. Further, he is responsible for the interface to our primary U.S. Regulators for information technology matters.

Anthony V. Turiano was appointed Head of Privacy and Compliance Services (PCS) in April 2019. In this capacity, Anthony leads the Cross Border Data Clearance, Export Control, Business Analysis, and Partner Support Services for OTRC risk and control platforms. This new organization is tasked with launching a Data Privacy function that will develop robust programs to support future global initiatives. Concurrently, Anthony is OTRC’s CAO and the Initiative Owner for the California Consumer Privacy Act (CCPA). Prior to Anthony’s appointment as head of PCS, he led the Regulatory Audit Management organization in coordinating and reporting Regulatory and Audit activities across O&T.
When you look at the contract, you may imply selling, so we were more conservative and said we don’t sell typically, but if you use one of these cards then we can have a discussion. That elevated our privacy transparency.

Bill: Thinking about how we may use and monetize data in the future, that’s going to be a very significant conversation in how we respond to things like CCPA and what the button says and what the product needs in terms of changing disclosure. Just to give you a sense, there’s a little over 1200 websites that have our privacy notice internally within the company. There’s only about 9 or 10 differences in the language, but it exists in that many places.

_Have you had to normalize the language to get past contract language?_

Anthony: Cookies have come up; although it wasn’t in the legislation, and the idea that a cookie may represent personal data and then selling that data for other purposes. So to be in front of it, we put out guidance that said if you do not want us to collect cookies, this is how you modify your browser to disable that activity. We are working with third parties now and maybe to come up with a cookie consent tool, so that when browsers come up, you can go ahead and say no, and that would carry through your engagement with companies.

With regards to CCPA, we built the infrastructure to collect the response to allow opt out at the least. We are creating underlying infrastructure, which we are starting to call “Good Privacy Hygiene”; we should know the data we have about our clients and where it’s kept. If you have multiple accounts and if you are appropriately allowed to request a response, we will respond to you with that information, not the data itself, but a description of the data. As it relates to California, we have a Graham Leach Bliley Act (GLBA) exception by which it means if the data is caught under that statute program, then we do not have to share with you. The next state that comes along may or may not have that. We are building with flexibility to achieve the high-water mark. Right now, we don’t need to leverage, but if Washington or New York or other states come along, and they require the data to be provided without the GLBA program exception, we will be equipped to do that. Then we as a company may determine that we aren’t going to manage each state separately; we may decide it will be the Citi standard for our privacy responses.

Is the technology prepared for what the regulation requires? Is that a challenge Citi has found? How have you managed the process of bringing in third parties to manage this?

Bill: The biggest technology challenge we had with GDPR is manual. Some of the processes are manual: we collect the requests manually; the letter processes is manual, etc. With the right to be forgotten and need to delete and not maintain data, we’ve had some significant challenges. We’ve had a number in the hundreds of applications that either did not have the capability to comply with the record retention policy of the company, which they should be. Secondly, they didn’t have the ability to delete an individual upon request. Over the last year plus, we’ve had a reasonably large program overseeing the storage limitation aspect of GDPR to make our technology applications compliant with all those requirements. Frankly, we need to change the way we architect applications, so we build that into all of them from the beginning. No one wants to continue doing this as a fire drill; it’s expensive, as you can imagine. That was a big gap for us, and we’re not there yet. We’re projecting to be there by the end of November 2020.
**Are there certain technologies that are interesting in this area?**

**Bill:** On the application side, a lot of is being baked in with our own code. Where we’re looking at technology to highlight unstructured data with PII (personally identifiable information). That is an area where, broadly, we have not solved for that yet.

**Anthony:** We’re partnering with our information securities colleagues. They are scanning other sites for PII and that can help us with where PII is kept outside of a structured system. I think when you circle back, one of the thing the company needs to do as we elevate our data knowledge is understand our processes, what those processes drive, and whether they handle PII. Hence, when the next regulation comes along we don’t have to go out and scan our inventory of queries as we’ll know it as part of our data hygiene to say ‘this is our data, this is how it’s kept, this is what’s being housed at a third party’.

**If we play out the near future, do you think the U.S. via an agency will create a national regulatory framework? How do you engage with the regulators?**

**Anthony:** We work with Global Public Affairs. They give us a lay of the land. When we update our steering committee, we’ll talk about the next states that are likely coming on, whether the legislature is in session, if there is a draft for discussion, and how we can action it.

**The regulators work on AI and AI ethics seems to be very regional at the moment. Does that come into your remit?**

**Bill:** Not as much. We are more involved in the data limitation side. I think about it in terms of products we offer and data assets and how we monetize it. I think that’s going to cause those smart data scientists to think about what we really want to add together. Where I do get concerned is how do we show regulators that the way we are dealing with the information and how we are using the information in a way that’s lawful. The typical thing that people raise is fair lending. How do we ensure we don’t use AI models, which we are going to figure out that certain neighbourhoods have harder time paying back loans, while we still have to comply with the 1977 Community Reinvestment Act. We can’t do that because the algorithms recognize what happens in human conditions, so we are going to have to look at that. That’s something that the US regulators have asked us about on a number of occasions, and I think — from a technology perspective and probably broader than that — how we are building the right rules about how you test something, how you show what the algorithm learned, and then you maintain it so you have a record of what you are doing and you can prove it to people, because over time this is going to be a challenge.

**On AI ethics and data privacy regulation, how do the two work together? Data privacy tends to be about data minimisation and AI is about having huge amounts of data. Can the two work together? What are the challenges?**

**Bill:** On the question of U.S. regulation, we don’t think it will get done in the near term, especially with this being an election year. States will come online and copy CCPA. CCPA is planning a new ballot called CRA in November, which might influence what happens next. It will have more teeth than the California Protection Act.
Lots of the regulation has an element around data localisation and not leaving geographical boundaries. What complication does that present for a global company that wants to transfer data across borders?

Anthony: It presents quite a big challenge. Our team is geared to help Citi businesses in the transfer of data. We are trying to look at things holistically and anticipate changes. The physical movement is not the key here, it's the access. If data is accessed from a disparate location, that's classified as a data movement. A subset of the countries around the world will require regulatory approval, a small number of 10 to 12, others require notification and others require some level of internal documentation.

Bill: Quite a bit of this impacts Citi’s business model. We want to be able to utilize, centralise, move cross-location as much as we can for our operations and data centres. As far as large data centres, we really only exist in three significant countries where we want the technology to be. Then there’s a bunch of places where we have it for regulatory reasons, so we are trying to minimize that. On the operational side, we sometimes get tangled up, not just in the privacy and cross border, but what we consider to be material outsourcing. That drives the complexity. As I think about this on a go forward, what I worry about right now is the way people are potentially passing requirements that are going to mean we need to unwind things to remain compliant. We had that with Turkey a few years back, we see it with Russia, and I think there are other potential examples around the world depending on where things end up. It causes us to have greater expense and complexity and then other regulatory compliance requirements get to be harder. A good example of that would be AML where we are trying to do the various surveillance requirements that are expected by the U.S. government. There are a number of places in the globe we are just not allowed to move the data for that particular reason.

Does it create a moral dilemma at all?

Bill: It’s a potential. What we are seeing in a number of countries is data localisation. The African countries have become a challenge in that regard; they want data maintained. In some case we’ve been able to convince them having a copy that isn’t used is sufficient, but we’ll see where that goes over time. If we have to start running data in local countries, that’s going to impact serving global clients. When I joined Citi, we had something like 240 data centres at the time. Today, we have a dozen and a few satellite ones that have any significant data content. It would be very challenging, and would create a different expense dynamic for us.
A Conversation with Dr Johnny Ryan
Chief Policy & Industry Relations Officer, Brave

Johnny, what’s your background?

I work at Brave, the privacy-focused web browser. My background is policy, more than anything else. Previously, I worked at a publisher, an AdTech company, startups, a university and a think tank.

What has prompted your interest in the area of privacy, especially within the online advertising industry?

People who spend time in the online advertising industry often suffer from a syndrome that I call ‘AdTech Remorse’. I’d say half of the employees at Brave suffer from ad tech remorse because they’ve been involved in some way or other. The person I report to, Brave’s CEO, may suffer from the most acute remorse, which is “JavaScript Remorse”, because he invented JavaScript. Back when he was at Netscape, he was involved in the very first picture tags and so forth. These technologies now prevail all over the web, and — though this was not the intention — are used to track every person who is active there.

You can’t think about life today without being concerned about this situation: we in the technology community accidentally built a surveillance system that encompasses every person and what they do.

Twenty years ago, if you had thought that what you read, watch and listen to online would be hoovered up by thousands of companies who you couldn’t name, and that an algorithm might decide not to shortlist you when next you are considered for a job interview as a result of what it knows you have read online, I don’t think that twenty years ago you could have imagined that. I think you would have assumed, if such a thing were possible, that it would have occurred only in some very strange alien nation of which you would not want to be a citizen. But that is the case today for everyone, all around the world.

What I have just described acutely disadvantages legitimate publishers. I used to work at a newspaper called The Irish Times. It was, and is, important to the life of the nation that I live in. I see things like “real-time bidding” (RTB), part of the advertising system, as a cancer at the heart of legitimate media. Not only is everyone being surveilled, but the system also undermines the businesses of legitimate media. At the same time, it enables not only algorithmic discrimination, such as the job interview example I mentioned, but it also enables micro targeting for misleading propaganda and so forth.

It is impossible to think about these things and remain content to continue to work in an industry that is essentially harming the fabric of society. That is long answer to a short question. The answer is: no thinking person can work in AdTech and not eventually suffer “AdTech Remorse”.

Two questions. One is, on the consumer side of things, and obviously having worked in AdTech you have a real in depth understanding of exactly what’s going on behind the scenes, but how much do you think consumers are aware? And, do they actually care about it? Have you seen any signs of behaviour changing as that awareness has changed?
It's possible that it is a dawning awareness. The God of AdTech Scandals is a generous god. These stories are featured on front pages of major papers now. The campaign that I have been leading to reform RTB ad tech has been on the front page of the Financial Times twice in the last half year. So it would be surprising if there were not a dawning popular awareness. The increasing use of ad blockers, now apparently used on a staggering 763 million devices, suggests that something is happening in the public consciousness.

However, whether people worry about privacy is not the important question. I suspect that people think or do not think about privacy in the same way that they think or do not think about pensions. If you have ever studied behavioural economics, pension planning is a burning problem. People put off until tomorrow what they should think about today. That is a fact of human nature. So, we do not think day to day about pension planning, even though the question of where our income comes from for the final (maybe) third of our lives may be the most important decision that we make for our standard of life. Because tracking pervades the Internet, “What we do today”, as the general in The Gladiator says, “will echo through eternity”. (It was Marcus Aurelius, in real life). Your online activities may very well affect your future. But it is very hard to keep that in the forefront of your mind.

It is very hard for the masses to understand this this amount of complexity, right? The huge masses don't have any concept of what's going on. Is that a fair premise in your mind?

It's an irrelevant question though, because this is not the responsibility of individual data subjects. (“Data subject” is what the European law calls this person.) It is not the responsibility of the data subject to constantly monitor what unknowable companies might know about them, and do with that information. In the European Union, the law is very, very clear; the responsibility is with those who process data about data subjects – particularly those who determine the purposes and means of that processing.

Now the same applies, increasingly, across the globe. Remember, the GDPR is not based on European ideas. It is based on American ideas from the '70s. The so called FIPs, the Fair Information Practice Principles, which emerged in 1973 and are enshrined in the 1974 U.S. Privacy Act. It is a long established concept that those who surveil you are responsible. Therefore, the notion that people should forensically cover their traces, and wear tinfoil hats, is not useful, although it is one that Google etc. constantly harp on. They like to talk about the “privacy paradox”, but in fact, the debate was decided before these companies were founded. Users can, and do, install Brave to protect their privacy — but for many the advantage of Brave is speed and convenience. Privacy as a feature may not be the most appealing thing.

Do you think they regain any of that debate with the Privacy Sandbox or is it too late to rewind that business model?

Well, there are a few questions here. What is the “privacy sandbox”? Is it something of merit? And will it ever exist? The answer to those questions is completely unknown.

One thing that they've been talking about recently is some plan, apparently two years hence, to block third-party cookies. Let me put that another way. Google, which owns the most popular browser, Chrome, is talking about introducing in two years’ time, what all other browsers are doing as a basic security measure.
Brave has been blocking third-party cookies for several years. Apple’s Safari does too, and Firefox and Microsoft Edge recently joined our party.

There wasn’t really any opportunity for Chrome to not do this, although they’re trying to delay doing so for several years. Google’s sandbox plan is potentially a set of useful but basic privacy protections. But I wouldn’t say it’s a wholesale one, nor does it protect one from Google’s own systems of surveillance. Note the words “third party” in the title of Google’s planned cookie measure.

Do you think what GDPR lays out is effective enough? Do you think companies have really reacted or they just burying their head in the sand or hoping that they’re not singled out from the pack?

When the GDPR was first applied in May 2018 it had been teased for 10 years by the European Commission. What is often forgotten is that the GDPR changes very, very little in established data protection law. What it does do is to make that law matter, by introducing sanctions and the power for data protection authorities to ban processing. It also introduces ideas about how multiple data protection authorities might be able to coordinate a little bit better. This made things that were already law tangible for companies.

By and large, in online advertising, very little has changed. Thinking about real time bidding, which is a large part of the problem in online advertising and its misuse of personal data: the Google approach to providing security of data is broadcast to thousands of companies is to tell these companies that they are responsible for notifying Google at their earliest convenience, if they plan to misuse the data they receive. Now, that sounds exactly like the GDPR description of a data breach in Article 5(1)f. Hundreds of billions of bid requests are continuously sent by Google’s RTB system to thousands of companies every single day, about what every person around the world is reading, watching and listening to. This is the biggest data breach of all time. Google’s most recently proposed GDPR measures are immaterial, because they do nothing to protect the data.

The same goes for the other main industry body called the IAB. It has what it calls a “transparency and consent framework”. You could accurately describe it as a hope and a prayer. It has no technical ability whatsoever to prevent the thousands of companies receiving data in the real time bidding system from doing whatever they would like to do. Instead, it merely sends a request that accompanies the data about what we are all doing online to thousands of companies: ‘Please do not let yourself use these data because you haven’t been told you can use them’. It is a “scout’s honour” system. Essentially, there’s a data free for all at the heart of the online advertising system. The Google and IAB RTB system broadcasts hundreds of billion times a day, sometimes sending out the most sensitive data you possibly could about people.

There has been no improvement since the introduction of GDPR of the security of these data. This is why I started a campaign of GDPR complaints, to force regulators to investigate and enforce Article 5(1f), the security principle. As a result, Google is now under statutory investigation from its lead GDPR authority, which is the Irish Data Protection Commission. The IAB is under investigation of a slightly different type, but for the same reasons, by its lead authority, the Belgian Data Protection Commission. We’ve had submissions to 16 regulators around Europe, submitted by 21 partners, NGOs, and individual activists.
Where RTB goes now is inevitable: it will be forced to stop broadcasting of personal data, because it’s entirely indefensible, and it’s quite an open and shut question.

The latest proposals from the IAB and Google are patently ridiculous. Google claims two things: first, it says it is going to try to attempt to audit thousands of companies that it broadcasts data to, to check what they do with trillions upon trillions of items that it has sent to them, and continues to send to them, without any investigatory powers whatsoever.

The second thing is, although at least there’s an acknowledgement perhaps that a little bit of auditing might have been useful in the past. The second thing is it says it will remove what are called contextual categories from its broadcast. So, if I’m on the website called gay.co.uk, which is targeted at homosexuals, and if I’m reading about a particular sex act, the bid request will have that URL in it, and is likely to also show in general terms where I physically am too. It may also have a tag that says gay life, for example. Similarly, it could say heterosexual life, if I’m on straightpeople.com. The same could apply to health conditions, religion, and so forth. Google says it will remove the contextual categories, from February on, but of course, it’s including plenty of data that allows you to query the URL itself and figure out what’s on the page. So, that makes Google’s removal of the contextual category meaningless.

The IAB has also made inadequate proposals, which came at the end of a six-month grace period from the U.K. regulator, the ICO (after which the ICO continued to do nothing). The IAB produced a 9,400 word document that contains only two proposals to address the data breach at the heart of our complaints against RTB. One proposal was to issue non-binding guidance to companies that use its RTB system. The second was to ‘start a work stream’ of discussion under the auspices of the transparency and consent framework, which is unable to make any impact on the security of data. So, zero progress from industry. It now falls to regulators to issue banning orders, so that the industry is no longer permitted to send personal data out in billions of bid requests every day.

On the point of the regulators, what are you seeing, in terms of the appetite to enforce? You mentioned the ICO review that’s been going on for the last, I guess, eight months or so. Do you think the regulators would go as far as banning processing? Are big fines enough?

We know from competition/antitrust enforcement that fines are immaterial if continued wrongdoing offsets the cost of the fine. I think that the regulators who matter understand that now. What matters is their power to ban processing. In other words, they have the power to force companies to change how they do business. That is the power that matters.

Now, I’ll say a word about the ICO. It hasn’t had this for eight months. It is 16 months since it received these complaints. 16 months since it had formal evidence from us about the biggest data breach that the U.K. has yet experienced. It has yet to lift a finger to stop that data breach, which continues every day.

The funny thing is though, that actually I blew the whistle back when I was inside an AdTech firm on 17th of January 2018, two years to the day before the ICO announced that it still wasn’t going to do anything, at the end of the six month grace period. That’s 24 months since the biggest data breach the U.K. has ever experienced, a breach that continues today.
I have very little confidence that the ICO is going to immediately move to do something sensible here. I fully expect that I and the complainants from the U.K., which Brave will be supporting, will be taking the ICO to the Information Tribunal, a Court that oversees the ICO. We will do whatever is necessary to compel it to perform its duties. If that means that we have to take a challenge against the U.K. Data Protection Act, if it should turn out that there is inadequate judicial oversight of the ICO, and that we have to then challenge the implementing law, which will then cause an issue for the U.K. adequacy decision post Brexit, we will do that too. This issue is not going away.

The U.K. regulator has failed to act for two years. If you're unable to act against a highly publicised data breach, which is without question the largest data breach your country has ever experienced, then you either there is a problem with the regulator, or a problem with the law. We will test both of those questions if forced.

Let me move on from the ICO. The Irish data protection Commissioner had already launched its statutory inquiry by the time the ICO published an interim report in June. That ICO interim report merely repeated back to us what was in the evidence that I had put together ten months previously. The Irish, however, I do believe will deliver. There is enough pressure behind this particular case that it would be very hard for a regulator not to do so. They would probably deliver of their own volition in any case. I'm not as close to the case in Belgium, so I can't give an opinion, but I would be surprised if any data protection authority could look at the evidence, which is self-explanatory, and not find that they have to come down against the data breach.

In terms of the Irish data protection commission delivering, do you have any sense of their timeframe around this particular matter? What it would involve in terms of delivering and could it end up in a really lengthy court proceeding? How should we think about this all unfolding?

The Irish regulator will apparently be issuing a decision on a Facebook WhatsApp case first, then I believe there will be a Twitter decision after that. I think the RTB decision may follow. Helen Dixon, the Commissioner, has says publicly that there will be a decision on all these issues in 2020. My guess is it'll be close to the first half year.

Now, what will then happen? It is entirely possible that there will be an appeal against that decision and that we will all end up with the European Court of Justice, which is fine by us. It will be apparent though, before we get to the ECJ that the industry has run out of road for its current approach. We've already seen Google blink and it's already starting to talk about some changes that are more material. I think once there's a decision from the regulators on this, it's going to be clear that the game is up. Right now, I think most industry participants are simply playing for time.

If we step forward a few years and you look at the future for a second, where does the business model of AdTech go? In the utopian future world, is there a situation where the consumer can be compensated for the use of that data?

The Brave model is partly that. Let's get into a few future models.

One possibility is that there may be a form of real time bidding where the broadcast continues to go out to thousands of companies to solicit bids from advertisers for the opportunity to show a person, but the broadcast no longer contains personal data. One would remove fine grained timestamps, fine grained location, and so on. That's an argument for computer scientists and statisticians because it's an open
question, unproven, as to whether combinations of data in a future RTB system could ever be non-personal. I’ve talked to one computer scientist who is very senior who believes they could be. I’ve talked to another who believes they couldn’t be. Both of these computer science professors work for our company. So even internally, we have two professors who don’t agree on whether it’s possible to have a safe RTB system. If it were possible, a safe RTB system that operated with just non-personal data could solve a few market problems.

First, you would not have a problem of audience arbitrage. Where if I go to The Guardian in the morning, and a bid request goes out about me that identifies me as person who wants to buy an expensive car, who visits the Guardian and reads car reviews, the advertiser could reach me at a huge discount on a crappy website later in the day. That’s what currently happens. The Guardian’s audience has been commoditized. If the website I’m on and I’m advertised to at a discount is a ‘bottom of the web’ website, The Guardian actually enables that ‘bottom of the web’ website to have a business model and that business model undermines their own. That’s a big problem.

The second issue is if a piece of software masquerading as me, goes to The Guardian, pretends to want to buy a car and then is bought as a purchased visitor by a corrupt website to make money through ads shown to me; all that is only possible because of the leakage of data in the real time bidding system. So, audience arbitrage and ad bot fraud are big, big problems. Also, the intermediaries who enable these discounts, which sounds like a very good thing to an advertiser, are charging opaque and very high fees ranging from 40 to 80% of the advertisers dollar. I’ll give you one example. The Guardian, which I used as an example previously, ran a sting operation. It masqueraded as a fake brand, or an advertiser, and bought ads on its own page using the real time bidding system with all the bells and whistles switched on. For every pound Sterling that it spent to buy ads on The Guardian website, it receives back from every pound, only 30p as Guardian the publisher, 70p went off into the intermediaries.

The model today is the broadcast of dangerous data, which are personal data. If feasible, it is still broadcast, but of supposedly safe data, if that’s possible. That would reduce but not eliminate fraud, it would reduce the AdTech tax because you no longer offer these discounts that arbitrage my audience, and it would reduce audience arbitrage. So there’s a lot of wins for publishers, for media consumers, and for advertisers. There’s a lot of loss for the intermediaries, who, today, are extracting huge rents, and reselling personal data under the table too.

In model two, by the way, I’m assuming that you have an RTB system that works without personal data in it, you would need the regulators to have enforced across the industry so there is no first mover disadvantage. If publisher one trips their bid request to make them safe, if such a thing is possible, but publisher two does not, then publisher one will be disadvantaged in the market, it will receive far less cash. You need regulators to set the market conditions by banning, at the RTB standards level. Those standards are controlled by Google and IAB, so it’s feasible to do this. The regulator would have to tell the IAB and Google ‘it is unlawful what you’re doing with the standard, you need to make them non-personal data only’, if that’s feasible. In that case, all publishers would now be sending supposedly safe bid requests and none of them would be disadvantaged because of that.

Model three is to have some deal where a user or visitor to a publisher has a relationship with the publisher, and the publisher, and maybe an intermediary, does a deal with an advertiser and the user is involved in that.
So, if I go to The Economist and The Economist and Audi comes to me, and The Economist says, ‘Hey, we’ve got some offer for you, for Economist readers’ and I say yes. Then you’ve only got two or three companies involved. You might have a few intermediaries, but I can know who they are. The data is still controlled; that's entirely possible. So that's a hyper premium thing that the publishers that are trusted will now be able to do in a way that possibly Google Facebook etc. might not be able to do. Certainly you won't have clickbait to be able to do that kind of thing.

Model four is, and Brave exemplifies this model, is to have an opt-in only system where a person can decide, if they want, to allow their web browser to build a profile of them that never leaves the device. In our case, Brave as a company can never access this profile. Your device functions like a Faraday Cage for your personal data. But your device can know you in a way that no third party ever can. Your device goes on dates with you, your device may even walk up the aisle with you when you get married. It may even be in the coffin with you when you die.

The device knows what the device knows. What's important is that no company should know what the device knows. You are asked when you install Brave, ‘would you like to switch this on?’ We certainly hope you'll say yes, but if you say no, that's fine. If you switch on Brave ads, then a highly intimate and knowledgeable profile is maintained and built on you over time on your device. Every day, in fact I think every few hours, everyone in a given geography receives a catalogue of ads, the same catalogue, but their device profiles are highly personal. So your device has profiled you and decides, well this person tends to read reviews of motorcycles between 11am and 11.20 most weekdays. There's an ad here for the particular category of motorcycle that they seem to read a lot about so maybe they want to buy that motorcycle. This ad is offering a driving test of this particular motorcycle and it might be that my user wants the opportunity to decide whether they would like that test drive. I know this because it’s on their device, and they said it's okay for me to store this internally, and never tell anyone, including Brave. Only the user's device knows this. Now, if that happens, the user then sees a notification at 11.05 on Tuesday saying ‘hey, Yamaha would like the opportunity to bring their bike over to your office tomorrow at 11.15. Here's the details of the bike, would you like to do that?’ If so you should send your availability, confirm that you're okay with them knowing where your office is, and you might have to send a photocopy of your driving licence.

I'm making this example up as I speak, so bear with me. Once the person using the device has said, ‘Yes, I want to know more about that offer. You have my permission device, to tell Yamaha about me’, then the advertiser can be talking to the user, the device can get out of the way. What the device has done there is to act almost as your personal assistant. It's receiving all of the spam, and it's deciding which spam might actually have some value to you. In Brave you have a little dialogue list where you can decide how many ads you want to see. Do you want to see zero ads an hour or do you want to see five. In that situation, the user will receive a cut of the payment that Yamaha has made.

There are different formats but the user always receives, at least as much as Brave itself. The user has a wallet, which they can switch on if they wish, and ads only work if you switched on the wallet so you'll always be accumulating this reward if you're seeing ads. The way that works is, there's a micro payment system based off a cryptocurrency called the Basic Attention Token (BAT). The Basic Attention Token can be cashed out by a person for real money, and in its very first seconds of life, a few years ago, something like $35 million was generated with the people buying BAT in the space of 30 seconds; it is now worth far more than that.
So, each BAT you receive, each item of that is worth cash. If Yamaha has paid $30 or $80 for the opportunity to show you that ad, you will receive a significant amount of cash by accepting that ad. It works from that perspective for the user too.

What proportion have opted in for this and allowed the profile to be built on the on the device and they've got the wallet switched on, etc.?

The answer is, I don't know. Only with the launch of version one in November did Brave extend what I described to multiple platforms. I think it started on Android and now it's on all platforms. The only figures I have are using a not very good industry metric to show how effective those seem to be. I would caveat that strongly. The industry uses click through rates, which measure how many people have clicked through from an ad to the thing being advertised, to figure out how effective an ad is.

The figures for Brave are pretty staggering. If you look at “Google Display”, which refers to the normal rectangle ads that you see on YouTube or any other Google property, the percentage of times people click through from an ad is 0.46%. Facebook’s equivalent is higher, 0.9%, but still less than a single percent. Search ads on Google are under 2%, 1.9%. Brave’s figures are 12%, which is literally an order of magnitude higher.

Now that might change. Brave is in the early stages; it only came out of beta in November, so who knows what the future model is as far as how you might advertise, which companies you deal with. What will not change is Brave will never allow data about the person to leave the device, unless the person makes an conscious and extraordinary decision for that to occur. This is the exact opposite to the normal situation.

If you think about all the various stakeholders operating within the online advertising industry at the moment, if the RTB model changes or there’s this sudden change in ability to use third-party cookies, what do you think the impact would be for the various stakeholders? So the publishers who, as far as I'm understand seem to be quite worried about it, the AdTech companies and the brands? How is it managed in a fairly orderly fashion if this is to try and help the publishers better monetize?

Third-party cookies have nothing whatsoever to do with how well publishers know their audiences. That's not what they are for. I take issue with the question and reject it 100% Third-party cookies are about other people knowing your audience if you're a publisher. If third-party cookies did not exist, publishers will be able to sell their audience.

Here is the problem. If third-party cookies are removed from the equation, but at the same time, Google, Facebook, etc., are still allowed to operate an internal data free for all where they get the benefit of their plugins, on apps and websites all around the web; those companies alone will have insight on the visits to someone else's properties. What we need is internal, not just external, but internal data protection enforcement. The principle in data protection law that needs to be pursued here is the purpose limitation principle. It's article 5 (1)b. Again, it’s an old American idea from the 1973 FIPPs that says data should be ring fenced by purpose, and not used beyond the purpose for which it was collected. It’s a very simple idea. Right now there is an internal free for all in Google, Facebook, and so on, that gives them an overwhelming advantage now, and will give them an even greater advantage in the absence of third-party cookies.
Are you not of the school of thought that seems to be fairly commonly discussed, that regulation around data protection and privacy seems to create more of an advantage for the Big Tech companies where they’ve got huge troves of data already?

That's absolutely wrong. Data protection problems in the online advertising market are many and varied. For simplicity, you can separate them into two categories, external and internal.

The external problems are where you have thousands of companies passing data around in an external data free for all. The internal problems are we have large companies, passing data around internally in an internal free for all.

Think about the situation where Google automatically combines data across all of its lines of business, even if it's totally unrelated. It combines these data, and cross uses them. It automatically signs in users of one service into all of the services. Now in competition terms that is tying, bundling, excessive collection and an offensive leveraging of personal data. If you were a competition academic, you would probably call it “platform envelopment”. The net effect of that is to entrench Google's dominant position, reinforce barriers to entry and exclude competitors. The same can be said of Facebook.

Now, that entrenchment, which comes from this internal free for all, will continue no matter what is done against the external free for all. It is true that entrenchment could accelerate if there is data protection law enforcement against the external free for all. However, that will only happen if there were no enforcement against the internal free for all. I'm using the word free for all many times here but the focus of the problem.

What purpose limitation does is it prevents a business from using and combining data internally. If you have plugins across the web, you do not automatically have the opportunity to use the data you get from those plugins to enhance your ad targeting business. If you have a system of mapping, for example, Google Maps or an email system, YouTube, any of them, you don't automatically get to opt everyone in. The second thing is, there are two other issues that prevent what in competition law would be called privacy policy tying, automatically bundling and tying all of your things together in a magical one size fits all for all of your data use needs privacy policy, which is what the companies currently do. Often what they do is they misclassify the data they're using as so called data or personal data. Actually those data are special category data. Now, Article 9 of GDPR defines special category data as any data that reveals ethnicity, political or opinion, religious philosophical beliefs, genetics, health, all that kind of stuff. Now, most of the data that is good for advertising is probably special category data and that means that for every single use of it, for each different type of thing you do with it, you should probably have what's called explicit consent. The user should probably then have ‘it should be as easy to withdraw as it was to give consent’. This would make these businesses essentially functionally broken up, internally; their ability to produce data would be defeated.

If we had enforcement of the GDPR, which we don't, internally within the companies, by the regulator, that would not be a thing that people are talking about. It is good that we’re now almost at a point where the regulators will be enforcing against the external issues, but they must, in the round, they need to be enforcing not only externally but internally.
What do you think will prompt them to do that? Will that come from the regulator on the data protection side or do you think that’s more likely to come on the antitrust, competition side?

We have a situation where an antitrust regulator, the German antitrust authority, attempted to use data protection law to prevent the cross-use of data between companies, between subsidiaries of Facebook. This was rejected by the regional court in Dusseldorf, and is now a matter of appeal before the German Federal Court, so we have to see what happens there.

The mistake, though, that caused the rejection on appeal from Facebook, was that the German authority did not make a complete case, it appears, in antitrust law. Had a Data Protection Authority made pretty much the same argument, it would have been able to force on the basis of data protection law. I think they were doing exactly the right thing; they may just have been the wrong player. Back to the U.K., you’ve got the U.K. competition and markets authority and you’ve got U.K. ICO. If they were to collaborate and the ICO led the way on enforcement, it would be within its rights to be able to enforce these things that would essentially fix the market.

Do you think there’s anything the companies and the large tech companies can do to try and hold the regulator off and placate them? Is there anything where the companies can be more proactive about these things so they can they can manage it rather than have it forced upon them?

What can industry do to hold this off? Well, it’s been doing a very good job for well over a decade, it just keeps doing the same thing.

Unless the ICO were to receive tremendous pressure, it’s entirely possible that it might let this issue keep going. But, it will receive that pressure. Complainants have the right under law to vindicate their rights, and they will have support from Brave in doing so. The ICO will eventually have to do its job.

I think the other regulators will do the job anyway. And if they do not, they will also come under the same pressure.

The ePrivacy regulation has been talked about for some time, and there was hope it would come in at the same time as the GDPR, and obviously didn’t. Is that a piece of regulation you think could still come down the pipe and have an impact or do you think that has died a death?

I don’t know the answer to whether it will or will not happen. I talk to a lot of people in Brussels. We just do not know, at this point, what the next step is, and if it will bear fruit.

There are two things that we are very enthusiastic about in the draft text that the European Parliament provided. First, the outright ban on so called cookie walls is important. Cookie walls are tools that prevent you from accessing a website or an app, unless you agree to use of your data that is unrelated to accessing that website or app. I think that’s clearly a mistake, or something like that. Right now the law does not allow it but the industry has been pushing very, very hard on the Council of Ministers for cookie walls. The second thing is both the Commission and the Parliament, in their text, has something called article 10, Privacy by default. Article 10 is where any service provider operating service, operating system or browser to give users a menu when they first run the system to allow users to say what level of data sharing they are happy with. By default, the answer is none or just essential. This will help extend the approach Brave is taking and we will be very supportive of that.
If they are going to tackle both the external and internal challenges around the data free for all, as you call it, is it something that Europe can do alone, or do you think this needs to be addressed at a global level? Do you think there’s appetite from other regulators globally or other markets to follow suit?

Let’s take a look at California for a moment. Forget about the CCPA, I think the focus should be on CPRA. When Alastair Mactaggart proposed the CCPA, they had in mind a very, very different law to what eventually was passed. Now he was convinced to let it go, instead of being on the ballot, to the normal legislative process and it got watered down. Now, the follow up, which he will be putting on the ballot in November is far tighter. It exists because he was not satisfied with what came out of his first invention. So CPRA, which will be put before California and voters in November, now has a legal definition of ‘cross context behavioural advertising’, for the first time ever. The definition of so called sale is now so broad there are no loopholes for AdTech. It is good to see that the California law is closing the margin of manoeuvre for an AdTech industry that has operated, in my view, in a very bad way.

Europe does not need other jurisdictions in order to solve problems within its own jurisdiction. The European digital single market is huge. It’s so vast that every large brand needs to be able to sell into; that market has its own rules. If you want to trade in that market, you must accept the rules. Irrespective of what a company like Google may do anywhere else in the world, it must abide by certain rules in the European market, with no ifs or buts. There's not confusion about that.

What is more likely, in my view, is that the number of jurisdictions around world that are now discussing, or actively enacting, GDPR clones is quite interesting. The GDPR is becoming an almost emerging de facto standard. So I think you’ll see an influence outside of Europe.

AI is being discussed and lots of innovation that requires big data, do you think that has an impact on how corporates think about innovation and where they should focus that geographically? Do you think it creates any disadvantage for Europe at all from that perspective?

If you're going to have a digital market that functions, that sustains legitimate media, then there clearly is a need for change. If we continue as we are, I don't think we're going to have a viable media market at all. That will have implications not just for the economy, but for society. You need a foundation of trust for the consumer in a market if the market is to be viable. But this is being eaten away at every day. There are profound problems that need to be fixed, because the status quo is not viable in the longer term.

The idea that is being floated about this single data lake within Europe, on the antitrust side of things, how does that work with GDPR? Are there areas of tension between the two coming?

I haven't looked into the data lake idea. My feeling is, it is very, very hard to have a large quantity of data that is truly anonymous. If that problem can be squared and who knows what's possible. I would be sceptical until I understood the detail.
Aside from Brave, are you seeing any other innovations that have been driven by the increase data privacy?

There’s a company called STIR in the Netherlands. It is the sales house of the national broadcast cooperation there. STIR has set up a system where, for the 90% of people who refuse to give consent for normal advertising, they advertising that is based only on context. What you’re seeing is revenue that they’re getting for that 90% of people using just contextual targeting, is actually incredible. This is a big win for their business. That is contrary to the normal narrative you get from the AdTech intermediaries.
A Conversation with Justin B Weiss
Global Head of Data Privacy, Naspers

What are the main challenges corporates are facing in managing and utilising data? Do you see any variation in these challenges across sectors and geographies?

I think the first big challenge that many corporates face is contending with the problem of changing (or new) requirements. The reality is that when you’re small, new rules or new operational requirements related to your data might end up being relatively straightforward to build into your processes, or to your systems, because there were not many existing processes in place to start with that need to be changed. In short, you don’t necessarily need to break something to fix it. That’s not to say that smaller companies don’t face challenges, but contrast that with a larger company.

When you’re big, updating something that is already there is one of the key challenges and is a price of being established. You’re relying on a lot of existing infrastructure, infrastructure that might have been built over time that you have grafted on through mergers and acquisitions, incorporating different types of systems that work in different ways. Data regulation can require us to tinker with a structure that’s in place.

When you are being called on to grant user rights, access or correct or delete information, to honour a right to be forgotten, to locate and map all of the locations of data that can be ever shifting within an organisation, and to maintain that organic map, these kinds of activities are very complex for a big organisation. I think among the companies that I work with and try to help support—whether small, medium or large—some of the biggest challenges for them come from what might seem like an easy task of doing a data mapping, or locating all of the different instances of a certain person’s personal information across many—maybe even hundreds—of systems. That’s a really big challenge. It’s a challenge that gets more complicated and sophisticated the larger the organisation.

Since GDPR has brought all this to the forefront, are there ways that corporates have gone about doing that, or is it still as big a challenge as it was a couple of years ago? Do you think corporates are managing to get on top of it?

I definitely think that the industry and corporates in general have made a lot of progress. We also benefit from the proliferation of third-party vendors who are now specialised in providing compliance support to big corporates. There’s some consolidation been happening in that sector and you may be familiar with that but there’s even some venture capital going into this new industry of automation for compliance. In the field of data protection, you have companies emerging that are offering technical solutions to make data mapping easier, to make user requests easier, to make consent management easier, to make cookie management websites easier, and more automated and streamlined. I think all of this development is very welcome.

When it comes to helping corporates drive forward in a state that allows them to be compliant, that part of it, I think, has been progressing nicely and will continue to do so in terms of companies relying on these types of vendors of third-party assets. On the other hand, some of the fundamental requirements do remain challenging, even for an organisation that has a certain maturity level.
These solutions can be very resource intensive and time intensive, from a human and technical perspective. They put strains on budgets; they require an expansion of budgets to an area that might not have been allocated previously. They put a lot of demands, not only on the time of those that work in the privacy office, but on individuals across the entire organisation.

Since GDPR, we’ve seen the proliferation of questionnaires: ‘let me know what data you’ve got in your department, what you’re doing with it, etc.’ Then humans must respond to those questions, often in large numbers given the diversity of activities in a company. Then you have a small number of individuals reviewing the answers and providing feedback and recommendations about how to look after the privacy and security implications of those responses. So there's a fairly complex web that’s being woven in order to support, not just compliance from a snapshot perspective, but how do you set up a process that's organic, can be flexible over time, that will update itself when there are changes. It's a very heavy list and it's quite operational in its nature. I think, particularly for those privacy professionals that have a legal background, some of the new skills they're having to develop very much go in the direction of project management. Managing a large number of multiple stakeholders, maybe managing vendors, managing systems and tools, and making sure that the ship continues to sail and run and update itself. Many lawyers in the field are having a trial by fire when it comes to learning how to implement these types of programmes.

*Do you think corporates are taking this seriously and providing the resource and finance needed for it to be properly tackled?*

At the board of directors’ level, there’s always been from a governance perspective an understanding of the range of risks that the organisation faces. A risk committee looks at these types of questions on a consistent basis. In the digital age, the importance, depth and breadth of digital topics on the board of director’s agenda has gained ground.

Material financial impacts tied to data protection topics might come from brand damage or mismanagement of security incidents in the cyber security context, as well as concerns about user trust. There have been very public and widespread privacy concerns emerging, particularly for the Internet and digital sector which, I think, boards recognise must be taken very seriously in a proactive way. Of course, the risk of fines, which has certainly increased and features as one of the key changes of the GDPR, brought the recognition that something akin to the level of competition fines could be appropriate in the context of data protection violations, and that is, of course, a motivator for those portions of the risk community and those advising boards of directors on risk.

*You work for a business that operates across lots of markets with very different approaches to uses of data and regulation. How do you think corporates can manage that evolving regulatory landscape? What are the main challenges?*

Corporates come in different flavours. From a governance perspective, you have some corporates that are extremely centralised and the decision making happens in a central office, and the world must follow along. Then there are other corporates that are extremely decentralised but allow their individual business units, and maybe sometimes at a country level, to look after themselves and be more local compliance oriented. Then I think you have organisations that are hybrids, depending on the area.
There are definitely organisations that are seeking to use GDPR as a global standard, either because it's perceived as the highest standard or convenient from an operational perspective to apply the same rule set to systems that are being used in multiple jurisdictions, like the ones you're talking about. I think other organisations are seeking to silo some of their businesses or business units or countries away from that, and take a very local compliance perspective, although it is becoming increasingly difficult given the proliferation of GDPR applicability among vendors; some of the major cloud providers are subject to GDPR. As a result, the GDPR framework often comes with those partners. Let's say you’re a corporation, based in South Africa or India or in China or in Latin America, you may well find that quite a bit of GDPR comes along with just doing business in the digital age with cloud providers and the types of vendors who offer support services for companies in that space. So there’s definitely a global impact, there's no question, whether it comes to direct applicability of GDPR, whether it comes through your vendors imposing GDPR requirements on you through contracts, even in developing markets, and as a result developing economies professionals know about GDPR. Many of their practices are having to align to it anyway.

On the flip side, although each country is on its own timeline, it's crystal clear that the march toward data protection legislation in many countries is well on its way. Most major economies have either individual pieces of data protection regulation or have proposals and bills to bring forward data protection legislation. The GDPR is absolutely a benchmark and a standard for markets to consider. When we look at Brazil, when we look at India, when we look recently in Japan, when we look at the United States, these economies are certainly noticing that data protection regulation is important in this age, from a trade perspective, from a governance perspective, from a citizens trust perspective, it's becoming a cost of doing business in the sector. There are really not very many places to hide from good data protection practice, for data intensive companies.

Do you think it’s the right approach for some of these markets in Brazil, India etc. for them to take GDPR as a template, do you think that works, especially if the government's desire might be around the uses of data and localisation etc.?

I think it's very important the major economies take GDPR as a referential, one desirable outcome no matter what the economy, is that the data can flow. Data by its nature wants to be cross border, wants to be able to go from point A to point B. One of the ways that those international data flows are supported, is through legislation that can reflect some common standards and common requirements. The peculiarities of each jurisdiction when it comes to the political or economic situation of the country, of course, must influence the local version of data protection legislation.

I do think it's wise, even from a basic digital trade perspective, to recognise that the European Economic Area is going to be a very significant partner for all of the jurisdictions, so to ignore it would be unsound as an approach. However, to mirror image it or copy it will not work in many jurisdictions. I think it is unrealistic, given the different political systems, the different government structures and the different capabilities of institutions, to expect all jurisdictions to mirror in a precise way the European way of doing things. However, I think that there is some cause for optimism with respect to interoperability between different countries.
When we look at the recent experience of Japan, Japan was able to successfully pass its own data protection legislation and negotiate with the European Union for an interoperable adequacy determination. I think that's very promising as it's a really large economy that is culturally very different from Europe. To see that example come forward, I think it's cause for some optimism, although we don't have enough examples like that to really show that parity can be achieved through different variations of data protection legislation. The predominant model that seems to have been implemented in most jurisdictions that have been recognised by Europe is indeed the European model.

I'm not naive to this tension. Yes, I think that we are on a path where regional variation, necessitated by local customs, local culture, local economics, local politics, will inevitably have to be accommodated in this global environment. Right now, each country's relationship is being taken as a case by case matter until we have a global treaty or a global instrument that can replace these bilateral negotiations. The more convergence we see among this variety of bilateral agreements, the more a multinational and the corporate can start to adopt a unified view on what their data protection regime and company approach needs to look like.

*Given governments increasingly realise the value of data as an asset and some have put in place the idea of data localization, not necessarily wanting it to move beyond their geographical borders. Do you think there is any hope of a global treaty at some point in the future?*

There are some common values and common standards in privacy that can be universal to countries. I do think that there will always be sticking points where there is regional diversity and regional variation. The real test is how to get through those sticking points. Can we get comfortable with differences and find flexible ways to achieve the same ends using different means, and sometimes different language? That's a challenge. It's always been a challenge. But I think that's the work of international data protection right now is to find those points of convergence and to bring up levels of data protection for all populations, whether they're rich or poor, or big countries or small countries. The sea level needs to rise in a good way from the protection stand point and I think it has been, frankly, over the past five to 10 years. Some of those improvements come from the tech sector itself, some of those improvements have come through regulation. I think it's really the combination of the two that's likely to bring the most globalised privacy protections for the most citizens.

*Have you seen much change in both how corporates are and can use data but also how regulators are looking to enforce it? While we're two years down the line, which in the investment world feels a long time, is this something that's going to take a long time to for the regulators to really get to grips with?*

I think the first comment would be that one, the historic drivers for enforcement haven't changed. The public regulators respond to public complaints. They respond to public priorities, whether it comes through media coverage of specific types of issues, working through just the raw number of complaints that they get, I do think regulators rightly prioritise their constituents. That's always been the case and it will continue to be the case. The penalties may be higher in certain jurisdictions now, but data subjects themselves drive that prioritisation.
The second point is I do think that many regulators have been investing in their own technical capabilities. Some of the better funded, larger regulators in the world have been investing in technical expertise, consultant groups, trying to really beef up their understanding of how technology works and what solutions may be out there to address privacy impacts of technology. That trend informs enforcement as well, because it strengthens the regulator’s capacity to negotiate good solutions with companies and also to call out bad acts and to own, and not be blinded by, any lack of expertise about the technology itself.

That’s a trend among larger, better-funded government agencies that are able to do that. They are sharing that information with the global community. The Global Privacy Assembly, which meets every year, is the group of regulators that through closed sessions brief one another on their learnings, on their observations, on their investigations. This information sharing is a big priority in the community. That informs not just investigations in one jurisdiction, but potential for investigations in multiple jurisdictions through coordination. So that’s the second thing that I think is interesting and important to observe.

The third thing that I would add, if we see for example CCPA, and we see some of the new provisions of GDPR, is the rise and role of third-party agents to help enforce the rules. The fact that either through litigation, class action or not for profit groups that want to make it easy for members of the public to make requests, the streamlining of processes enabling others to advocate on behalf of larger populations in the area of privacy is happening. I think it’s also informing privacy enforcement. There are an increasing number of cases originating not just from one subject, but groups of individuals represented by organisations that are purporting to bring forward complaints, which are garnering attention of regulators.

Have you seen any change in terms of consumer interest or behaviour and how data is used by corporates? Do you think there is a real appetite from consumers for more transparency?

Different companies and groups will have different metrics for figuring that out. One metric that I think is immediately observable is that when you open up rights to individuals in a specific jurisdiction to make requests to access, correct, delete, and suppress their data, you can measure the number of requests that come in. We do have a regional variation in volumes of such requests. Certain countries may have a very high spike in the number of requests that come in, and then it tapers down. The citizens are constantly exercising those rights. In other jurisdictions these rights are not really taken up or used much.

This regional variation must reflect, we presume, two things: cultural variation and/or the success of publicity campaigns to make people aware that they have these rights and easy to use techniques to exercise them. In some jurisdictions, government leaders themselves spend a lot of time promoting these rights and making sure that the public knows about them. Other third-party organisations have campaigns to make sure people are aware. Some companies make it really easy to exercise the rights through self-service tools. If you offer those types of self-service tools, you have metrics about how many people are using them on a daily basis. That’s definitely a cornerstone for objective criteria you use to measure the level of engagement from the public on this topic. You know how many rights requests companies are managing, but that’s also got a dependency on general public awareness, awareness of the rights and mechanisms to exercise them. I do observe increasing awareness. Every time there’s a story in the news, there’s increased awareness among members of the public who follow it.
Such engagement is an important thing to continue to measure and track, whether the number of those requests go down over time, or increases over time would be a function of many different factors, something I think people should continue to research and take very seriously.

Are there any particular geographies that stand out either way, where there’s been a huge spike?

It's a difficult question for me to answer in a succinct and specific way because we monitor so many jurisdictions. I'd literally have to go through all our jurisdictions, because they are all important. In general, I would say that European citizens post GDPR seem to have a pretty good understanding of their rights. It's interesting because they've enjoyed these same rights or very comparable rights since 1995. I think it's the publicity and the awareness raising that happened thanks to GDPR, reminding people that they may be exercising rights they didn't realise they'd had before. Because of publicity, and because there are new buttons and features to enable those rights on websites then sometimes maybe even out of curiosity a person might click something as they'd like to know what a company has on them. It can be driven by genuine concern. No matter what, I do think the awareness level in Europe is pretty high. Other countries around the world that are quite European oriented, maybe because they have a European language as the principal language, in South America or in Africa, for example, have a pretty good sense that something is afoot in terms of rights in Europe. Awareness has also grown outside the European Economic Area and there are a lot of users of European products around the world that have been playing with these tools and exercising rights as well.

There are large markets where it's not clear that there's significant engagement yet, but it could change with future laws. If you look at Brazil and India, two interesting markets with big populations, Brazil's passed its law. Those rights come into force for Brazilians in the summertime this year. In the case of India, a very serious bill that would grant very similar comparable rights is on the table and expected to pass in some form, which brings another one of the largest populations in the world into that mix. It'll be really interesting to see culturally in India, what the Indian people care the most about, as reflected through what we can observe, and the number of requests.

Do you think this creates an advantage or disadvantage for those geographies with more stringent laws around data protection from an innovation perspective? And thinking about where capital might flow for your tech investment?

Compliance with these types of requirements comes at a cost. I mean, you have to invest the human and technical capital to implement systems that allow you to comply with these types of requests and to locate data inside of innovation. If you don't have those requirements, obviously you can spend that money somewhere else. If you don't have those requirements, some of those companies may choose to voluntarily pursue them anyway as a matter of self-regulation and as a way to align to the expectations set by their competitors. Some of the companies may have to implement these requirements anyway as cost of business because of the partners they use who are subject to that type of regulation.
I see the tech sector gradually maturing, the expectations governing data protection in the tech sector are increasingly just a cost of doing business. In terms of strategic areas for investment, are there certain areas that feel more data risky, or might come with more compliance burden that's been impacting investor’s decisions?

That's a super interesting question. It touches on how companies value and assess through due diligence the opportunities and risks presented in various types of data intensive investments. I think that one of the notable things that has happened since GDPR is found in the M&A space. Asking what is the strategic opportunity presented by data? Also, what are the compliance costs associated with implementing fully compliant systems and ensuring that customers maintain their trust levels before, during and after various kinds of transactions or mergers and divestitures or whatever that may be? The focus on data and data protection is one of the key factors in the success of M&A deals in the tech sector. It has been elevated by the GDPR.

Certain business models clearly come with more sophisticated privacy challenges than others and, as a result, when companies are evaluating the cost of integrating compliance solutions, you're going to have variability based on the business models. Business models that are inherently non-transparent, business models that happen in the background, business models that are extremely complicated, tend to face a lot of headwinds when they have to comply with data protection legislation. Whereas other types of digital businesses operate very comfortably in a strict, explicit consent type of regime.

If you make your regulation too business model specific, you end up impacting and influencing markets and that may not be desirable over time. Principles-based policy legislation that can be flexible, applied to different business models in different contexts, definitely feels pragmatic, particularly for a country, an economy, for a region or a planet that wants to benefit from the diversity of these models. I think we've seen even glimmers of that recently in the digital space. Everybody should be working to figure out what is the right way to pursue healthy data protection and, in the context of business models, if the business model is viable and socially desirable. So if a business practice is perceived as socially undesirable, then we need to be thinking about more than just the data protection piece, we should be thinking about whether this is a business that should be legal at all. If instead we think that there are social benefits from certain digital business models, then we should be working hard to find a way to adapt to support them.

I was going to ask you about AI and ethics. That seems to be quite a big topic of conversation within the privacy industry. How do you think about the regulation of AI or how it should be regulated? What are regulators looking to do in that space?

Governments and businesses alike are enthusiastic about the opportunities for AI. Those same communities are increasingly mindful of the potential negative repercussions of unbridled deployment of these technologies. There’s rough consensus that guardrails and bridles are needed in order to pursue a lot of these benefits. Even in the scientific community itself, there seems to be this recognition and this feeling that technologies can be good, but they can be used for bad, so we should have rules that prevent that.

The data ethics dialogue is a very, very healthy one, to recognise and really start brainstorming what are some of the bad things that can happen? What are some of the great things that can only happen if we use this data in this way? We have to put those uses on the wall and as a global community, we have to whiteboard them and we have to discuss pros and cons and options.
And I think that dialogue is very organic right now. Consensus is forming around certain key principles, but I don't think that the industries that are deploying AI have necessarily gone all the way to operationalize and give us a list of actual techniques that can be used to implement those ethics, principles and guidelines. We're still at the higher level, grappling with what the principles should be.

It's an environment that's not yet ripe for very prescriptive regulation. This is one of the reasons we've seen some countries holding back on trying to codify much more than high level principles and guidelines, lots of public discussion documents, hundreds and hundreds of ethics and AI material that you can refer to. For those of us that are trying to think over it and say, right, what are the common elements? What are people talking about? Where should we be focusing our attention? I think that work really is happening. It's happening in AI teams and in AI teams in partnerships with legal and policy support inside of companies. Good practices are already emerging that respond to those prompts and those provocations from an ethics perspective.

I think we need to keep good stories around and say, how do you address bias? Here are some testing techniques that you can use, or here's an opportunity to be transparent about how an algorithm works in this context. Here is our user interface, we think it's interesting as a way to create visualisations. All of that privacy ethics design within the context of AI is afoot, but much more of it is needed.

That doesn't mean you do nothing from a regulatory perspective. I think you tread very carefully and deliberately. Guardrails are needed on the banks of a river that is allowed to continue to flow. You don't want the water to be over running the banks and hurting people or structures. You want to be managing that innovation. That's an age-old problem for all kinds of regulation. Artificial intelligence, machine learning, data science and analytics has been around for much longer than people recognise in various forms, but it's having a moment now. It's having a sunshine moment. That sunshine brings a lot of scrutiny and hopefully this will be helpful in preserving a lot of these benefits, but also helping us protect the “shores” from excessive flooding.

*Do you think any of the existing data protection regulation could limit, in some way, the use of AI given the broad definition of personal data and how that can be or can't be used? Do you think that presents challenges for companies where they're trying to implement and deploy AI more and more?*

Data protection legislation does put pressure on artificial intelligence applications that reflect pure research, where you don't necessarily always know at the beginning the insight that you're going to discover that may be useful. I think data protection puts a lot of pressure on really knowing the purpose for our research project and presuming that's going to stay static, as opposed to research by its nature being really dynamic. What I've seen certainly on some of the projects is that you have it all set up, you do an appropriate disclosure to the public, make sure that people know how that data is going to be used. You describe your purpose and then you discover something really interesting and you want to either share it, or you want to work more in that area and start to leave the initial framing of what you described to people. Data protection usually wants you to go back and ask again. I think that is a tension that comes up whenever you're trying to apply a notice and choice regime to something that's very organic like research. So, we definitely struggle with that. We think about it a lot. There are techniques for trying to improve the way research works in a way that's consistent with public expectation, but still allows researchers to generate a lot of value.
This continues to be a work in progress, but that very topic is at the heart of a lot of the discussions happening through think tanks and public policy within organisations that are advocating, in dialogue with government, for flexibility that is often required in order to extract the maximum benefit out of research initiatives. I think that's going to continue to need to be an area of focus and scrutiny. Provided that appropriate safeguards and risk mitigations for people are put in place, overly prescriptive notice and choice regimes may have to give way to broader opportunities to use and optimize data, to use data for short periods of time. Other techniques that are privacy techniques, super robust security as a way of ensuring that research initiatives are done in a way that is really protecting people's interests as individuals, but at the same time collecting some vital data out of data sets that are often very large and involve large numbers of people. Frankly, if you don't have contact information how do you ask for permission? Our communities are still working through some of those problems and cases. But they're not insurmountable as it is?

They can't be, right? You have to forge a social consensus, both with customers and members of the public, but also with regulators and with government officials. There's probably some middle path that emerges over time, but we have to have enough good experiences and few enough bad experiences so we can get socially comfortable with this methodology. Otherwise, if there's some really, really bad misstep, then I think reactionary regulation could indeed emerge. In my view that would be unfavourable for everybody. So the more people that can be working in this balanced mind-set and really transparent about what techniques they're trying to bring to make the research ethical, I think the better.

Are there any other topics of conversation that the ePrivacy industry is particularly focused on at the moment that we haven't touched on, or is perhaps not apparent from an outside perspective?

Usually when I hear the phrase ePrivacy, I think of the European E-privacy directive or regulation, as opposed to just referring more broadly to the tech sector, or the Internet sector. The famous cookie directive that imposes those pop up banners on websites comes to mind. Technology that is used online to enable personalization for websites to remember people's choices over time, whether they're cookies or browser based techniques, whether they're mobile based, all of those technologies get labelled as trackers or tracking tools. They've been very challenging for the regulatory community for several years. I think there's been a recognition that they're about economic business models, as well as personalization benefits in the use of these technologies. They are relied upon by the largest Internet companies and success stories in the world. While at the same time, their inherent lack of transparency, the fact that they happen in the background, and we don't usually look at the clockwork in the back of the clock when we're reading the time, is consistently problematic for that sector.

There's been a long-standing attempt to address some of those concerns through self-regulation, through forms of co-regulation, through targeted legislation, such as in the case of the E privacy directive in Europe. Yet I think the broader community doesn't seem to have cracked the nut. So what we're seeing proposed now are potentially business models impacting changes to some of the architecture of the way websites are working. The dependence on browsers and the predominance of a number of browsers, from Chrome, Firefox, Safari, and Microsoft etcetera. Anything that those browsers do to modify the handling of technologies that flow through them have a huge impact.
When one of the browsers announced that they’re going to make a change to the way third-party cookies are handled, or the do not track discussion, which has been going on for several years, every tiny little change impacts legions of other small, medium sized and large companies.

I think that the relationship between what historically we’ve talked about as first party and third party has been fraught and heavily technology dependent. With the turn of the screws, and the pressure on those technologies, and the pressure to change the approach of those technologies, you feel the collateral impact in the case of the marketing and advertising sector and its relationship to the publishing sector, which has struggled to monetize when you look at big newspapers and big platforms, who have to choose between a subscription model versus an ad supported model, free services versus pay services. All of those big industrial factors are tied into these issues of how the technologies work, how the browsers relate to websites, how tracking and personalization works, the ability to micro target versus target more demographically, or in the aggregate. That whole range of topics is really fascinating. These issues have been present for many years. It’s not a new topic, but it’s one that seems to continue to be challenging for both industry and for regulators.

I do think that recent announcement about browsers deciding how they will modify the mechanisms that underpin some of those business models will be very impactful. And whether that leads to changes in the business models, or just changes in the market, with respect to where power centres and advertising dollars go, consolidation around various parties, remains to be seen. That's something I think everybody will continue to watch very closely.

It’s also very interesting because I’ve certainly read and seen discussions about the intersection of competition and data protection. That's one of those areas where there are questions about overlap and the intersectionality. Also what it means to be a repository of big centralised data is something that is very front of mind. For all sectors, it ties to our ability to engage in really robust artificial intelligence, it ties into the ability to personalise services, it ties in the ability to scale and get network effects and there’s so many different factors that flow from the ability to aggregate data when there are few players that have that ability. Of course, markets are impacted and questions about competition and data protection and the role of data and competition won’t go away. There have been academic discussions about the concept of considering data in the context of competition inquiries, but they’ve not been particularly mature. I think that both the US Federal Trade Commission and the European Commission have signalled an interest in the topic and so I would predict that intersectionality is going to be explored further.

There's perhaps tension between data protection, or the idea of data minimization, versus maybe the competition angle. In the EU there is talk about a single data lake, trying to break down the walls of the walled gardens. How do you think the two can ever be married? How do governments and regulators try to level the playing field when you’ve got a huge data sets already sat with a handful of operators?

I think if I knew the perfect answer to this question, I would be the king of everything. It's a very, very sophisticated challenging question. I'm not an antitrust or competition expert, so this is very armchair commentary from my perspective. Putting data aside, through history there have been really, really successful companies because they’ve had a great product, meeting users’ expectations and users’ demands, they have the collateral benefits that come with that.
One of the prices of success, in general, and this is true of the major Internet success stories that we’ve got on the planet, is that all that success is not always predictable from the outset. Ten years ago, or even five years ago, the policy and business model and technical infrastructure that was put in place at the beginning, in some cases the very infrastructure that enabled them to achieve a level of success, may not be the infrastructure people are comfortable with once the business has matured and reached a very large scale.

As I said at the beginning of our conversation, going back and changing things after you’re big is really hard. It’s really, really hard. I think that's the experience of a lot of Big Tech players right now, where the scale and the data has been achieved because of the structures being the way they are. If the privacy or security concerns are tied explicitly to the very structure that made you successful, then you’ve got a real problem. It is scary messaging to hear. Unless, and until, really successful players can show that they can operate the models in an environment of trust, and transparency and accountability, those pressures will not go away. It's one thing to state publicly our commitment to trust and accountability and getting it right and working with people, but at the end of the day, the customers and the citizens and the government representatives are the judge of whether the businesses are functioning in a sustainable, acceptable way. That’s true of all industries and all sectors.

If you were to contrast with heavily regulated sectors in the energy space, or in the financial services space, regulation emerges when there are harms to people or fears of the harms to people. You have to address those things proactively in order to end up being a so-called heavily regulated industry. I think the tech sector is of course young by sectoral standards; it may not quite be ready for heavy industrial regulation, but certainly the conversations about what is the right kind of regulation for the sector are something that always would have to come inevitably. I think that those discussions are happening real time.

*Do you think the tech industry as a whole is doing enough to address the concerns that the government might have? Or do you think it’s all a bit surface level and not much really changes underneath? It’s hard to know when you’re sat outside of it, and reading articles about it.*

I think that, as in all things, there are leaders of the pack, middle of the pack and back of the pack. There are organisations that are extremely serious, respectable, responsible and accountable in data governance, and they recognise that it is a strategic imperative for the success of their business going forward. There are very brilliant, mature, committed executive teams that talk about these issues in a very serious way, and expose the boards of directors to them.

There are, of course, other organisations that back away from this kind of approach because there’s not really clear guidance. This sector is operating in a lot of grey areas from a regulatory guidance perspective, and different organisations, and cultures of different organisations, respond differently to those grey areas. I could say that there’s quite a few very serious, engaged multinationals that are contributing significantly to the thought leadership that’s needed in order to mentor the sector.

There are also organisations that are not participating who should be. Sometimes organisations respond to sticks, sometimes they respond to carrots. We’ve seen increasingly that there are more sticks of late, and that's brought more people to the table, which at the end of the day is what we need.
Do you think there are any sectors that could be seen as trusted custodians of database from a consumer/citizen perspective, as well as from a regulatory or government level?

I think data by its nature is risky for people. No matter who holds it, wherever they may be, some of the common risk like cyber exist, whether it is the government, or whether it's a private sector. Cyber security is a global issue that all sectors face equally. Anyone who tells you that they are fool proof and completely secure, and zero risk, is not telling the truth. Those types of risks are horizontal risks that everybody needs to be thinking about in a flexible way, in a proportionate way. I think that the cybersecurity industry is engaging increasingly, both through the proliferation of technology and the offering of insurance, as well as becoming a board level agenda item.

Overall, my answer to your question is no, I don't see any sectoral safe haven for people's data. I think there are risks, no matter where it may be, that comes with the business and are necessary to support many of the benefits we enjoy as a society. I do think that certain sectors traffic different categories of data that different populations may consider to be sensitive. That's not to say they shouldn't be using this type of information. In the healthcare sector it is incredibly important to use sensitive health information, and to be very data driven. Lots of research and discoveries in the medical care sector are predicated on people's information and often very, very sensitive information about diagnoses, risk factors and so forth. In the financial services sector including Fintech players, really sensitive information about people's financial accounts, the ability to withdraw and transfer monies and remit it around the world absolutely brings risk, but that's inherent to the nature of that line of business, and is necessary to bring banking to the unbanked, and to offer people credit.

The broader tech sector, including the more Internet-oriented search engines and communications, ecommerce sites and so forth, are clearly providing social value to people. They're all trying to be very responsive to consumer demands and expectations. I don't think that any one of them is inherently, because of their business model, necessarily more or less secure, but there are certain Internet businesses that are less reliant on online tracking technology versus more authenticated user experience. This known identity user experience versus the anonymous or anonymized user experience has different impacts for people.

Some might say that where you don't have to provide identifiable information it's safer. Others might say, ‘oh, all of these cookies and personalization things, even though they don't know who I am, they figured out who I am. That's creepier.’ There are diverse views about the privacy intrusiveness of different kinds of business models and for that reason, I don't attribute certain degrees of security or safety, from a privacy or data perspective, to any one sector or business model. It remains highly contextual. We have to account for differences in culture. We should look at population, the user base, and the experiences that people have that are presented by these different companies, and do our best to meet everyone’s expectations, while offering a product that is fundamentally useful to people and meets societal needs constructively.
A Conversation with Danny Weitzner

Founding Director of MIT Internet Policy Research Initiative and Principal Research Scientist at the MIT Computer Science and Artificial Intelligence Lab

Do you think the current data privacy policies, such as GDPR and CCPA, are fit for purpose? What do you think the regulators have got right and what do you think they have got wrong? It’d be interesting to hear your thoughts.

I think the GDPR does represent the global state of the articulation of individual privacy rights. It’s thoughtful, comprehensive framework, but it really doesn’t provide the specific guidance that more data intensive businesses need. In particular, I think there’s going to be some tension between, just in the European context, the approach that has been taken with the GDPR, which does tend to be a maximalist approach to privacy, on the one hand, but on the other hand, what gets called AI policy. To me, that's happening in real time; the new Commission came and identified this as a high priority during a 100-day process right now to get a framework for AI policy.

When you look at really what is meant by AI policy, it’s nothing more than data policy. It’s policy about what data is going to be available, how it can be used, what are the regulatory requirements and, of course, a huge amount of that data is going to be personal data. I think that the approach to those questions, particularly in the European context, is going to be interestingly different from the GDPR in that a lot of the high visibility efforts in framing the GDPR, were frankly dictated by the trajectory of the U.S.-dominated Internet industry. The big Internet platforms consumed a huge amount of attention in privacy thinking in the period that lead up to the GDPR. There are people who have these views about nefarious behavior by Europe to try to cut down the U.S. Internet companies. I don't tend to give those views a whole lot of weight; I think the Internet platforms are a pretty legitimate and, in many ways, urgent targets of privacy regulation. They also represent a particular framework and they have a particular political context about them in that they're not European.

I think it’s going to be really interesting about the AI policy debate that is going to merge with the GDPR, inevitably, is that a lot of the key industries that will be affected by European AI policy are really European champions, whether it's in financial services or insurance or automotive or pharmaceuticals and Life Sciences. These are all markets where you have very substantial European competitive positions, unlike the Internet platforms, and where there’s obviously a real stated interest from European policymakers to maintain and extend that economic frame. In simple terms, if you're a European policy maker, making policy about BMW, Munich Re., any of the European based pharma companies, all of whom are making major steps in in the AI arena, could be a very different dynamic than the GDPR dynamic.

Here's where I think the rubber hits the road on questions about GDPR interpretation. One of the things that I think is actually quite smart, and, and agile about the GDPR framework, unlike the CCPA, is that GDPR, while some people caricature it as being all about consent and explicit consent, it also has a lot of flexibility in providing other, what are called lawful bases for processing data. The revolution, if you will, of the GDPR was to say that you can’t process data without some legal basis. That’s almost the polar opposite of the U.S. framework, which has a lot of rules about specific kinds of data processing, but is essentially a set of negative regulations about specific things that one cannot do with data.
Whereas the GDPR takes really the opposite approach and says, 'all data has to be processed, under control', number one. Processing data under control, means that you have to have a legal basis for any processing activity. Consent is the one that gets a lot of attention because it's in a certain way the clearest. I think from a business perspective, and actually from a consumer perspective, it's often the least satisfying one, in that we know that businesses are having to spend a huge amount of effort on rethinking their concept frameworks, number one. At the same time, some studies that we've done in my lab, that people have done all over, consent doesn't work very well as a privacy mechanism. People don't understand what they're consenting to; they make decisions that are, arguably, not really, in their long run interest in order to get a short term benefit, while at the same time being very frustrated about what's going on and confused. It's not really, I think, a privacy approach that is suited to a high intensity data analytic environment that we are already and that we will be in even more when you have machine learning driven applications and services.

What's interesting to me about the GDPR is you have these other legal bases for processing. The one that I think is really important, but underexplored, is this idea of legitimate interest. I think it's actually a very powerful and potentially agile tool for companies that have a lot of personal data and want to make use of it for advanced analytic purposes in machine learning context. What's interesting is the DPAs have been very cautious about what they're saying in application of legitimate interest beyond just the advertising context. You look at just the text of the regulation and the recitals and the commentary, it is clear that legitimate interest is meant as a particular burden-shifting mechanism. It's an opportunity for a data controller to say, 'here's what we're doing with personal data, and here's why we believe we have a legitimate interest in doing it, and here's why it doesn't impact fundamental rights of the data subject.' The burden is on the data controller to get that right.

I think the DPAs do not seem eager to lay out a big framework for analyzing legitimate interest, because I think they don't want to necessarily take the burden of the controllers, just yet. I think that it does create a real opportunity for, again, the more sophisticated analytics-driven services that are going to obviously develop in lots of different sectors to find a different path for the legal basis for processing data. By the way one of those paths may well the codes of conduct that are contemplated in the GDPR. This idea that industries and sectors can get together and propose detailed codes of conduct for certification by the DPAs, which in some sense allows a sector to explore the question of what is legitimate interest in a particular context with regulators and get some assurance.

That would have to be part and parcel of the AI policy?

My view is that a lot of the progress that we've made on privacy over the last couple decades has come through that particular code of conduct mechanism. I worked quite a bit when I was in the Obama administration on the Consumer Privacy Bill of Rights proposal that President Obama made, not accepted by Congress, but we took as a centerpiece, the idea of enforceable codes of conduct that industries could get together and say to the regulators 'here's how we propose to comply with a given set of privacy rules.' We thought at the time that that was a more agile approach to making sector specific privacy rules more rapidly than you could do with a traditional regulatory rulemaking process.

That idea was picked up in the GDPR, again as a way to have a more agile response to privacy needs, in new business models basically. If you look at the evolution of privacy rules in online advertising in the U.S., it's actually really been through that kind of Code of Conduct process.
Those haven't been perfect, I think they've actually allowed the online advertising marketplace to evolve under some oversight from the Federal Trade Commission. This created a structure where a lot of business can develop in a stable, accountable way from a consumer perspective.

**Do you think public policy should evolve in the area of data protection the US from where it is now? I know we've got CCPA, but a federal level?**

I think it has to. First of all, there is going to be a growing number of state level privacy statutes. The CCPA is just the first one, there's one that's very active in the state of Washington, at this point, New York, Illinois, and other states are considering the same, so there'll be a need for a consistent set of rules nationally in the U.S. I also think it's awkward for the entire U.S. information marketplace to have its principal regulator be sitting in Brussels. It's no criticism of the Europeans. I saw Boris Johnson had on the list, liberation from the GDPR. That's pretty interesting. I think, the U.S. will need we will need to have our own privacy framework.

**Will that be the Federal Trade Commission (FTC) here?**

Right now the Commission is significantly underpowered both from a resource perspective and an authority perspective; that'll have to change. There is talk about, some people are calling for an independent privacy commission of some sort or another. I think the experience of creating the Consumer Financial Protection Bureau, which I think has had considerable ups and considerable downs, I think may suggest that the best thing to do is to add to an existing agency rather than try to create a new one. The only reason you'd really create a new one, I think, is if you needed a different governance structure.

I think first of all, privacy questions are hard to separate cleanly from competition issues, from the AI policy issues, security issues, etc. I also think that, in the European context where there had always been an omnibus approach to privacy, it makes some sense to have independent regulators with that same scope. In the U.S. we are always going to have a highly sector-based set of privacy rules. We're going to still have health privacy rules and financial privacy rules and, if you keep track of the lobbying activity of your colleagues’ guy in Washington, they are holding on fiercely to GLB. They don't want to be swept into some other regulatory environment. I don't think we'll have a new independent commission; I think we'll add to what the FTC does, is my guess.

If we were to get in a democratic president and the President was on the left hand side of the left, then maybe we would go with something more.....

I think the Consumer Financial Protection Bureau (CFPB) was great from the founders perspective when it was controlled by their appointee and then it went pretty much 180 degrees to the more conservative side with the partisan shift. Unless we have a very dramatic partisan shift, both in the White House and in both houses of Congress, it's hard to imagine anything other than a bipartisan piece of privacy legislation passing, which means it's going to have to take into account these different regulatory perspectives. I think that enhancing the status quo, as opposed to creating something totally different, is just much more likely.

**On AI policy, who picks up the U.S. version of what the European Commission is doing with AI policy? Would that be more of an industry led thing through some of the work being done on the West Coast or is it more the FTC?**
One of the interesting political dynamics here is that one of the major multilateral initiatives that the Trump administration has supported is the OECD’s AI policy recommendation. Key to those recommendations is the idea that, ultimately, you have to look sector by sector to figure out what the right approach to AI regulation is. There are some common themes about fairness and discrimination and robustness. Ultimately, what you do about AI policy in transportation is pretty different than what you do about AI policy in finance or in criminal justice or in the other sectors. I actually think what they’ve done on AI policy is pretty is pretty thoughtful. There was a recent Executive Order that says to all of the regulatory agencies, ‘go figure out what makes sense to do in your sector.’ With a different administration, there’ll be different body language about that, but it's ultimately still what makes sense. When you look at what to do about, for example, machine learning applications when you’re in consumer finance, you come back to basic questions about fairness and discrimination and market efficiency and access to credit and access to capital and all that. The agencies that know how to address those issues are the ones that we have. They have to add technical expertise, certainly, but there are some governments around the world that have gone and created a ministers or ministries for AI. I just don't think that's going to make sense in the long run, really anywhere, and certainly not what we're going to do in the U.S.

*When you refer to it as AI policy, when I am a layman when I think of the AI, it's a very different part of what I think AI is versus what we're talking about here. Can you just unpack that a bit and explain why everyone uses that shorthand to explain what we're talking about?*

It's a really broad set of questions. It's everything from workforce questions, to military strategy questions, to R&D policy, to the more core consumer regulatory issues.

I think it's actually a muddle, but it is in a certain sense, shorthand for really advanced data analytics. How do governments take advantage of that and leverage that as much as possible, develop strategic advantages with respect to that technology, and how do governments deal with the regulatory challenges that come up? It is a policy theme, not a policy sector from my perspective. It's really about data; it's nothing more than that.

*If you're right that the U.S. goes down the path of a sector by sector approach, but, Europe and, maybe others, has this monolithic approach, do you think we run into issues? As soon as you said that I immediately thought well, Amazon as a retailer but now they sell advertising, and Google is now doing autonomous cars, and you go down the list and there’s almost been bizarre bleeding of a very discreet industries that the Internet has indeed spawned, which makes these historical white lines far less clear?*

I think it is an issue but I think the reality is we know how to deal with it from the last decade or so of looking at the Internet marketplace. You had these, seemingly sui generis services. Take Uber or Airbnb; for a while we thought they were a separate category, then in the end the traditional regulatory frameworks managed to reassert themselves. Hopefully they don’t do it in a mindless way, but if you’re driving cars on the street, you got to conform to the regulatory interest that has been traditionally established there.

I think that there’s no question that there’s a set of unique policy issues associated with the behavior of the other Internet platforms, but, as they start to become more and more present in the real world, if you will.
When it was only electronic commerce, it was kind of easier and simpler, but it just isn't now and everyone understands that. I look at the difference between the U.S. approach, and the European approach in some ways as just a question of where you start. Europe likes to start with a set of high level general principles that probably just reflect the European legal structure. In the U.S., we like to start more kind of case by case and sector by sector. In some sense, in the end, they converge on questions about what do you do about a particular kind of service or a particular kind of application. It doesn't matter so much. For a while I think the platforms tried to say 'we're only platforms, don't apply any rules to us.' I think no one is offering online banking service, and saying, 'it's online so we don't have to implement any of rules; around the edges you have the cryptocurrencies, but even they realized that when they touch the traditional world they have to acknowledge it.

I think constitutes a meeting in the middle, in a way, between a sector by sector approach and in this omnibus approach. In the end you have to deal with sector specific behavior, regardless of where you started. That's why I think this intersection between the codes of conduct, and the legitimate interest ideas is where the rubber will meet the road.

Is that shorthand for CCPA world?

No, no, not at all. I think the CCPA is a very, cautious piece of privacy law. In particular because it doesn't cover quite a lot of the behavior that the GDPR covers. It's limited to commercial sale of personal data and particular kinds of uses of consumer data in a personal context. It has a bunch of broad notice and access policies, but it doesn't reach anywhere nearly as broadly as the GDPR reaches. As with the GDPR, there's a lot we don't know about it. The Attorney General still has to finalize regulations, interpreting a lot of that statute. I don't really see it as a business model changing law in many ways, unless the State Attorney General is going to get very aggressive with enforcement and, I think that would probably face some pushback.

I'd like to get your thoughts on how you think privacy regulation as it is, or as it may evolve, could impact these innovations and the ability for them to be deployed? Whether the various approaches across the globe, whether that creates any advantage or disadvantage from a technological innovation perspective?

Let me actually talk about the technical side of that question, which I think is in the long run the most important. Whichever legal framework you're talking about is, you have this trajectory where the regulators are expecting that companies that use personal data be able to exercise more control, have more predictability, give users more control, over their personal data than has been the case in the past. You can characterize a lot of these, whether it's the GDPR or others, as saying to data users, to data controllers, 'you can do a fair amount but you have to do it under control; that is you have to be able to guarantee that the data won't be misused, you have to be able to guarantee that people can correct or remove their data, if they want.'

One of the things that we're learning about this is that, in most cases, companies that are subject to these rules, whether it's right to be forgotten rules, or rules that say you have to be able to account for all the different uses of personal data. In fact, we do not have the technical infrastructure today, in any organization to actually comply with those rules in a serious way. I haven't studied Citibank's systems but my guess is that if someone said, 'okay, find me all of Danny's data, and all the things that Citibank has done with Danny', my guess is that it would be hard.
I think that's the fault of computer science, to be really honest. We have not built systems that allow the kind of accountability and control over personal data that privacy laws are actually now starting to demand.

What we've seen is that, whether it's areas such as database architecture or particular kinds of machine learning applications, machine learning algorithms, applied cryptographic techniques; we have a lot of the pieces for enabling better control over personal data. There's still some pretty substantial research that has to be done to actually put them together into an environment where an enterprise can really say, 'we now have control over personal data.' I would say that, in answer to your question, in the long run, the limiting factor on what can be done with personal data is partly going to be what will society accept and what's responsible, but also the ability to guarantee that all these different more innovative aggressive uses of personal data, can be done in a way that is under control. That is fundamentally an organizational requirement, but it's got to be supported by the right technology.

The database administrators, the people who are actually running the information systems and enterprises, were saying, 'look, we actually don't know how to, for example track purposes to a database.' So GDPR says you can use data as long as it's used consistent with the purpose for which it was collected; that's key to the legitimate interest idea. No one know how to that, technically. GDPR requires that the company's respect right to be forgotten request. No one knows how to find me. Poor Facebook, they've been trying for two years to create a delete history feature so users should be able to delete evidence of their activity on Facebook. They can't find all the history because of the way their systems are architected.

We're getting together a group of people in computer science at MIT and companies around the world that want to look at these problems. I think there's actually a lot of really interesting innovation opportunities here, because I think there are new capabilities that enterprises all around the world, and governments all around the world, are really going to need and we just don't have today. Number two, I think they're critical to having a curious, but agile privacy regulatory environment. People don't want to completely shut down the use of personal data, but they want to know that surprises won't happen. That's what I pick to be a really core privacy principle; just don't surprise people, but right now we don't have systems that can handle that.

When we're talking about the data landscape, I describe the U.S. as the Wild West. We have very little constraint over cell phone tracking, geo location, that kind of stuff. When we look at the global map, the US stands out as being the Wild West until CCPA came along. Was your statement about the sectors and the regulators of each sector, your intuition tells you that as we go through the coming years, whatever happens with politics, the regulators of those industries will have to rein in those activities? So, the telcos will be regulated, therefore our geolocation data will be better controlled, the Internet companies will get anti-trust regulation? Am I capturing that right?

I think that's the trend. I don't think it's going to happen immediately or all at the same speed, but that's what's going to happen because we're developing cultural sensitivities about a growing set of data uses and misuses.

The public understanding of the uses and misuses will slowly improve. The industries, which they interact with, will slowly have to up the game, and at some point, we'll get to a state of normalcy where users can control the data. My question about the future was if you flip it on its head, would that give us an advantage in that we can then sell our data into these exchanges, and I can get compensated for the use of my data?
I think that those who can convince their regulators that they have control over the personal data they handle, are going to have an advantage over those who just say, ‘give it to us, don't worry, nothing bad will happen.’ I think we're past that nothing bad will happen.

The choice is either, at the extremes, an onerous set of rules that say, you can't do this, you can't do that, you can't do the other thing; just to be protective. Or, something that's more nuanced that says, 'here are the harms we are concerned about as a society' and, to certain extent, this is where I think the codes of conduct become really important, 'you tell us as an industry, how you're going to protect us against those harms.' That's basically the model that you guys in the financial services industry live in with respect to financial regulation, more or less. There's a whole lot of difference to the professional expertise of the banking sector, obviously not without oversight. I would say that, in some insignificant part, because we have these things like, accounting rules and public market disclosures and all that kind of stuff that allow a subtle assessment of whether things are in balance. We don't have that as personal data yet, but I think that's where we have to get to.
A Conversation with Andres Wolberg-Stok
Head of Strategy, Office of the CTO, Citi

What is your role at Citi?

I worked in digital product management or digital strategy on the consumer banking side of Citi until very recently — always on the business side, first for Citi Latin America, then for Citi’s U.S. businesses, and since 2012 globally. A little over three years ago, we realized that regulators and policymakers in the U.S. and elsewhere were really looking to modernize the environment. They needed people they could work with on the bank side, people who would have that hands-on experience of how things actually work, how things actually get done and where the friction and the hidden challenges might be. I started to argue that we should set up a function like that, and of course, eventually I was told ‘OK, you do it’. So I shifted from digital product management and strategy into what we ended up calling Digital Policy for Citi Global Consumer Banking.

That work got started in an innovation unit tasked with driving a mobile-first digital reinvention for Citi’s consumer bank. When my work outgrew that unit, I moved under the Head of Technology for Consumer Banking. But it started to connect also to our Institutional Clients business and to take on a broader dimension, so now I am moving into the Office of the CTO as Head of Strategy, which includes expanding our the digital policy work to serve all of Citi.

So maybe we’ll just start with like the current state of affairs. Obviously, GDPR went live in 2018 and then CCPA went into effect earlier this year. When you reflect on where we’ve come from and where we’ve got to go, where do you think we are in this regulatory story of tightening up data privacy regulations around the world?

We are at a very early stage. If you zoom out, the fact that so much of the focus has been on GDPR and CCPA masks the breadth of the spectrum of cultural environments in which very different approaches to consumer data control and privacy are going to end up emerging — or not emerging. I would put us on guard against looking for convergence and coincidences between GDPR and CCPA, and thinking this is where the world is going. There’s going to be a lot more, and it might be very different.

Just drill down there a bit. You say cultural differences; obviously we’ve an authoritarian state of the world and then there’s a more democratic state of the world, is that what you mean by cultural differences or is it more about the nuances of like Gen Z and millennials interacting versus baby boomers?

There are probably several dimensions of difference and you just named two of those: One is the nation-state approach (or supra national, as in the case of the EU), the other one would be generational. Probably closer to the nation state one, but still distinct, is the angle of just general expectations in society which also vary. They may be closely related to the nation-state approach, but not necessarily. As an example of a disconnect, there’s the societal expectation, which is going to be very different if you look at consumers in China and consumers in the U.S. Then there’s the reality of what the nation-state does. With all respect for some fundamental philosophical differences, the practical differences may be fewer or less marked than members of society would expect. For instance, facial recognition is in use, broadly, at US borders now. If you are enrolled in the Global Entry program that
makes it faster for pre-vetted travellers to enter the United States, you may have noticed something in some U.S. airports in the last year: You will approach the machine but you do not even get to scan your passport, because it has already recognized your face and it just spits out your entry ticket. It’s incredibly convenient, magical even, but I do not recall being asked or having had a choice. That is just one example of how, if you talk to people in a society, their engrained attitudes towards privacy might be very different but the practical side may not be that different. When you look at what the nation-state does in its governance function, some of those things may be converging more than we know. There was a lot of debate earlier this year around a company which has been amassing pictures of people in the U.S., and it sells an AI-based facial recognition mapping identification service to government agencies. On one hand, members of the U.S. House Financial Services Committee are up in arms and are demanding an investigation because people have not signed up to have their facial features used in this way (and because it’s also being used for financial services). At the other extreme from this congressional committee, you have law enforcement agencies, essentially saying “can you please not talk about this because we are using the service to find minors who are being trafficked into the sex trade or abused, and this is saving lives.” Both of these extremes are debates that are happening only at a very specialist esoteric level, and they’re not part of a broad social conversation yet. Probably at least those three dimensions are worth noting, that we’re talking about. The generational one, I’m not even clear that that is a unidirectional axis, and that there isn’t a lot of contradictory data. Secondly, the dimension around expectations and tolerances in each society; and thirdly, what nation states and governments are actually doing.

So how do you think, how to regulate some governments go about managing data being used for good data, not being used for goods if you have one all-encompassing approach?

I would begin from the perspective that there are going to be some legitimate differences between what you expect of government and what you expect of private business, in terms of privacy and rights to choose and control your data and your information.

You are going to have to factor into that different perspectives on what is legitimate and what, and varying political connotations. If it is happening in a country of which you are not a fan, it doesn’t become any less tolerable than if it’s happening in a country with values which you espouse without reservations.

We have to operate in this patchwork of different government attempts to tighten up control over personal data. I think we still see the U.S. as being somewhat the Wild West, outside of CCPA. Do we assume as a bank that we have a gold standard and that Citigroup will define our own gold standard that will appease and address all of the global regulations? Is that how we have to go as a global multinational?

That’s a very important question. To your point about the U.S. being the Wild West, one of the concerns for many companies, not just in financial services, but large brands that operate nationwide, is about the fragmentation. It is about replicas of CCPA popping up in dozens of different states, and it is about the horribly costly complexity of dealing with a profusion of state-based requirements. This is the main reason why many large companies have been advocating for a federal approach to this that sets the bar wherever it needs to be, but at least it sets it once and there is one bar, rather than having to potentially worry 50 times over about meeting specific requirements in specific ways for state based-legislation. So that’s one angle.
Then, if you zoom out and think, not just U.S. nationwide but globally, there’s a question facing companies that operate around the world, which is: are they better off complying in each jurisdiction with the minimum that they need to do to get there, locally? Or are they better off just embracing globally some of the more stringent requirements and maybe gaining some efficiencies from that? It’s a complicated equation as to whether what they would be leaving on the table from the delta between the higher global standard they’ve adopted and the less stringent local standard in some jurisdictions. Is the failure to capitalize on that delta going to outweigh the global efficiencies of just taking care of it once, and having one global standard? One of the variables in that equation is going to be that they are not just competing with each other as multinational companies. They are always going to be competing on the ground with local rivals, who will not be embracing those potentially more stringent global varieties of data-privacy frameworks. Then does that potentially, over time, affect the multinationals’ ability to compete effectively in local markets with less stringent requirements?

On a related point, I just want to get on my soapbox for a moment about the general notion of deleting data, or people’s so-called “right to be forgotten.” Personally, I find that very Orwellian, like rewriting history. It just seems outright sinister. I know that’s a very big word, but it just seems very wrong for someone to be able to bleach history away, delete records of things that have actually happened in the past and are part of the factual record, just because they don't like the reputational implications, or what it does to them. That kind of approach, I think, is just philosophically wrong.

In addition, in the 21st century, you absolutely cannot be advocating seriously at the same time for immutable blockchain records, and for people’s rights to delete past events that they dislike. Those two things are completely incompatible.

**How do corporates deal with the idea of data minimization from some of the pieces of regulation we have seen, with the idea of AI and deploying blockchain? How do corporates manage to deal with the tension between both of those?**

There is still a big gap between promise and fulfilment or fulfilled reality. Therefore, it is probably too early to draw any long-term conclusions. At this stage, we are at the decade mark of the general concept of blockchain, Bitcoin and DLTs — and we are not conflating Bitcoin and blockchain — but overall it's very difficult to point to really large-scale, live implementations that have changed the world. The theoretical benefits continue to outpace those in the real world. This could be mistaken for deep scepticism. I just mean that realizing the potential has not kept up with the hype and practical challenges at the “zero trust” end of the spectrum. The upside seems a lot more abstract and hypothetical than it may be in some much more limited, down-to-earth permissioned environments in which participants can somewhat trust each other. If they completely distrust each other and expect that they would have no recourse whatsoever, then I would say, you still have a bit of a fundamental contradiction in terms. Because in order to trust that environment of complete lack of trust, you will still need to trust the infrastructure — the DLT. Someone, or a collection of unknown actors, is always going to own and operate that infrastructure. I do not think we have even begun to scratch the surface of that question so far. By the way, I am also very curious about how you actually use a Blockchain to certify the origin of precious stones, or produce, or anything. That leap from the physical world to the DLT world is always going to involve human action…
Do you think the infrastructure is in a position to handle what’s required from the regulation, or do you think there’s a big gap there as well?

You have to sympathize with how regulators are struggling with these questions. When we talk about financial services in particular, one of the reasons they are challenged is that on one hand, they are hard-pressed to name any successful real-world, live-in-production DLT-based initiatives that matter. Yet on the other hand, they are under huge pressure to look into issuing digital/crypto central bank digital currencies, because of initiatives like Facebook’s Libra that have lit a fire under everyone.

In a survey published in January this year, the Bank for International Settlements reported that of 66 central banks that responded, more than 80% were engaged in “some sort of work” on central bank digital currency (CBDC). It is a challenge for regulators and central banks to have to think about how they would get in front of these things without first having had a chance to see the technology operate at scale in a private environment.

So do you think it is potentially really slowing down the pace of these sorts of initiatives and technology development?

There are two layers to that. One is that Libra sparked this chain reaction; China decided to say that it would accelerate its work on a digital Yuan. The Europeans, looking to the east, think to themselves “Digital Yuan, plus Belt and Road, equals not good — so let us get the ECB to get cracking on a digital crypto Euro of some sort.” Next thing you know, you have U.S. lawmakers pushing the Federal Reserve to explore a digital dollar. When you ask why, they say “because the Europeans are experimenting with it.” So you get this kind of three-way arms race, and it’s not clear people focus enough on big, big questions such as, “are you thinking about a wholesale or retail central bank digital currency, and in either case, what are the implications for funding of commercial banks.” Obviously, as a result, what are the implications for credit and lending overall in society? These enormous questions are not being debated a lot in public. There has been some interesting academic work on this recently, and it reveals that there are even bigger questions behind that, and no real obvious answers at all. It might be a good moment to sit back, cool down, and think things through. Even with the publicity over the competition between central banks to out-announce one another on how they are looking at a digital cryptocurrency, that is probably not a great move in terms of setting expectations. Aside from the fact that we have a lot of fiat money in use all over the world today, and the distinction between a digital currency and cryptocurrency is not standardized, and it’s not always clear what central banks are talking about when they talk about CBDCs.

Do you think the FTC or maybe the CMA end up in an antitrust situation where the regulators look at a couple of companies and say you have so much personal data that you’re using to generate ad revenue, we’re now going to somehow curtail that through a different way of approaching the puzzle?

I don’t know whether that stage is approaching. In some cases, a lot of the data is plausibly argued to have led to political outcomes, on the ground. Any decisions on that front are going to carry huge political connotations and are not going to be just about technical competition. Regardless of the outcome, there is going to be a political reading, and that might create a pendulum, which swings every four years in one direction then the other, depending on who feels most aggrieved or benefited by the way the data ended up being used.
It becomes a political football; it already is. You have the left-right political football. You have the “European distrust of American Big Tech” political football.

Another consideration, squarely back in the domain of data: In the European Union with PSD2 the Second Payment Services Directive; and in many places in Asia, Australia, Hong Kong, and so on, there are a growing number of mandatory open-banking regimes that will force banks to allow consumers to enable third parties of their choice to come and get their account data from their banks. With the honourable exception of Australia, which has decreed a broad “consumer data right,” it's striking how elsewhere not much is said, on behalf of consumers, along the lines of, “well, it's great that you're forcing my bank to allow them to give someone else access to my bank data. But I would also love access to my health data for someone else to crunch. I would love to be able to give someone access to my telecoms data and my energy consumption data and my location data and my social media data, and my online shopping browsing data,” and all of these things where you can very easily imagine how it would snowball exponentially when you can begin to combine all those data sets. Innovators could get really creative, and consumers could really benefit. The focus is inordinately on banks and bank data, which is very important and makes a lot of sense. Why aren't we focusing on all these other domains, which are arguably equally important or have comparable potential to improve people's lives if we truly want consumers to be able to leverage their own data?

One concept that may come in handy is the major difference between portability and real-time access. If we think of portability as the switching mechanism that you need in order to be able to move your financial life from Bank A to bank B or to FinTech X, or to move your cell phone number from carrier A to carrier B, that's “portability”. Continuous data access is a very different thing. It's not about “okay, I no longer want Google to have my data or I want to work with a different search engine provider, but I want them to be able to know me as well as Google does after three years of mapping my searches, so I'm going to exercise my right to one-off portability.” That is very different from saying “I want this new app to be able to always tap into what I'm doing on Google, or Facebook, forever,” as we apply to Fintech with PSD2; it is not about portability, it is about continued access and transactionality.

The potential value to consumers if they could combine those vast datasets and then tap into them on an ongoing basis is very different from the one-time portability question. And yet, even in specialized circles, you will often hear the term “portability” used to refer to both things without distinction.

I'd love to get your thoughts on AI ethics and how we start to think about how this data is used safely and soundly? Any thoughts on AI ethics or how the firm's thinking?

Over the last few months, we have developed a set of Citi AI/ML ethical principles. They incorporate the top elements of this sequence of frameworks that we have seen pop up around the world, and they are now working their way through the organization for validation. I think Singapore kicked it all off a year ago at Davos, where they rolled out what they billed as the world's first national ethics framework for AI/ML. Next came the group of EU experts. Then it was the OECD, so you went from one country to more than two dozen in the EU, to 43 in the OECD. It has kept happening — there are now dozens of frameworks. Once again, just as we discussed here about data and privacy, we are going to start seeing variations and fragmentation, and it is probably going to be most noticeable through omission of certain things that certain jurisdictions will not say or for which they will not sign up.
I plan to track closely how this evolves over the next year and whether some global convergence begins to emerge in AI ethics. There are real reasons to make sure that as societies we harness as much as we can of the constructive potential of AI and machine learning. To do that, we will need to be incredibly deliberate about controlling the downside, which is potentially very real and harmful for a longer spectrum of potential issues that range from developer bias to outright malicious intent, to more mundane problems of data quality at the input, or opacity of previous passes through some kind of AI/ML rig that may already have shaped your data when you take it in. Many, many new ethical questions will confront practitioners. If anyone interprets the term “ethical” to mean “abstract and academic”, they couldn’t be more misguided. These are very practical questions, with very real implications for real people.

Some of the same questions we discussed at the beginning of this conversation are going to apply. If you are a multinational company, you sign up for a more stringent set of principles. Actually when it comes to AI/ML, I do not know if you could do anything else. It just is not the same as a broader use of data for marketing or other purposes. I do not know whether you will be able to keep AI/ML in airtight compartments that prevent contagion of any issues. There is a lot at stake.
Citi Global Perspectives & Solutions (Citi GPS) is designed to help our clients navigate the global economy’s most demanding challenges, identify future themes and trends, and help our clients profit in a fast-changing and interconnected world. Citi GPS accesses the best elements of our global conversation and harvests the thought leadership of a wide range of senior professionals across the firm.

All Citi GPS reports are available on our website www.citi.com/citigps

- **Technology at Work v5.0**
  *The New Normal of Remote Work*
  June 2020

- **5G and Mobile Operators**
  *Is Virtual Becoming Real?*
  April 2020

- **Closing the Loop on Global Recycling**
  *Finding a Global Solution to the Patchwork of Recycling*
  February 2020

- **Disruptive Innovations VII**
  *Ten More Things to Stop and Think About*
  February 2020

- **Building a TCFD With Teeth**
  *What the Markets Need to Price Climate Risk*
  February 2020

- **Banking the Next Billion**
  *Digital Financial Inclusion in Action*
  January 2020

- **Experiential Commerce**
  *A Brave New World*
  January 2020

- **Investment Themes in 2020**
  January 2020

- **Corporate Finance Priorities 2020**
  January 2020

- **Car of the Future v4.0 – 2020 Update**
  *The Race for the Future of Networked Mobility*
  January 2020

- **The Global Art Market**
  *Drivers of Evolution*
  December 2019

- **Education: Power to the People**
  *Exploring Opportunities for Private Capital in Education*
  November 2019

- **Digitizing Governments**
  *The Journey to Enacting a Digital Agenda*
  October 2019

- **Electric Aircraft**
  *Flightpath of the Future of Air Travel*
  September 2019

- **Energy Darwinism III**
  *The Electrifying Path to Net Zero Carbon*
  September 2019

- **For Better Or Worse, Has Globalization Peaked?**
  *Understanding Global Integration*
  August 2019
<table>
<thead>
<tr>
<th>Title</th>
<th>Issue/Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factory of the Future</td>
<td>Flexible, Digitized, and Sustainable</td>
</tr>
<tr>
<td>Video Games: Cloud Invaders</td>
<td>Bracing for the Netflix-ization of Gaming</td>
</tr>
<tr>
<td>Bank X</td>
<td>The New New Banks</td>
</tr>
<tr>
<td>Investment Themes in 2019</td>
<td></td>
</tr>
<tr>
<td>China’s Belt and Road Initiative</td>
<td>A Progress Report</td>
</tr>
<tr>
<td>Migration and the Economy</td>
<td>Economic Realities, Social Impact, &amp; Political Choices</td>
</tr>
<tr>
<td>Disruptive Innovations VI</td>
<td>Ten More Things to Stop and Think About</td>
</tr>
<tr>
<td>UN Sustainable Development Goals</td>
<td>A Systematic Framework for Aligning Investment</td>
</tr>
<tr>
<td>ePrivacy and Data Protection</td>
<td>Privacy Matters: Navigating the New World of Data Protection</td>
</tr>
<tr>
<td>Technology at Work v4.0</td>
<td>Navigating the Future of Work</td>
</tr>
<tr>
<td>Managing Cyber Risk with Human Intelligence</td>
<td>A Practical Approach</td>
</tr>
<tr>
<td>2019 Corporate Finance Priorities</td>
<td></td>
</tr>
<tr>
<td>Car of the Future 4.0</td>
<td>The Race for the Future of Networked Mobility</td>
</tr>
<tr>
<td>Feeding the Future</td>
<td>How Innovation and Shifting Consumer Preferences Can Help Feed a Growing Planet</td>
</tr>
<tr>
<td>Rethinking Single-Use Plastics</td>
<td>Responding to a Sea Change in Consumer Behavior</td>
</tr>
<tr>
<td>Putting the Band Back Together</td>
<td>Remastering the World of Music</td>
</tr>
<tr>
<td>Electric Vehicles</td>
<td>Ready(ing) For Adoption</td>
</tr>
<tr>
<td>Sustainable Cities</td>
<td>Beacons of Light Against the Shadow of Unplanned Urbanization</td>
</tr>
</tbody>
</table>

Note: The issue/year information is provided next to each title.
IMPORTANT DISCLOSURES

This communication has been prepared by Citigroup Global Markets Inc. and is distributed by or through its locally authorised affiliates (collectively, the "Firm") [E6GYB6412478]. This communication is not intended to constitute "research" as that term is defined by applicable regulations. Unless otherwise indicated, any reference to a research report or research recommendation is not intended to represent the whole report and is not in itself considered a recommendation or research report. The views expressed by each author herein are his/her personal views and do not necessarily reflect the views of his/her employer or any affiliated entity or the other authors, may differ from the views of other personnel at such entities, and may change without notice.

You should assume the following: The Firm may be the issuer of, or may trade as principal in, the financial instruments referred to in this communication or other related financial instruments. The author of this communication may have discussed the information contained herein with others within the Firm and the author and such other Firm personnel may have already acted on the basis of this information (including by trading for the Firm's proprietary accounts or communicating the information contained herein to other customers of the Firm). The Firm performs or seeks to perform investment banking and other services for the issuer of any such financial instruments. The Firm, the Firm's personnel (including those with whom the author may have consulted in the preparation of this communication), and other customers of the Firm may be long or short the financial instruments referred to herein, may have acquired such positions at prices and market conditions that are no longer available, and may have interests different or adverse to your interests.

This communication is provided for information and discussion purposes only. It does not constitute an offer or solicitation to purchase or sell any financial instruments. The information contained in this communication is based on generally available information and, although obtained from sources believed by the Firm to be reliable, its accuracy and completeness is not guaranteed. Certain personnel or business areas of the Firm may have access to or have acquired material non-public information that may have an impact (positive or negative) on the information contained herein, but that is not available to or known by the author of this communication.

The Firm shall have no liability to the user or to third parties, for the quality, accuracy, timeliness, continued availability or completeness of the data nor for any special, direct, indirect, incidental or consequential loss or damage which may be sustained because of the use of the information in this communication or otherwise arising in connection with this communication, provided that this exclusion of liability shall not exclude or limit any liability under any law or regulation applicable to the Firm that may not be excluded or restricted.

The provision of information is not based on your individual circumstances and should not be relied upon as an assessment of suitability for you of any particular product or transaction. Even if we possess information as to your objectives in relation to any transaction, series of transactions or trading strategy, this will not be deemed sufficient for any assessment of suitability for you of any transaction, series of transactions or trading strategy. The Firm is not acting as your advisor, fiduciary or agent and is not managing your account. The information herein does not constitute investment advice and the Firm makes no recommendation as to the suitability of any of the products or transactions mentioned. Any trading or investment decisions you take are in reliance on your own analysis and judgment and/or that of your advisors and not in reliance on us. Therefore, prior to entering into any transaction, you should determine, without reliance on the Firm, the economic risks or merits, as well as the legal, tax and accounting characteristics and consequences of the transaction and that you are able to assume these risks.

Financial instruments denominated in a foreign currency are subject to exchange rate fluctuations, which may have an adverse effect on the price or value of an investment in such products. Investments in financial instruments carry significant risk, including the possible loss of the principal amount invested. Investors should obtain advice from their financial advisors before making investment decisions. Although the Firm is affiliated with Citibank, N.A. (together with its subsidiaries and branches worldwide, "Citibank"), you should be aware that none of the other financial institutions mentioned in this communication (unless expressly stated otherwise) are (i) insured by the Federal Deposit Insurance Corporation or any other governmental authority, or (ii) deposits or other obligations of, or guaranteed by, Citibank or any other insured depository institution. This communication contains data compilations, writings and information that are proprietary to the Firm and protected under copyright and other intellectual property laws, and may not be redistributed or otherwise transmitted by you to any other person for any purpose.

IRS Circular 230 Disclosure: Citi and its employees are not in the business of providing, and do not provide, tax or legal advice to any taxpayer outside of Citi. Any statements in this Communication to tax matters were not intended or written to be used, and cannot be used or relied upon, by any taxpayer for the purpose of avoiding tax penalties. Any such taxpayer should seek advice based on the taxpayer's particular circumstances from an independent tax advisor.

© 2020 Citigroup Global Markets Inc. Member SIPC. All rights reserved. Citi and Citi and Arc Design are trademarks and service marks of Citigroup Inc. or its affiliates and are used and registered throughout the world.
NOW / NEXT

Key Insights regarding the future of ePrivacy & Data Protection

**INNOVATION**

We are seeing an emergence of privacy-focused tech which could provide consumers with alternative means to take more control over their data. / **A global attempt at creating an AI Ethics framework could trigger an acceleration of innovation in data privacy technologies and solutions.**

**REGULATION**

The EU’s GDPR was the first large-scale regulation to address data privacy and protection, but the current patchwork of regional data privacy requirements make for a challenging environment for global corporates. / **Global harmonization of data privacy rules will be tricky given cultural, political and societal backdrops.**

**TECHNOLOGY**

Big Tech players from the U.S. and China currently dominate and large pools of data sit in the hands of a small number of cash-rich platforms. / **A battleground is heating up as world leaders recognize the development of artificial intelligence (AI) — which requires large pools of data — as being an important force in driving economic growth and social good.**