DATA: THE NEW BLACK GOLD?

Seminar on the use and ownership of data

Thursday 2 February 2012

Sponsored by
SNR DENTON

Below is a summary of comments made at the seminar.

Chair's Introduction – Adam Singer

The Chair welcomed guests and thanked Intellect, who were co-hosting the event with BSAC, as well as SNR Denton for providing the use of their facilities.

This event was very timely in terms of the current regulatory debates on data in the US and Europe. The title of the event 'Data: The New Black Gold?' was a good metaphor as data was similar to crude oil, it was flammable, it was capable of great pollution, and it required vast refining to create value. Data was the crude oil of the information age; it was the fuel of the silicone engines. The event would touch on what data made possible, the regulatory issues of storing and using data, and how businesses could take advantage of this new informational power. The 'data extraction' sector felt like the early stages of the oil industry, automotive or cinema industries at the beginning of the twentieth century, it was new, inevitable, powerful, but it was developing very rapidly. The currency of Google and Facebook was data. Users paid Facebook with personal data which was extremely valuable. A question for the seminar discussions was who benefited from data? Was this a Faustian pact where users relinquished their data for a bauble of convenience? Or, alternatively, did a voluntary subscription of data to these organisations through users' activities simply represent an increase in the quality of users' connected lives? Google, Twitter and Facebook were essentially data traders and users were data generating units. This was a world where the concept of privacy was replaced by levels of information accessibility. The data world brought new risks similar to the arrival of the automobile. For the convenience of personal transportation, we were willing to pay the price in terms of road construction and fatalities. The price of data would be inevitable spills. Music spilt data when it met piracy and Napster. Governments spilt in terms of lost laptops and WikiLeaks. Sony spilt data when it leaked a torrent of credit card information pertaining to PS3.

The Chair introduced the panel on 'Citizen Concerns and is the Legal Framework Fit for Purpose?' This was aptly timed as the EU had recently announced plans for a new Data Protection Directive.

Panel: Citizen Concerns and is the Legal Framework Fit for Purpose?

Chaired by Derek Wyatt

Nick Graham, Partner SNR Denton

Nick explained that current data protection legislation was based on the European Directive. New regulation was published on 25 January which would significantly impact on to the current privacy regime. The current regime was quite fragmented in terms of different rules in different countries within Europe. In the UK, one could be fined £500,000 for breaching the Data Protection Act, but in Belgium there was no fine. The US had an entirely different regime which was an issue as data did not respect borders. The US regime was sector-focused and much narrower but there were more data breach regulations and notification. In addition to non-harmonisation, in terms of audiovisual and digital media, one of the biggest issues was transparency and simplicity in explaining what companies were doing with data. Consumers wanted to use apps but it was quite tricky to explain in simple terms what data was being processed in connection with that. Nick thought that sometimes Europe got it wrong in terms of the rules governing data use. For example, regulation in relation to cookies and the fact that they now required consent. Nobody wanted to be snooped on, Phorm had run into difficulties around online behavioural advertising. However, requiring consent meant that home pages would be covered with pop ups. This was akin to designing a new electric light where consent had to be acquired before it could be switched on. Currently, there appeared to be a privacy battle in progress between privacy and public advocates. It seemed increasingly that one had to side with one group or the other. This binary approach to assessing the issues was perhaps not the best way to find solutions. On jurisdictional issues in terms of where the consumer went to seek redress, rules varied from country to country. In the UK, the regime was policed by Christopher Graham, the Information Commissioner, who possessed powers to fine and serve enforcement notices. Some of the activity was a reputational issue. Google had a problem with StreetView which had been signed off by the Information Commissioner's Office (ICO), however, as a result of negative reputational issues they agreed to a voluntary order, the results of which were published in August. In the US, an individual would have a right to sue for negligence, however, it was very difficult to prove a loss and so receive compensatory damages.

Rob Reid, Senior Policy Advisor, Which?

Rob would cover what Which? considered to be key consumer concerns around privacy and data use online, focussing on research that they had undertaken around the use of tracking for profiling and online behavioural advertising. Which? had conducted multiple surveys into consumer understanding and awareness of online privacy issues over the past two years. The results had been reported in the magazine, online and they had gathered further consumer input through comments on blog posts on the subject. This research had raised some clear consumer concerns around personal data and online privacy. These concerns could be split very broadly into three categories:

- Transparency. 53% of the people Which? had spoken to said that they were concerned that their data was passed to unknown companies without their permission. 48% said that they were concerned that they didn't know how their data was being used.

- Choice. In order to exercise choice, people needed to know what choices they had. Less than 50% of people Which? had spoken to were aware of online behavioural advertising. Interestingly, given the choice, 64% said that they would prefer to receive adverts of relevance to them but that they objected to the covert way in which the information was gathered. 70% felt that tracking cookies were an invasion of privacy.

- Control. Users wanted greater levels of control over how their data was collected and how it was used by companies. 62% of people had taken some action to control how their data was collected, such as deleting cookies, for example. Taking control did not necessarily mean that users would refuse requests to collect and use their data. The desire for control in the surveys that Which? had done involved opting into cookies and 73% of people wished to be given this choice. However, when people realised the impact this could have on their web browsing activities, only 49% of people wished to opt into cookies. Some work needed to be done on allowing people to consent.

Rob concluded by saying that from Which's research it was clear that consumers were largely unaware of how and why data on them was collected and used. When they learnt of the processes at work, they often imagined the worst. For example, almost a third of people believed that financial details were routinely logged by cookies. There was a real need for industry to fill this gap with clear, unbiased and accessible information. Consumers would not wholly object to the collection of their data if they were given a choice. There was a danger that if the industry did not engage with consumers in providing information and a choice on data collection, once consumers learnt that these practices were occurring they were more likely to reject them entirely.

Donna Whitehead, Government Affairs Manager, Microsoft

Donna said that Microsoft was interested in the issue of privacy. Bill Gates had launched a department called Trustworthy Computing ten years ago, and across Microsoft globally, where of 9000 employees, something like 4000 of them had privacy as part of their job commitments. Microsoft dealt with customers: they were a publisher via MSN, they ran an ad network, Microsoft Advertising, and they were a browser manufacturer. When it came to data privacy, had the perception of policy concerns started to outweigh the reality of citizen's concerns? The technology revolution had ended but it had also only just started. The majority of people were now used to having technology in their lives, many people now had smartphones. The use of laptops would decrease and tablets would increase. However, in some ways, what the technology revolution could do was only just being touched on. Governments were trying to create legislation in an unknown environment. There were some political slogans in use such as 'the right to be forgotten'. What this meant was ambiguous but should refer to a re-branding of processes which already existed to ensure that citizens could manage their own data. Companies that made use of data were going from strength to strength, with Facebook reportedly worth $100bn, and geo-location applications such as Foursquare were increasingly popular across devices. The BBC Homepage could be modified by users according to their interests. These were all fantastic uses of data.

This issue needed to be put into perspective. It was important that a citizen could manage their own data within a solid framework, and at that point, there would be no further need for debate.

Steve Taylor, entrepreneur

Steve said that he had been Director of Innovation at Aegis Media and had left 18 months ago. He had seen firsthand the excitement that the possibilities of data had on that industry. A recent McKinsey report had estimated that there was an unmet demand for more than 150,000 data analysts and specialists in the US alone. There were parallels to be drawn between the effect that digital had on the music industry and the burgeoning uses of data. Regulation, legislation and litigation failed to effect change in consumer behaviour. Steve Jobs had changed consumer behaviour by making digital content easy to access, cheap, and integrated into a supply chain. He felt that the issue of data was in a similar ferment currently. Another statistic from the same McKinsey report estimated that intelligent use of big data would save US healthcare $300bn a year. Was there a case for a benign contract between consumers and the owners of big data resulting in a win/win situation? This was about much more than monetisation. There were opportunities to extract value from data which went beyond the commercial sphere and which had the potential to improve quality of life in many ways for people both in the developed and the developing world. It had been suggested that there needed to be a central repository where the citizen could access all of their data, commercial, social etc.

Q&A

Philip Virgo, Conservative Technology Forum commented that dark market websites offered the most cost effective way of finding out about an individual.

Nick Graham said that it was true that there were databases where one could buy data on anyone in the UK including name and address. Google-ing someone was the quickest way to do this. However, those databases would not contain things like credit card details. There had been websites such as B4Usearch.com, which a few years ago was a legitimate series of online services but branched out to provide extra data. Nick had searched for himself to see how the site worked and been horrified to see that his name, contact details, and a map showing where he lived, had been available. The site had been enforced against by the Information Commissioner's Office. At the time, there had been limited enforcement powers so the site had been penalised on grounds other than their processing of data.

Luke Crawley, BECTU commented on Steve Taylor's question concerning whether it was possible to achieve a benevolent settlement between the user and the service concerning use of data. Luke thought that this was not possible because state bodies as well as private companies that held data seemed extremely bad at keeping hold of it.

Nick Graham agreed. Big databases were an issue because of the security risk. If there was a problem, the risk was much greater due to the volume of data. It was important to consider this on a global scale because data did not respect borders. One of the issues would be understanding privacy expectations in different countries and particularly with European legislation, the focus was on the individual's rights, as opposed to attitudes in the US which were more focused on the fact that businesses owned the data. This was a reason why a benevolent settlement could be difficult. Allowing consumers to opt-in at the point of access to services was problematic as it involved explaining complex data flows and data sharing arrangements. However, this may be a transitional issue as consumer awareness increased with

technological advances, for example, everyone knew what a cookie was now, but they would not have done a few years ago. The other point was the validity of consent. To be valid the law stated that consent must be freely given. Therefore, there had to be a mechanism to allow consumers to withdraw consent at a later point.

Patrick Robinson, Yahoo commented that the vast majority of data that media businesses and services would be collecting on consumers was not 'sensitive'. It would be very rare to hold people's medical records, tax information, or credit card information. Luke had raised an interesting point about data security. However, it was not necessarily the main focus for media companies, which was information that was not necessarily to do with an individual as a specifically identifiable person, but was about an individual online. The kinds of information Yahoo collected from users in relation to advertising concerned the pages they viewed online. These things were not necessarily sensitive; they did not identify an individual. What they did was build a profile about what an individual did online. The challenge was doing that in a transparent way that allowed people to control the collection and use of that data. When consumers were given access to the data held on them, they wanted to make sure that the information was accurate. Consumers used Yahoo's advertising preference manager to provide information about their likes and dislikes or what demographic they belonged to. People were willing to control their own information once they were informed and able to make proper choices. That was the challenge for industry.

Lavinia Carey, British Video Association said that the British Video Association (BVA) had talked to Which? about how to explain to the public how extensive the availability of online video services was. One of the first questions that Which? asked was how those services ranked in terms of their code of practice in using consumer data. The BVA did not have access to that information but wished to promote legal video services as opposed to illegal services. As a consumer, Lavinia found it very difficult to control what data was collected and used by services without boycotting them altogether. Providing consumers had some control over use of their data, there were substantial benefits for both parties.

Rob Reid agreed with Patrick in terms of providing control and transparency to consumers. However, he disagreed that the advertising sector was doing a good job in informing people. The IAB self-regulatory scheme was not working in this respect. The reason people were likely to opt out of receiving targeted advertising was because they did not like the idea that a profile was being built up on them across various internet domains. That was not prevented when one opted out using the IAB scheme, one opted out of receiving a targeted ad, the profiling could still continue. The scheme could be improved. Rob agreed with Lavinia. When Which? met with the BVA, they had asked if Which? could help promote sites that were providing cheap, affordable legal downloadable content. One of Which's concerns was that these sites tended to gather and hold a lot of data on their customers, and there could be issues with how secure the sites were. There had been a study in the US on use of data by Netflix, for example, where it was possible to identify single individuals even though the data was anonymised.

Steve Taylor disagreed with Patrick's point. He did not think that the facility to opt out or to change preferences was made easy. Last week, LinkedIn had created a new option for their users to opt out of having their social data used to target advertising. However, it had not been well publicised and involved a very complicated process. It signified an attitude of sticking to the letter of the law but not the spirit of it.

Donna Whitehead said that from Microsoft's point of view as an advertiser, they were constrained by the way that the internet was set up. If a consumer opted out, there would be blank spaces on their homepages. Facilities like Yahoo, Microsoft or parts of Google were intermediaries for the web and pulled content from many different places. They were attempting to create tools for consumers. Consumer education was crucial in order for people to understand how they could take control of their own data. However, tools were available via the self regulatory scheme. There were also tools available via Internet Explorer that consumers could use to control privacy settings, such as 'in private' browsing, which stopped any collection of data at all.

Lord Merlin Erroll said that legislation very rarely did what it was intended to do. He was not sure how much good legislation in this area would do. There was also likely to be a lot of unintended consequences of legislation, for example, imposing strict penalties for data breach would make companies less likely to report it. Very few people understood this area. Many thought it was possible to control the internet and others thought it would be possible to set a territorial boundary without extremely far-reaching consequences. The internet did not suit regulatory systems as it was a complex world of networks where putting rules in one place would produce an unexpected flow elsewhere. It was impossible to predict the outcome. France had tried to form a regulatory system which went beyond their boundaries. Even if this had been successful across Europe, the problem was that there were countries that would not cooperate for whatever reason such as Russia on piracy.

In terms of consumers managing data, one problem was that someone else may use one's computer, such as a family member, to search for something, which would then be added to one's online profile and may not be desirable to see against one's name. Another problem was that consumers were sometimes forced into providing personal details in order to use services, if one was not willing to provide one's date of birth, the service was denied. This forced the consumer to consent. Laws could be passed but it depended where the site was hosted as to whether the laws would apply.

David Elstein, BSAC said that there were two types of information being gathered. There was anonymised aggregated information for research or other purposes and there was personalised information gathered and sold on by a relatively small number of large organisations making billions of dollars from it. In terms of personal information, there was an argument that companies should be paying for this. There could be a mechanism whereby companies could not use consumer data unless they could demonstrate that they had paid for it, or if not, that they could not use it outside their organisation. This would allow a cash nexus which was appropriate to the value of the information. If Facebook was worth $100bn and it had 1bn users, did that mean each user was worth $100? When ITV sold advertising breaks for *Downton Abbey* the individuals watching were unknown, it was just ten million people with a broad profile. This meant that there was a lot of wastage for an advertiser. Targeted advertising eliminated that wastage but it was only possible with much more precise profiling which was worth money. Regulating so that companies must pay for that information was better for consumers and organisations than a consent system.

Paolo Siciliani, BBC Trust said that he had previously worked at the Office of Fair Trading which had done a market study around targeted advertising and targeted pricing. They had found that consumers were happy with targeted advertising but not with targeted pricing. The distinction between targeted advertising and targeted pricing was very blurred. For example, there could be a certain advertisement for a particular product variety which was more expensive because a consumer lived in an upmarket

area. If the advertising had not been targeted, the consumer may have been happy to pay less for a different product.

Rob Reid said that there were business models in use that paid consumers for their data. Consumers gave information about things they were interested in and received a small amount of money in return when advertisers profited. Targeted pricing was something that Which? would be concerned about if they found evidence of it.

Donna Whitehead said that Microsoft did not have the capability to link what a user searched for with that individual user as Lord Erroll had suggested. There was the capability to find out the type of user and link it to what they searched for but it was not matched to an IP address and then to a post code and then to a property. In answer to David Elstein's point, the internet was advertising funded so one used Facebook for free. If users wanted to receive cash for providing Facebook with their data, they might have to pay Facebook to access their accounts.

Steve Taylor thought David Elstein's point had been that if his data was then sold on beyond the media owner that somehow the user should share in that. The idea of what data one owned was fascinating. Knowing the height and age of an acquaintance was a piece of data about them. Did they own that data?

Nick Graham said that data protection was founded on the basis of privacy so consumers had rights in a legal sense. What privacy advocates would say was that it was based on respect. A regime where a consumer could sell their data as if they were a commercial entity was difficult as they would lose all rights to that data which was at odds with the respect principle.

Mark Selby said that the debate had parallels with the arguments around CCTV. CCTV only exploded in the way it did because of the James Bulger tragedy which that caused public opinion to change. It would be interesting to see when consumers would wake up to the implications of the algorithms analysing their data. It was possible to analyse individuals using only location based data, it was possible to tell if they were visiting a clinic regularly, which was interesting for medical insurers. What might seem like an innocent piece of data such as the data related to smart metering, for example, could be more sinister. Smart metering was stopped in The Netherlands when it was realised that people would know when houses were empty. The campaign that stopped smart metering in The Netherlands was if it had been available then, it could have been used to find Anne Frank. The value of data changed with time and the analytics available.

Richard Dikstra, Belle Media asked about how jurisdiction was defined when an individual set up a social network or a service with global appeal? What protection did someone setting up a global service in the UK have that no other jurisdictions would claim that they had broken the rules in their country?

Nick Graham said that this was a difficult area. Local law would be determined by the local legislation. In Europe, the data protection law applied for services established in Member States or equipment used there. There was a subtle debate about whether cookies dropped on a PC qualified as use of equipment. Whereas, consumer law applied based on targeting services at the jurisdiction and was about protecting the individual. There were differences between Europe and the US. Massachusetts had a very broad privacy law that protected Massachusetts residents and a service did not have to be established in Massachusetts for the legislation to apply. While there were some common trends, an

international business setting up had to consider how many touch points to have in local jurisdictions and ensure compliancy with all of the relevant local laws.

Simon Milner, Facebook said concerning jurisdiction, all Facebook customers outside of North America had a contract with Facebook in Ireland which was where they were regulated. The Irish Data Protection Commissioner undertook a three month audit of Facebook and all its practices, and produced a 200 page report in December which highlighted many things Facebook did well and some they needed to improve. Facebook recognised its incredible responsibility given the amount of personal data that people stored with them. Facebook did not sell data to anybody. The data provided was used in order to try and target advertising to users which paid for the service. Facebook were pleased with the Commission's plans in terms of the philosophy of a 'one-stop-shop' for companies operating across the EU.

Adam Singer said that the discussion reminded him of debates in the music industry after the invention of MP3 and before iTunes, that copyright could be protected and legislated for in the world as it was currently known. There seemed little difference between the copyright debate and the privacy debate. They were both about information access regimes. Technology had meant that the concept and effectiveness of copyright had been up for debate for the last few years. There was no permission to believe that copyright rolled forward. The industry wanted it to, it was important. The privacy debate was the same issue. The debate was focused on the current situation. The complex set of laws relating to copyright had proved very difficult to educate the consumer about, the same applied in terms of privacy.

The Chair asked the panellists for any final thoughts.

Donna Whitehead said that going back to the comment on location based data one had to specifically allow mobile phone apps to use one's location data. The Windows phone used a 'one way cryptography hash tag' which anonymised the data so when someone opted in to location based services, the data could not be linked to a person. One note to leave the discussion on was a comparison between Mark Zuckerberg who had taken masses of data and made a fantastic business model, and Julian Assange, who had given masses of data away for free.

Rob Reid said that he agreed that issues around data collection and data use were likely to grow rather than diminish. As more was collected and the technology that was used to collect that data evolved, consumer awareness did not seem to be keeping pace. Consumer awareness was vital to the industry going forward. Which? was considering whether it was feasible for them to develop an accredited privacy policy, a simplified version which consumers could understand.

Nick Graham said that the new EU regulation had not been discussed much. Industry needed to be cognisant of the fact that it was difficult to explain privacy regulation in simple terms and the evidence was that people did not read privacy policies. The industry needed to be cleverer about how it got the message across.

Panel: How are Businesses Adapting to the New Environment?

Chaired by Adam Singer, BSAC Chairman

The Chair said that privacy was an area where there seemed to be a generational moral shift as his children and grandchildren seemed to be trained in hiding data in public. The ability to say things about yourself in public which only others understood if they had been given the key of context seemed to be a new skill that had developed. The panel would be looking at how to take advantage of these opportunities. The power of advertising was about to shift dramatically with the increasing ability to use data to target advertising. Those with the power to mine data successfully would do very well. In this world, statistics became cool. There would be an arms race between advertisers, media owners and content owners in terms of how data was used. It seemed likely that this would cause another tumble down the value chain for established media, especially established radio and television broadcasters. Radio had not been adroit at utilising the return path, for example. Traditional industries, who had built their models on broadcasting a one-way traffic to an audience, were not genetically disposed to understand the power of the return path. David Abraham, CEO of Channel 4, had said 'if broadcasters only rely on external data sources rather than building their own they risk disintermediation'. The question was whether broadcasters could acquire sufficient information from data streams to be able to maximise advertising revenues? The real test of broadcasters' ability to mine data would be through services such as YouView launching later in the year.

David Boyle, Head of Insight, Zeebox

David said that currently businesses were not deriving as much benefit and value from data as they should be. The 'data as oil' metaphor was interesting in terms of how to improve the power of the data. Oil had to be refined so that it was accessible to people. As a consumer, oil was part of your everyday life. Google did a wonderful job for consumers of packaging data up and making it available. Companies often did not do that. Data was locked away forcing consumers to go out of their way to get hold of it. Data was not processed so that it was understandable to users. Another comparison was the demands that businesses put on data. Oil was measured in miles-per-gallon and the same was true of data. One could get incredible mileage out of data if one asked the right questions of it. However, too often data was not used in a powerful and efficient way. In the TV world, instead of thinking about the average audience for a show, or the average age of a viewer, broadcasters should be asking what all the different types of viewers were that engaged with the show. For example, the type of viewers that loved it and the type of viewers that were not that interested. This was when data became more powerful. Finally, it was important to remember that deriving value from data was also dependent on human skill and judgement.

Theo Bertram, UK Policy Manager, Google

Theo said that Google had a service called Google Trends which allowed users to search for different terms and identify trends and patterns. They had found that they were able to predict epidemics occurring before the WHO could, due to the frequency of particular terms. Businesses would need to be adaptive in this new environment. Moore's Law described how growth in the power of computer data was not linear but exponential. It described how the ability to place transistors in integrated circuits doubled every two years, and with their increased speed, chip performance doubled every 18 months.

MIT professor Erik Brynjolfsson illustrated this using the image of the second half of a chess board. A Chinese Emperor was challenged by many people to a game of chess and eventually one man beat him and the Emperor offered him a big pile of rice as the victor. The man refused and said that he would prefer to fill the chess board by putting 1 grain of rice in the first square, 2 grains of rice on the second, 4 on the forth, and so on. As the amount on each square kept doubling, by the time the end of the chess board was reached, the earth's surface would be covered 5 ft deep in rice. Drawing a comparison between this story and the exponential expansion in the power of computers, the second half of the chess board was about to be entered where there would be extremely large leaps in the power of data processing. The impact of this would not just be for personal data but for all kinds of data processing. It was not just about tech companies, it was about all companies as it would enable decisions to be taken based on very fine granular detail rather than on gut instinct. The McKinsey Report had predicted a demand for data analysts and data savvy managers across industries. It would be important for businesses to acknowledge that there was not only one business model that would work in this new environment. There needed to be multiple and changing business models. The question of scale was a particularly important one for the UK. A successful, or even moderately successful, business in the US reached 200 million people. In the UK, that figure was 30 million. 200 million put a business on a global scale, but 30 million did not. This presented a challenge for the UK and meant there was all the more reason for the UK to compete globally. The Daily Mail and The Guardian, for example, had far bigger global readerships than they did in the UK. What made data really valuable was when it was converted into insight. For example, a cod liver oil salesman had discovered through using Google Analytics that there were a high number of searches for cod liver oil in Berlin and around the Canary Islands. He realised clubbers were buying cod liver oil as a supplement for the day after they had been out. He needed to understand the data that was available and have an insight into it to realise that there was a new market for his product.

Louisa Wong, General Manager, EMEA, Aegis Media, AMNET

Louisa said that she ran a trading desk, which was a new emergence in the advertising landscape, and could be described as similar to a hedge fund. The people employed within her team were statisticians looking at data constantly and bidding for media on advertising exchanges. A lot of the media buying was being transacted in an online or a digital environment, a change which was driven by an explosion of data in the marketplace. Terence Kawaja from LUMA Partners had created sector landscapes mapping the digital ecosystem. At the bottom of the chart were data exchanges or data suppliers. Data came in different guises. In the digital world, it was cookie data, however, in terms of the advertising landscape 'data' was anything that provided insight into how consumers behaved, for example, engagement with a set top box. That data was captured and started to allow an understanding of how consumers behaved within the electronic programming guide. Data also referred to behaviour online, use of social media, use of search engines, browsing across devices etc. Anything that had an IP address enabled an understanding of consumers' behaviour. This did not mean personally identifiable information but an understanding of how big groups of users were behaving which allowed targeted advertising to be served. Whilst there was an explosion of data, there was an over supply of media as well. Data powered where investments were made, on TV, in the press or in social media. For example, the press was not just measured through readership but circulation relating to geographical locations. This linked in to advertising in other forms of media and planning where to make investments. Consumers were exposed to multiple media channels throughout their daily lives across devices. If a consumer looked at the BMW website for example, with the idea of upgrading their car, the IP address that they used and the geolocation becomes an identifier of that consumer and what they may be

interested in. BMW could then see trends of people in particular areas looking at their website which enabled them to serve a targeted advert to specific IP addresses with a special offer on customer service at BMW for example. Although an IP address was not personally appended to the consumer, it became a common denominator. This enabled BMW to become more effective in terms of where investments were made as a deeper understanding was gained of what customer demographics looked like and were interested in.

Q&A

The Chair asked whether harnessing the power of data about a large group of people actually enabled the prediction of trends going forward which were highly saleable to advertisers and major organisations?

Theo Bertram said that Google Trends showed the probability of particular search terms across regions and time. HM Treasury, for example, had used it to see if they could find correlations between the cycle of the economy and searches for jobs and what that meant for unemployment figures. However, these were predictions or probabilities, not certainties. Google could sell this information but that was not their business model. There were lots of people doing research with this data but the results depended on whether they were able to convert that data into something meaningful.

David Elstein said that data was not just an economic issue; it was also a political and cultural issue. The citizens' ability to respond to the opportunities that data collection represented was just as important as businesses making money out of the information that was being gathered. Google had funded a project designed to anticipate complex civil wars or outbreaks of genocide by enabling individual citizens to upload data which was interpreted by experts for significant patterns with the idea of allowing key governments to intervene to prevent conflict.

David Boyle said that he had worked in US and UK politics. There were good and bad consequences of the increasing volume of data available. In the US, it was possible to buy almost any piece of data on almost anybody in the country which was scary. The UK had much stricter privacy laws. However, there were also some positive consequences to data availability in the US. The environmental groups, women's groups, the labour unions and the Obama campaign worked together to share messages and to share contacting individuals. If someone was a member of a women's group, an environment group and had also signed up for the Obama campaign, the groups might coordinate who would manage the relationship with that person during the election. The progressive forces for good in the world could better spread their communication and talk to the right people on the right issues during the election.

Louisa Wong said that there were issues with the idea of governments using data to make the world a better place. Limiting the use of social media, for example, as had been suggested after the London riots was dangerous. Countries such as Thailand censored what its citizens could say online. That was not a situation that the UK should be in.

Paul Malyon, Experian said that businesses were becoming cleverer about how they used their data and how they use their money to advertise. What businesses were trying to avoid, using the oil analogy, were leaks of money. Returning to the BMW example, the company could now register the details of potential customers who had used their website and avoid sending out an expensive glossy brochure in the post to people who did not fit the demographic that they could sell to. Taken further, one could start

to see organisations trying to limit their risk around how they used data and only advertising to certain demographics. This was something that needed paying attention to.

The Chair agreed. There would be algorithmic issues as businesses adapted to new possibilities. New businesses were likely to come in and start using this data in a new way, and by definition they would push the boundaries of acceptability. It could be argued that Facebook had pushed the boundaries of acceptability. When Facebook had first arrived, he had been shocked by the extent to which people were prepared to make their lives so open. There were new opportunities in this space which were not being exploited at the moment and businesses would emerge to take advantage of them.

Louisa Wong said that advertising budgets were not increasing; they were staying flat. Advertising agencies and brands had to be smarter about where they made investments. Algorithms and technology were constantly referred to but the key was in the human application of data. No one in the industry had really managed to grapple with the exponential growth of data yet.

Theo Bertram said that Larry Page, the CEO of Google, believed that the job of computers, of data, was to create space for humans to do the things that they enjoyed. Computers helped to organise life but to live, human choice and human control were necessary and computers should be in the background. There was an element of truth in the idea that the algorithm could not choose between one individual and another, which was one reason that Google argued against the idea that search engines could regulate the internet. Algorithms could not replace the scrutiny of a judge. Checks and balances were necessary of which algorithms were incapable. With that in mind, Google was always working to improve its algorithm. They had an advertising preference manager which allowed users to change or delete details about themselves which were linked to advertising they were served. For every 8 people who visited that page, 7 did not make any changes, and only 1 either deleted details or edited them.

David Boyle said that with the exception of outdoor advertising which was geographical in nature, the concept of neighbourhood ceased to matter. If the data was available, advertisers and companies should be able to find customers in any area. In politics there were some neighbourhoods that were solidly Republican and Democratic canvassers would be sent there to knock on doors but they would be targeted to knock on the right doors. In the TV world there were some shows which were unpopular. However, if a business knew the type of person that watched them and connected with that show, then that was an opportunity to engage with them.

Lord Errol said that one problem with targeted advertising was that it was not sophisticated enough to know when someone had bought something, which was irritating for the customer and a waste of money for the advertiser. Another thought was that given the exponential growth of data, did the panel think that the consumer would end up being able to hide in plain sight due to the anonymous mass of unsorted and unstructured data?

David Boyle said that one trend, as the growth in the power of data and data analytics accelerated, would likely be that consumers would want greater safeguards and controls, for example, searching on Google without being logged in and without anything being linked back to an IP address. Consumer trust would become increasingly important for internet brands. Data would enable everything to done more quickly and enable more personalised results. This could have the consequence that consumers became blinkered. David thought that there were two kinds of search. One was to do with convenience and immediate information, for example, searching for the closing time of the nearest John Lewis

where Google could work out before the user had finished typing the location they were likely to be looking for. The other kind of search was the serendipitous exploration of things that one did not expect or know about. David was not sure if anyone had quite worked out how to advertise for that inspirational part of the web. In terms of targeting TV advertising, Zeebox tried to understand what was happening in the show, the subject matters and the issues, not all of which were directly related to the point of the show, and allow users to click on those and explore the random and interesting things that happened to crop up during the show. Data that they generated often showed that it was side topics or random facts that people were engaging with much more than the core of the show itself.

Jonathan Davies agreed that the capacity of the person doing the research to ask the right question was key, not just the availability of data. In analysing vast quantities of data based on the questions asked, it could be serendipity or just wisdom that one came up with a completely unexpected fact which distracted from the original purpose of the research. The issue was about the extent to which the manageability of the data had increased and the extent to which people were getting better at posing the questions which were a function of the information they seek. This would result in the increased ability to analyse these data.

Joshua Green, Arts Alliance said that in terms of businesses targeting advertising in finer and finer granularity, it forced consumers to be tied to a profile. There were some systems where in order to use a service, one had to enter personal details and this meant a profile would be built up. This could lead to prejudice as instead of watching a broadcast programme on the BBC when one might unexpectedly discover something that one liked, a consumer would be fed something that they were already likely to like. This would happen in every aspect of one's life. Ultimately, government regulation would have to deal with the issue of requiring data from consumers.

Louisa Wong said that there was a framework developed by the IAB, along with the Incorporated Society of British Advertisers (ISBA) and the European Advertising Standards Alliance (EASA), to address the EU Privacy Directive around the right to anonymity, or the opportunity to opt out. Consumers could opt out of receiving targeted advertising through youronlinechoices.co.uk. However, she agreed that current legislation meant that mechanisms for opting out were quite flimsy. Consumers were not able to fully opt out and consumer education also had a role to play.

Theo Bertram said that it was possible to search on Google without being logged in or one could use Chrome Browser in 'incognito mode' which meant that no information was stored. There were ways to use products and services without any exchange of personal data. However, many internet businesses gave away products and services for free that users liked and which were paid for through advertising and through targeted advertising mostly. Without targeted advertising, it would be much harder to provide those services. As a user, refusing to provide any information yet wishing to continue to receive services for free was not a sustainable model for innovation.

An audience member said that anonymity did not particularly interest him. He was concerned about the poor service and poor quality of data. An average ad campaign was served to the same person with five different cookies. There was a huge amount of inefficiency. If he went onto Google's preference profiles on his iPhone, iPad and computer, he was apparently three different people. Youonlinechoices.co.uk as the first step by IAB was useful, however, presenting an average consumer with a list of companies tracking them, many of which that person may never have heard of was not

effective enough. Consumers wanted better service and the chance to correct companies who had the wrong information about them.

Steve Taylor said that he used to work for Aegis. Most of the sources of data that Louisa was working with on the trading desk belonged to third parties. What was to stop BMW purchasing the same data sets, hiring the same people, and doing the job themselves? Would agencies be disintermediated in the future?

Louisa Wong said that the value lay in the people working at Aegis. They did not own the data but they facilitated it on behalf of the client. If a brand decided to take that function in-house, they would find it very difficult to buy at scale and apply that logic at scale. There were brands such as Amazon and P&G that had tried to take these functions in-house. P&G had struggled and now were working with their agency again. Amazon had been very successful because their front door was online. Louisa did not think that agencies would be disintermediated. Within Aegis the number of people working on one account and across multiple channels, press, radio, outdoor, print, digital, meant that brands could not take that function in-house.

David Boyle said that to the point on the imprecision of data, he agreed that the targeting of advertising was often very crude. Firstly, people running advertising campaigns often did not ask interesting questions and analysts did not make clear the options for marketing personalisation. When David had worked for the Labour Party during the 2005 election, they did some polling which was targeted at a micro level and tried to convince the party to do targeted individual mail shots to voters. The idea had been shot down. 10 variants of the mail, targeted for different segments from the polling, had been agreed instead of many personalised ones. All of which had gone through a sign-off process of various people in the election campaign. In the end, they had all been averaged back down to the same language.

Theo Bertram said that data would become increasingly targeted. However, he did not think that it would ever be possible to identify the individual nor should it be. As the ability to target and the power of data increased, the control that the individual user had over that data did too. Facebook, Yahoo and Google were all providing users with more control which was of value to those businesses as it helped make it more accurate. There was an interesting parallel piece of work going on in the public sector. The Midata initiative was trying to put the individual in charge of their own public data. Currently, there was a large amount of erroneous information in the public sector and one of the ways to remedy that was to give the individual control of their own personal data to make those changes. Focussing data more sharply would always involve giving the user more transparency, control and choice.

Fiona Clarke-Hackston, Chief Executive, BSAC, thanked the panellists for their interesting and insightful comments, and Adam Singer and Derek Wyatt for their lively and engaging chairmanship. She also thanked SNR Denton for their generous sponsorship, and Intellect, BSAC's co-hosts, for their help in organising the event.

Guests had received a briefing paper on data protection concerns prepared for BSAC by SNR Denton which would also be circulated via email and would be available on the BSAC website.

In May, BSAC would be holding a Conference entitled 'Get Creative: Making the Most of the UK's Creative and Digital Sectors', bringing together leaders from publishing, music, TV, film, technology, games and new digital media businesses to consider what the UK could do in order to compete more effectively. BSAC was running the Conference in partnership with Oliver & Ohlbaum Associates.