Which?

Britainthinks
Insight & Strategy

POLICY RESEARCH REPORT JUNE 2018

Control, Alt or Delete?

Consumer research on attitudes to data collection and use

Contents

This document includes a research summary written by Which? and the full research report written by BritainThinks.

Which? Research summary

BritainThinks Full Research Report

Executive summary

Data-dependent technology, in its various forms, has become fully integrated in society: our quantitative research found that 9 in 10 (91%) people are online every day and that they use various data-dependent products and services. For example, three quarters (75%) of people use online maps, two thirds (67%) social media, a similar percentage (64%) apps, and half (53%) streaming services.

Our qualitative research surfaced how much people enjoy the benefits of data-dependent technology and how it has a transformative effect on their lives. We asked consumers how their lives have changed in the last 10 years and they spontaneously talked about the positive impact of technology; from the functional – managing money and bills online, and travelling with the help of navigation services – to the social – keeping in touch with family and friends through social media. In addition, people see it as empowering them as consumers, increasing choice in retail markets not just through online shopping, but also by allowing them to better research their options, for example through using price comparison websites.

However, while technology is at the forefront of people's minds, consumer data is not. We have undertaken comprehensive qualitative and quantitative research to explore what people know and how they feel about their consumer data being collected and used by commercial organisations.

Our comprehensive research programme consisted of:

- A quantitative telephone survey of a nationally representative sample of 2,064 UK consumers, with a separate boost of an additional 150 interviews in Scotland, between 18 and 28 January 2018. The survey was cognitively tested prior to fieldwork to ensure comprehension. Data was used to develop a segmentation of consumers.
- Six focus groups, lasting two hours, with nine to ten people in each, between 20 and 27 November 2017. Locations were London, Nottingham and Colne, Lancashire. Participants were recruited to ensure a spread of gender, age, ethnicity, self-assessed knowledge about data collection and level of comfort with data collection and sharing.
- 21 face-to-face depth interviews with vulnerable consumers[1] in London, Nottingham, Colne,Newport, Leeds, Perth and St Albans, between 20 and 27 November 2017 and 7 and 27 February 2018.

1 Vulnerable consumers were defined as: older people aged 80 years and over; people belonging to a lower SEG group (DE); people with a long-term physical or mental health condition/ disability; and people who do not feel confident speaking, reading or writing in English

- Four deliberative workshops, each one lasting 1.5 days, between 7 and 27 February 2018. Each deliberative workshop consisted of 24 people and locations were Newport, Leeds, Perth and St Albans. Participants were recruited to ensure a spread of gender, age, ethnicity, self-assessed knowledge about data collection and self-reported confidence online.

The following themes emerged from our research programme:

1. **The data ecosystem is invisible to consumers, limiting their knowledge of it.** This means that their attitudes are mostly formed on only a partial understanding of how their data is collected, what precisely is collected and known about them, and how it's used.
2. **People believe, incorrectly, that data transactions are 'bounded'.** People conceptualise the collection of consumer data as a series of bounded transactions – where individual pieces of data are 'given' to an organisation in order to receive a specific product or service. They are very rarely aware of the extent of third-party sharing or that their data can be amalgamated to form an individual-level profile.
3. **People judge the acceptability of data collection by what impact it has on them.** Without telling people how the use of their data may impact them they do not have the necessary information to assess the acceptability of data collection.
4. **People are primed to 'accept' data collection as having a positive impact, because it is easier to identify and conceptualise benefits than harms.** When informed about data use, people tend to assess acceptability by evaluating whether there is a benefit to their data being collected and whether there could be tangible detriment. However, detriment is often hard for people to conceptualise (other than financial harm); in contrast benefits are easily identifiable.
5. **People are pushed into operating in a space of rational disengagement.** Where the cost of trying to engage (eg understand what data is being collected and attempt to control this) is so much greater than any benefits they receive from doing this, there is little reason for them to do so. It is often perceived that there is little benefit to engagement, because there seems to be a lack of alternatives to the desired product or service.
6. **People feel powerless to engage with organisations who collect and use their data.** There exists a power imbalance between consumers and organisations. This results from: 1) how dependent people have become on technology in their day to day lives; 2) consumers' lack of knowledge about the full extent of data collection and use by organisations; and 3) a lack of alternatives if they want to stop using specific companies whose data collection practices they might be concerned about. This means that people are often left feeling powerless to try and shape their engagement with organisations who collect and use their data.
7. **People want meaningful control over their data:** When people learn about the data ecosystem they tend to feel unable to control what data is collected and how it's used. Sometimes they may want to have direct personal control over their data. However, in some instances (for example when not enough information is given for them to make a decision), the remedy may not be to make it the consumer's responsibility. Instead the action should

be to better control the ecosystem, by ensuring good governance and that when things go wrong (such as breaches) clear accountabilities exist and recompense is given.

We have identified eight specific findings that we believe give insight into **consumers' attitudes and behaviour** towards data collection and use:

1. **Concern about data collection and use varies widely**; it should not be presumed that consumers are either all apathetic or all anxious. Nor is concern fixed, some people's level of concern increases when they find out about the data ecosystem. Our segmentation quantifies how consumers differ in their attitudes to data and their digital behaviours.
2. **Vulnerable consumers are more likely to be concerned about data collection and use**, because they perceive that tangible detriment could result from it.
3. **The majority (81%) of the population are concerned about organisations selling anonymised data to third parties.**
4. **People are often surprised that there isn't more regulation of data collection and use.**
5. **People are often pragmatic about data collection and use**, if they see the relevance or benefit to them.
6. **Attitudes and behaviour are not necessarily congruent.** Our segmentation shows that very different behaviours can underlie the same attitude.
7. **A person's concern about inferences being made and third-party selling does not necessarily translate into taking action to restrict what data can be observed about them.** Our analysis showed that concern about inferences and third-party selling is not a predictor of being 'data restrictive'.
8. **Parts of the population will respond differently to policy recommendations**, depending on their perceived need for change and their willingness to take action themselves.

We have analysed how people *think* about data collection and use, how they *feel* about it, how they *behave*, and how all of this differs depending on the type of consumer. These insights offer policy makers an opportunity to engage with how people may react to policy proposals. For instance, our segmentation can be used as a tool to understand who, and what percentage of the population, may respond positively to a recommended change, and who may fail to engage.

Policy makers can engage with our segmentation on its dashboard (https://consumerinsight.which.co.uk/data-dozen) and explore the various groups and see how they are demographically spread across the population.

Importantly, we have also explored *why* consumers are thinking, feeling and behaving in this way. A combination of an invisible ecosystem, people's cognitive limitations in conceptualising potential detriment and a lack of alternatives create an environment where people are primed to 'accept' their data being collected and used. When considering policy interventions, knowledge of these factors should help to evaluate whether recommendations will succeed.

Having an intelligent and robust understanding of consumers is fundamental to developing successful policy. By adding our insights into the collective evidence base, we hope that it helps policy makers to develop impactful recommendations and facilitate positive change.

TOLERANT 35%

Somewhat more likely to be accepting of personal data collection and use, apart from it being sold to third parties.

CONCERNED 29%

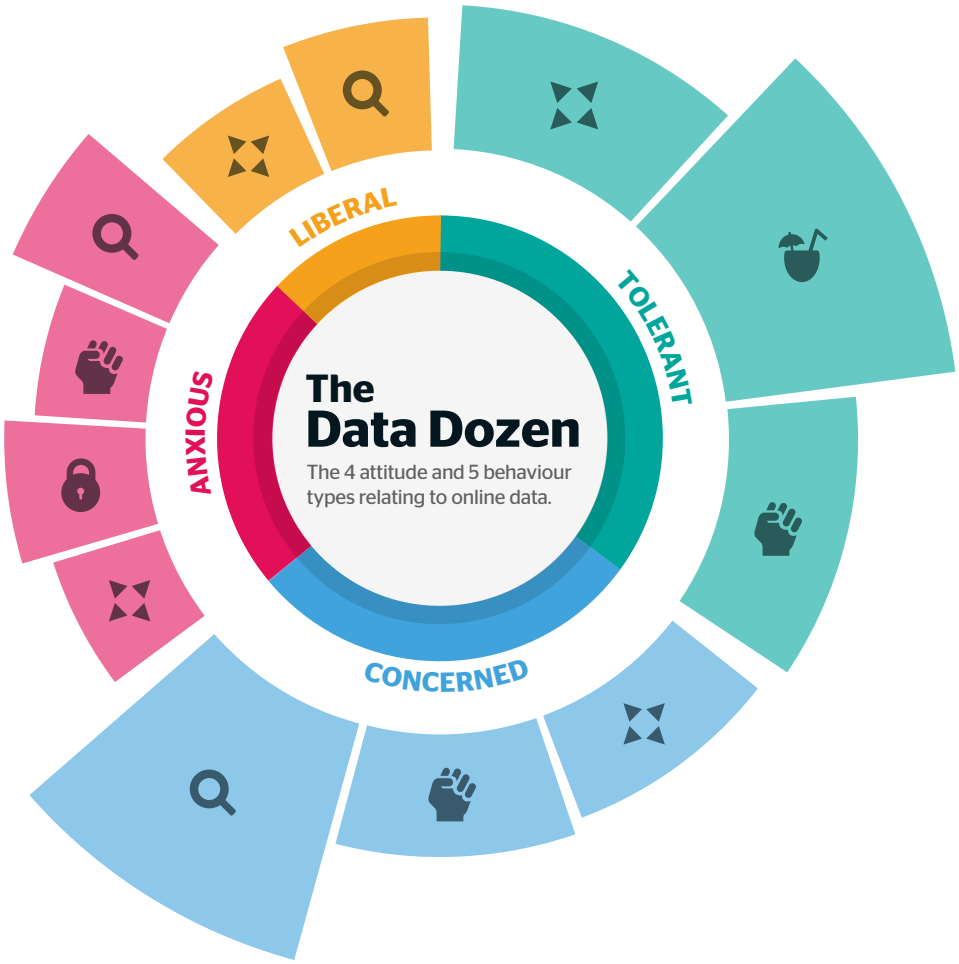
Very worried about having inferences made about them, but most feel confident they know how to control what data they share.

ANXIOUS 23%

Worried about what organisations collect and do with data about them, and less likely to feel confident that they know how to control what they share.

LIBERAL 13%

Not worried about data collection and use, including third party selling of data. They are most likely to say they don't care what organisations do with their data as long as they get what they want.



ACTIVIST
Frequently online, very restrictive with their data.

CASUAL
Online fairly often, unrestrictive with their data.

PROTECTOR
Sometimes online, sometimes restrictive.

BROWSER
Hardly online, unrestrictive with their data.

MAXIMISER
Often use online shortcuts.

For more information visit **which.co.uk/data-dozen**

Which?

Research themes

1. The data ecosystem is invisible to consumers; limiting their knowledge of it. This means that their attitudes are formed on a partial understanding of how their data is collected, what is collected and how it's used.

Our research found that people's insight into uses of their data is related to what they can see: targeted adverts, recommendations and unsolicited marketing (as a result of third-party sharing), which they believe come from known organisations sharing their data. However the scale of the data ecosystem is not visible to them, and they therefore have little awareness of the amount of actors involved and the extent to which they can be profiled and have inferences made about them.

Because consumers perceive the ecosystem to be 'hidden' from them, their view is that it is a purposefully opaque world of data collection and use, and they don't fully understand how they benefit from it. When they learn more about the ecosystem many become more concerned about the collection and use of their data.

> *'The most surprising thing I've found is actually how much information gets taken without your permission. I think we all probably knew it happened, but to get a realisation that from almost every app and every website you use information gets taken. That's surprised me and concerned me a little bit.'*
>
> Workshop participant, Perth

> *'It's just unbelievable how much information sites can get from you without you even realising it.'*
>
> Workshop participant, St Albans

People's perception of the benefit of data collection and use varies across the population, and is dependent on the weight they attach to costs and benefits. We conducted a statistical segmentation[1] and found that there are four attitudinal groups ('Liberal', 'Tolerant', 'Concerned' and 'Anxious') in the population who have different levels of concern about data collection and use.

In our segmentation, 23% of people are in the 'Anxious' group – defined by being less likely to feel comfortable with data collection and more likely to feel concerned about inferences. Along with not feeling confident they know how to control what data they share. This group are more likely than the rest of the population to feel that they didn't benefit from sharing their personal

2 https://which.co.uk/data-dozen. For methodology see appendix.

information with organisations. Some 72% of the group felt this way, compared to less than half (46%) of those who were in the 'Liberal' group (13% of the population) – defined by being more likely to feel comfortable with data collection and less likely to be concerned with inferences and third-party selling.

When 'hidden' aspects of the data ecosystem are revealed to people, the following become evident:

1) People can become more positive about data collection in some instances. When they learn about *why* data is collected they can see that in some cases it is necessary for their product or service to work, or that it has a benefit to them. In this way people sometimes move from a negative perception of data collection to a more positive one, dependent on the context.
2) It can also make people feel more concerned when they learn that there are some outputs which they were not previously aware of, because they are invisible – in particular personalised pricing and personalised information. The 'hidden' nature of personalised pricing and information means that people worry that they could be negatively affected, for example being shown higher prices or limited information, but not know. And their concern is that, if they don't know, how can they do anything about it.

> *'You think you've looked up the best prices, been savvy and that, and you don't even realise what you've not been shown.'*
>
> Workshop participant, Leeds

2. People believe, incorrectly, that data transactions are bounded: where individual pieces of data are 'given' to an organisation in order to receive a specific product or service. They are not aware of the extent of third-party sharing or that their data can be amalgamated to form an individual-level profile.

As discussed, in general consumers are not knowledgeable about data collection or the data ecosystem. Our qualitative research found that, before being given more information on the subject, they tend to perceive data collection as a series of single-bounded transactions, where individual pieces of data are 'given' to an organisation in order to receive a specific product or service.

Within this concept of a bounded transaction there is an acceptance, by some more informed consumers and those in the 'Liberal' segment, that data collection is the 'price' you pay for a free product or service.

However, bounded transactions reflect an incomplete picture of the data ecosystem: consumers have little awareness of what is done with their data, including that these data points can be combined with data collected by other organisations, and that a detailed profile of behaviours and preferences could be potentially made of them.

> *'It [making inferences] feels a bit sneaky! I don't feel as though I have been giving my consent for this to happen.'*
>
> Workshop participant, Perth

In addition, people's default position is to only consider the data that they proactively share, for example information entered in forms (including financial details for online payments) and posts on social media. When prompted to think about other ways in which information may be collected, people recognise that information is observed about them, because they have seen the outputs in the form of targeted adverts and other marketing. However they do not know what information is collected in this way, how it is collected and by whom.

When we explained the data ecosystem to them, including the fact that consumer data flows beyond the bounded transactions they imagined, they are unpleasantly surprised that their data is allowed to change hands continually, that it can be bought, and that this can all be done with profiles, not just individual pieces of data.

> *'I am really surprised by [learning about] data brokers. I didn't even know they did that!'*
>
> Workshop participant, St Albans

> *'The idea that my information, my whole profile…me as a whole rounded individual is sold off… the loss of privacy. I feel quite strongly about being able to remain anonymous and we're losing that.'*
>
> Workshop participant, Perth

3. People judge the acceptability of data collection by what impact it has on them. Without telling people how their data is used and how this may impact them, they do not have the necessary information to assess acceptability.

Without being given information on how data collection may impact them, consumers struggle to determine whether particular examples of data collection are acceptable or not.

By not giving consumers contextual information on use and impact, it is forcing them to make decisions that cannot be meaningful. This is an important point considering that the discourse around data is usually around collection and, at best, generalised use.

> *'I didn't actually mind the personal profile assumptions that they made about me… as long as it just stayed there for a bit of fun. But when [assumptions are] used for different things… like if I wanted to get a mortgage or credit… that would worry me a lot.'*
>
> Workshop participant, Newport

4. Consumers are primed to 'accept' data collection as having a positive impact, because it is easier to identify and conceptualise benefits than harms.

When people learn more about data collection and use they decide whether the practice is acceptable by considering the impact it may have on them – what potential benefits they may receive and what tangible harm could occur.

> *'It's acceptable when it's having a beneficial effect on me. I don't think I'm that bothered… unless it's having a negative impact on me'*
>
> Workshop participant, Leeds

However, their evaluation is skewed because:

- The tangible benefits of data-dependent technology are easier for consumers to bring to mind, because they experience them so often (sometimes daily). They are therefore easy to identify and are meaningful to people. In contrast people find it hard to conceptualise tangible detriment (other than financial), because it is not clear to them what non-financial detriment would be.

> *'There are advantages [to data collection and use], but we don't know the disadvantages.'*
>
> Workshop participant, St Albans

- The benefits that come from data-dependent technology are now a necessary part of life, rather than a 'nice to have'. For example, the benefit of being able to keep in touch with people via social media is now perceived to be a standard part of life, and people feel that if they were to leave social media they would become isolated from friends and family and miss out.

> *'It's just the way it is, the way life's going and you have to conform with it, otherwise you'll be left behind.'*
>
> Workshop participant, St Albans

Therefore, when consumers are informed about how their data is collected and used, they have on the one side direct, known and certain benefits of technology, which have become a necessity in life. And they weigh these up against intangible, unknown potential detriments of having their data collected and used. It is in this context that people appear to be accepting (via their behaviour) of data collection. But in reality, they are primed to accept it because of cognitive biases and a lack of alternatives.

5. People are pushed into operating in a space of rational disengagement. Where the cost of trying to engage (eg understand what data is being collected and attempt to control this) is so much greater than any benefits they receive from doing this, there is little reason for them to do so. It is often perceived that there is little benefit to engagement, because there seems be a lack of alternatives to the product or service that is desired.

Research[3] often says that people are resigned to data collection and use, ie that they ultimately accept the situation, because they feel they can't do anything about it. We believe that a more accurate reflection is that people are pushed into a space of 'rational disengagement' – where the lack of benefit of engaging means that they don't bother to. We feel that this is an important distinction because it indicates that people are not ultimately accepting it; they are forced to accept it as they feel there is no alternative.

3 Ipsos MORI (2016) *Digital footprints: consumer concerns about privacy and security*; Illuminas for Citizens Advice (2016) *Consumer expectations for personal data management in the digital world*; Turiw J, Hennessy M, Draper N (2015) The trade-off fallacy

'I think we live in an age where this is the future. This is only going to get bigger, it's only going to get massive and I just think, unfortunately, unless laws are going to get changed, we're resigned to it.'

Workshop participant, Newport

An example of rational disengagement came from our research when participants spontaneously spoke about terms and conditions. We found that people do not want to read T&Cs because the 'cost' of their time reading and understanding them was too high, and they felt that there was no benefit in doing so as there wasn't anything that could be done if they did not like them. In areas where consumers feel they have little effective choice, clear and concise T&Cs will not be enough.

6. People feel powerless to engage with organisations who collect and use their data, due to their lack of knowledge about data collection and use, dependency on data-driven technology and a lack of alternatives. This informs rationally disengaged behaviour and consumers' evaluation of outcomes.

Our research highlights that consumers feel powerless to engage with organisations who collect and use their data. This affects their behaviour and attitudes.

This disempowerment results from:

- People's lack of knowledge about the full extent of data collection and use by organisations, which is not visible to them. Their perception of 'bounded transactions' means that they don't always think about who else is involved and where their data is going. And they tend to focus on the value of what they are receiving, rather than the value of what they are giving (in the form of their data).

  'I know that… data companies do have a lot of information from us, but I suppose I didn't realise to what extent they have it.'

  Workshop participant, St Albans

- How central data-driven technology has become to people's daily lives and their consumption of products and services, which means that they feel they cannot give it up.

  'For the world to function in today's society I feel that data does have to be given; whether we like it or not it's going to be given.'

  Workshop participant, Newport

- A perceived lack of alternatives if consumers want to stop using specific companies whose data collection practices they might be concerned about. Because technology is so integral to life – affording ease and convenience to leisure and practical activities – even if people do have concerns they feel that they are left with Hobson's choice: either use the service or don't. And, if they don't, they suffer the losses from disengagement. People feel they don't have alternatives to allow them to take action and resolve their concerns, without disconnecting from technology.

'I didn't realise that Google makes assumptions based on what I've looked at already but – what can you do?'

Workshop participant, Leeds

This power imbalance informs rationally disengaged behaviour and consumers' evaluation of outcomes.

7. People want meaningful control over their data.

67% of people, when they have not been informed about data collection methods and use, say they feel confident they know how to control what data they share. However, when people are given information about the data ecosystem, they realise that their data is being collected and used in ways that they weren't aware of. With this realisation often comes concern that they are unable to control their data (both collection and use).

'Sometimes I am worried because I don't know what is out there about me, and I don't know what to do about it.'

Workshop participant, Newport

Consumers also become more concerned once they understand that the data they have given as individual pieces in different contexts can be used in ways they cannot control: 1) accessed by parties whom they feel they didn't consent to; 2) amalgamated to create a profile of them; 3) potentially used to make inferences about information they may not want to be known.

'... it's how personal those assumptions are that I'm most uncomfortable with. They're making assumptions about whether I'm married, whether I'm single... my sexual orientation...'

Workshop participant, Newport

Previous research[4] has identified that consumers don't feel in control of their data. We concur, and offer a nuanced analysis of what this means:

- Sometimes people may want to have direct personal control over their data and, in some cases, this can beneficial. However, we cannot presume that in all cases this is suitable because: 1) the ecosystem is too big and complex for them to keep control; 2) people are unlikely to perceive that the benefit is worth the cost of engaging, because concerns are mostly intangible at the moment and detriment is hard for them to identify; 3) cognitive and behavioural biases, such as hyperbolic discounting and bounded rationality,[5] may limit the effectiveness of many measures.

4 Ipsos MORI (2016) *Digital footprints: consumer concerns about privacy and security*; Illuminas for Citizens Advice (2016) *Consumer expectations for personal data management in the digital world*; Communications Consumer Panel (2011) *Online personal data: the consumer perspective*

5 *Bounded rationality*: consumers are unable to contemplate the multitude of consequences that could result from sharing their data due to cognitive limitations. This leads to decision-making based on heuristics and simplified models. *Hyperbolic discounting*: refers to the idea that people do not discount distant and close events in a consistent way. There is therefore a tendency to trade-off privacy costs and benefits in ways that may be inconsistent with individuals' initial plans, and which ends up "costing" future selves in favour of immediate gratification.

'My data belongs to me, and I should have a complete say about whether it's shared, and who it's shared with.'

Workshop participant, Perth

- When people say they want more control, we believe they often mean there should be more control in the ecosystem, consisting of good governance and clear accountabilities and recompense when things go wrong (such as breaches). This is reflected in the fact that people already think (often wrongly) that there are regulations which protect them, and are concerned when they hear there are not.

'For me the most concerning thing about data collection is that it's not strictly regulated yet. So for financial services and industry… it's very strictly regulated, but the data collection industry… I'm not too sure it is that tightly regulated yet.'

Workshop participant, St Albans

Specific findings

1. Concern about data collection and use varies across the population; we have quantified this in our segmentation.

Consumers' initial reactions to data collection, when given no information about it, are mostly negative. Our quantitative research found that:

- 66% of people would not be comfortable if organisations use information that they have worked out or guessed about them from their observed shopping habits or browsing history.
- 67% of people are not comfortable with organisations using information an individual has shared publicly (for example on social media).
- 68% aren't comfortable with organisations using information they have gathered from observational methods (for example tracking browsing history).

Our segmentation demonstrates that comfort with data collection and level of concern about use does, however, vary across the population:

- 'Tolerant' and 'Liberal' groups are more likely to be accepting of different types of data collection (they make up 35% and 13% of the population respectively), and less concerned about inferences being made.

  'An acceptable amount of data collection is pretty much anything really, I'm really really relaxed with it.'

  Workshop participant, Newport

  'I really don't have a problem with it at all, until the day comes that it actually does harm to me.'

  Workshop participant, Perth

- In contrast there are two groups – 'Anxious' and 'Concerned', who make up 23% and 29% of the population – who are defined by being less comfortable with data collection and more likely to be concerned with inferences being made, and their online behaviour being observed or public posts (eg on social media) being used.

  'I find it's quite concerning that they have all this information about me, a lot of information, information that I probably wouldn't want people to have… That they can see what I'm doing on social media…even though I think that my settings are quite private, they can still get this information.'

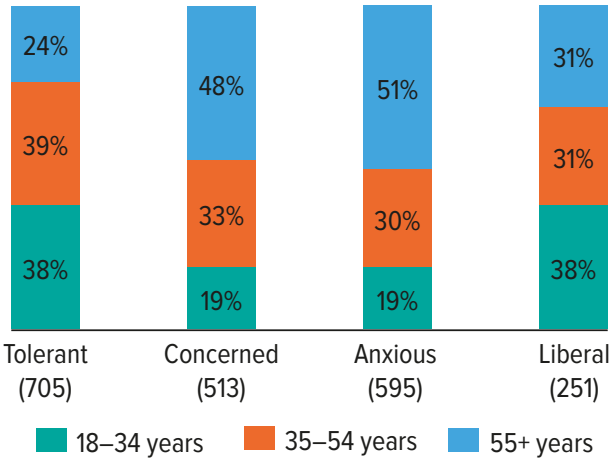
  Workshop participant, St Albans

'I'm quite against it [data collection]. I think it's very invasive and intrusive. I do understand there are benefits of it, but it's a hard sell for me. I like to think I'm quite a private person and data collection is just something I'm not completely au fait with.'

Workshop participant, Newport

It is worth noting that while older adults are generally less likely to be comfortable with data collection, there are some older people for whom this is not the case – 21% of adults aged 55+ years are in the 'Accepting' segment and 11% are in the 'Liberal' segment. And there are some younger adults who are less likely to be comfortable with data collection: a fifth (20%) of 18 to 34 year olds are in the 'Concerned' group and 16% are in the 'Anxious' group.

Age ranges within each segment



2. Vulnerable consumers are more likely to be concerned about data collection and use because they perceive that tangible detriment could result from it.

People are concerned about the lack of control they have over their data and the resulting unknown of what might be done with it and how this might impact them. However they are rarely able to identify tangible detriment, other than that which is financial (eg fraud, stolen card details). In contrast when we interviewed vulnerable consumers we found that they are very concerned about the tangible impact that the sharing of their data could have. This includes that 'irrelevant' data could be used 'against' them – for example stigmatising them based on health conditions and being charged a higher price for a product or service as a result.

'I think they [organisations] can stigmatise you if they know what groups you like or that you are looking for help for your condition.'

Person with a long-term health condition, St Albans

3. The majority (81%) of the population are concerned about organisations selling anonymised data to third parties.

Third-party selling is a concern across the majority of the population and it is seen as a murky and morally dubious practice. Our survey found that 8 in 10 (81%) consumers would be concerned if organisations were selling anonymised

information about an individual to a third party. The only people who are not concerned are those in the 'Liberal' segment (13% of the population).

When we told people about the extent of sharing within the data ecosystem these perceptions are intensified. They frequently feel that:

- They don't have control over where their data goes.
- It is made purposefully difficult for them to opt out of their data being shared.
- Data they consented to give in one context is being used in another, which they wouldn't have given consent for if asked.

> *'When they're sharing your information with companies, you don't know who's getting that information, you have no control over it and the worry is whoever they're sharing it with… what are they doing with it?…We're still in the dark about it, we're not being told…'*
>
> Workshop participant, Leeds

4. People are surprised that there isn't more regulation of data collection and use.

People generally assume that there are regulations against the widespread sharing and use of their data. When we provided them with information about the data ecosystem many were surprised that their data was allowed to change hands so many times, and some assumed that regulations would not allow such practices.

> *'That's surely illegal [sharing data with third parties without consent]… I don't think they are allowed to do that because of data protection!'*
>
> Workshop participant, St Albans

In fact, we found that after the news earlier this year that up to 87m Facebook users data was improperly shared with political consultancy Cambridge Analytica,[6] the top concern of consumers was that there seemed to be little regulation of what Facebook and Cambridge Analytica were doing with people's information.[7] This was true across all segments (although those in the 'Tolerant' segment were jointly concerned about the fact that people's personality and beliefs were given to a third party without their consent).

When people think of the security of their data, they are generally content to put their trust in the organisations collecting their data to keep it secure. There is therefore surprise when people find out that their data had been stolen from 'reputable' providers who hadn't told their customers that they had suffered a data breach.[8]

However, surprise is not necessarily accompanied by concern if the individual has not experienced tangible detriment from the breach. As discussed, the only tangible detriment that people can identify from their data being stolen

6 We define 'more vulnerable' in this research as (i) older people aged 80 and over; (ii) people belonging to a lower socio-economic group (DE consumers); (iii) people with a long-term physical or mental health condition/ disability; and (iv) people who do not feel confident speaking, reading or writing in English.

7 See *Consumers and their data: Research review*, Which? (2018)

8 Populus, on behalf of Which?, interviewed 2,064 UK adults by telephone, between 18-28 January 2018. Data was weighted to be representative of the UK population

is related to their financial data (eg ID theft, fraud). But this is not particularly worrying for them as they believe that banks, payment providers and other financial organisations will reimburse them for any financial losses.

> *'There's a risk in anything you do, but really, what is the worst that could happen?'*
>
> Interviewee, Leeds

5. People are often pragmatic about data collection and use, if they see the relevance or benefit to them.

Our focus groups found that people would switch from talking positively about data-dependent technology and online services, to immediately being negative about data collection. They weren't making the connection between certain data collection being necessary for the technology or the online service to work. Instead they defaulted to a wholly negative view of data collection being part of a murky, morally dubious world.

In our deliberative research we gave examples of what data is collected and how it is used. From this informed perspective, people spontaneously evaluated whether or not it was acceptable based on whether the collection was necessary for the product or service to function and what impact the use of the data had on them, ie whether it led to a benefit and whether it led to tangible harm.

> *'It's acceptable when it's practical and has a benefit for the user, for example, being told that a road you often use is being closed by a satnav, or getting a personalised offer from your bank, or a Fitbit helping you to improve your health.'*
>
> Workshop participant, Leeds

Giving people information on use can work in an organisation's interest, as people tend to start with a negative perception of data collection, assuming that the organisation is collecting it for their own benefit. By providing them with information about what is collected and why, people can understand why it may be necessary and may move to a more accepting position.

> *'If they let you know why they're collecting it, and if they let you know who they're passing it on to, and why they're passing it onto the parties that would be okay. If we were given all the information in the first place then I think it would be acceptable.'*
>
> Workshop participant, Leeds

6. Attitudes and behaviour are not necessarily congruent. Our segmentation shows very different behaviours can underlie the same attitude.

Our segmentation[9] shows that there is a relative lack of a relationship between attitudes and behaviour. It found that the population can be split into four attitudinal groups; however nested within each of these are groups of people who

9 The book *Networks of Control* (Christl and Spiekermann, 2016), (http://crackedlabs.org/en/networksofcontrol), points out that: 'Apparently, hashing is in fact pseudonymisation rather than anonymisation.' In CMO, Adobe's digital marketing magazine, 'leading privacy lawyer' Ruth Boardman suggests that 'marketers should stop trying to convince themselves they are working with anonymised data, rather than personal information'.

may behave very differently, despite holding the same attitudes. For example, in the 'Concerned' segment (people who are very concerned about inferences being made), there are people who display the following behaviours, while still holding the same attitude:

- 'Activist' behaviour: they are taking action to restrict what data can be observed about them, and in addition are more likely to be 'dirtying' their data by putting incorrect information in forms.
- 'Maximiser' behaviour: they are taking advantage of the shortcuts afforded to them online, for example saving their bank details in forms and logging in to other services using their social media.
- 'Browser' behaviour: they are online relatively little or not at all and are less likely to take action to restrict what information can be observed about them.

7. A person's concern about inferences being made and third-party selling does not necessarily translate into them taking action to restrict what data can be observed about them.

Our analysis[10] shows that being concerned about inferences being made or third-party selling are not significant predictors of whether an individual is more likely than average to be 'data restrictive',[11] for example, by clearing their cookies, restricting permissions and checking privacy settings.

Instead the following factors are predictors:

- Time spent online – those who go online for more than five hours a day are 1.7 times more likely to be 'data restrictive', compared to those who are only using the internet for one to two hours a day.
- Going online for leisure: those who are high leisure users[12] are 1.5 times more likely than average to be restricting their data.

In exploring the hypothesis that being comfortable with data collection and feeling in control are a result of taking action to restrict data collection, we found that:

- Those who are more likely than average to be comfortable with data collection methods are 1.5 times more likely than the rest of the population to be 'data restrictive'.
- Those who are more likely than average to say they are confident they know how to control what data they share and what can be seen about them online are 1.6 times more likely than the rest of the population to be 'data restrictive'.

10 Logistic regression analysis was used in determining the likelihood of being above average in taking action to restrict what data can be collected about you. See appendix for full methodology and results

11 'Data restrictive' is defined as being at least one standard deviation higher than average on our summary 'data restrictive' measure derived via Factor Analysis. The survey questions that loaded strongly against this measure were: checking privacy settings on social media and email; clearing browsing history or cookies; and restricting permissions on what information apps and websites can access

12 Leisure use is a summary construct derived from factor analysis of a battery or survey questions relating to consumer data behaviours. The questions that most strongly related to the construct related to reporting using the following in the past three months: consoles or websites for gaming online, streaming services, mobiles apps, social media platforms, messaging services and public wi-fi. High leisure users are those who scored above average on this construct

We also found that those who use online shortcuts more than average (eg allowing a website to remember payment details, logging in through social media) are 1.8 times less likely than the rest of the population to be 'data restrictive'. In our segmentation, this behaviour is associated with being a 'Maximiser' (24% of the population).

8. Parts of the population will respond differently to policy recommendations, depending on their perceived need for change and their willingness to take action themselves.

Our segmentation indicates that policy recommendations which encourage people to change their behaviour will be more successful with some consumers and less with others. For example:

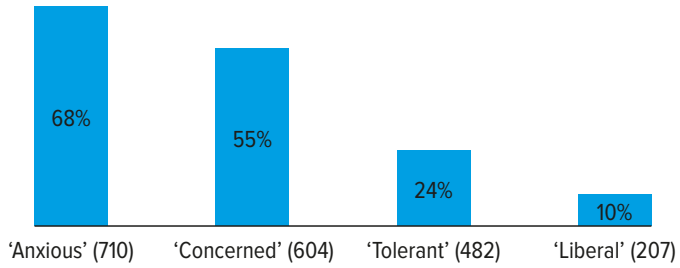
- Those who are in the 'Liberal' attitudinal segment may be most resistant to change their behaviour, as they are not concerned about data collection and use.
- 'Maximisers', who like to use shortcuts afforded to them by technology and are more likely to be early adopters of technology, may weigh up what impact any change may have on how they want to be able to use technology.
- 'Anxious Activists', who are concerned and trying to do as much as possible to control what data is collected, may be the most likely to respond positively to behaviour change recommendations.

There is already some evidence to support this from a general public poll Which? conducted soon after the Cambridge Analytica and Facebook news story.[13] We found that:

- Only 10% of people in the 'Liberal' segment said that, in the light of the Cambridge Analytica and Facebook story, they were concerned about what organisations can do with information about their personality and beliefs. Only 6% said that they got a lot more concerned after the story broke.
- In contrast 68% of those in the 'Anxious' segment said they were very concerned (in the light of the story) and 48% said they became a lot more concerned after the story broke.
- Those who are in the 'Anxious' and 'Concerned' segments are more likely to say they have reduced their use of Facebook since the news story broke. Around a quarter (28% and 26% respectively) said that they were using it less, compared to only around 1 in 10 of those in the 'Tolerant' (10%) and 'Liberal' (13%) segments.

13 Populus, on behalf of Which?, surveyed 2068 UK adults online between 26 and 27 March 2018. The data has been weighted to be demographically representative of the UK population

Percentage of each segment who say that, in light of the Cambridge Analytica/Facebook news story, they are very concerned about what organisations can do with information about their personality and beliefs



However, as discussed, those exhibiting the same attitudes don't always demonstrate the same behaviour. For example those who are in the attitudinal segment 'Anxious' vary in their response to the news story; those who are 'Activists' tend to be more likely to decrease the number of shares (38%) compared to those in the 'Maximiser' (24%), 'Unprotected' (21%) and 'Browser' (13%) segments. It therefore cannot be presumed that people with the same attitude will all take action to the same extent.

Appendix

Consumer data segmentation method

We adopted a data-led hierarchical clustering method for this segmentation. It was conducted in two stages to first derive high-level clusters based purely upon attitudes, then a second stage based upon the behaviour related questions. The second-order behavioural clusters were nested within the initial attitudinal clusters.

We used hierarchical clustering to determine the optimal number of clusters, then k-means clustering to determine cluster membership. The inputs to the clustering were summary constructs created via two factor analyses of batteries of survey questions. These resulted in summary 6 measures from 23 questions relating to attitudes and summary 3 measures of behaviours from 12 survey questions, explaining 55% and 47% of the variance in the data respectively.

The survey questions which loaded most strongly on the constructs are shown in the following tables 1 and 2.

Table 1: Attitude constructs and strongest related survey questions

How concerned would you be about inferences?	• How concerned would you be if organisations can categorise someone based on age, income and gender? • How concerned would you be if organisations can predict an individual's personality? • How concerned would you be if organisations can predict an individual's lifestyle?
How concerned would you be about organisations selling information?	• How concerned would you be if organisations can sell anonymised information about an individual to a third party? • How concerned would you be if organisations can sell information to third parties about an individual which hasn't been anonymised?
Feel they don't benefit from data collection	• I am very cautious about sharing my information with other people and organisations (agree/disagree) • I don't like seeing adverts on websites and social media based on my previous searches (agree/disagree) • I don't benefit from sharing my personal information with organisations (agree/disagree)
Don't care about data I share	• I don't care if people see what I post on the internet (agree/disagree) • I don't care what organisations do with the information I share with them, as long as I get what I want (agree/disagree)
Comfort with data collection methods	• How concerned would you be if organisations can categorise someone based on age, income and gender? • How concerned would you be if organisations can predict an individual's personality? • How concerned would you be if organisations can predict an individual's lifestyle?
Feel they have personal control of their data	• I feel confident that I know how to control what information I do and don't share with organisations (agree/disagree) • I feel in control of what is seen about me on the internet (agree/disagree)

Table 2: Behaviour constructs and strongest related survey questions

Online shortcuts	• How frequently do you allow a website to automatically save and remember your payment card details? • How frequently do you share your location on social media platforms such as Facebook or Twitter? • How frequently do you use your social media account to log in to other services?
Dirtying data	• How frequently do you deliberately give incorrect information on a form, such as an email address, date of birth, or phone number? • How frequently do you use a different email account for websites which you don't want to receive communication from?
Data restriction	• How frequently do you check your privacy settings on social media and email platforms? • How frequently do you restrict permissions on what information apps and websites can access on your device? • How frequently do you clear browsing history or cookies?

The attitudinal clustering derived four segments ('Tolerant', 'Concerned', 'Liberal', 'Anxious') and within each of these, the second level of clustering for behaviours derived 12 sub-segments that we grouped together post-hoc into five types, based on our own assessment of similarity in the behavioural construct statistics for those groups. We named these 'Maximiser', 'Protector', 'Casual', 'Activist' and 'Browser'.

Logistic regression method: Drivers of 'data restrictive'

'Data restrictive' is one of the summary constructs relating to consumer data behaviours used in our segmentation. It was derived from factor analysis of a battery of 12 survey questions. The questions that most strongly related to this construct are given in table 3 below.

Table 3: Strongest related survey questions to data restriction

Data restriction	• How frequently do you check your privacy settings on social media and email platforms? • How frequently do you restrict permissions on what information apps and websites can access on your device? • How frequently do you clear browsing history or cookies?

We conducted a logistic regression to see which independent measures in our survey most strongly related to being above average for 'data restriction'. For the purposes of this analysis those that scored more than one standard deviation higher than average for data restriction were classed as 'high data restriction', and all others comprised the rest. We restricted the sample to just the 1,884 subjects that reported to having used the internet each day.

The table below summarises the logistic regression results for variables predicting above average in data restrictiveness.

Predictor	β	**S.E.** β	**e** β **(Odds Ratio/ Liklihood)**	**Significance (p value)**
Age (for each year increase)	-0.01	0.01	0.99	*
Gender: Male (compared to Female)	-0.07	0.15	0.94	
Social Grade E (comparison group)				
Social Grade A	0.17	0.37	1.18	
Social Grade B	0.46	0.31	1.58	
Social Grade C1	0.20	0.30	1.22	
Social Grade C2	0.53	0.31	1.69	.
Social Grade D	0.03	0.36	1.03	
Time online: More than 5 hours (comparison group)				
Time online: Less than an hour	-1.27	0.34	0.28	**
Time online: 1 to 2 hours	-0.53	0.23	0.59	*
Time online: 2 to 3 hours	-0.37	0.22	0.69	.
Time online: More than 5 hours	-0.15	0.20	0.87	
Leisure net use – Above average compared to rest	0.38	0.22	1.47	.
Practical net use – Above average compared to rest	0.04	0.18	1.04	
Smartphone/portablePC – Above average compared to rest	0.13	0.22	1.14	
Blackbox/digTV/games – Above average compared to rest	0.08	0.23	1.09	
Loyaltycard/TPS/desktopPC – Above average compared to rest	0.63	0.19	1.88	***
Knowledge score – Above average compared to rest	-0.29	0.32	0.75	
Service use score – Above average compared to rest	0.07	0.20	1.07	
Concern Inference – Above average compared to rest	0.28	0.21	1.32	
Concern Sell info – Above average compared to rest	-0.05	0.36	0.95	
Confidence Control – Above average compared to rest	0.45	0.23	1.57	.
Comfort data collection Methods – Above average compared to rest	0.39	0.18	1.48	*
Don't Benefit – Above average compared to rest	0.14	0.35	1.15	
Don' t care – Above average compared to rest	-0.51	0.36	0.60	
Dirtying Data – Above average compared to rest	0.71	0.19	2.03	***
Data shortcuts – Above average compared to rest	-0.61	0.24	0.54	**
Constant	-1.70	0.01	0.99	***

. p < 0.1, *p < 0.05 **p < 0.01 ***p < 0.001

β: Coefficients for predicting the dependent variable from the independent variable, in log-odds units.

S.E. β: The standard errors associated with the coefficients e β: Odds ratio for the predictors Nagelkerke R Square: 0.121

BritainThinks full research report

Britainthinks
Insight & Strategy

1. Key Insights

Spontaneously, data collection is not a top-of-mind concern for most consumers, who instead focus on the benefits of using products and services that collect information about them.

- Overall, these types of products and services are perceived to make consumers' lives easier and increase consumer choice.
- Few consumers feel that they have experienced any detriment as a result of using these types of products and services to date, and most struggle to see how data collection and sharing could cause them any direct, tangible harm in future.

Quantitative research demonstrates that starting perceptions of data collection, and the extent to which consumers feel in control of their data, varies significantly across the population.

- The public groups into four attitudinal segments: just over half of the population are "Concerned" or "Anxious" about organisations collecting and using their data, while the remainder is either "Tolerant" or "Liberal".
- Qualitatively, even those who are more concerned about data collection and sharing often feel that there is little they can do about them. For the majority, these practices are perceived to be causing too little direct harm to warrant sacrificing convenience and access to services they are accustomed to using for 'free'.

Most consumers are operating with an incomplete picture of what data is being collected about them and what happens to this information, and are basing their view of what is and isn't 'acceptable' in relation to data collection on relatively limited knowledge.

- Most realise that they are 'giving' away some form of data about themselves, but awareness tends to be limited to specific transactions with products and services.
- False assumptions and myths are rife, including the commonly-held belief that devices such as smart TVs, Siri and Alexa are actively 'listening in' and recording consumers' whole conversations.

When consumers learn more about data collection and their 'picture' of data collection becomes more complete, their levels of concern tend to grow. There are a series of key 'penny drop' moments that have a major impact on their perceptions of the issue:

- Learning that organisations are not 'just' collecting bounded or discrete pieces of information about consumers, but aggregating their data to create detailed, individual profiles.
- Realising the extent of inferences being made about them based on their information, which they fear may be incorrect or reductive.
- Understanding the extent to which third party sharing takes place, and the role of data brokers in the data ecosystem. There is particular surprise about the extent to which these types of organisations are able to monetise consumer data.
- Learning that in addition to advertising and recommendations, information and pricing is personalised and targeted based on consumers' information.

For the majority of consumers, data collection becomes unacceptable when one or more of four things happen:

- When they feel that what happens to their data is out of their control, such as when they are not fully aware of what is happening to their information or they don't have the opportunity to opt out.
- When the information collected about them doesn't feel relevant to the specific context and it is unclear how the data collected is required for a product or service to function. Or when data is seemingly collected for one purpose (e.g. social media) but is used for another which is considered very different (e.g. determining prices).
- When there is no tangible benefit either to the consumer or to society of this information being collected and used. This benefit might be the very service or product that data sharing 'allows' them to access (e.g. accessing 'free' public WiFi by sharing their email address), or a benefit that is intrinsic to sharing their data (e.g. being able to monitor their health and fitness by sharing health data).
- When there is a risk of tangible harm coming either to the consumer or to other groups in society. These harms often only 'materialise' for consumers once they become aware of the full complexity of the data sharing ecosystem. Some, but by no means all, consumers place a high premium on protection of their privacy to avoid these harms depending on their personal views and circumstances.

Even from a more informed position, consumers often feel that they have too little power to take action themselves, that the power imbalance is 'weighted against them', and their sense of resignation about data collection endures.

- There is a strong sense that the horse has already bolted and that it is too late to resume control once their data is already 'out there'.
- Actions to change permissions and privacy settings are often seen as surprisingly easy to enact, but consumers feel that there is no guarantee that

organisations will not find workarounds or continue to share their data with others.

Consumers expect to see government, regulators and consumer bodies working on their behalf to hold organisations collecting their data to account, and to make the first move in 'breaking the stalemate'.

- They would like to see government, regulators and consumer bodies taking action to ensure that: consumers are informed and in control of what data they are sharing and how this is being used; the data being collected is relevant to the context in which it is being gathered; that there is some tangible benefit of collecting and sharing this data, either to the individual consumer or to society more generally; and that these practices do not cause consumers any direct, tangible harm.

2. Background, objectives and methodology

2.1: Research background, aims and objectives

To respond to the increasing prevalence of data collection and usage, Which? has conducted a large-scale policy review to understand the following:

1. Do consumers understand how their data is collected, traded and used, and how this may affect their choices?
2. Once they understand more about how their data is collected and used, is it possible for consumers to take more control of their information?
3. How far will upcoming policy changes rebalance data collection in consumers' favour?

As part of this review, Which? commissioned BritainThinks to conduct primary research with consumers to understand their awareness and perceptions of data collection and the potential impacts that it has on consumption.

Specifically, the consumer research aimed to explore the following themes:

Consumers' understanding of what data is collected about them

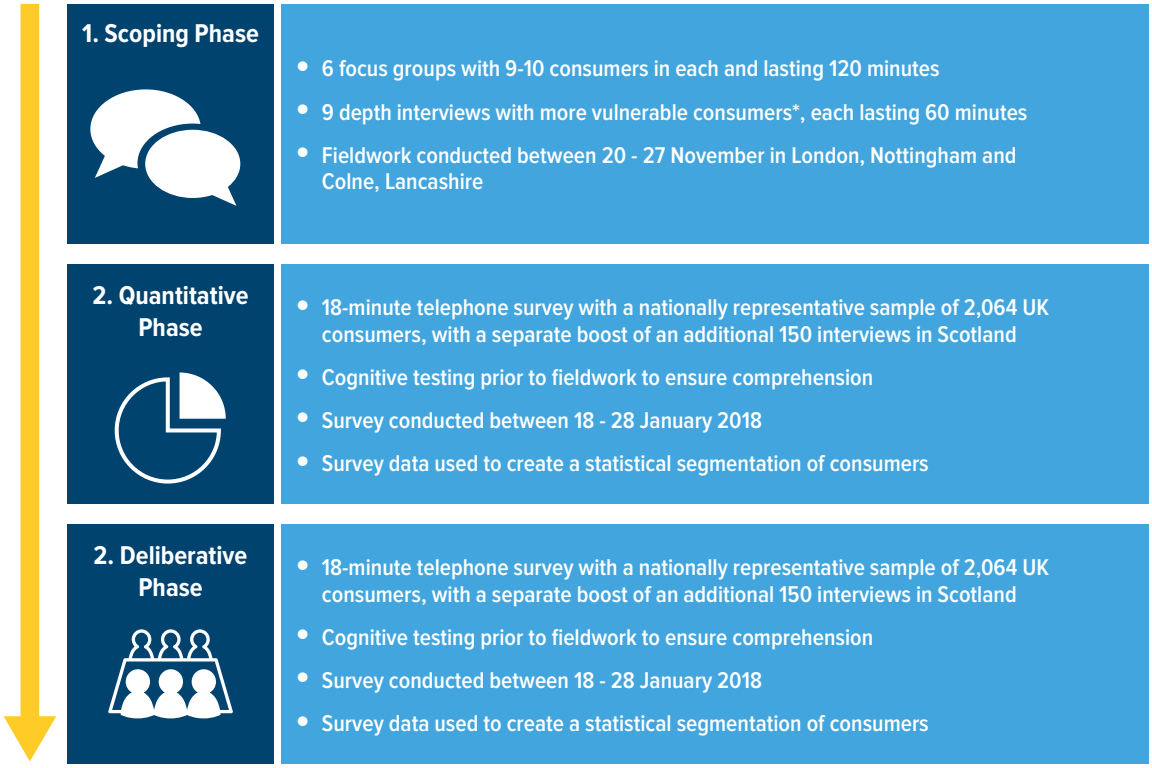
- **Data collection:** Are consumers aware of the types of data collected and the types of organisation involved? How much does this bother them?
- **Data Sharing:** Are consumers aware that their data is traded and how do they feel about it? Are they aware of and do they trust dominant brands in this space? Do they know how to opt out?
- **Stolen Data:** How seriously do consumers take the risk of their data being stolen? Do they know how to find out if this has happened? How do they feel when made aware of breached?

Understanding of the potential impacts of data collection on consumption

- **Information Targeting:** How aware are consumers that they are being targeted in different information channels? Do they value this or are they bothered by it? Do they feel that targeting limits their access to information?
- **Pricing, Offers and Exclusions:** How aware are consumers of personalised pricing, offers and exclusions based on data held about them? What value judgements do they make about it? How aware are they that their 'digital reputation' can affect the offers they receive and products they can access to information?
- **Individual Data and Consumer Choice:** Do consumers make the connection between 'free' services and their data being sold to advertisers and brokers? Do they mind? Are they aware of the innovation that comes from the use of individual data?

2.2: Research methodology

The consumer research comprised three phases:



**Please note that we defined 'more vulnerable' consumers in four different ways in this research:*

1. *Older consumers, aged 80 and over.*
2. *Consumers belonging to a lower socio-economic group (DE consumers).*
3. *Consumers with a long-term physical or mental health condition/ disability.*
4. *Consumers who do not feel confident speaking, reading or writing in English.*

This report combines the findings from all three phases of the project, showing consumers' spontaneous perceptions of data collection and the extent to which these change as they are provided with more information about the issue, as in the deliberative phase.

In the deliberative phase of the project, participants discussed and were provided with more information about the following themes.

Further detail about this phase of the project, including the sample frame and the information participants were provided with is available in the appendix.

SESSION	OUTLINE
The role of technology	Participants shared the technology, websites and apps that they use the most.
Data collection: what is being collected	Participants were provided with examples of everyday products and services, and the consumer data that they collect.
Data collection: how data is observed	Participants were provided with information about and discussed cookies, fingerprinting, and a demonstration of Ghostery, a web browser extension which allows users to identify (and block) trackers on websites.
What happens to this data: inferring and building an individual profile	Participants were provided with information about and discussed inferences, psychographic profiling, data being combined to make an individual profile, and the idea of the 'digital self'.
What happens to this data: third party sharing	Participants were provided with information about and discussed the data sharing ecosystem, including the role of brokers.
Security of information and stolen data	Participants discussed how likely or not they felt it is that their information could be stolen, and how much of a concern this is.
Tailoring and targeting of adverts, recommendations, prices and information	Participants were presented with examples of and discussed targeted adverts, recommendations, personalised pricing, and tailored information. They also discussed the relative acceptability of different forms of targeting and considered the extent to which these practices have positive and negative impacts on consumer choice.
Choice in markets and innovation	Participants considered the extent to which the amount of choice available to them as consumers has changed over the past decade, and were provided with information about the companies/apps owned by Google, Facebook and Amazon.

Throughout the deliberative fieldwork, participants also weighed up the following trade-off questions based on options which are currently available to consumers, or which have been proposed by organisations but not (yet) taken forward:

1. It is possible to turn off all trackers (using software such as Ghostery) and/or targeted advertising (using Google and Facebook settings). Would you choose to or not? What if this meant no longer being able to store passwords, or use the 'autofill' function when competing forms?

2. If third party sharing and profiling didn't happen, it is possible that you would need to pay for your email services and social media. Would you prefer to pay for these services than have your data shared?

3. You could get cheaper insurance but to do so you would have to allow insurance providers to access your Facebook. Would you do this? (Please note that this trade-off question is based on a proposal from an insurance company which was not taken forward.)

4. The professional social network LinkedIn (owned by Microsoft) is the largest of its kind and far ahead of its competition. Last year a judge in the United States ruled that it must allow a third party company to 'scrape' data publically posted by users, allowing it to compete. LinkedIn said that it was 'disappointed in the court's ruling', and that they would continue to fight to protect their members' ability to control the information they make available on LinkedIn. Was the judge right or wrong?

5. Companies are using individual data to develop a range of innovations, but risks of privacy breaches might increase. Is the balance too far towards innovation or too far towards privacy?

Illustrations in this report have been produced by Sally Pring (spring-boards.co.uk).

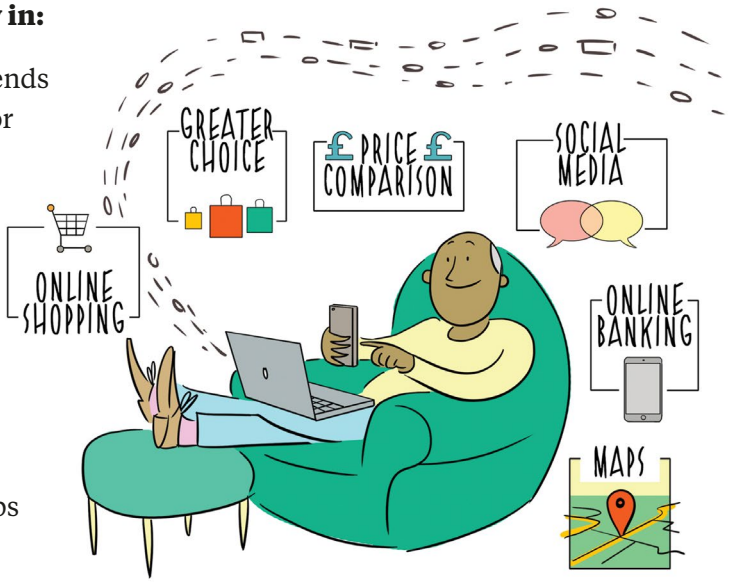
3. Understanding consumers' mind-sets

3.1: Consumers' starting points on data collection

Data collection is rarely a top of mind concern for consumers and is an issue that most, spontaneously, have little awareness and understanding of. When thinking about data-dependent devices and services, consumers rarely mention concerns about the actual data they are collecting, and instead tend to focus on the benefits these devices and services confer. The benefits consumers describe spontaneously can be grouped into two key themes:

1. Greater convenience, particularly in:

- Communicating with family and friends through social media, particularly for those with family overseas.
- Managing money, bills and utilities through online and mobile banking.
- Shopping for products and services through online shopping.
- Getting around, through navigation services such as SatNavs, online maps and route planners.



2. Increased consumer choice, particularly in:

- Shopping, not just through online shopping websites and apps, but also related services including price comparison websites and apps. Consumers also describe doing their own research on the best prices and products by shopping around online.
- Online entertainment, with services such as Netflix meaning that consumers have access to a wider range of films and programmes available to them to watch, anytime.

"I feel like I have so much more choice – you can shop around, you can really research things before you buy them."
Workshop participant, Perth

"I can't remember the last time I went into a bank for example. Things like mobile banking make life so much quicker and easier."
Workshop participant, Leeds

When consumers do talk spontaneously about some of the downsides of technology, innovation and data-dependent devices and services, they are less likely to focus on concerns related to their own data, and more likely to focus on abstract, societal concerns affecting other people. These include concerns about an over-reliance on technology in society and on digital rather than face-to-face or verbal forms of communication. In particular, they have concerns about the impact that these drawbacks might have on younger generations.

> *"I think it has caused a disconnect in society…relationships are now more shallow."*
>
> Scoping Phase participant, London

> *"Sometimes it feels a bit like I'm addicted to my phone and that I can't function without it. If I leave it at home it's like I've lost a limb or something."*
>
> Workshop participant, Perth

> *"The younger generation don't know how to look anything up for themselves. In my day you had to go to a library."*
>
> Workshop participant, Leeds

Most consumers struggle to think of specific negative impacts of data collection. For many, the most front-of-mind impact is the volume of marketing emails and phone calls they receive after sharing their email address or phone number (which most attribute to the sharing of data, rather than its collection). They also talk readily about their data being shared with 'third parties'. However, very few make the connection between these outcomes and the issue of 'data collection' spontaneously.

> *"If you put your email address into anything you just immediately get bombarded by emails."*
>
> Workshop participant, St Albans

> *"I get lots of email rubbish, and how have they got my email? I know they shared it with someone else."*
>
> Workshop participant, St Albans

In qualitative research, when consumers engage with the issue of data collection, rather than being something that they are actively concerned about, data collection is instead seen as a 'fact of (modern) life' that they have to engage with, for four key reasons:

1. The benefits of technology and data-dependent devices and services are usually seen to outweight the negatives - and, in many cases, it's hard to find an alternative
2. Most can see (at least some) legitimate reasons for data collection and believe that this has always happened in one form or another
3. Consumers are often working with an incomplete understanding of how their data is collected and used
4. Data collection is simply 'how it is', and something that they struggle to imagine ever changing

"I think it is just one of those things, you know you have to give up that information."

Workshop participant, St Albans

"It's difficult to live your life without sharing your data. You can go to Pret and buy a coffee and if you use your card then they know that you have bought it. There's not really any way of stopping it unless you want to stay at home all the time."

Scoping Phase participant, London

"These are not new things that are happening, years before the internet we had the census forms – you could tick or fill in everything you wanted to, but we had a choice of filling information in. Businesses always want information, this is now just an automated way of doing it."

Workshop participant, St Albans

The perspective of vulnerable consumers: more vulnerable consumers tend to have a more overtly negative starting point when considering data collection. These consumers (e.g. older consumers) may not be participating in technology to the same extent as 'mainstream' consumers and so may not be enjoying the benefits. In addition, they may feel greater distress at some of the visible impacts of data collection, with the volume of marketing emails they receive often feeling invasive or unmanageable.

"I don't really use much technology…no. I use my pad [tablet] for playing Scrabble but that's it really."

Deliberative Phase participant, older consumer, Perth

"They just want to send you more emails, it starts to get a lot and you're having to delete them constantly."

Scoping Phase participant, consumer with a long-term health condition, London

3.2: Segmenting consumers across the population

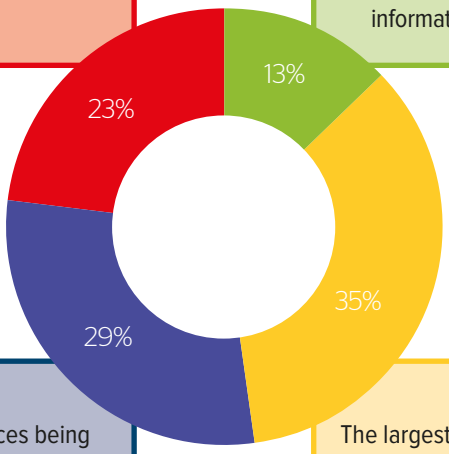
Levels of concern about data collection and the depth of resignation about this issue vary across the population. Consumers' attitudes towards data collection, before receiving information about the issue, can be split into four segments:

ANXIOUS

They are somewhat uncomfotable with data collection and concerned about inferences being made. Most likely group to not feel confident that they know how to control what data they share with organisations. In addition they are the only group where the majority don't trust organisations to not share their data if they don't give them permission.

LIBERAL

The smallest group in the population. They are similar to the tolerant group in that they are more likely to be comfortable with data collection and less likely to be concerned about inference being made. However, in addition they aren't concerned about third party selling (in contrast to other groups) and they are most likely to say that they don't care if people see what they post online and don't care what organisations do with their information as long as they get what they want



CONCERNED

They are very concerned about inferences being made. Nearly all say that they're very cautious about sharing their information; however, the vast majority are confident that they know how to control what data they share with organisations. More likely to believe they don't benefit from sharing their personal information with organisations.

TOLERANT

The largest group in the population. They are a little more comfortable about data collection than average and less likely to be concerned about inferences being made. However, they are concerned about third party selling.

Importantly, attitudes towards data collection do not necessarily dictate how consumers behave with regards to protecting their data. Within each attitudinal segment there are a variety of different behavioural characteristics present to varying extents:

1. **"Maximisers" (24%)**: This group likes to use shortcuts such as logging into other services through their social media and saving their details. They are also higher than average on taking action to restrict what data can be observed about them.
2. **"Casual" (16%)**: This group is lower than average on trying to restrict what data can be observed about them, and are online an average amount.
3. **"Protector" (8%)**: The smallest group and only seen within the "Anxious" attitudinal segment. They are somewhat more likely than average on taking action to restrict what data can be observed about them.
4. **"Activist" (19%)**: This group frequently takes action to restrict what data can be observed about them, and in addition are more likely to be "dirtying" their data by putting incorrect information in forms.
5. **"Browser" (33%)**: This group include people who are online relatively little or not at all. They are significantly less likely to take action to restrict what is observed about them.

This means that individuals who share the same broad outlook towards data collection may be arriving at these attitudes for very different reasons and based on very different experiences.

Below we set out each segment (and behavioural characteristics within each segment) in further detail, starting with the members of society who are least concerned about data collection and use, to those who are most concerned:

"LIBERAL" – 13% OF THE POPULATION

Just over one-in-ten members of the public are 'Liberals'. These consumers tend to feel in control of their data and are unconcerned about the potential uses of their data and the selling and sharing of data with third parties. These consumers are more likely to be male than female (64% versus 36%) and have an average age of 48, making them younger than the concerned segments.

Within this segment there are **two behavioural groups**: "**Browsers**" and "**Maxmimsers**".

- Just under two-thirds (62%) are "**Browsers**". They are skewed towards an older age group (49% are aged 55+) and spend little time online compared to the average consumer. They are less likely to use social media and apps than the average person.

- Qualitatively, these "Liberals" tend to feel in control of their data as they are rarely online. They feel that little of their data is 'out there' and therefore that data collection doesn't affect them personally.

- Meet **Margaret, 78, from Perth**: she retired ten years ago and lives alone. She is not 'online' and doesn't own a computer or a smartphone as she doesn't feel she has any use for them. She buys groceries from her local corner shop and supermarket, using her loyalty card and debit card to make purchases. Because she doesn't use the internet, Margaret doesn't think that organisations are collecting information about her: she feels this isn't relevant to people in her situation.

"I don't go on any of that, no…it's all the young people these days who are using computers who need to watch out."

Browser Liberal, Perth

- The remaining two-fifths (38%) of "Liberals" are "**Maximisers**". They are skewed toward a younger age group (59% are aged 18-34) and spend an above average amount of time online. They are more likely to be using shortcuts afforded to them online, such as logging in to services through their social media, and are more likely to say they like to upgrade to new technology when it comes out.

  - In focus groups and deliberative workshops, "**Liberals**" who display "**Maximiser**" behaviour are much more likely to describe the selling of their data to third parties and other uses of their data as the 'price' of using free services and to say that they feel unconcerned about them.

  - Meet **Michael, 27, from Leeds**: he is a busy young professional who values convenient access to his favourite apps over putting excessive restrictions on his data. He says that he would find it a huge hassle if he didn't let apps and websites save his passwords and his payment details, and he says that it is 'fair game' that he is 'paying' for this convenience with his data.

"I love the fact that I can just go on the app, look at what I want, and buy it – it's three taps and I have bought a new shirt."

Maximiser Liberal, Leeds

"It's the way it is though, right? They have to know what I am doing on the app to make it better, so it doesn't bother me."

Maximiser Liberal, Leeds

"TOLERANT" – 35% OF THE POPULATION

Just over a third of the population fit into the "Tolerant" segment. These consumers are more likely to feel in control of their data and comfortable with inferences than the average consumer, but have high levels of concern about third party data-sharing in line with the majority of the population (excluding the "Liberal" segment). These consumers are almost evenly split between men and women (52% male versus 48% female) and have an average age of 44, making them the youngest of the four segments.

Within the "Tolerant" segment there are **three different behavioural groups**: "Casuals", "Activists" and "Maximisers".

- Almost half (46%) of 'Tolerant" consumers are "**Casual**". The majority of this group is aged 35+ years and go online for an average amount of time, both for "practical" uses such as online banking, as well as "leisure use" such as social media. They are less likely to be taking action to restrict what data can be observed about them, for example restricting permissions on apps and clearing cookies.
- Just under a third (29%) are "**Activist**" – this group is mostly aged between 18 years and 54 years and going online for an average amount of time. Unlike the "Casual" group, this group are more likely than average to be taking action to restrict what data can be observed about them, for example app users are restricting permissions. In addition, they are more likely to "dirty" their data by putting incorrect information in a form, and to try and control communications from organisations by using different email accounts for those they don't want to receive emails from.
    - Qualitatively, those who are "**Activist**" tend to be better informed than other consumers about data practices overall and consequently feel equipped and aware of actions they can take to protect their data, giving them a sense of control and comfort about data collection.
    - Meet **Linford, 47, from Leeds**: he is a mechanic and lives with his wife and two teenage children. He feels like he is in control of his data and able to protect himself from being included on marketing lists and getting 'spammed' as a result. He says that one of his 'top tricks' is creating a separate email address to use for accounts he doesn't really care about and when accessing free services, such as public WiFi. He uses a false name on these account and has given an incorrect date of birth.
- Just over a quarter (26%) are "**Maximiser**" consumers – this group is skewed towards the youngest age group (18-34 years). This groups spends the most time online, particularly going online for leisure: nearly everyone is using social media and messaging services. They are more likely to use the shortcuts afforded to them online, for example logging in to services through their social media, and they are more likely to say they like to

upgrade to new technology when it comes out. Their behaviour doesn't show them to particularly be trying to restrict what information can be observed about them; they are no more likely than average to clear their cookies and check their privacy settings on social media, although app users are more likely to restrict permissions than the average social media user.

    - Meet **Maria, 36 from St Albans**: she is a busy mum of two, and uses the internet more or less everything, ranging from online banking to streaming TV. She really likes the recommendations she receives on Netflix as she can always find something to watch within minutes of getting some time to herself. Maria says that she has a 'million and one things to think about' and that any concerns about her data are not top of her priority list.

"I don't really know why they would want my information, I'm just a regular mum!"

Tolerant Maximiser, St Albans

"CONCERNED" – 29% OF THE POPULATION

Just under a third of the population fit into the "Concerned" segment. These consumers generally feel in control of their data, but are concerned about some of the potential uses of it and in particular are more likely than the wider public to feel uncomfortable about organisations making inferences about their personality and lifestyle. These consumers are more likely to be female than male (58% versus 42%) and have an average age of 55.

Within this segment there are **three different behavioural groups**: "Browsers", "Activists" and "Maximisers".

- Over half (55%) are "**Browsers**". Skewed towards the older age group (63% are aged 55+ years), this group spends very little time online compared to the average consumer. They are particularly less likely to use social media and apps than the average person. However when they are online, they are less likely to be taking action to restrict what data can be observed about them, for example by clearing cookies and restricting permissions.

    - Qualitatively, the sense of control among those who are "**Browsers**" is often derived from their being offline, meaning that they feel little of their data is "out there".

- Just under a quarter (24%) are "**Activists**" – this group is evenly spread across age groups. They are spending an above average amount of time online and are using messaging services, social media and apps. They are more likely to upgrade to new technology. They are the group with the highest percentage of app and social media users taking control of what data they share through their settings. They are also more likely to be

'dirtying' their data by providing incorrect information on forms and more likely to use multiple email addresses to control what communications that they see from organisations.

  - In focus groups and deliberative workshops, consumers in this segment tended to feel in control because they are taking proactive steps to protect their data.

- Just over a fifth (21%) are "**Maximiser**" consumers – this group is spread across age groups. They spend an above average amount of time online and are going online equally for leisure and practical use. They are more likely to say they like to upgrade to new technology when it comes out and they like to use the shortcuts afforded to them online, for example logging on to other services through social media.

  - Qualitatively, this group often feel in control because of low levels of what data is being collected about them and how this data is being used. On learning more about this in focus groups and workshops, their levels of concern often grew, leaving them feeling less confident and in control.

  - Meet **Diane, 56, from Leeds**: she has a big family and is always on her phone, catching up with friends and sharing photos of her family on Facebook. Diane feels relatively in control of her data because she restricts her Facebook privacy settings so that only people she knows can see her posts, but she feels less comfortable when companies seem to suggest things she might like to buy and doesn't know how to stop this happening.

"I am always careful with what I post on Facebook, as I don't want people to know if I am on holiday…it's just annoying when they suggest things, it's like they know me."

Concerned Maximiser, Leeds

"ANXIOUS" – 23% OF THE POPULATION

Just under a quarter of the population fall into the "Anxious" segment. Similarly to the "Concerned" segment, these consumers are concerned about the potential uses of their data, and in particular inferences being made about them. However, unlike the "Concerned" segment, this group are much less likely to feel in control of their data. These consumers are more likely to be female than male (56% versus 44%), and have an average age of 56, making them the oldest of the four segments.

Within this segment there are **four different behavioural groups**: "Browsers", "Protectors", "Activists" and "Maximisers".

- Just under two-fifths (39%) are "**Browsers**" – the vast majority of this group are older adults (74% are 55+ years old) who spend little time online compared to the average consumer. Only half have a smartphone and only a quarter are on social media. When they are online, their behaviour seems to indicate they are less likely to be taking action to restrict what data can be observed about them.

- Just over a third (35%) are "**Protectors**" – the majority of this group are aged 35 and over. They are spending an average amount of time online and their use of social media and messaging services is consistent with the general population, though they are less likely to be using apps. Although they aren't confident that they know how to control what information they share, their behaviour shows they are more likely than average to take some restrictive actions, for example clearing their cookies and restricting permissions on apps if they use them.

- Just over a tenth (13%) are "**Activists**" – spread across age groups, this group is spending an above average amount of time online, more for practical than leisure use. Although they are less likely than average to say that they feel confident that they know how to control what information they share with organisations, in reality they are more likely than average to be taking protective action, both by restricting what data can be observed about them and dirtying data.

    - Meet **Jack, 71, from Perth**: Jack considers himself to be a relative latecomer to technology, but has in recent years become a "convert" after being given an iPad by his grown-up children. He now can't imagine life without it, and uses it regularly for online banking, social media and emailing friends and family. Jack describes himself as a cautious user of technology and conducts a lot of his own research about how to keep his information secure online, particularly his bank details.

- Just over a tenth (13%) are "**Maximisers**" – this group is skewed towards those younger than 55. Nearly all of them are using social media, messaging services and apps. They spend an above average amount of time online and like to use the shortcuts afforded them online.

    - Meet **Steph, 22, from Newport**: Steph is a student with her own make-up blog. She is an avid user of Instagram and loves having easy access to her social media accounts. Steph says that she is concerned about her information 'falling into the wrong hands' – particularly her photos – and she doesn't like brands knowing much about her beyond the carefully curated image she portrays on social media. However, she doesn't know how to do anything about this without sacrificing access to the websites and apps she loves so much.

"Sometimes I am worried because I don't know what is out there about me, and I don't know what to do about it."

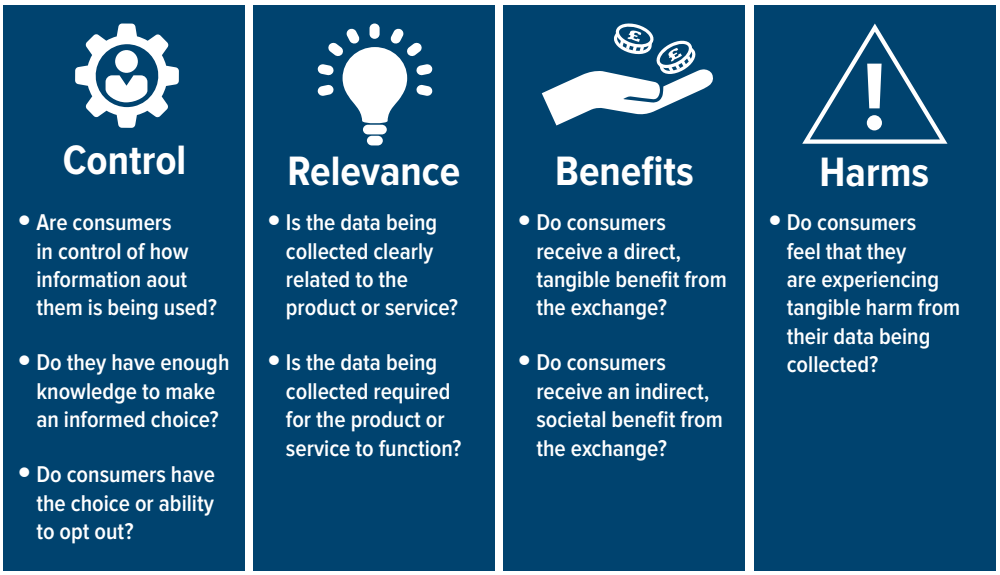
Anxious Maximiser, Newport

3.3: Considering acceptability

When consumers participating in the deliberative workshops were provided with more information about data collection, there were a series of 'penny drop' moments, irrespective of their initial starting point. These moments tended to expand their understanding of data collection and also led them to question the acceptability of certain practices. These included:

- Understanding that data collected about them are not 'just' discrete, anonymised pieces of information, but that these can be aggregated to create a detailed, individual-level profile.
- Being exposed to the scale and complexity of the data ecosystem and data sharing, the role of data brokers, and data sharing for contexts and purposes other than marketing and advertising.
- Learning that beyond advertising or recommendations, information and prices can be targeted and tailored based on consumers' information.

There are four key factors which emerged across the workshops as important to consumers of all attitudinal and behavioural types in weighing up the acceptability of data collection practices, and which caused these specific examples above to 'cross the line' for many consumers:



"It's definitely benefitted me; having my details saved so I don't have to repeatedly enter them. It's convenient."

Workshop participant, Perth

"Acceptable data collection to me is where they're doing data collection in order to make things better"

Workshop participant, Leeds

"It's acceptable when it's practical and has a benefit for the user, for example, being told that a road you often use is being closed by a SatNav, or getting a personalised offer from your bank, or a Fitbit helping you to improve your health. Or when it's being used in an emergency, for example by the police or the emergency services, or monitoring terrorist posts on Facebook, or stopping children from viewing pornography."

Workshop participant, Leeds

"I think the red line is when the information is attributable and identifiable - then you lose privacy."

Workshop participant, Perth

"The red line is when I don't know about it."

Workshop participant, Leeds

As a general rule, the more consumers learnt about data collection and the more complete their picture of data collection and sharing became, the more their levels of concern and perceptions that aspects of these practices are unacceptable tended to grow. This additional knowledge had a particular impact on how consumers conceived of their sense of control and potential harms in relation to their data. The majority began the workshops feeling relatively in control of their data until they were exposed to more information about practices which are taking place without their knowledge (particularly the existence of data brokers monetising consumer data), and unable to identify any tangible harms of data collection and sharing until they learnt more about how their data can and is being used (particularly in the context of tailoring information and pricing).

"I am surprised that the older generation really struggle to understand permissions and what apps know about you – it's not that surprising to me. We have been teaching each other today – I teach them how to fix permissions, they have been teaching me about insurance!"

Workshop participant, Perth

> *"The more you think about it, the more worried I get."*
>
> Workshop participant, St Albans

> *"We know that everyone knows our data, but what was new [in the workshop] was where it goes and how many people are seeing our data"*
>
> Workshop participant, Leeds

Strikingly, relatively few consumers focused explicitly on concerns about infringements of their privacy over the course of the workshops. For some participants, this was important, and even a vague sense of an infringement of their privacy constituted a potential 'harm' or detriment in principle. However, these views were not widely or deeply held and 'privacy' did not emerge as a major focus for most consumers when thinking about their data. The consumers for whom privacy really did matter and for whom this theme dominated discussions were often those with circumstances that might make them more 'vulnerable', such as long-term health conditions. For these consumers, infringements of privacy went beyond high-level discomfort to being perceived to have the potential to cause them genuine detriment, particularly through discrimination.

The perspective of vulnerable consumers: reflecting their more negative starting point when considering data collection, more vulnerable consumers were often quicker to feel that organisations were crossing the line when it comes to collecting and using their data. This is especially true when considering perceived potential harms related to infringement of privacy. Examples included:

- A participant with a mental health condition who felt conflicted about 'liking' support groups on Facebook related to her condition because she feared that this would indicate her condition to Facebook and her contacts. She felt that there was nothing that she could do to stop Facebook gaining access to this information.
- A participant who had been a victim of domestic violence in a past relationship who felt very concerned about the amount of data that is collected about her, and her ability to control who has access to it. She felt that there was little she could do to feel confident that virtual information collected about her physical location couldn't be accessed by her ex-partner.
- An elderly participant from a lower socio-economic grade was glad to be able to receive a lower price for her energy after switching providers because she qualified as a 'vulnerable customer'. However, this participant saw this as a result of relevant information she had willingly shared directly with her energy provider, rather than information that they had observed or deduced about her, and so this didn't affect her generally critical view of data collection.

"I like some support groups on Facebook. It's personal information for me and I don't want the whole world to know. But you can't change it from happening, can you?"

Consumer with a mental health condition, St Albans

"There's all these new ways that people can get your information now. I'm very careful about it."

Scoping phase participant, consumer with a disability, Colne

4. Consumer attitudes towards data collection

4.1: Consumers' awareness of data collection

The vast majority of consumers are aware (albeit to a limited extent) that they are sharing data and that data is being collected about them when they use different data-dependent products and services. However, most have low awareness of the full spectrum of ways in which data can (and is) being collected about them, and many conceive of data sharing as taking place in 'bounded transactions' with their providers in order to use a product or receive a service.

Awareness of proactively shared information

Most consumers are aware that they are proactively sharing (or 'giving away') some information about themselves when using data-enabled products and services. The most front-of-mind examples of this include:

- Entering an email address or other contact information to access a service such as free public Wi-Fi.
- Posting and sharing information on social media.
- Allowing apps such as Google Maps to access information such as location.
- Creating accounts and sharing payment information when shopping online.
- Entering details into price comparison websites.

> *"You can't order something online without giving your address and other details away."*
>
> Workshop participant, St Albans

> *"It [sharing your information] basically speeds up everything you are buying! They have your information so you just click."*
>
> Workshop participant, St Albans

> *"Sometimes you cannot advance without putting that information in, for example it comes up red in the form saying you have to fill this in and provide that information, even if you don't want to."*
>
> Workshop participant, Perth

For many consumers, when they think about data collection the tendency is to focus on the specific transaction and context in which that data is being collected (e.g. sharing their data in order to access a specific service). Most consumers, bar the most informed (typically those exhibiting super-protective behaviours, especially in the concerned segments), tend not to think very far beyond these specific, transactional examples without prompting. For example, few spontaneously consider:

- Information that others are sharing about them, such as what their friends and family are posting about them on social media.
- Information that might be collected when they aren't using a specific device or service, such as apps collecting information in the 'background', such as apps collecting their location data when they are not in use.

Awareness of observed data

Many consumers have a sense that some information is also observed about them, referencing visible indicators that this is happening:

- Seeing targeted adverts on social media pages for products that they have searched for and viewed on other webpages or devices.
- Receiving recommendations for products or services that they might be interested in based on previous purchases (e.g. Amazon or Netflix recommendations).
- Awareness of 'cookies' due to references in pop-up notifications (although for most consumers, there is little understanding of what cookies actually are).

> *"I think they do watch what you are doing. Google Maps says 8 mins to home, even though I didn't put my home details, so they know my address somehow."*
>
> Workshop participant, St Albans

The language that consumers use to describe data collection also suggests that they recognise that some information is observed about them, with different kinds of consumers demonstrating different attitudes and levels of comfort with this.

For consumers who feel most uncomfortable about data collection (and often those who feel vulnerable), this is often framed negatively and suspiciously in language around surveillance i.e. being 'watched' or 'listened to'.

For consumers with more moderate levels of concern (the majority), this is often rooted in references to targeted advertising 'following them around' the internet. For some, this is a source of discomfort, while others are just bemused.

For consumers who feel more comfotable and savvy, there tends to be more recognition that they are allowing themselves to be tracked e.g. by accepting cookies, and that this is part of the 'deal' they have made to be online.

"I don't want assumptions being made about me, I don't want to be followed. I don't want to be watched."

Consumer who speaks English as a second language, Newport

"I bought some trainers, and then there they were on my Facebook, Instagram, Google! It was like the trainers were following me around!"

Workshop participant, Newport

"Where it says cookies, you're so used to just clicking accept without even thinking about it."

Scoping phase participant, London

Appetite to find out more about data collection

Beyond these starting assumptions, many consumers have relatively limited interest in finding out more about data collection. This disengagement is partly because even consumers who have some concerns about data collection do not tend to hold them deeply and struggle to see how data collection is causing them tangible harm. It also reflects the complexity of the subject matter, which feels complicated, technical and inaccessible. Many consumers believe that even if they could find out more, they would be unable to understand it and it would be weighted in the favour of organisations collecting data, rather than consumers. In addition, for some, an overriding sense that they do not have any alternative means of accessing products and services, plays an important role in their disengagement.

For many, this limited interest in finding out more about data collection is strongly related to their experiences of first signing up to data-dependent devices and services. They believe that information about data collection is intentionally hidden in long and complex terms and conditions no 'normal' consumer can ever be expected to read, and that the onus is on them to actively 'opt out' of sharing their data. Most believe that if they ever challenged an organisation holding their data, they would be told that they had given consent for them to do so by agreeing to terms and conditions or failing to notice and 'untick' a checkbox when first signing up to that service.

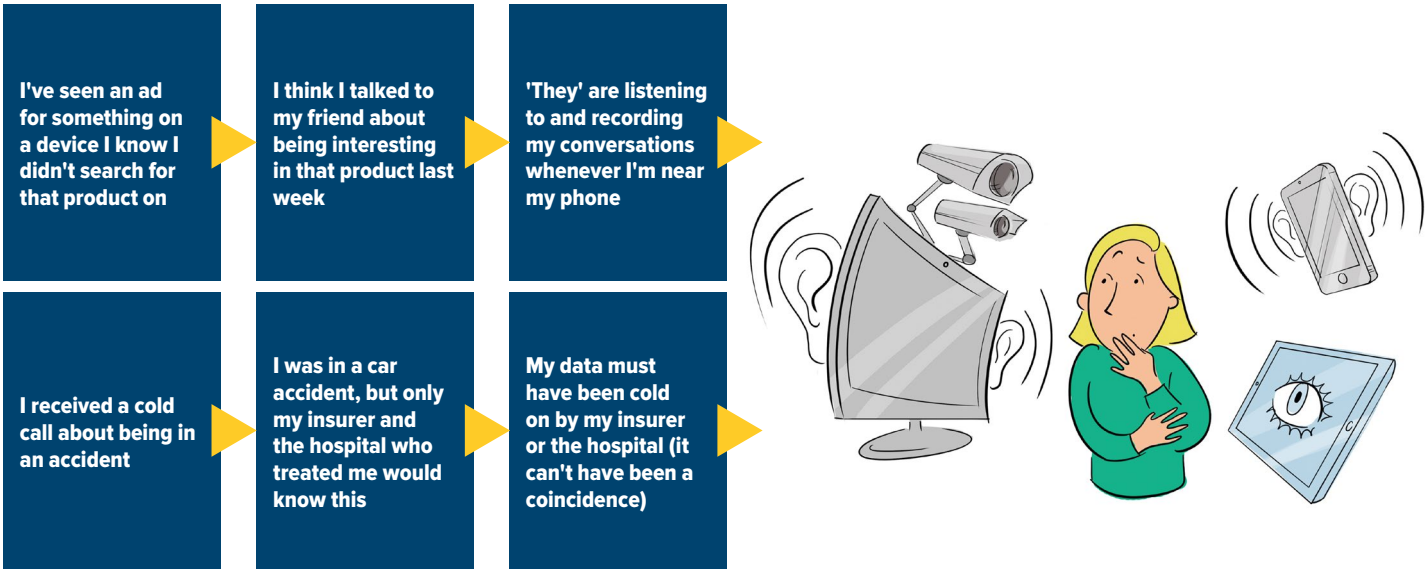
> *"I've heard that the terms and conditions on iTunes are longer than the average Shakespearian play!"*
>
> Workshop participant, Newport

> *"Some privacy settings are not clear and terms and conditions are so complicated, so some people don't know if they are sharing it with the whole world."*
>
> Workshop participant, St Albans

As part of the deliberative workshops, participants were encouraged to engage with their privacy settings and read some of the privacy statements provided by major collectors of data including Facebook, Google, FitBit and other apps that they use regularly. All but consumers in the "Activist" segments tended to find this exercise surprising and illuminating. Most were surprised by the sheer amount of information being collected by each of these services and apps (particularly outside the direct 'transaction' with that service or app, e.g. when it is not in use), but many were also surprised by the relative accessibility of these settings and privacy statements, with many changing their settings as a result.

Prevalence of myths and false assumptions

Consumers' limited awareness and understanding of data collection means that they are often making false assumptions about what is happening to their data, regularly missing important pieces of information and making their own conclusions about what is happening and why. The most prominent of these assumptions is the belief that smartphones, laptop webcams and smart TVs are always 'listening' to and 'recording' their owners' conversations. For many consumers, news articles and information that these devices might be 'listening out' for a trigger word (such as 'Siri' or 'Alexa') only served to confuse the issue further). These exchanges are two common examples of consumers' reasoning on this topic:



"It's not acceptable when they are listening to what you are saying out loud, it makes you feel there's no red line."

Workshop participant, Leeds

4.2: Responses to information about data collection

When provided with more information about data collection and specific examples of products and services collecting their data, consumers' responses reveal the importance of the themes of perceived control, relevance, benefits and harms in determining acceptability. It is important to note that at this stage in the deliberative workshops, consumers were engaging with each of these examples in isolation, rather than how information from each might be used or shared more widely:

EXAMPLE	CONTROL	RELEVANCE	BENEFITS	HARMS
Smart Travelcards	Consumers feel that there is generally a choice between using a paper ticket or a smart travelcard, and between registering smart travelcards or using them unregistered.	Consumers can see a clear rationale for collecting information relating to their location in the context of travel-related products and services.	Consumers can see a clear societal benefit if information collected is used to manage traffic and improve customer journeys. A small number had personally benefitted from receiving targeted information about delays or planned closures to their favoured routes.	Information collected largely feels generic, un-sensitive and therefore unlikely to cause consumers harm. However, some consumers raised concerns about potential links to their payment data.
Activity Trackers	Consumers feel they have a clear choice over whether or not to buy an activity tracker. Those who do have trackers feel that they have a choice in how they use it – e.g. whether or not they wear it to sleep (and therefore allow it track their sleep).	Consumers feel most information collected is relevant to health and fitness. However, some questioned information that felt less directly relevant to health and fitness, and where they did not understand its role in the functioning of the product (e.g. IP addresses). There is particular resistance to the idea that the provider of a relatively expensive product (c. £100) might be collecting any information that is not strictly necessary to the functioning of the product.	Consumers who own activity trackers could see a clear benefit of collecting this information in supporting people to manage their health and fitness. A small number had benefitted from discounts on their life or private medical insurance premiums by demonstrating healthy lifestyles and behaviours.	Most consumers are unable to identify any tangible harms from sharing this data: while it felt personal and sensitive by nature (because it relates to their health and fitness), many could not see this information having any value to other people and organisations.

Mobile Apps	Most apps are seen as optional rather than essential and therefore a choice – though increasingly necessary to participate in modern life. However, there is some belief that with most apps, you have no choice but to accept the T&Cs (requiring you to give away your data).	Most consumers could see some 'legitimate' reasons for apps to collect data in order to work better/at all (such as information about their device and internet connection). However, they feel that a line is crossed where: ● Consumers can't see a link between the data being collected and functionality. ● The app is collecting information even when it isn't in use (e.g. 'always' having access to location services).	For many consumers, the ability to use mobile apps at no financial cost is felt to be a sufficient benefit of sharing their data. However, many conceived of mobile apps making their money through advertising (particularly as paid-for, advert-free versions of many apps are now available), and few engaged with the full picture of the value of their data at this early stage in the workshops.	Most of the information being collected by apps feels relatively generic and is therefore seen as unlikely to cause consumers harm. However, some felt this crossed a line when information might be gathered from friends and family who may not have signed up to the app in question (and therefore hadn't given 'permission' by proxy).
Smart TVs	Smart TVs This example is felt to offer users a sense of 'false choice' by: ● Allowing them to use the device only if they accept the T&Cs – meaning that the device is useless if you don't accept. ● Seeming to withhold this information at the point of sale, so that only make this 'choice' after spending money. Some consumers are uncertain what does and doesn't constitute a smart TV, and believe that it may not be possible to buy a 'non-smart' TV (either now or in the near future).	Most of the information feels relevant and useful to allow the development of recommendations. In addition, very little of this information feels 'new' – most consumers believe that viewing figures have always been collected in some way, and that this is simply a more modern method of doing so. For some more informed consumers, who increasingly think about a smart TV as similar to a smartphone, tablet or laptop, the information being collected feels intuitive and obvious.	The collection of data to create tailored recommendations is generally welcomed as making consumers' viewing experiences increasingly personalised. For many consumers, this is one of the most front-of-mind benefits of data sharing and collection when they do engage with this topic.	Few consumers could point to any tangible harms of sharing this data, and the potential for recommendations to reduce consumer choice is not front-of-mind. However, references to collecting information using voice searches can stoke fears among more uncomfortable consumers that they are being 'listened to' or 'watched' by their devices. This feels invasive and unacceptable in principle for these consumers. These concerns are very challenging to dispel once raised.
Public Wi-Fi	This example is felt to afford consumers very little control by collecting information about their location even if they aren't using or signed into public Wi-Fi (whereas those signed into and benefitting from the public Wi-Fi are generally perceived to be 'fair game').	Consumers can struggle to see the relevance of information about their location in a public setting, when they aren't using public transport services. However, a minority rationalised this by likening these forms of data collection to existing forms of observation, such as CCTV.	Some consumers can see a potential societal benefit of collecting location data, including collecting and sharing data for public safety or crime prevention reasons. Few can see a direct, personal benefit.	Most consumers are unable to point to a tangible harm resulting from the collection of this data. However, because this example was felt to afford consumers so little control, those who were more concerned about privacy were particularly likely to see this example as invasive and in breach of individuals' privacy.

"I'm really negative about the observed data – it's like being watched. I know a smart TV can do that – it's like being watched, it's really concerning."

Consumer who speaks English as a second language, Leeds

"It's okay when there's a personal benefit – especially a health benefit."

Workshop participant, Perth

"It's not okay when it's hidden in the T&Cs, and when you've paid and you still have to give it away."

Workshop participant, Perth

"When they are using it to fight crime and terrorism and stop other people from doing bad things, then it is fine to collect data."

Workshop participant, St Albans

"I've said all along that it's acceptable when it's having a beneficial effect on me. I don't think I'm that bothered about where it's going, unless it's having a negative impact on me"

Workshop participant, Leeds

5. Consumer responses to what happens to their data

While most consumers have at least some awareness of the data that is collected about them, consumers' understanding of what then happens to their data is much more mixed. When provided with more information, this leads to important 'penny drop' moments for most consumers, causing them to question the extent to which these practices are acceptable.

5.1: Data profiling

As explored in the previous chapter, most consumers recognise that they are proactively sharing information about themselves and that information about their behaviour is also observed online. However few know specifics about how this is collected, what specific information is being collected, and how this data is then used.

When prompted, the majority envisage that this information is relatively generic, anonymised, and specific to a single transaction with a product or service. For many, this does not 'cross the line' of acceptability because:

- They feel relatively **in control** of whether they choose to use that product or service (though some say that not engaging with these products and services is not really a viable option in the modern world).
- They assume that the information collected is probably **relevant** to that specific transaction and that it may be required to make a product or service work.
- They can usually see some personal **benefit** of using these products and services overall, and often of sharing information specifically, for example by making their user experience more personalised and targeted.
- They cannot conceive how collecting this information could cause them to come to any harm if it is anonymous and generic.

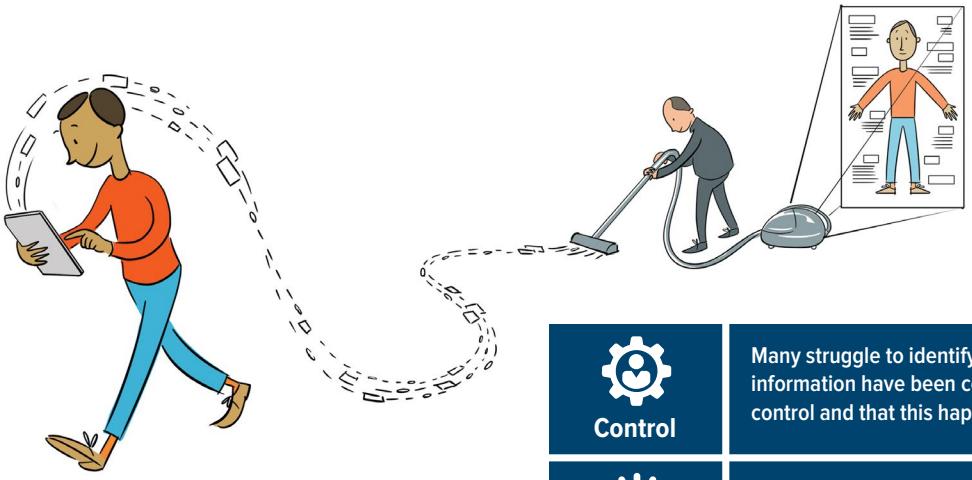
This means that, for all but the most informed consumers, the idea that this information is combined, collated and aggregated about individual consumers (and can be de-anonymised) is an important 'penny drop' moment. This impacts on consumers' levels of concern about a number of issues related to data collection, including:

- **The security of their information**: without knowing that their information is combined and profiled, many consumers envisage that only discrete pieces of information can be stolen (for example, their email address). This leads many to be relatively unconcerned about the risk of their data being stolen, with the exception of their financial information (explored further in Chapter 6). Most do not see the value of this non-financial data or the detriment that it could have on them if it's stolen.
- **Their privacy**: even information that might otherwise be considered to be private or sensitive (for example, health or location data), is perceived to be somewhat innocuous in isolation, and relatively few consumers mentioned privacy as a concern in relation to their data in the early stages of the workshop. Learning about profiling leads consumers to question these assumptions because they had not previously considered that their full name and other information might be connected to characteristics which they consider to be 'sensitive'. This meant that, for some consumers, privacy became more important at this stage in the conversation.

> *"I guess because I expect it to be just my email address or telephone number it is ok to an extent."*
>
> Workshop participant, St Albans

> *"It's okay when it's generic information that doesn't feel personal or private, but it's different when this is put together."*
>
> Workshop participant, Perth

When learning more about data profiling, most consumers are surprised about the extent and detail of their 'digital self'. For some consumers, this crosses the line of acceptability by making them feel they are not in control of information about them, they are uncertain about how this amount of information could be being used in their best interests, and that their privacy has been invaded. At this point in the workshops, many of the participants whose starting perceptions fit with the "Tolerant" and "Concerned" segments started to become more negative.



Control	Many struggle to identify how these different types of information have been collected, causing them to feel out of control and that this happened without their consent.
Relevance	The level of detail included in many data profiles goes far beyond what most consider to be relevant to the functioning of a specfic product or service.
Benefits	Beyond basic demographics and 'consumption' information which might be used for marketing, advertising and tailoring, people question how this information can be used in their best interests.
Harms	Some see potential for discrimination on the basis of 'sensitive' information including ethnicity, religion, sexuality and health data when this is deannonymised and appended to other data.

"[Creating data profiles for individual consumers] that's cloning!"

Workshop participant, Newport

"When it's really specific – that's where it becomes a problem."

Workshop participant, Leeds

"A key turning point on acceptability of sharing my data in my mind are things that are attributable – so where I am identified. But where it is not attributable and identifies me as part of a wider population, that is acceptable. It becomes questionable when it identifies the individual, as it becomes intrusive, and can be used inappropriately and illegally."

Workshop participant, Perth

"It's not okay when it's used to target vulnerable people - for instance targeting gambling addicts, and when they're using these assumptions for things like insurance premiums."

Workshop participant, Perth

The perspective of vulnerable consumers: in line with the majority of 'mainstream' consumers, vulnerable consumers are shocked to learn that the data collected about them can be combined to make an individual level profile. In particular, they tend to be concerned by the level of detail that might constitute their 'data self' and the potential impact that this could have on their privacy. While concern is particularly high about information relating to the factors that might make them vulnerable (e.g. information about a health condition or disability) and which they believe that organisations might use 'against' them, for many, the concept of a data profile was also felt to infringe their privacy in principle.

> *"I'm not comfortable with them having anything that can trace back to me – I don't think it's right."*
>
> Older consumer, St Albans

> *"They'll never stop pilfering your information, they will always find ways to get more of it."*
>
> Consumer with a long-term health condition, St Albans

> *"Sometimes I think an organisation might want to know about me specifically, but then I wonder if it might just be me being paranoid."*
>
> DE consumer, Leeds

5.2: Inferences and assumptions

Just as they have a 'sense' that organisations are observing their behaviour online, the majority of consumers also suspect that organisations are making assumptions about them based on their information. For many, this spontaneous awareness of inferences and assumptions comes from a combination of exposure to:

- **Targeted advertising**: consumers recognise that they are being shown certain products or services because an organisation has assumed that they will be of interest to them. Until consumers are exposed to the concept of a data profile, this is often perceived to be based on fairly high-level factors such as their sex or age bracket.
- **Product recommendation**s: such as loyalty card vouchers or Amazon recommendations for specific products, which consumers believe are driven by assumptions about their tastes based on information about their past purchases and searches.
- **Insurance risk profiles or access to credit**: more informed consumers also highlight the reliance on assumptions for these essential consumer products, based on detailed information including age, income, health data and their credit history.

> *"I get adverts for ladies clothing on my Facebook, things that they know I would like."*
>
> Workshop participant, St Albans

> *"It's good where it helps businesses get a better idea of the audiences they have so it gives them a better idea of the products they should be making."*
>
> Workshop participant, Perth

> *"I guess it's useful for businesses to have that information. I used to work in business and would not want to speak to customers that are not relevant to us – they need to have good profiles."*
>
> Workshop participant, Perth

> *"It's like car insurance. You get a cheaper premium if you are older."*
>
> Workshop participant, Perth

Consumers tend to describe these inferences and assumptions as organisations 'guessing' or 'working things out' about them, and rarely, if ever, refer to assumptions, inferences and algorithms spontaneously. However, on probing, most consumers are generally aware that these estimates of their preferences are computer-driven. For example, upon consideration most consumers are aware that:

- Recommendations on services such as Spotify and Netflix and based upon assumptions about the type of music or TV series they like, partly based on what they might have watched or listened to in the past, but also on what other people in their broad demographic grouping have chosen to watch.
- Targeted advertising on services such as Facebook can be based on assumptions about a person based on demographic information, such as their age and gender. Until they are provided with information about data profiles, consumers generally conceive of these categories as being very broad and unspecific (e.g. 'women in their 20s and 30s').

Based on these starting perceptions, the majority of consumers see it as relatively acceptable and unavoidable that organisations might make assumptions about them based on their information. This largely reflects the perception that the outcome of these assumptions do not have a detrimental impact and, instead, can be beneficial to the consumer. Examples include targeted recommendations for products or services that consumers then go on to buy or enjoy, or lower prices on credit products because they have built up a credit history.

However, on learning more information, the extent to which inferences are made is a surprise which crosses a line for most consumers. In particular, many are shocked about the sheer range of characteristics that organisations

can make inferences about and the relatively limited set of information which organisations can use as the basis for making inferences (such as just a handful of Facebook 'likes'). This often leads to concerns about how frequently organisations could be using inferences to make decisions about consumers and the range of decisions which could be informed by or entirely based on inferred information.

Significant proportions (particularly those in the "Concerned" and "Anxious" segments, and more vulnerable consumers) express particularly strong concerns about organisations inferring things about them that they would not want to share or to be collected, including assumptions made about:

- Factors that they consider to be personal or sensitive such as their sexuality, political views and religion. Concern about this varies significantly between different consumers, with consumers who might identify with minority groups generally (though not always) more likely to be concerned.
- Factors that feel subjective or which seem to verge on 'value judgements' such as inferences about an individual's intelligence. For some, this raised real concerns about why organisations would want this information about individuals and made them question how this type of inference could or would be used.

> *"It's not okay when it's something that you wouldn't share in real life - for instance if you are not open about your sexuality in real life."*
>
> Workshop participant, Perth

> *"I was struck by the data profiles and information in them, it makes you think – what use is some of this data? The other thing is accuracy, people making some assumptions might get things wrong!"*
>
> Workshop participant, Perth

Underlying these concerns is scepticism about the extent to which inferences are likely to be accurate, which for some consumers is a real concern, while for others it is a hopeful sign that organisations don't know all that much about them. In deliberative workshops, participants were given the opportunity to experiment with the information that organisations are likely to assume about them based on their Facebook and Twitter profiles (if they have them) using the 'Apply Magic Sauce' tool[3] . Some were actively relieved to see the algorithm under-estimating factors such as their age, and found it difficult to believe that a computer could accurately infer their preferences. Conversely, some felt that if organisations were using this information to make decisions about individuals, they could be treated unfairly or receive suboptimal outcomes as a result of incorrect inferences.

3 Apply Magic Sauce is an online tool which creates predictions based on a consumer's online activity. The tool predicts consumers' age, gender, sexuality, personality traits and preferences, amongst other things, using behavioural data from their Facebook and Twitter profiles. Information collected by the tool is not shared outside of academia and they do not keep prediction profiles of individuals or personally identifiable information. Please see the Appendix for a link and further information about this tool.

"They make a wild assumption and it goes on and on and gets out of control."

Workshop participant, Perth

"I don't mind because I'm not proactive on Facebook – most of it is a bit stupid. And if they make wrong assumptions about me then it just goes to show Facebook don't know as much as you think they do!"

Workshop participant, Perth

Consumers identify a number of reasons why inferences may be reductive or inaccurate, often relating to concerns about the principle of being stereotyped or 'put into a box':

Concerns about 'stereotyping'	Concerns about historic information	Concerns about a lack of nuance/individuality
Assuming that all people in a particular 'category' behave in a certain way, with potential implications for discrimination and consumer choice.	Inferences being made based on attitudes which may have changed over time – with particular implications for information consumers may have shared at a younger age.	Computer-driven algorithms failing to recognise individuality, nuance and subtlety and taking information at face value.

As a result, some people describe real concern about organisations using inferences about them to target information and prices

"Just because I am black, they might think I like chicken and rice, but that's not true!"

Workshop participant, Leeds

"It's not true – what people say, what people look at – it's not necessarily that person."

Workshop participant, Newport

"To me, it becomes completely unacceptable when the information is incorrect. If it's correct on any of the assumptions, or near enough, then fine. But if it's incorrect and they use that information that's incorrect, that makes me really annoyed."

Workshop participant, Leeds

The perspective of vulnerable consumers: these consumers tend to be even more concerned about the extent to which inferences and assumptions are made about them. In particular, those with a disability or health condition expressed about assumptions being made about their condition and being stigmatised as a result. These participants tended to feel that there was very little they could do to stop these assumptions being made about them, other than disengaging completely from products and services on which they are often very reliant because of factors relating to their vulnerability, including limited mobility.

> *"I find it uncomfortable and uneasy when assumptions are made about me."*
>
> DE consumer, Leeds

> *"I think they [organisations] can stigmatise you if they know what groups you like or that you are looking for help for your condition."*
>
> Consumer with a long-term health condition, St Albans

5.3: Data sharing with third parties

While awareness and understanding among consumers about what information is being collected about them, and how this is being used is relatively limited among many consumers, most suspect that at least some of their data is being shared and 'sold on' to 'third parties'. This suspicion is largely driven by personal experiences of direct marketing, particularly from claims management firms, and third party data sharing is rarely understood in any detail beyond this. However, 'third parties' are organisations that many consumers instinctively feel negative about because they associate them with 'nuisance' or 'scam' phone calls, emails and texts, and the vast majority of the population say that they feel negative about third party data sharing at face value.

> *"I guess this [data sharing] wasn't really a surprise, but I understand now why I get so many calls! We don't know really who or what is behind it though."*
>
> Workshop participant, St Albans

> *"When I get a PPI call, I do think 'where did they get my number from!?' I know I certainly didn't give it to them!"*
>
> DE Consumer , Newport

"I think it's unacceptable that organisations are able to sell my information on to third parties. I don't mind organisations having it when I've agreed to it, but I don't trust them when they start selling it on to other organisations"

Workshop participant, Leeds

"Selling your details on, I just don't like. I feel like they're infringing upon your personal data protection"

Workshop participant, Leeds

The perspective of vulnerable consumers: these consumers are especially likely to express concerns about data sharing and are more uncomfortable with their personal data being shared without their knowledge or permission. Their circumstances and potential barriers to communication may also mean that they feel less able to manage some of the manifestations of data sharing, such as unsolicited marketing calls. Some were actively shocked to learn that data sharing is not an illegal practice.

"That's surely illegal…I don't think they are allowed to do that because of data protection!"

Consumer with a long-term health condition, St Albans

Consumers tend to operate with an incomplete picture of data sharing and third parties, which means that most are surprised to learn about the extent of the data sharing 'ecosystem' and data sharing practices. Particular trigger moments include the realisation that:

- **Whole profiles are being shared**: until informed otherwise, most consumers believe that only discrete pieces of information about them are being shared with third party organisations, such as their telephone number or email address in isolation. Many are shocked to find that their full data profiles – including inferences about them based on this data – can be and are shared with third parties.
- **Consumer data is shared for reasons other than marketing**: participants in the deliberative workshops were surprised to see references to their information being shared for purposes including 'analytics' and to see references to data users including law enforcement, lawyers and private investigations.
- **A whole industry of data brokers exists by sharing and selling on consumer data**: most consumers had no sense that such organisations exist and felt automatically on the 'back foot' when they learnt about data brokers, assuming that this industry is murky and untransparent if it is currently unknown to most consumers like them. As a result, there was particular surprise to see brand names consumers had previously viewed as 'reputable' classified as being 'third party' organisations and 'data brokers', such as credit referencing agencies.

"I don't mind sharing when I choose to give it [my information]... they're holding you to ransom, but I don't know what information about me is being shared."

Workshop participant, Newport

"I am really surprised by data brokers. I didn't even know they did that!"

Workshop participant, St Albans

"What I've found most surprising today about data collection is the data brokers...I hadn't realised that organisations I trust, like Experian, are also a data broker, although it does make sense...I wouldn't say I'm naïve, I think this is just the way of the world now"

Workshop participant, Leeds

"My information could be sold without me knowing, and I am getting no material gains from it. How is that fair?"

Workshop participant, Perth

"I don't even know 650 things about myself, how do the data brokers have all that information on me?! They'll probably know my shoe size and what I am eating!"

Workshop participant, Perth

For many consumers, the full extent of data sharing 'crosses a line' against each of the factors they tend to use to determine acceptability:

Control	Data sharing is perceived to offer consumers very little control: • The sheer scale of the data sharing ecosystem emphasises that data sharing is happening without consumers' knowledge and consent. • And consumers' own experience suggests to them that organisations are deliberately making it difficult for them to opt out of data sharing as possible
Relevance	The level of detail included in many data profiles goes far beyond what most consider to be relevant to the functioning of a specfic product or service. • E.g. sharing information about sexuality simply because the broker has it. • This fuels concern among vulnerable consumers that 'irrelevant' data could be used 'against' them - such as stimatising them based on health conditions
Benefits	Many consumers struggle to see any benefits of data sharing unless prompted with specific examples of products and services which are dependent on data sharing to function, such price comparison websites and fraud prevention services. A small number of more informed consumers and 'data liberals' see data sharing as part of the 'price' they pay for free services such as apps
Harms	Consumers see greater potential for them to come to harm as a result of data collection when that information is then shared, particularly without their consent. For some, these are relatively low level concerns about being targeted and 'hassled' by third parties who have gained access to their data, while more vulnerable consumers are concerned they could be unknowingly stigmatised

"I share my details to get a credit rating. It's almost deceitful that they share my information beyond that."

Workshop participant, St Albans

"When that company came into the mix and my financial details were shared, that made me uncomfortable."

Workshop participant, St Albans

"I don't care how much they collect. It's how they distribute it is what concerns me. They can take everything from me if they want, as long as they use it appropriately and fairly and not against me, that's all that matters."

Workshop participant, Perth

"I don't understand what all the fuss is about to be honest – that's how we are getting these services for free and this is no big surprise to me."

Workshop participant, Perth

The perspective of vulnerable consumers: when thinking about data sharing, vulnerable consumers are especially likely be concerned. Most feel that data shared with third parties could be used 'against' them by providers who might not necessarily disclose that they hold this information and who might deny them or charge them a higher price for a service as a result. For many of these consumers, the uncertainty of what might be happening with this information was concerning in and of itself.

> *"I just don't understand why they want that [health] data. What are they doing with it, why are they taking it?"*
>
> Consumer with a long-term health condition, St Albans

For many, learning about data sharing and the number of organisations involved in the data sharing 'ecosystem' further emphasises just how small a player they are as an individual consumer compared to the organisations collecting, using and sharing their data. The fact that a whole industry exists and profits purely from the selling and sharing of data (i.e. data brokers) is a particular 'penny drop' moment that leads many consumers to believe that data sharing may be even less transparent than they originally believed.

Participants considered the following trade-off: if third party sharing and profiling didn't happen, it is possible that you would need to pay for your email services and social media. Would you prefer to pay for these services than have your data shared?

Despite significant concerns about the fairness and transparency of third party data sharing, the majority of consumers said that they would not be willing to pay for the services they use in order to stop their data from being shared, for some or all of the following reasons:

1. **They feel it is too late to take back control of their data**: their privacy has already been compromised and their information is already 'out there' and they feel powerless to try and re-establish control.
2. **Financial costs outweigh any potential benefits**: they would be unwilling to pay for a service that they are currently receiving for free, and feel that the financial cost would be more tangible – especially given that it could exclude consumers who are financially stretched.
3. **Lack of trust in organisations**: they feel that many organisations would still share their data with third-parties, even when paying them a fee to ensure that this does not happen.
4. **Worth the 'risk'**: some accept that this practice occurs and feel they have not personally experienced any detriment from their data being shared, so would be unwilling to change the status quo.

"Because I expect it and I know the way things are, it is ok to an extent. But I still don't like that it is sold and shared."

Workshop participant, St Albans

"Even if we pay for it, it will still probably get shared…it's not worth it."

Older consumer, Perth

Among the minority who were willing to pay for these services to avoid having their data shared, many were reluctant to pay on the basis of trust in an organisation alone. These consumers cited specific conditions that would have to be met for them to pay for a service:

- The ability to see what data organisations held about them;
- The ability to have that data amended or deleted;
- Control over which organisations are given access to their information – allowing them to pick and choose who they share information with;
- Having legal assurances that data will not be shared without their consent; and
- Giving consumers the option to 'monetise' their data in return for reduced costs, for example sharing information about themselves with one specific organisation could reduce their bill by 5%.

"It's acceptable if it is regulated and there are things in place – I don't mind paying if it is safe and regulated by an ombudsman."

Workshop participant, Leeds

6. Security of information

Security of information from risks including data breaches is rarely front-of-mind when consumers think about their data and data-dependent devices and services, and particularly the collection and sharing of information which is non-financial. This reflects:

- A belief that the impacts of data being breached or stolen in this way are relatively limited, with **little to no long-lasting harm** to the consumer, While consumers do recognise that there could be negative impacts of their financial data being stolen, these impacts feel abstract and low-level in the context of the belief that financial firms will always reimburse customers who have fallen victim to fraud (for many, supported by personal or second-hand experience). When prompted to consider their non-financial data, most consumers simply cannot see how criminals could profit from this information. Even introducing information about potential security risks to non-financial data through the increase of 'connected home' devices in deliberative workshops had limited effect on consumers' concern: most struggled to understand how a criminal could obtain this information and why it might be of any value.
- Growing fatalism among consumers and a perception that securing their data is **beyond their control**. Coverage of high-profile data breaches of organisations ranging from Talk Talk to the NHS have only served to reinforce the perception that breaches are inevitable and will happen regardless of what consumers do. Most consumers struggle to understand what they could have done differently in these contexts – including consumers who are engaging in protective behaviours from a privacy perspective.

> *"There's a risk in anything you do, but really, what is the worst that could happen?"*
>
> Consumer speaking English as a second language, Leeds

> *"Let me give you an example – 27th October, I went shopping in Iceland near where I work, and bought bread and milk. A little later, my bank sent me a message saying call this number. I called it and they told me that transactions were being made in Buenos Aires. They knew I couldn't have travelled from my office to Argentina in that time! So they helped me get my money back…I was so impressed that they knew about it and warned me."*

Workshop participant, St Albans

"You don't bat an eyelid now – you're just used to it."

Workshop participant, Newport

"Companies get breached and [your] information gets stolen…Do you remember Santander? That was a big one as well."

Workshop participant, St Albans

In this context, most consumers say that they are content to put their trust in the organisations collecting their data to keep it secure, and particularly larger organisations including banks, major retailers, and major payment system providers such as PayPal, Visa and Mastercard, who they expect to be investing in the latest security measures to keep their customers' data safe because it is in their best interests to do so. Trust in smaller organisations and businesses to invest in the same level of security protections is generally much lower, though notably, these organisations tend to be more trusted to respect consumer data from a privacy perspective: for example, they are less likely to be associated with 'selling on' customer data.

"I would be reluctant to buy from a website if they don't have PayPal, as I would like to have that security and I don't think I would trust it."

Workshop participant, St Albans

"When you have bigger companies like Tesco or John Lewis, I don't think you mind as much when you share your information. But when it's a clothing shop I haven't used before I like to check it out first to see if it is secure."

Workshop participant, St Albans

"Nothing is 100% safe, but if you go through PayPal and banks, it is more secure."

Workshop participant, St Albans

Some consumers are relatively open to the fact that they are doing very little to protect their data and may not be following even what they consider to be the most basic cyber security advice, such as having a strong, separate password for their most important accounts. Given that consumers believe that organisations holding data about them will be 'hacked' no matter what, are uncertain exactly what this data is and how it is being used and shared, and struggle to see any potential negative impacts of being hacked, many believe that nothing they can do will be worth the inconvenience of consistently implementing secure behaviours. Those who are consistently taking protective measures, such as 'dirtying' their data, are in fact often driven more by concerns about privacy than they are by security.

"With my bank details I am more cautious. I don't input them on everything."

Consumer with a long-term health condition, Newport

"What can you do other than limit what you put out there?"

Workshop participant, Newport

In deliberative workshops, participants were given the opportunity to use the tool 'Have I Been Pwned?'[4] to see if their data had been breached. Despite their sense of fatalism and growing awareness of high-profile data breaches, most consumers found that they had under-estimated the amount of information about them that had been breached. In particular, consumers were surprised to find that:

- **Information had been stolen from 'reputable' providers who hadn't told their customers that they had suffered a breach (either directly, or indirectly, through the media)**: for example, there was some surprise in seeing references to accounts held with organisations such as LinkedIn and Yahoo being compromised.
- **Information about them had been stolen from unfamiliar organisations that they had not heard of before and who they had not consented to hold their data**: this included data brokers, organisations which had only become known to the majority of consumers over the course of the workshop.
- **Information beyond email addresses and passwords had been stolen**: including home addresses and IP addresses, and in some cases, all of the above.

"Wow! I didn't know all of that had been stolen…I might go and change my password right now!"

Workshop Participant, Perth

"It's great to know there are ways and means to review what has been obtained and what is shared, that's good but was surprising."

Workshop participant, Perth

"I would pay to ensure my information is kept more secure, but I wouldn't pay for just one app, I would want to pay for a package like Sky TV or something, and pay by type, so I can pick and choose."

Workshop participant, Perth

4 'Have I Been Pwned?' is an online tool that allows internet users to check if their personal data has been involved in a data breach, by cross checking their email address against a large collection of database dumps and pastes. The name is based on a hacker jargon term 'pwn', which means to 'comprise or take control, specifically of another computer or application'. Please see appendix for further information.

However, while exposure to this information in the setting of a deliberative workshop increases concern for consumers, it rarely spurs them into action and instead appears to reinforce a sense of resignation about the security of their data. Many believe that once their information is 'out there' (through cyber crime, but also legal practices such as data sharing) there is very little they can do to protect it. In deliberative workshops, relatively few consumers mentioned taking protective measures, such as changing passwords for breached accounts. While some consumers recognise that they probably could find out more about how to protect themselves after a breach, the perceived limited consequences of data breaches did not make this feel important or urgent.

The perspective of vulnerable consumers: while workshop participants tended to focus on concerns that their financial information could be stolen, vulnerable consumers were more likely to worry about the safety of their non-financial data. This feeling was primarily driven by personal experiences, and the perceived risks associated with their life situation, such as limited mobility, or factors which make them feel particularly vulnerable to harm in general. These more vulnerable consumers often felt that in the event of a data hack, other details, over and above their financial data, could potentially compromise their physical safety and wellbeing if they 'fell into' the wrong hands.

> *"Having suffered domestic violence, I think this stuff should be private – it's concerning because I see that nothing is really private."*
>
> Consumer with a long-term health condition, Greater London

7. Targeting and tailoring based on consumer data

7.1: Awareness of targeting and tailoring of adverts and recommendations

As explored in previous chapters, most consumers have some awareness that at least some of the adverts they see and services they use are targeted, tailored or personalised to them based on their information. The most front-of-mind examples of targeting and tailoring are:

- Personalised recommendations on services such as Netflix or Spotify.
- Targeted advertising online, in particular on services such as Facebook.
- Receiving personalised vouchers and discounts as a result of having a supermarket loyalty card.

Based on these front-of-mind examples, and most consumers' starting point that this targeting is taking place on the basis of relatively generic categorisations, most consumers tend to feel that targeting and tailoring is positive, saying that they enhance services by increasing the relevance of the content they are shown. For many, targeted advertising is seen as preferable to non-targeted advertising because it reflects their interests, and personalised discounts tend to be preferred to generic ones because consumers feel that they are more likely to use and benefit from them.

> *"I want the recommendations and I like the choice, the majority of the time they do get it right."*
>
> Workshop participant, Leeds

Without prompting and based on their starting perceptions, consumers do not tend to consider the broader impacts of targeting and tailoring on society as a whole, and there is relatively little concern that targeting and tailoring might impact negatively on consumer choice. The majority of consumers – bar the "Anxious" segment – tend to believe that they are ultimately in control. Those who are more "Tolerant" overall say that they can simply 'ignore' recommendations and targeted advertising if they wish to do so, while a significant minority of those who are "Concerned" are engaging in protective behaviours such as ad-blocking.

Consumers who do raise concerns about tailoring and targeting of advertising and recommendations are often more likely to reference 'irritations' than active concerns. This includes frustration when recommendations, advertising and discounts are 'irrelevant' to them, or when others in their household can deduce their search history from the targeted adverts they receive. However, as with many other aspects of data collection, consumers struggle to pinpoint any detriment they might experience as a result of these 'gripes': for many, the most tangible harm they can imagine is a spoilt surprise for a friend or family member.

> *"My partner found out about the birthday present I was going to get him because he saw the adverts I was getting and guessed it."*
>
> Workshop participant, Leeds

> *"It's just a suggestion. I'm not so impressionable I think I need to buy something straight away if I see an ad for it."*
>
> Workshop participant, Perth

> *"You'd have to be a hermit to only watch TV you'd been recommended on Netflix. You talk to other people and get recommendations from them too, so it's not like you're just going to watch what Netflix tells you to watch."*
>
> Workshop participant, Leeds

> *"Data collection has benefitted me when information has popped up, and I've gone on to sites that I wouldn't necessarily have been on initially, but it's given me access to products that I wouldn't have found before"*
>
> Workshop participant, Leeds

The perspective of vulnerable consumers: consumers who might be identified as vulnerable are far more likely to conceive of potential harms as a result of targeted advertising and recommendations without significant prompting. These included:

- Concerns that addictions, such as gambling or alcohol addictions, might be manipulated by cynical advertisers.
- Concerns that sensitive information that they would rather keep private (even from family and friends) might be revealed to others using the same device through targeted advertising. Suggested examples included receiving adverts for support for mental health conditions such as depression, or adverts based on reductive assumptions (e.g. adverts for HIV tests based on the assumption that someone might be gay). This concern was particularly pronounced among vulnerable people who were in employment who believed that colleagues might be able to see this kind of advertising on their devices.
- Concerns that consumers with some mental health conditions or learning disabilities may not be able to recognise and 'filter out' targeted adverts and recommendations. One participant recounted an example of a family friend who had got into debt after regularly exceeding her credit limit when shopping online, and who she believed had suffered as a result of 'irresponsible' advertising.

> *"She couldn't stop herself, she bought everything. It was bad because they clearly didn't take into account who they were advertising to."*
>
> Consumer with a long-term health condition, Newport

> *"I think that it's disgusting that the gambling industry could single someone out like that using technology because they know they've got an addictive personality."*
>
> Workshop participant, Leeds

7.2: Acceptability of targeting and tailoring of adverts and recommendations

Personalised recommendations and targeted advertising – which at surface level are generally seen as actively positive at best and irritating but innocuous at worst – are generally considered to be acceptable. However, targeting and tailoring of advertising and recommendations is felt to 'cross the line' for some consumers when it happens in public. Consumers quickly understand that, if the information this tailoring is based upon is accurate, recommendations and adverts shown to individuals in public could reveal information about consumers that they may not want to share about themselves. Having learnt about inferences earlier in the workshop, some were also concerned that these forms of advertising could lead to false, inaccurate and 'reductive' assumptions being visible to other people:

EXAMPLE	CONTROL	RELEVANCE	BENEFITS	HARMS
Netflix recommend-dations (currently a widespread practice)	Consumers tend to feel that they are aware that these recommendations are personalised and feel able to choose whether to engage with them or not.	The information consumers understand to be collected to make recommendations function (i.e. information about their consumption habits) feels relevant to the service.	Consumers see the recommendations they receive as a clear benefit of their data being collected. Many consumers are able to describe examples of TV shows and films that they do not necessarily believe they will have discovered without these types of recommendations.	Because recommendations feel private (i.e. only visible when logged in to an account, with potential to create different user profiles within the same household), most consumers cannot conceive of any tangible harm of tailored Netflix recommendations.
Targeted adverts shown on Facebook (currently a widespread practice)	Whilst many consumers do not feel able to opt out of these adverts (with only a minority referencing ad blockers), they tend to feel that they are in control of whether or not they choose to buy the products or services they are being advertised.	The information which consumers expect to be collected in order to target advertising – a combination of consumption data and demographic information – feels relevant and necessary in order for targeted advertising to function.	While most consumers are reluctant to describe advertising as 'beneficial', consumers tend to describe targeted adverts as more beneficial than non-targeted adverts because they are more likely to reflect something that they are likely to be interested in. Indeed, many took the view that if targeted advertising is a 'fact of (modern) life', they would like these adverts to be as personalised and accurate as possible.	Most consumers can generally only see a potential harm if they share devices or accounts with partners, family and friends. In this instance, they believe that there is a risk that targeted advertising may reveal aspects of their browsing history or other preferences that they would prefer to keep private. Very few understand spontaneously that these adverts might be visible across devices to people who share the same IP address (i.e. using 'fingerprinting' technology).
Facial Recognition Software – Piccadilly Lights[5] (not currently a widespread practice)	Consumers did not believe that they would be able to opt out of this form of targeting as they expected that it would happen without their consent or knowledge. Most could not see how advertisers could pragmatically obtain consent from consumers when targeting them in large crowds in this way.	Because information is collected about individuals in the context of a crowd, the information required to target advertising to a large group of people feels relatively generic and high-level. Most recognised that information about factors like the average age of a crowd could feasibly be used to target advertising, though they questioned the likely accuracy of this targeting.	Consumers believed they were unlikely to benefit from these adverts because they believed that they would be reductive and inaccurate, based on information from a wide group of people in the vicinity, as well as their own information. Some felt that this defeats the purpose of targeted advertising.	Most consumers cannot see direct harms unless facial recognition technology develops to allow targeting by more sensitive factors. However, unlike the vending machine example below, most recognised that targeting in a crowd means that it would be very difficult to deduce specific information about specific individuals.

5 Workshop stimulus about Facial Recognition Software at 'Piccadilly Lights' was based on the following article: https://www.theverge.com/2017/10/16/16468452/screen-london-picadilly-circus-cars-targeted-ads-landsec. See appendix for further information.

Face-scanning vending machines[6] & facial scanning software in Tesco till adverts[7] (not currently a widespread practice)	Consumers were surprised to learn that information about them could be gleaned from face-scanning and facial recognition technology, and were unsure whether they would be made aware of this practice if it became widespread in the future. If consumer consent is not obtained, and consumers aren't made aware that face-scanning technology is being used, consumers felt that they could lack control and may be unable to opt out. Most consumers were relatively cynical and expected that they would not be asked for their consent in practice.	The information consumers expected to be collected to make recommendations function in this instance (such as their demographic profile), felt relatively generic and high-level. However, some questioned the relevance of this to their consumption habits and preferences. Some were reluctant to believe that information about someone's sex could be used to develop an accurate recommendation about their shopping preferences.	Consumers struggled to see how they would personally benefit from these types of recommendations because they believed that they would be reductive and inaccurate, based on simplistic assumptions about them based on factors such as their sex. They expected that these adverts or recommendations would be less accurate (and therefore beneficial) than those shown on Facebook or Netflix because these examples also incorporate their actual tastes and consumption data.	Few consumers could see a direct, personal harm related to this technology because they understood it to be based on relatively generic information that most people would be able to 'guess' about them (e.g. their broad age group, and their sex). However, on discussion about how this technology could develop in the future, a small number became concerned that it might have the potential to reveal more 'sensitive' information about them to other consumers in the vicinity of the vending machine or till, such as their mood, mental state or inferences about their health.

"As far as knowing my name and giving me suggestions, and knowing what I need to look at – this is good! It saves time! It's so convenient!"

Workshop participant, Leeds

"I don't think it's intrusive when they tell me what movie to watch, that's fine and I like that."

Workshop participant, St Albans

"Some people naturally have a sad face. It's not fair targeting them by their face."

Workshop participant, Perth

"Being targeted based on how I present myself physically is wrong. Technology cannot replace a human's perception of you, they don't know my mood."

Workshop participant, Perth

"The way technology is moving forward in a commercial sense is worrying. We haven't given our permission for the Piccadilly lights to record and advertise to us have we?"

Workshop participant, St Albans

6 Workshop stimulus about face-scanning vending machines was based on the following article: http://theweek.com/articles/489132/japans-facescanning-vending-machines. See appendix for further information.6

7 Workshop stimulus about Tesco till adverts was based on the following article: http://www.bbc.co.uk/news/technology-24803378. See appendix for further information.

7.3: Awareness of targeting and tailoring of pricing and information

While most consumers have seen some form of targeted advertising or personalised recommendation, just a fraction described any spontaneous awareness that information and pricing can be similarly tailored. Many were surprised to learn that different people will be shown different results, in a different order, when they enter the same search term on a search engine such as Google, and that prices might vary according to consumers' behaviour online. The closest examples most consumers could bring to mind spontaneously related to dynamic pricing of flights and hotels (which they often related to wider market forces and peaks and troughs in supply and demand), and tailoring of insurance premiums according to consumer behaviour (e.g. black box insurance 'rewarding' safe driving).

> *"I thought everybody, when they Google something, they get the same results in the same order… I didn't realise that Google makes assumptions based on what I've looked at already, but what can you do?"*
>
> Workshop participant, Leeds

7.4: Acceptability of targeting and tailoring of pricing and information

Consumers tend to have more fundamental and deeply held concerns about personalised pricing and information than they do about targeted advertising and personalised recommendations. This largely reflects their lack of awareness about what is currently happening in relation to this practice, how it is likely to develop in the future, and a perception that if you aren't even aware that you are operating in a 'filter bubble' there is little you can to do protect yourself.

> *"You think you've looked up the best prices, been savvy and that, and you don't even realise what you've not been shown."*
>
> Workshop participant, Leeds

> *"It's really damaging for young people who haven't voted before. How can they truly make up their minds in elections if Google decides what they see?"*
>
> Workshop, Leeds

This concern is consistent across the majority of consumers and is a key 'penny drop' moment for even otherwise "Tolerant" and "Liberal" consumers. Many consumers raise particular concern about the impact of these practices on potentially vulnerable consumers, children and younger people who have grown up with technology and may be unable to find offline alternatives.

The perspective of vulnerable consumers: as with most issues related to data collection, consumers who are more likely to be vulnerable are particularly likely to raise concerns about the impact of personalised pricing and information. In the context of widespread data sharing and inferences being made about 'private' factors including their health, several of these participants believed it was a very real possibility that they could be penalised on the basis of price because of factors such as a disability.

Most consumers also believe that the 'stakes' are higher in relation to personalised pricing and information than advertising and recommendations, with far greater potential for detriment. For many, personalised pricing undermines their sense of control and consumer choice: they believe it is impossible to act as a savvy consumer – e.g. by shopping around and looking for the best prices and deals – if you have no way of knowing if you are being shown the lowest prices. Concerns about personalised information include the potential for manipulation of public opinion by powerful groups and organisations, particularly in relation to news, politics and elections.

Underlying this is a continued concern among many consumers about the accuracy of the information being used to tailor and target prices and information. For those who have limited faith in the accuracy of algorithmic inferences, and feel demographic information is likely to stereotype them, they are often concerned that these forms of personalisation will lead to unjust outcomes which they are unable to challenge or rectify because they don't know what assumptions organisations are making about them.

> *"You can be denied services because of things like this [inferred data] and it's completely wrong."*
>
> Workshop participant, Leeds

> *"It's not fair that I get a bad price because I have a Hotmail account even though I'm very sensible."*
>
> Workshop participant, Newport

Consumers' specific reactions to current and potential future examples of these practices were as follows:

EXAMPLE	CONTROL	RELEVANCE	BENEFITS	HARMS
Admiral Insurance – Social Media based premiums[8] (based on a proposal which was then not taken forward)	Consumers feel they would have some control over this service if it became available and it was opt in, and they understood the information that decisions were based upon. However, if information from social media was used without their consent, many feel that this would violate their sense of control.	Most consumers struggle to understand how information gathered from social media could be an accurate predictor or relevant in determining insurance premiums. Basing calculations on information such as posts shared by friends felt 'judgemental', subjective and irrelevant to many.	Some consumers could see a clear personal benefit in that they might receive cheaper insurance if this kind of product became available (particularly those who were relatively cautious about what they shared on social media). A smaller number could also see a potential societal benefit in allowing younger consumers to use their social data in order to build up their credit history	Many consumers thought it was more likely that they would receive a higher than a lower insurance premium if they allowed a provider access to their social data, which they saw as a clear harm if they did not have any choice or access to alternatives. A minority could also see a potential societal harm whereby consumers who do not use social media might face barriers to accessing cheaper insurance prices or totally unable to access insurance products in the future.
Personalisation of Google search results[9]	Significant proportions of consumers were unaware that their search engine results are individually tailored to them, and some were actively shocked about this practice. Consumers' low levels of knowledge about this practice made them feel that it is happening without their consent and outside their control.	Many consumers do not believe that the information that might be seen as relevant for targeting advertising and recommendations (such as demographic information and consumption data) should necessarily be driving tailoring of non-product related information. There was some discomfort with assumptions about a consumer's broader values (e.g. their political leanings) being inferred from this data.	On face value, most consumers could not see any benefits of this personalisation and felt that the harms outweighed the benefits. However, a smaller number described potential benefits of this information being personalised, including information being targeted to individuals' local areas (e.g. automatically showing the closest branches when a consumer searches for a certain shop brand).	Most consumers felt that they could be harmed by deliberately or unintentionally being denied access to relevant information. Several also identified a potential societal harm for younger people in particular, who may be forming their beliefs purely as 'digital natives' and without the same claimed ability to seek out balanced perspectives.
Insurance premiums based on email addresses	Consumers feel that they are unaware that premiums can and are personalised in this way and therefore unable to take action to ensure they are not personally disadvantaged. Many felt opposed to this in principle, rather than being concerned by the specific amounts by which they might be 'over-' (or under-) paying.	Most consumers struggle to see how their email address could be an accurate predictor of their risk profile and therefore felt it was irrelevant to the price they should pay for insurance. This partly reflected a poor understanding of how risk is calculated in general.	Perhaps unsurprisingly, consumers tend to feel much more positive about this practice if they believe that they might personally benefit (by receiving a lower quote).	Most felt that consumers could be harmed if they are charged a higher price for their insurance without any transparency about the factors driving that decision that might enable them to 'improve' their behaviour and therefore reduce their premiums.

8 Workshop stimulus about the Admiral insurance proposal was based on the following article: https://www.admiral.com/black-box-insurance/. See appendix for further information.

9 Workshop participants were shown a film produced by alternative (non-personalised) search engine 'Duck Duck Go' to explain personalisation of Google search results: https://vimeo.com/51181384. See appendix for further information.

Personalis-ation of job adverts based on gender[10]	Consumers have low awareness of this practice taking place at the moment, and are concerned that if this was happening more widely they would most likely be unaware of it and therefore unable to take action against it.	Consumers feel that the information being collected and used as the basis of targeting – gender – is in this instance irrelevant to whether an individual should be shown a job advert.	Very few consumers could see any benefit of this type of targeting.	Most consumers could see a clear societal (and often, personal) harm of this type of targeting in perpetuating discrimination over time and denying women access to opportunities. On prompting, many felt that it was important that algorithmic decision-making should not reflect the 'prejudices' of the past and/or the individual programming it.
Cambridge Analytica [11] [12]	Consumers feel that they are currently unaware of targeting based on assumptions about political preferences, and therefore unable to take action to counteract it. There is broader concern that this form of personalisation denies them control over the forms of information they have access to.	Consumers tend to feel that this form of personalisation is dependent on collecting data about them that may not accurately reflect their values and political beliefs. There was particular concern about seemingly 'irrelevant' consumption data being used to infer potentially personal or sensitive information, such as consumers' sexuality, religion and political preferences and beliefs.	Consumers could not see any benefit to themselves or to society of this type of targeting.	Most consumers felt that this practice could present serious harm at both the individual and the societal level. In the context of ongoing press coverage of international interference in major elections, many felt that it was entirely possible that citizens' beliefs and voting patterns could be 'manipulated' by governments and major organisations. This was a particular concern for young generations, who might not have a history of forming their views of major issues and political parties before this information was tailored and targeted.

10 Workshop stimulus about the personalisation of job adverts was based on the following article: https://www.theguardian.com/technology/2015/jul/08/women-less-likely-ads-high-paid-jobs-google-study. See appendix for further information.

11 Please note that all research pre-dates mainstream media coverage of Cambridge Analytica in mid-March 2018.

12 Cambridge Analytica: https://www.youtube.com/watch?v=lBgHrn-TrD8

Reflecting these concerns, the trade-off participants considered in relation to this issue represented the only example across the workshops where consumers opted for greater control of their data, even if this might come at a (small) personal financial cost:

Participants considered the following trade-off: you could get cheaper insurance but to do so you would have to allow insurance providers to access your Facebook account. Would you do this? (*NB this question is based on a proposal made by an insurance company that was not taken forward*)

A majority of consumers in deliberative workshops said that they were uncomfortable about using social data to determine their insurance premiums. Consumers felt that:

- **Information from their social media profiles is irrelevant to their insurance products** – most consumers struggle to see how information gathered from their Facebook or Twitter profiles would be relevant to their risk profile and therefore feel that any premium they received as a result of this would be inaccurate. Some feel, more generally, that it is inappropriate for organisations to repurpose information shared for fun on social media for other purposes.

  "It's irrelevant and inaccurate. How many exclamation marks someone uses shouldn't make a difference."

  Workshop participant, Newport

- **They would lose control of their ability to secure a good deal and improve their risk profile** – most consumers do not understand what data could be used to make inferences about them and therefore feel unable to affect or control the decisions that are made. This particularly applies if information taken into account of premium calculations includes content shared on social media by their friends and contacts.

  "Why should I be judged for what my friend might have tagged me in?"

  Workshop participant, Perth

However, a minority of "Tolerant" consumers were more open to this idea, believing that:

- **Their data profile would not disadvantage them personally** – because their behaviour, and that of their friends, is generally 'sensible' and might see them offered a lower premium.

  "I don't put anything on Facebook anyway – I'm pretty sure I'd be fine."

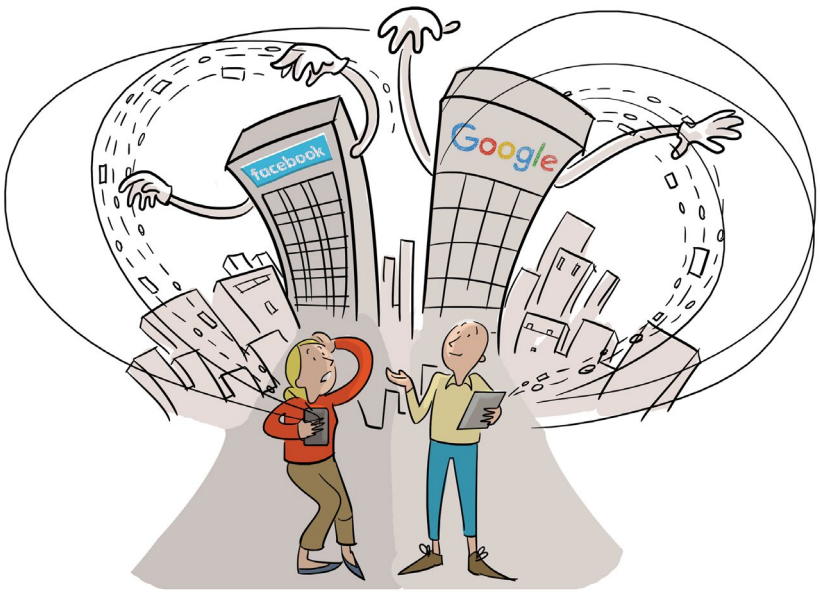
  Workshop participant, Newport

- This service would be opt in – allowing consumers choice and another potential avenue through which to save money and find a good deal.

  "If it's a choice then it's okay – as long as you can choose which insurer you want to go to."

  Workshop participant, Perth

8. Individual data and consumer choice



There are a small number of companies which are front-of-mind for consumers when they consider and discuss data collection and sharing in detail: Google, Facebook, and to a far lesser extent, Amazon and other major technology firms such as Apple and Microsoft. Google and Facebook seem to be so dominant in consumers' minds on this issue because of a low-level awareness that these businesses 'only' deal with data: in consumers' minds, they don't 'make' or 'sell' products (unlike Apple and Microsoft, who are conceived of making most of their money from selling hardware and software).

When exposed to information about the size of these companies and their ownership of other providers, in deliberative workshops some consumers were surprised that they had underestimated the dominance of companies which they already consider to be very large and powerful. Despite this, the vast majority of consumers struggle to connect the size and dominance of these companies to their access to consumer data. Even when they consider the topic in some depth, and are exposed to information about data-dependent business models (e.g. data brokers solely engaging in collecting and sharing consumer data), consumers struggle to understand the value of their data and the different ways in which their data is being monetised. As a result, most find it difficult to understand how access to consumer data could give a company a competitive advantage, beyond the huge advertising profits they might be able to garner.

> *"There are 4 or 5 massive American companies that are influencing the rest of the world, and it's getting less and less. There seemed to be dozens of search engines at one time and now there are two or three."*
>
> Workshop participant, Leeds

Because they are starting from a place of very limited understanding, many consumers 'fall back' onto their broader worldview and pre-existing views on free markets and competition when considering the implications of the dominance of a small number of firms for consumer choice, generally expressing either one of two perspectives:

- **Those who broadly accept the dominance of a few large companies** – for many, this is considered a necessary part of free market competition and a 'fact of life' that some organisations succeed while others fail. For some, this opinion is driven by the belief that the services and prices provided by these large organisations are likely to be better than their smaller competitors.
    - Some cite Google's search engine as an example of 'legitimate' success in a competitive market. These consumers perceive Google to be so dominant in its 'sector' because its search engine is more efficient and more aesthetically pleasing than its competitors'.
    - Others point to changes in the social media market as evidence that consumer choice is still prevalent, arguing that the declining popularity of Facebook, recent rise of Instagram and Snapchat, and 'extinction' of MySpace and Bebo demonstrate that consumers can 'vote with their feet' if they are unhappy with an online service.
- **Those who are pro-competition in principle and are concerned about the future technology 'landscape'** – this, smaller, group of consumers express the view that competition is important in principle to ensure that companies remain accountable to their customers. This perception is reinforced by concerns about the corporate behaviour of dominant technology firms, for example in relation to corporate tax.
    - This group often feel that it is a very real possibility that in a decade there could be an even smaller number of dominant firms as they acquire one another over time, with very negative impacts for consumer choice.

The perspective of vulnerable consumers: generally speaking, vulnerable consumers tend to fall into the second 'camp' and were concerned by the dominance of a small number of larger organisations. This is particularly strongly felt by those with health or mobility conditions, who may be more dependent upon internet-enabled devices and services. Some of these participants spontaneously raised the concern that their dependence on technology and higher-than-average use of these services would lead to a particularly accurate data profile being aggregated by those providers.

> *"Wait, so that means if I use WhatsApp and Instagram, then Facebook can see all that stuff too? That's so scary – they must know absolutely everything about me."*
>
> DE consumer, Leeds

Consumers' consideration of a potential trade-off in this area emphasised the difficulty for consumers in judging what 'good' looks like in the context of major organisations collecting and sharing consumer data:

Participants considered the following trade-off: The professional social network LinkedIn (owned by Microsoft) is the largest of its kind and far ahead of its competitors. Last year a judge in the United States ruled that it must allow a third party company to 'scrape' data publicly posted by users, allowing it to compete. LinkedIn said that it was 'disappointed in the court's ruling', and that they would continue to fight to protect their members' ability to control the information they make available on LinkedIn. Was the judge right or wrong?

In considering this trade off, consumers are more likely to agree that the judge was wrong to make the decision to allow the data scraping. Indeed, some felt concern around the fact that this judge alone had the power to make this determination that affected their data.

Consumers tend to take this view for one of two reasons:

1. **It means that consumers lose control over who has access to their data** – a concern fuelled for many by the perception that third parties and data sharing is in principle negative and rarely benefits the consumer. This viewpoint drove some participants who were otherwise concerned about dominance of a small number of technology companies to decide that the judge was wrong to rule in this way.

"We are for competition, but in this instance the judge ruling is wrong. The individual has the right to share information with just LinkedIn only."

Workshop participant, Leeds

2. **It is unfair to ask a business to give away their competitive advantage** – because more consumers fall into the mind-set of feeling positive about free market competition overall, a significant proportion take the view that it is unfair to ask a business to give up their competitive advantage in order to stimulate competition. Some felt that LinkedIn was being punished for its success.

"LinkedIn have built up the business and it's been taken away from them. It's a disgrace!"

Workshop participant, Perth

9. Conclusions: what next?

9.1: The benefits of innovation versus consumers' concerns about their data

As explored at the very start of this report, when consumers reflect back on how their lives have changed as a result of data-dependent devices and services, they almost always focus on the huge benefits that these devices and services have brought to their lives. But they also talk about the rapidity of the pace of change, and the fact that technologies that they would have thought impossible just a decade or even a few years ago are now commonplace. For the "Liberal" and "Tolerant" segments (equivalent to just under half of the population in total), this pace of change is exciting and leads to a sense that the more data you share, the more benefits you are likely to receive in return.

> *"If you give more, you get more. For example, Facebook. You share information and at the same time you get something back, communicating with people."*
>
> Workshop participant, St Albans

> *"It's a drawback [data sharing] but it is to be expected as we get it for free. Facebook wouldn't have made millions and millions without people sharing their data…the thing is some of these apps are a necessity for me now as well"*
>
> Workshop participant, Perth

> *"I realise that I'm using my data as like a fee to use these free services, and I don't really mind. I find most of this data collection quite acceptable because I don't mind paying with the currency of data"*
>
> Workshop participant, Leeds

For other consumers, and particularly those from "Concerned" and "Anxious" segments (equivalent to just over half of the population in total), the pace of innovation can feel overwhelming. They can find themselves 'catching up' with innovation and engaging in products and services which they don't necessarily have a choice but to use if they want to participate fully in modern life and society. These consumers still love the benefits that technology and innovation has brought them, but say that they might choose to engage with organisations like Facebook and Google differently now had they been aware ten years ago of the implications of data collection and sharing.

> *"Privacy should always be under review, there should be greater accountability."*
>
> Workshop participant, Perth

"Privacy is being compromised, people need to be made more aware of what is happening to their information. We need more education – people need to be told about what is going on."

Workshop participant, Newport

Participants' consideration of emerging, innovative technologies such as facial recognition software illustrated where the majority of consumers fall out on this issue:

Participants considered the following trade-off: companies are using individual data to develop a range of innovations but risks of privacy breaches might increase. Is the balance too far towards innovation or privacy?

While consumers agree that innovation is important, most felt that the balance is currently too far in favour of this at the expense of concerns about their data, such as their privacy.

In particular, learning about the advances in and the use of facial recognition technology led to concern about what information about them will be collected without their consent in the future, seriously limiting their ability to control who has access to what data.

As a result, consumers called for greater regulation of the amount of data that is collected and shared about them, while some also hoped that innovation could lead to improvements to data security systems.

"At the moment, it's too far balanced towards innovation. It needs to be rebalanced."

Workshop participant, Perth

"In the future, I just think that data sharing will go nuts, it will be channelled further. Those lights in Piccadilly are the start of what the future will be."

Workshop participant, Newport

"There's nothing bad to know about me, but it's still scary – there's nothing about you that is private. I don't like it, and I think it's just going to get worse. That's progress for you. I don't like it."

Consumer with a disability, Leeds

9.2: Taking action

For most consumers, there is little sense that it is up to them to redress the balance, and little belief that the system is likely to change. Some consumers – particularly those who initially fell into the "Anxious" segment – may feel empowered to take action, such as monitoring privacy settings on social media, Google and apps, or using tools such as Ghostery to monitor and potentially block trackers on websites. However, these actions feel relatively small-scale compared to the scale of the issue and extent to which data collection takes place. The most negative and disempowered also question how effective these measures really are and believe that organisations will likely find a way around them to access their data.

There are three key barriers to (concerned) consumers taking action themselves and breaking out of their sense of resignation when it comes to data collection and sharing:

- It is **hard for consumers to imagine what 'better' looks like** when they believe that they understand only part of the picture when it comes to data collection and sharing, and when they are unwilling to give up the data-dependent devices and services they now rely upon and have become accustomed to accessing for free or at low cost.
- They believe that their **information is already out there and that there is little that they can do to regain control**. There is little spontaneous understanding that data becomes outdated and therefore depreciates in value and that there may still be merit in protecting new or more recent information.
- The **power balance feels weighted against the consumer**, right from the way in which consumers believe that they are asked to give permission to share their data, to the scale and complexity of the data ecosystem (particularly the existence of a whole sector of organisations profiting from consumer data the average person is felt to know nothing about).
    - Strikingly, even by the end of the deliberative workshops, when participants had been exposed to a lot of information about data collection and sharing over an extended period, the majority still struggled to conceive of the 'value' of their data. This exchange felt fundamentally unequal and untransparent when consumers as individuals cannot monetise their own data.
    - Moreover, for some consumers there is a sense that the organisations that they might expect to hold organisations to account and protect their best interests – government, regulators and consumer bodies – have been behind the curve with regards to data collection practices. Virtually no participants had any awareness of upcoming changes to data protection legislation which might lead them to re-evaluate these views.

"We need to ensure that the law reflects the reality. Current laws about this feel outdated – like they're from the 70s."

Workshop participant, Perth

"It's not strictly regulated. If you look at the financial services sector, it is strictly regulated. I don't think this area is tightly regulated yet."

Workshop participant, St Albans

Within the context of these limitations, workshop participants developed a set of recommendations for ways in which the dial could be shifted to ensure that data collection and sharing does not 'cross the line' of acceptability:

Control	Ensuring that consumers feel informed and in control of the data being collection about them by: • Raising awareness of existing tools to help consumers to manage their data, from managing privacy settings to using ad blockers to using non-personalised search engines • Shifting the emphasis from data collection being 'opt out' to 'opt in' by re-framing terms and conditions and the way in which consumers initially grant access to their data
Relevance	Ensuring that the information being gathered about consumers always has some relevance to the context in which it is being collected. • Where technical information is required for a product or service to function (e.g. access to IP addresses), consumers wants, in theory, to be able to access simple information about why this type of information is necessary.
Benefits	Ensuring that there is always some tangible benefit of data collection and sharing: • Either directly to the consumer, e.g. by giving them access to a product or service for 'free', or by improving their user experience • And/or to society more broadly, such as using data for the 'public good', including public safety
Harms	Ensuring that data collection and sharing cannot cause consumers tangible harm: • Consumers expect government, regulators and consumer bodies to 'catch up' with major technology firms and other organisations collecting their data to hold them to account, ensure that they aren't actively operating against consumers' best interests and draw consumers' attention to practices which could genuinely cause them tangible harm.

10. Appendix

Sample frame outlining the spread of participants attending the workshops:

- **Sex:** Minimum of 12 male participants and 12 female participants at each workshop to ensure a spread.
- **Age:** An even spread within the following three age groups at each workshop: 18-34; 35-54; 55 and over.
- **SEG:** A spread across all socio-economic groups at each workshop, reflective of the local population.
- **Life-stage:** A spread of respondents who are pre-children, with children living at home, with children living away from home, and who have never had children.
- **Ethnicity:** At least 10 BAME respondents attended each workshop.
- **Location:** At least 6 respondents in each workshop live in small towns/ rural areas.
- **Perceived levels of knowledge of data collection:** A spread of perceived knowledge within each age group.
- **Levels of comfort about data sharing and collection:** A spread of levels of comfort within each age group.

Detailed information about the deliberative phase workshop agenda, and information that participants were provided with:

SESSION	OUTLINE
The role of technology	Participants shared the technology, websites and apps that they use the most, and discussed the things that they are able to do today that they were unable to do a decade ago as a result of technological innovation.
Data collection: what is being collected	Participants were provided with the following examples of everyday products and services, and the consumer data that they collect: • Mobile apps/ Facebook • Smart TVs • Public Wi-Fi • Smart travelcards • Activity trackers Participants discussed this information and considered the extent to which different types of data collection are acceptable or unacceptable.

Data collection: how data is observed	Participants were provided with information about cookies, fingerprinting, and a demonstration of Ghostery https://www.ghostery.com/). Ghostery is an extension for web browsers that allows the user to see whether there are any trackers on the website that they are visiting. The extension also blocks many third-party data-tracking technologies and can anonymise the user's data to further protect their privacy. Participants discussed their thoughts and concerns about data being observed about them in the light of this information.
What happens to this data: inferring and building an individual profile	Participants were provided with information about inferences, psychographic profiling, data being combined to make an individual profile, and the idea of the 'digital self'. They were also encouraged to use the Apply Magic Sauce tool to understand how inferences might be made about them (https://applymagicsauce.com/demo.html#). Apply Magic Sauce is an online tool which creates predictions based on a consumer's online activity. The tool predicts consumers' age, gender, sexuality, personality traits and preferences, amongst other things, using behavioural data from their Facebook and Twitter profiles. Participants outlined their responses to this information and considered the extent to which they felt inferences and profiling is acceptable or unacceptable.
What happens to this data: third party sharing	Participants were provided with information about the data ecosystem and three examples of data sharing: • Android apps sharing information about their users with other organisations • Price comparison websites and the data they share about users to retrieve quotes • Data brokers and the data profiles they create and sell on to third parties Participants discussed their responses to this information and considered the extent to which data sharing is acceptable or unacceptable.
Security of information and stolen data	Participants discussed how likely or not they felt it is that their information could be stolen, and how much of a concern this is. They were then encouraged to use 'Have I Been Pwned?' website (https://haveibeenpwned.com) to see if their information has been breached, and reflect on their results. 'Have I Been Pwned?' is an online tool that allows internet users to check if their personal data has been involved in a data breach, by cross checking their email address against a large collection of database dumps and pastes.
Tailoring and targeting of adverts, recommendations, prices and information	Participants were presented with examples of targeted adverts, recommendations, personalised pricing, and tailored information, including: • Targeted advertising on Facebook • Netflix video recommendations • Push notifications on smartphones • 'Piccadilly Lights' using facial recognition technology to target adverts (https://www.theverge.com/2017/10/16/16468452/screen-london-picadilly-circus-cars-targeted-ads-landsec) • Tesco's plan in 2013 to install facial-scanning technology to target adverts to its petrol station customers (http://www.bbc.co.uk/news/technology-24803378) • Vending machines in Japan using facial recognition technology to recommend products to users (http://theweek.com/articles/489132/japans-facescanning-vending-machines) • The proposal (later abandoned) from Admiral Car Insurance to use social data when determining car insurance premiums (https://www.admiral.com/black-box-insurance/) • In-app purchases based on assumptions about likelihood of spending money • Tailoring of search engine results and the 'filter bubble' (https://vimeo.com/51181384) • Tailoring of political campaign adverts/ Cambridge Analytica (https://www.youtube.com/watch?v=IBgHrn-TrD8) Participants discussed the relative acceptability of different forms of targeting and considered the extent to which these practices have a negative impact on consumer choice.
Choice in markets and innovation	Participants considered the extent to which the amount of choice available to them as consumers has changed over the past decade, and were provided with information about the companies/apps owned by Google, Facebook and Amazon. They considered the role that data collection has to play in innovation, and assessed the extent to which this innovation is 'worth' the potential risks to consumers' privacy.

Which?, 2 Marylebone Road,
London NW1 4DF
Phone +44 (0)20 7770 7000
Fax +44 (0)20 7770 7600

For more information please contact:
Harriet Pickles – harriet.pickles@which.co.uk