2018

Quantified Self Report Card

AUTHORS:

Rochelle Fairfield
Heather Mann

Human Data Commons Foundation | www.humandatacommons.org

Project Overview

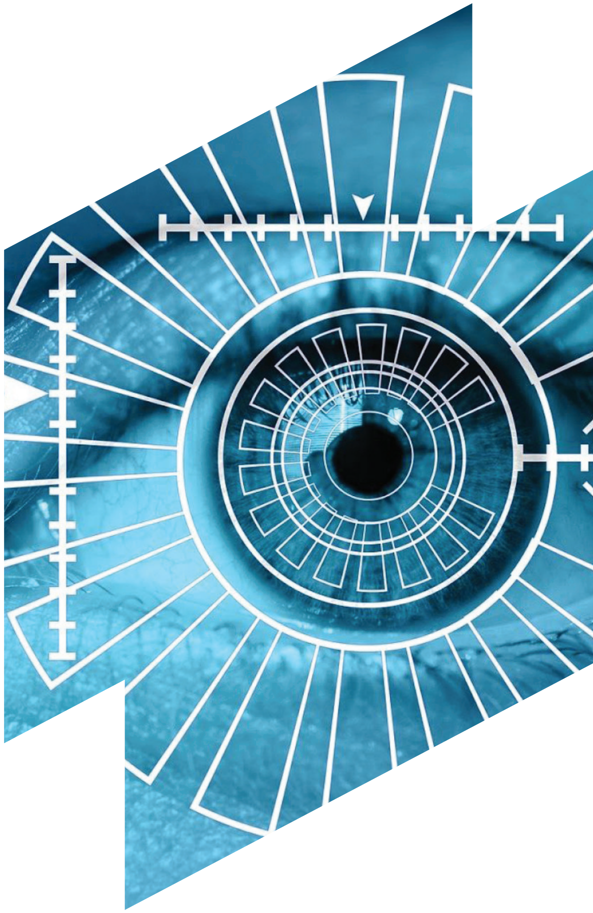
As digital technology steadily transforms business, communications and day-to-day life, we are producing data at an unprecedented rate.

The Quantified Self movement highlights how digital technology and data allow us to track our very selves – our activity, sleep, fitness, and fertility – offering the promise of "self knowledge through self-tracking".[1] Numerous markets have arisen to harness the power and promise of the Quantified Self, from wearable devices to applications, integration platforms, and middleware. But what happens with all this data? What do users need to know when it comes to their privacy and their digital rights?

The 2018 Quantified Self Report Card is meant to inform the everyday user about common industry practices around the collection and storage of their physiological data, including bio-markers for emotion and brain wave data. It provides consumers with a summary of our research into the Terms of Service and Privacy Policies of 57 companies in the Quantified Self arena, from startups to major conglomerates. We rated companies on three categories: User Rights, Data Collection & Storage, and Third Party Sharing, and ranked them within their industries.

We first set the context for the report card by covering significant developments around data privacy matters and digital rights for 2018, including the the Facebook / Cambridge Analytica privacy scandal and the European Union's General Data Protection Regulation. Adding colour to our ratings, we provide excerpts from policies and discuss pertinent issues relating to digital rights, from what constitutes "personal data" to how companies are using and abusing their power over customers.

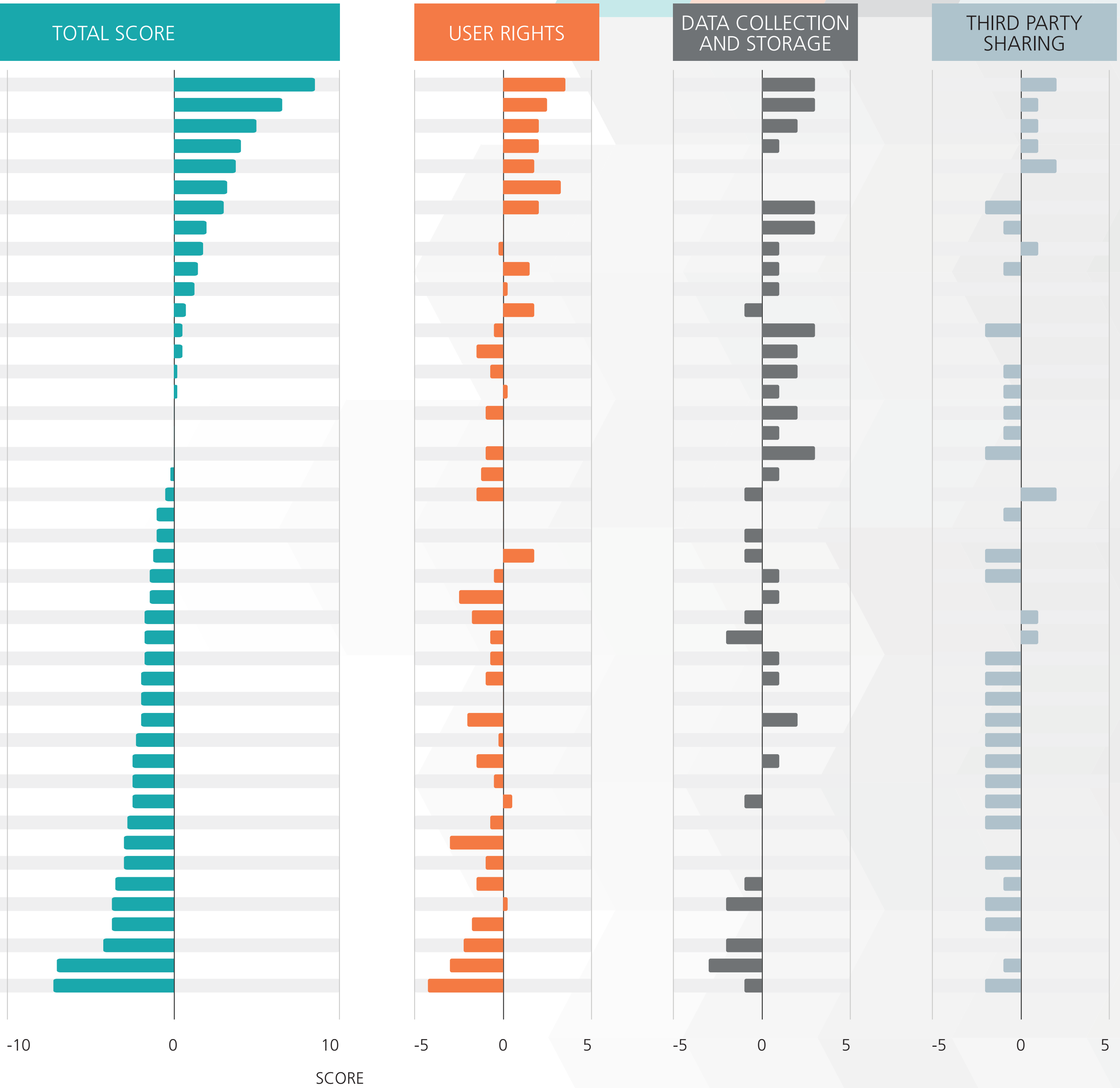
We hope that after reading this report card, users will have greater awareness of how they are giving data away, to whom, the implications of that, and some of the things they can do about it.



Results

Data Privacy Rankings > **Summary**

COMPANY	TOTAL SCORE
Clue	8.5
Human API	6.5
Decadoo	5
Runtastic	4
Tom Tom Sports	3.75
Google	3.25
Fitbit	3
Motivation Science	2
BioBeats	1.75
Nutrino	1.5
Affectiva	1.25
Nokia	0.75
23andMe	0.5
Huawei	0.5
Michael Kors	0.25
FX Development	0.25
LibreView	0
X2AI	0
MyZone	0
Pacifica	-0.25
SAM App	-0.5
Interaxon	-1
Microsoft	-1
Emotiv	-1.25
PredictBGL	-1.5
Samsung	-1.5
Xiaomi	-1.75
Exist	-1.75
Beyond Verbal	-1.75
bloomlife	-2
Consumer Sleep Solutions	-2
Mio Global	-2
Fossil Group	-2.25
Garmin	-2.5
GlobalBeat	-2.5
MAD Apparel	-2.5
Suunto	-2.75
HeartMath	-3
Moov	-3
Qualcomm	-3.5
Peakswre	-3.75
BLOCKS	-3.75
Wahoo Fitness	-4.25
Soleus Running	-7
Mobvoi	-7.25
*Apple	n/a



*Apple's Privacy Policy and Terms of Service weren't able to be rated as they pertained to their website, not their wearables.

Table of Contents

Introduction

If you've heard the term Big Data bandied about but realize you don't really know what the Big Deal is, you're not alone.

You probably know that whenever you engage with digital apps, tools, and platforms, you produce data. However, you are likely more focused on using these technologies than on how they might be using you.

The Quantified Self is a movement inspired by how technology provides new opportunities to track our very selves – our activity, sleep, heart rate, and more. The promise of this movement is "self-knowledge through self-tracking".[1] Quantified self applications, platforms, and devices (often called "wearables") produce data (i.e. information), and lots of it: from biological info, to location info, to web-browsing behavior, and beyond.

Who cares where you are and how fast your heart is going? Why is that valuable data to anyone besides you? One reason is that data is required for Artificial Intelligence (AI) to learn on. Data is like food to a growing baby – AI algorithms need lots of it to grow and get smarter. The more data you feed to AI, the more quickly it learns and becomes more reliable and useful and potentially misused. Currently AI is in a massive growth spurt globally. It's is a boom industry, so there is lots of hype, money to be made, industry dominance to be competed for, prestige and notoriety to be gained. This business climate and culture makes it challenging for industry to be the best stewards of your data that they could be.

When it comes to industry practices regarding user data, unless you work in tech, it would be expected that you feel in over your head. How can the average person keep up with everything there is to know, especially something that moves at the breakneck speed of technology? It would feel good to have some sense of what the real deal is behind the buzz with Big Data and AI – and how these concepts relate to the wearables you might be using or thinking of buying.

This report card is designed to widen your awareness of common practices regarding your data in the Quantified Self industry. We present a summary of our ratings of 57 companies, including devices, applications, integration platforms, middleware analytics companies, and conglomerates (a fifth category that overlaps across the other four categories).

After presenting the ratings of selected companies, we discuss the implications of common industry practices around user rights and data privacy. To be sure, there are many exciting, humanity-serving possibilities for technology, big data, and AI. Indeed, the Quantified Self movement is keen on exploring "what new tools of self-tracking are good for".[2] Overall, industry provides exposure to many of the pleasant possibilities of advances in AI and tech, so the good news won't be repeated here. Instead, we'll focus on areas for improvement and areas of concern.

Introduction (continued)

Too Long; Didn't Read

> *"One thing people dislike more than their privacy being invaded is trying to understand how it can be invaded."*
> – Greg McMullen, personal interview

You probably guessed it – most people simply don't bother to read the contracts they sign online. A 2017 study by Deloitte of 2000 American consumers found that 91% of consumers said they did not read the terms of service and privacy policy for applications, software updates, or online services before agreeing to them. For Millennials aged 18-34, that number was 97%.[3]

We suspect these numbers may be underestimates. Way back in 2008, one research group estimated that it would take an average web surfer 244 hours to read through all the privacy policies that they visit in a year.[4] If you read the policies for 8 hours a day, that's one whole month of reading privacy policies every day! And this was before the rise of smartphones with their many applications, as well as most wearables.

Ironically, companies now have the capacity to assess with reasonable accuracy whether their users actually read their contracts. They have the ability to track how long you spend on web pages, what you click on, and in some cases, eye movements to know where you are looking on a page. These capacities are primarily used for commercial purposes, for example to improve the services offered, and to sell more customized advertising. Companies could use these capacities to know with reasonable accuracy whether you actually read the legal agreements before signing, as a means of ensuring the integrity of the contracts. However, we have yet to hear of a single company using tracking technology in this way.

We see an opening here for revolutionary ethical leadership from industry. Companies could ensure their contracts were read, and not allow customers to use the product or service unless they really do read the policy they are claiming to have read. The first company that starts doing this will be setting a high and highly ethical bar for industry, and taking a bold first step towards fixing the broader power imbalance problem inherent in much of current industry practice.

2018: The Year of Data Privacy Awareness?

2018 was an interesting year for data privacy, marked by two key events. First, in March, news broke that Cambridge Analytica had harvested personal data from from tens of millions of people's Facebook accounts without their consent to use for political purposes. Next, in May, a milestone policy went into effect: Europe's General Data Protection Regulation, or GDPR for short.

Data Leak! The Facebook-Cambridge Analytica Data Scandal

The Facebook-Cambridge Analytica scandal can be traced back to 2014, when a researcher named Alexsandr Kogan developed an application for Facebook called "This is Your Digital Life". This application was a personality quiz, and it was downloaded by about 270,000 Facebook users. As was the case for any Facebook application at the time, the app allowed Kogan to access data from the profiles of those who downloaded the app as well as the profiles of all their Facebook friends. This meant that data was collected on millions of Facebook users without their consent.

Rather than deleting this data, Kogan kept it stored on a private database. Cambridge Analytica, a voter profiling company that worked with Donald Trump's election campaign as well as Britain's Brexit campaign, paid Kogan to acquire the data. The company then used the data to build a software program that could predict people's voting tendencies and influence them with personalized political ads.[5] In 2018, Christopher Wylie who formerly worked at Cambridge Analytica, informed the press of these activities, setting off a justified media storm.

Introduction (continued)

The Facebook-Cambridge Analytica scandal has been described as a watershed moment in data privacy awareness. Facebook lost $100 billion off the value of its shares in a matter of days, and Facebook CEO Mark Zuckerberg was summoned to meet with US lawmakers to discuss Facebook's culpability in the matter.

The GDPR Raises the Bar for Digital Rights

While the Cambridge Analytica-Facebook privacy scandal received much media attention, the average North American citizen was not likely paying much attention when the European Union's General Data Protection Regulation (or GDPR) went into effect in May, 2018. Some people may have noticed an influx of requests to agree to updated terms of service and privacy policies. However, many assumed this was in response to the Facebook-Cambridge Analytica breach or another kind of scandal. Few were aware that it was largely a response to the GDPR becoming enforceable on May 25, 2018.

Yet, this legislation is quite significant: it has been referred to as "the most important change in data privacy laws in 20 years".[6] This piece of legislation did two important things. First, it created a common standard for handling data across European Union countries. Prior to that, each country had different laws; the GDPR harmonized them. Second, and most importantly for the individual, it made an people's autonomy and right to privacy as it relates to their personal information much stronger.

Although the GDPR only applies to EU citizens, any company that deals with data from EU citizens companies must adhere to its policies. The GDPR is setting a new standard for the tech industry globally, and not a moment too soon. For more on the GDPR: https://eugdpr.org/.

What Constitutes Personal Information?

The question of what constitutes personal information is central to discussions around data privacy and security. In reviewing policies for the Quantified Self Report Card, we had considerable discussion on this question as well.

The GDPR offers the following definition of personal data:

> *"'[P]ersonal data' means any information relating to an identified or identifiable natural person ('data subject')."*[7]

We see that as a good definition, but one that poses a significant challenge: namely that the word "identifiable" is open to interpretation. Importantly, the landscape has changed when it comes to how easy it is to identify someone from their data.

Before the era of big data, data was generally considered secured if identifiers such as name or birthday were removed. However, with big data, removing direct identifiers is no longer sufficient to render an individual's data anonymous. As noted in MIT Technology Review,

> *"What modern data science is finding is that nearly any type of data can be used, much like a fingerprint, to identify the person who created it: your choice of movies on Netflix, the location signals emitted by your cell phone, even your pattern of walking as recorded by a surveillance camera."*[8]

Technology is changing the playing field for how easily people can be identified from their data.

Adding some nuance to this discussion, the GPDR distinguishes between *pseudonymized* and *anonymized* data. Pseudonymous data is data in which direct identifiers have been masked. The GDPR considers pseudonymous data to be "personal data" because it relates to identifiable persons, and therefore falls within the scope of the policy.

Introduction (continued)

One step beyond pseudonymized data is anonymized data, which the GDPR defines as "data rendered anonymous in such a way that the data subject is not or no longer identifiable."[9] The GDPR considers anonymized data as outside its scope. However, with the proliferation of big data, such as users generate through wearable devices and apps, it has become increasingly feasible to indirectly identify individuals from their data. While there is still some debate over the matter, it seems that scholars and privacy experts in tech generally agree that true anonymization is no longer realistic in the age of big data. According to Paul Ohm, a law professor at Georgetown University, "data can either be useful or perfectly anonymous but never both."[10] Perhaps the term "anonymization" has become somewhat misleading!

We believe pseudonymization is a good start in protecting personal information, but it is important that companies consider pseudonymized data as containing personal information. For practical purposes, we recommend companies treat any user data as personal data.

In our review of companies' privacy policies and terms of services for this Report Card, we got the sense that many companies continue to define "personal data" much more narrowly than makes sense in this day and age. In the past, personal data was data that included an individual's direct identifiers, such as their name or birthday. At the extreme, some companies appeared to consider data without direct identifiers outside the definition of "personal data" – and outside the scope of their privacy policy.

Adding some complication to our research, it was often difficult to determine what definition of "personal data" companies used. We opted to give companies the benefit of the doubt in terms of what data they considered within the scope of their policies, yet we note this as a red flag and a limitation to our research. If companies operate by a conservative definition of "personal data", they may be abiding by the policy for some user data, while taking liberties around their treatment of other data generated by users.

Research Methodology

Selection of Companies

In selecting companies for the 2018 Quantified Self Report Card, we began with the list of companies included in the 2017 Report Card. To ensure the 2018 list reflected current trends, we also visited Wareable (https://www.wareable.com/), a website authority that reviews wearable devices in the Health & Fitness industry. This website posts "Best Product" lists, identifying the most notable devices in several product categories. We drew companies from the following lists on Wareable: best smartwatches, best fitness trackers, best GPS watches, best heart rate monitor, best sleeptracker, and hot wearables. This search yielded 10 companies listed in the 2017 Quantified Self Report Card and 11 new companies, which were added to our list for the 2018 Report Card.

Through web searching, we identified four additional companies for the 2018 report card: one application (Runtastic), two integration platforms (Qualcomm and Exist), and one middleware company (Vivametrica). We also elected to include 23andMe, a prominent company for coding genetic data. We classified 23andMe as an application; see the following section on classification of companies.

From this broad list, we elected to remove a total of 14 companies for a variety of reasons, including: the company no longer existing, redundancy with other companies on this list, or being somewhat outside our scope. This left us with a list of 57 companies.

Classification of Companies

Companies were classified into four categories: Devices, Applications, Integration Platforms, and Middleware. These classifications were similar to the 2017 report card, with two notable differences. First, we elected to separate the "User Platforms" category into two categories: Applications and Integration Platforms. Second, we elected not to treat conglomerates as a discrete category, but as an additional designation that could apply to any company across the four categories.

We defined the four categories as follows:

Devices – Companies that primarily manufacture equipment and tools used for self-tracking.

Applications – Companies that primarily offer applications (software) and/or websites which allow the user to engage in self-tracking, including through self-reported data, but which do not sell hardware for this purpose.

Integration Platforms – Companies that allow the user to integrate his/her/their health & fitness data from a variety of sources.

Middleware Analytics – Companies that offer analytics and algorithms to other companies who build user-facing services.

In addition to these categories, we designated companies as Conglomerates if they met the following definition:

Conglomerates – Companies that were well-established before venturing into the Quantified Self market, and still provide products and services outside of biometric and self-tracking devices and platforms.

In this year's report card, companies from any of the four categories could also be designated as a Conglomerate.

Research Methodology (continued)

Questions and Rating Criteria

After some discussion, we identified three broad domains by which to evaluate companies on their terms of service and privacy policies concerning their treatment of user data:

User Rights
How "user-friendly" are the companies' policies (Terms of Service and Privacy Policy), and what rights are users given around how their data is processed and used? This domain was also a major area of focus of the GDPR.

Data Collection and Storage
What practices do companies follow in collecting and storing user data? How easy is it for users to follow up with questions around the security of their data, or to address matters in court?

Third Party Sharing
To what extent do companies share user data with other companies, and how clear is their policy on third party sharing?

Our rating form included 10 questions: 5 questions fell under User Rights, 3 questions under Data Collection and Storage, and 2 questions under Third Party Sharing. Multiple response categories were provided for each question (see Appendix A). We reviewed companies one at a time by visiting their website, and used a Google form to rate them according to our scorecard.

The ten questions we used to assess companies were:

Category	Question
User Rights	1. Are policies (Terms of Service and Privacy Policy) easy to find? 2. Are users notified of changes to policies? (Terms of Service and Privacy Policy) 3. Are policies (Terms of Service and Privacy Policy) written in clear and readable language? 4. What are users' rights regarding data access and ownership? 5. Are users given the right to be forgotten?
Data Collection and Storage	6. What contact information does the company provide in case users have questions or concerns related to how their data is processed? 7. Does the company use encryption when transfering data? 8. Is it clear which jurisdiction governs the contract?
Third Party Sharing	9. On reading the policy, how well-contained does users' information seem? 10. On reading the policy, how clear is it on third party data sharing practices?

A link to our research tool, including options and points awarded is included in Appendix A.

Research Methodology (continued)

Rating Scheme and Scoring

Our rating scheme awarded points for user-friendly policies/practices, and deducted points for unfriendly practices. Companies could receive a maximum score of 1 point on each question and a minimum score of -1 point. Responses were tallied and scored after all companies had been reviewed and rated by both raters. The raters' scores were tallied across the ten questions, and each company's final score was computed by averaging the two raters' scores. This meant the maximum overall score a company could achieve was 10, while the lowest possible score was -10.

After rating the companies independently, we tallied our responses. For each question, we noted any items that were highly discrepant between raters (i.e. more than 1 point apart), and reviewed our responses on these items along with the relevant parts of the policy. In some cases, one (or both) of us elected to change our response. In other cases, we stood by our responses in spite of the discrepancy. This highlights that the response options do not represent solely objective categories, but also involves subjective assessment around the meaning of the information contained within the policies.

Exclusion Criteria

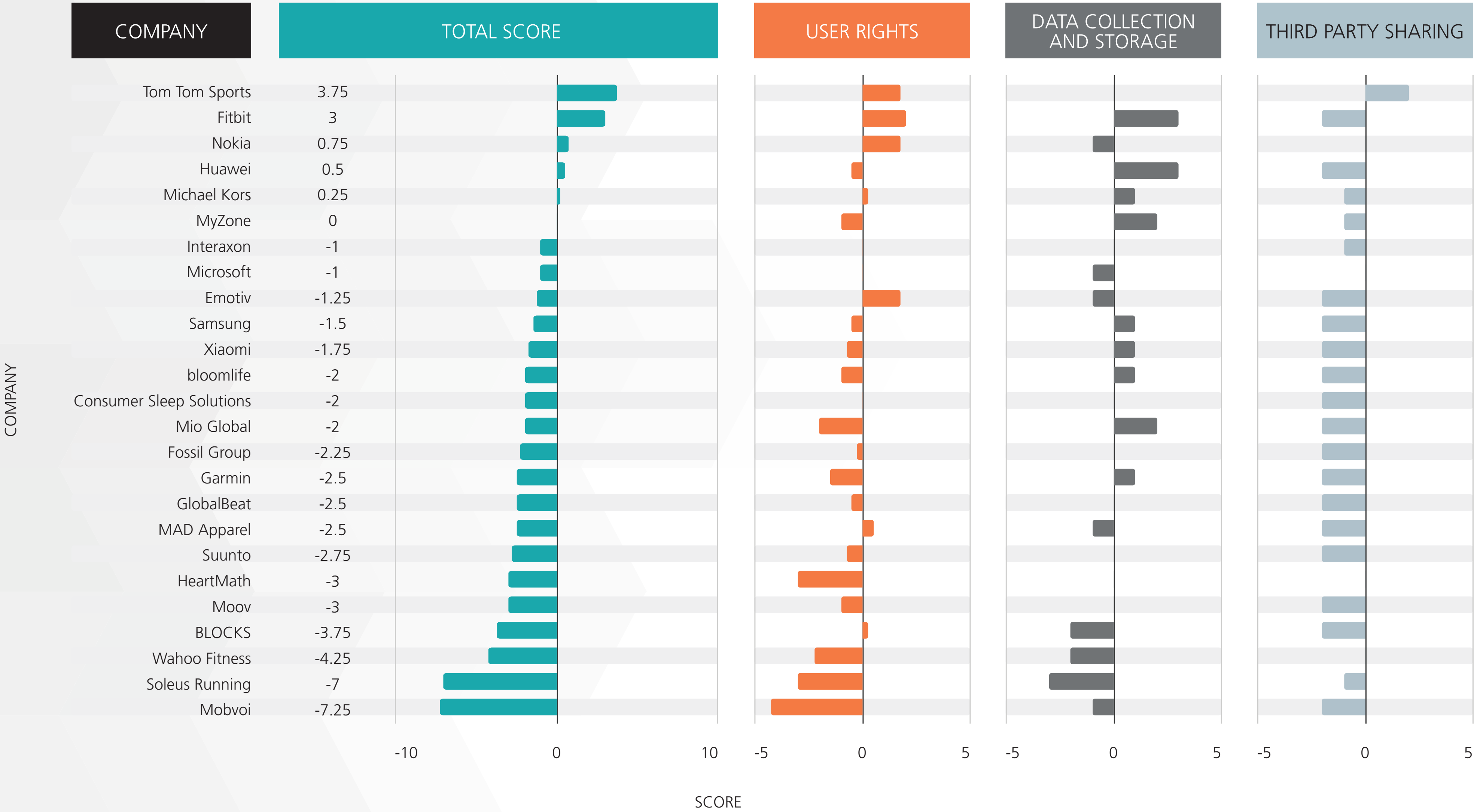
During our research, we found that for a number of companies, the terms of service and privacy policies available on the website referred only to user data collected through the website, rather than data collected through the company's products or services. We opted to flag these companies and exclude them from our analysis. In addition, we flagged one company where we were unable to find links to the policies on the website, and another where the links did not work.

In addition, when attempting to rate Apple Health, we found it difficult to identify which policy on the website pertained to this application. Apple posts a number of policies, pertaining to different operating systems and applications. For this reason, we opted to flag and exclude Apple Health from our analysis.

For a full list of the companies excluded and the reasons for exclusion, please see Appendix B. We note that our ratings were completed during a particular time frame (March-May 2018), and that companies may have updated their website or policies since the time of our ratings.

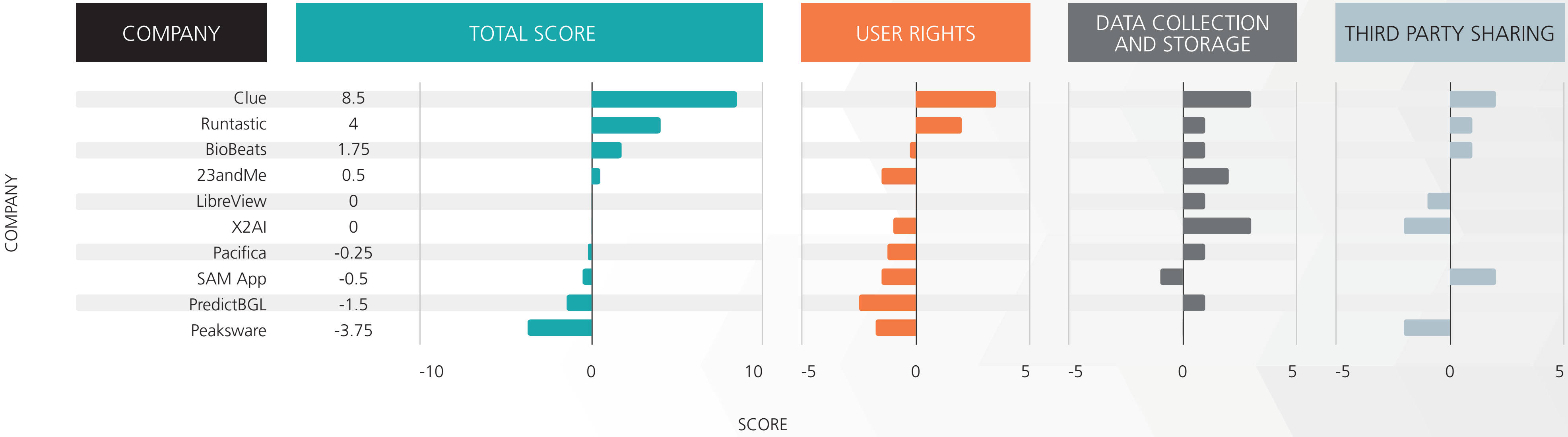
Results

Data Privacy Rankings > **Device Companies**



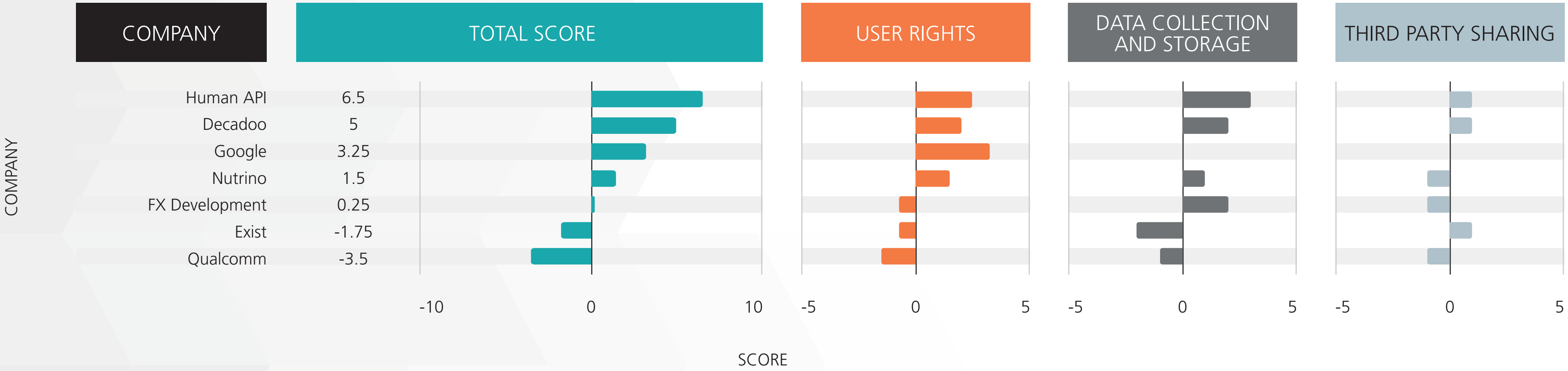
Results

Data Privacy Rankings > **App Companies**



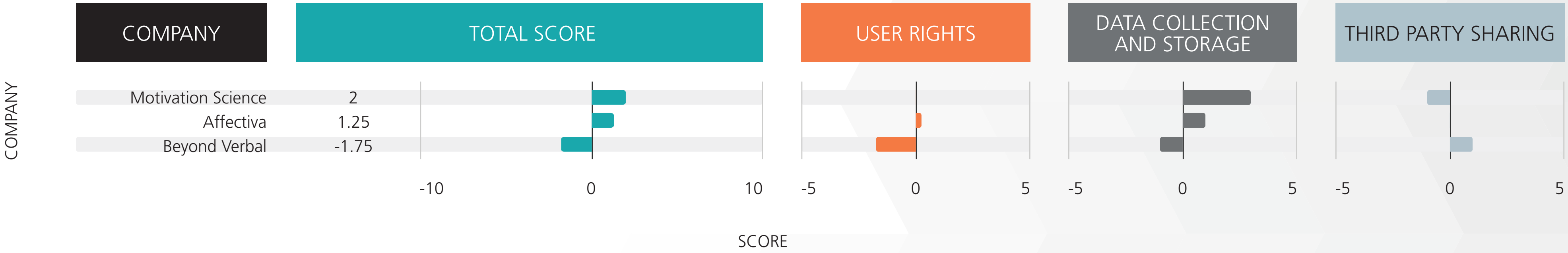
Results

Data Privacy Rankings > **Integration Platform Companies**



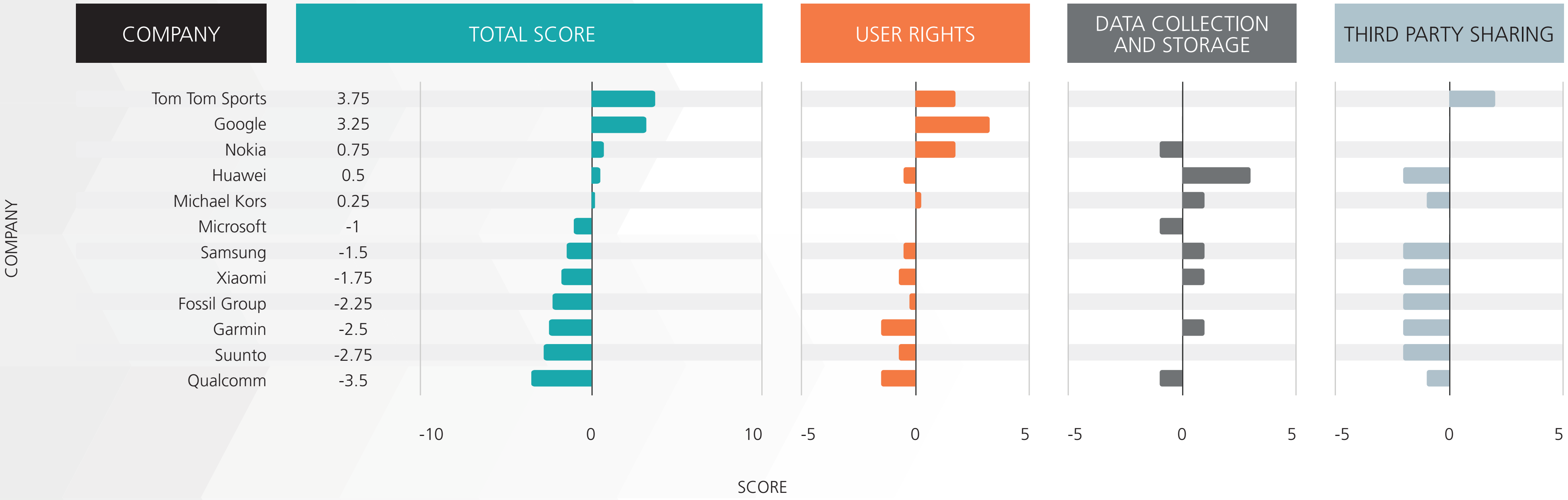
Results

Data Privacy Rankings > **Middleware Companies**



Results

Data Privacy Rankings > **Conglomerates**



Discussion

As the ratings indicate, companies vary in what measures they take to protect your rights, safeguard your data, and disclose who your data is shared with.

What does this mean for what is expected of the average person navigating these policies? What does it show us about the current state of industry praxis relating to privacy, user rights and autonomy, corporate reach, privilege and power? And how can users and industry help move towards a brighter rather than a more dystopic future?

In the following pages, we consider these questions in the context of our research.

Getting Real on "Informed Consent"

We first consider a simple question with a rather elaborate answer: What is the demand placed on you, as a user, in order to knowledgeably and meaningfully provide informed consent on the vast majority of these contracts?

1. Given the international reach and fast pace of tech, you'll need to be fairly well versed in and up-to-date with international law. This background is needed to assess the policies of the latest third party providers your app is sub-contracting to, to ascertain which country's laws govern the contract between the third party and the company you contracted with, and to be knowledgeable about the pertinent laws in relevant jurisdictions.

2. You will also need a good understanding of commerce and technology. In order to keep up with the cascading legal and business implications, you'd be advised to track the latter on at least a bi-weekly basis to keep up with advances in technology, AI, and so on.

3. Once you've updated yourself on tech, law, business and their interactions at the international level, you'll need to regularly re-read most of the terms of service and privacy policies for each application, software program, social media platform, biometric tracking device, navigation system, etc. that you use, since many of them put the onus on the user to check back for changes to the policy. In 2008 this would have taken about one month per year. With tech's exponential growth in the past decade, the estimate would now be significantly higher. It's likely that it would now take you months of doing nothing but reading policies in order to stay up-to-date!

On a brighter note, industry has an unprecedented opportunity to put its considerable capacities towards making policies more user friendly and also fair for users. Industry certainly has a role to play in making "informed consent" more than a nice idea. In our review of companies in the Quantified Self domain, we found that privacy policies were in general more readable and user-friendly than terms of service policies. However, since users are asked to agree to both, terms of service policies need to "catch up" to privacy policies, and they have a long way to go.

We also note that some companies are showing real leadership in this area. We direct the interested reader to Clue by Berlin-based technology company Biowink GmbH, a menstrual tracking app. Clue was our top scorer, and provides an example of a company making an effort to stand up for user rights: https://helloclue.com/privacy.

A Power Analysis of Privacy Policies

Steven Lukes has a useful analysis of three layers, or what he calls three faces, of power.[11] The first face is decision-making power, the second face is agenda-setting power, and the third face is ideological power. The first two faces are particularly relevant to this discussion. While the first face is fairly straightforward, the second face tends to operate somewhat covertly, so it can take a bit of practice to see. The second face is the power to set the agenda of what is even on the table for deciding on.

We see this second face of power at play in the use of privacy policies and terms of service contracts. While companies typically allow a few decisions to be made by the user (for example some of privacy settings), the choices the user gets to make are decided entirely by the company. In other words, one party unilaterally sets the agenda.

Discussion (continued)

Balancing power would mean that users and companies have equal power to determine what issues or topics are included in the agreement. For example, who owns data is often not expressly addressed. It would be in users' interests to have data ownership explicitly addressed and mututally decided on as part of these contracts.

The Exercise of Power: Spotlight on Contract Language and Trends

In our reading through over a hundred policies (including both terms of service and privacy policies for all the companies we rated), certain clauses and phrases stood out. To save you from reading through them all, here are some excerpts that help illustrate areas of interest and concern, and occasionally of delight for their sheer unexpectedness. We provide links to the policies quoted, noting that the companies may have changed their policies' language and wording since the time of our writing.

Note: Many companies changed their policies significantly in response to the GDPR, which came into effect during the last month of our research. The difference between language quoted and some of the updated policies demonstrates the impact of this regulation on raising the bar for consumer rights and protections.

An example of contract language spelling out an imbalance of power:

> *"You may not assign this Terms of Service without the prior written consent of Moov, but* ***Moov may assign or transfer this Terms of Service, in whole or in part, without restriction."*** [12]

This one from Mobvoi was highly unexpected, as it covers not just the world, but the Universe! An example of poetic overreach:

> *"By posting content on our Site, you expressly grant Mobvoi a non-exclusive, perpetual, irrevocable, royalty-free, fully paid-up worldwide, fully sub-licensable right to use, reproduce, modify, adapt, publish, translate, create derivative works from, distribute, transmit, perform and display such content and your name, voice, and/or likeness as contained in your User Submission, in whole or in part, and in any form throughout the world in any media or technology, whether now known or hereafter discovered, including all promotion, advertising, marketing, merchandising, publicity and any other ancillary uses thereof, and including the unfettered right to sublicense such rights,* ***in perpetuity throughout the universe."*** [13]

An example of a company benefitting from you, without having to pay you:

> *"You acknowledge and agree that any questions, comments, suggestions, ideas, feedback or other information about the Service ("Submissions"), provided by you to* ***Moov are non-confidential and Moov will be entitled to the unrestricted use and dissemination of these Submissions for any purpose, commercial or otherwise, without acknowledgment or compensation to you."*** [14]

Discussion (continued)

An example of a typical user content clause (from Bloomlife):

> *"By posting or otherwise making available any User Content on or through the Service, you expressly grant, and you represent and warrant that* ***you have all rights necessary to grant, to bloomlife a royalty-free, sublicensable, transferable, perpetual, irrevocable, non-exclusive, worldwide license to use, reproduce, modify, publish, list information regarding, edit, translate, distribute, syndicate, publicly perform, publicly display, and make derivative works of all such User Content and your name, voice, and/or likeness as contained in your User Content****, in whole or in part, and in any form, media or technology, whether now known or hereafter developed, for use in connection with the Service and bloomlife's (and its successors' and affiliates') business, including without limitation for promoting and redistributing part or all of the Service (and derivative works thereof) in any media formats and through any media channels."* [15]

Some policy language and clauses made us smile.This one stood out for its thoroughness in capturing all of "what could go wrong":

> *"Please note that the Beyond Verbal's Site, as with most Internet applications, are vulnerable to various security issues including without limitation various* ***eavesdropping, electronic trespassing, sniffing, spamming, nuking, hacking, spoofing, "imposturing", breaking passwords, harassment, fraud, forgery and system contamination including without limitation use of viruses, worms and Trojan horses causing unauthorized, damaging harmful access and/or retrieval of information and data*** *on the User's computer or the User's information and data on Beyond Verbal servers and other forms of activities that may even be considered unlawful, and* ***hence should be considered unsecured****. Information and data may also not reach their destination or reach an erroneous address or recipient."* [16]

Relative to last year's report card, it was more common this year to see admission that privacy is nigh impossible to guarantee. A typical example:

> *"Bloomlife cares about the integrity and security of your personal information. However,* ***we cannot guarantee that unauthorized third parties will never be able to defeat our security measures*** *or use your personal information for improper purposes.* ***You acknowledge that you provide your personal information at your own risk."*** [17]

Discussion (continued)

User Beware! Power Misuse and Over-reach

Accessing Your Contacts

It is typical for apps and services to request access to your contacts or friend list. While this is common practice in industry, in effect it violates your friends' and contacts' privacy and autonomy. These contacts have not directly been asked to have their information collected, nor have they given their permission for the service to do so. We note that this practice was at the heart of the Facebook-Cambridge Analytica data breach.

Because your contact information is likely in numerous friends' and colleagues' personal contact lists, that information could be flowing through numerous companies without your knowledge or consent. Though common practice in industry, this is the kind of overreach that makes people rightfully mistrustful.

We see a big opportunity here for industry to show leadership by preemptively solving the problem of this kind of overreach. Rather than waiting to be caught by regulators, like children sneaking cookies from the proverbial cookie jar, industry players could take greater responsibility for setting a higher ethical bar. Restraining from overreaching around permission is one practice, for example, that companies could begin to explore and implement that would be a great service to society.

Limiting Legal Recourse

In an industry where regulation has been relatively scant and unable to keep up with the pace of technology, litigation is one avenue for holding tech companies accountable. If consumers can seek legal recourse when companies mess up, companies are at least somewhat incentivized to be accountable for their actions, and to learn from their mistakes.

However, companies are moving further and further towards limiting a consumer's legal recourse. Most require you to give up any right to class action lawsuits or arbitration against them when you sign their contracts.

Why try to neutralize the minimal mechanisms that provide consumers with safeguards? This is not simply a rhetorical question – it may point to a systemic double bind that companies face, of being profit-driven yet ideally not predatory about data collection. To sincerely engage the question may provide a stepping stone on the way to resolving this double-bind. In an industry geared to problem solving, this is an opportunity to better fulfill on their mission of 'making the world a better place' while at the same time genuinely respecting user rights and autonomy.

Some examples of contract language limiting users' recourse:

> *"YOU MAY NOT BE ABLE TO HAVE ANY CLAIMS YOU HAVE AGAINST US RESOLVED BY A JURY OR IN A COURT OF LAW." [emphasis theirs]*
>
> *"No Class Actions: You may only resolve Disputes with Fitbit on an individual basis, and may not bring a claim as a plaintiff or a class member in a class, consolidated, or representative action.* ***Class arbitrations, class actions, private attorney general actions, and consolidation with other arbitrations are not allowed under our agreement.****"*[19]

Disconcertingly, it is also becoming more common for companies to go so far as to exempt themselves from international agreements on responsibilities towards users. Clauses like "The application of the United Nations Convention on Contracts for the International Sale of Goods is expressly excluded." are creeping in to these contracts.

Lastly, one challenging thing about "data" is that it is possible to easily make a copy of it. How feasible is it in practice for companies to delete users' data on request, even if their policy says they will do so? This brings up the broader question of how to ensure compliance around data protection policy. While this question is beyond the scope of our research, it certainly came up in the process of reading the policies.

Discussion (continued)

Lying by Omission

Policies often tell a partial truth about why they collect the information they do. Companies typically communicate what they plan to do with your data in broad strokes. The focus is usually on "what's in it for you", such as how they will use your information to improve app functionality. Less emphasized, if mentioned at all, is the extent to which companies collect your data for their sole benefit. Companies' other motives for collecting user data, such as furthering customer and market segmentation, training AI, or for as yet undefined purposes, should be spelled out in their contracts.

We believe trust is built on telling people things they wouldn't know to ask. This is why lying by omission erodes trust, and why we believe the benefit to the company should be clearly spelled out. This kind of transparency creates an environment that encourages deeper accountability and supports reaching for higher ethical standards.

The Impacts of Personalization and "Privacy as a Privilege"

Privacy policies claim they may use information they collect directly from you, along with information other companies have from or about you, to better tailor their ads or services to you. Here is wording typical in many policies

> *"To serve you better, we may combine information you give us and information about your product interest and purchases with information from third parties, including demographic information and information that is publicly available. We may also combine this information with information from our affiliates. We use that combined information to enhance and personalize your shopping experience with us, ..."*

Many companies reserve the right to acquire numerous data sets on you from various sources so that they can reconstruct your particular personal preferences as closely as possible. This allows them to supply you with ads you like for things that are more likely to appeal to you. This may seem nice on the surface – but what's the downside?

For one, it puts the company in a conflict of interest regarding their commitment to your privacy. The more data they have about you, the more can be leaked. This is reflected in a new trend in privacy policies warning users that their data is never really fully secure.

Secondly, going beyond individual risk, using data to segment or "profile" customers is vulnerable to furthering bias against numerous segments of society. Showing ads for only certain products and services to only certain segments of users contributes to the cascading effect of further privileging the privileged while further disadvantaging the already disadvantaged. Amazon recently scrapped an HR hiring algorithm after noticing that it was biased towards hiring men and not hiring women.[20] Algorithms tend to pick up on patterns of difference and magnify them. Information itself may be neutral, but its use always includes a dimension of influence. An industry standard of being transparent and explicit about this influence is necessary for integrity in the AI industry.

Businesses participating in the information industry have a bold opportunity to demonstrate leadership by becoming accountable for how this influence is used – by whom, on whom, and to what effect. To begin with, companies could adopt a standard procedure of inquiry into the ethics at play throughout the conception, design and production phases of their product or services. Secondly, companies could make sure to meaningfully include in this process those who are at risk of being excluded from using a product or service, and incorporate their feedback into the product design.

Privacy as a Privilege

Data privacy is increasingly becoming a socio-economic issue. As industry explores offering users paid accounts with premium services, it isn't hard to imagine a future where privacy is a privilege of those with enough disposable income to cover fees. Those who can't afford to pay may be pressured to sell their data, or may have fewer rights regarding their data in exchange for using products and services for "free". While a future where subscriptions cost as little as $5-$10 each per month may not seem prohibitive, the costs could quickly add up considering the number of apps and platforms that are becoming the norm to use – many of us have 20-50 applications or more on our smartphone alone.

The threat of privacy becoming a privilege is not in keeping with the spirit of the United Nations' Sustainable Development Goals to leave no one behind in poverty, and to create equality. Is industry up for the challenge of making money from customers without sacrificing equality and people's privacy? We believe they are. However, turning the tides may require some bold moves from key players to counter the current blind spots in industry practices.

Recommendations: For Legislators, Industry, and Users

Legislation has a hard time keeping pace with technological development. It is likely that the GDPR will need to evolve sooner than we think; in fact, it may already be outdated. As mentioned in the Introduction, AI is predicated on machines consuming vast amounts of data and algorithms updating based on patterns in the data. In other words, algorithms no longer follow rules, but follow patterns in data that are too complex for humans to grasp. AI, and especially deep learning, clashes with several GDPR regulations, such as collecting data only for specific purposes, and being transparent around how data will be used.[21] Furthermore, the right to be forgotten is technically difficult to impossible when a user's information is in a blockchain.

Does this imply that technological developments will be thwarted by data protection regulation? Or does it imply that our ethics will cave to our technological capacities?

We believe it is important for regulators to strike a balance between the pragmatics of big data and AI, and users' rights to privacy and transparency. Not having any regulation in this arena will lead us down a frightening path. It seems that, with the possible exception of the GDPR, regulations around data protection have been largely reactionary. Furthermore, regulation is slow, while the pace of technological development is lightning fast. Governments and regulators will need to figure out how to keep up with the pace of technological advancement.

Discussion (continued)

The onus however should not be entirely on regulators to ensure ethical practice. As society matures, the culture of entrepreneurialism can too, and begin to raise the bar on ethical standards. Some companies we reviewed clearly had different policies in place for EU citizens compared to other users of their products, implying that they adhere to the GDPR for legal reasons as opposed to ethical reasons. While it may be seen as a tall order to expect companies to go beyond what is legally required of them, we see some that already do. Furthermore, If we don’t put more of the onus on companies to be ethical in their treatment of user data, industry practices are likely to fall short. The opportunity for industry to take the lead with regulation and system level ethical principles would be a significant evolutionary step not just for the industry, but for humanity as a whole.

Lastly, it is important that users pay attention to matters of data privacy. That includes you! We hope our research has made you more aware of the current state of data protection practices, and maybe even inspires you to do a little of your own research. We have included a Useful Links section below, with websites and articles that shed more light on this topic. In cases where it is possible to request a copy of your data, why not give this a try?

Acknowledgments

Special thanks to Greg McMullen for his insights into GDPR regulations as well as general legal concepts around data. Any misrepresentation of either is the sole responsibility of the authors.

We are grateful to Scott Nelson for his unwavering support for this Report Card.

We thank Eric Pinfold of Phitted Design for his expert design work.
eric@phitted.com | *phitted.com*

Websites and Articles Cites

[1]Quantified Self. http://quantifiedself.com/

[2]Quantified Self Institute. http://qsinstitute.com/about/what-is-quantified-self/

[3]"2017 Global Mobile Consumer Survey" by Deloitte. Published 2017. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/technology-media-telecommunications/us-tmt-2017-global-mobile-consumer-survey-executive-summary.pdf

[4]"Nobody reads privacy policies – here's how to fix that" by Florian Schaub. Published in The Conversation. http://theconversation.com/nobody-reads-privacy-policies-heres-how-to-fix-that-81932

[5]"Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach". Published 3.17.2018 in The Guardian. https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

[6] EU GDPR Portal. https://eugdpr.org/

[7] General Data Protection Regulation, Art. 4 – Definitions. https://gdpr-info.eu/art-4-gdpr/

[8]"Has Big Data Made Anonymity Impossible?" by Patrick Tucker. Published 5.7.2013 in MIT Technology Review. https://www.technologyreview.com/s/514351/has-big-data-made-anonymity-impossible/

[9] The General Data Protection Regulation. Recital 26. https://gdpr-info.eu/recitals/no-26/

[10] "Big Data and Data Anonymization: Is Anonymization an Illusion?" by Dick Weisinger. Published 2.9.2017 in Formtek. http://formtek.com/blog/big-data-and-data-anonymization-is-anonymization-an-illusion/

[11]Power: A Radical View, Second Edition, by Steven Lukes. Published 2005 by Palgrave MacMillan. http://voidnetwork.gr/wp-content/uploads/2016/09/Power-A-Radical-View-Steven-Lukes.pdf

[12]Moov Terms of Service. https://store.moov.cc/pages/terms-of-service

[13]Mobvoi Terms of Service. https://www.mobvoi.com/ca/pages/terms-of-service

[14]Moov Terms of Service. https://store.moov.cc/pages/terms-of-service

[15]Bloomlife Terms of Service. https://bloomlife.com/terms-of-services/

[16]Beyond Verbal Privacy Policy. http://www.beyondverbal.com/privacy/

[17]Bloomlife Terms of Service. https://bloomlife.com/terms-of-services/

[18]Moov Terms of Service. https://store.moov.cc/pages/terms-of-service

[19]Fitbit Terms of Service. https://www.fitbit.com/en-ca/legal/terms-of-service

[20]"Amazon scraps secret AI recruiting tool that showed bias against women" by Jeffrey Dastin. Published 10.9.2018 by Reuters. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

[21]"AI Has a Big Privacy Problem and Europe's New Data Protection Law Is About to Expose It" by David Meyer. Published 5.28.2018 by Fortune. http://fortune.com/2018/05/25/ai-machine-learning-privacy-gdpr/

[22]"AI Has a Big Privacy Problem and Europe's New Data Protection Law Is About to Expose It" by David Meyer. Published 5.28.2018 by Fortune. http://fortune.com/2018/05/25/ai-machine-learning-privacy-gdpr/

Useful Links

We are not the first to review privacy policies. We refer the interested reader to Terms of Service; Didn't Read, a user rights initiative to rate website terms and privacy policies: https://tosdr.org/

Interested in understanding the privacy policies of the products and services you use on a regular basis, but don't have time to read them? Polisis created Pribot, an AI-powered tool that provides a summary snapshot of the content contained in any privacy policy when you provide the policy link: https://pribot.org/polisis

Articles of Note

"Gartner picks digital ethics and privacy as a strategic trend for 2019" by Natasha Lomas. Published 10.24.2018 in TechCrunch. https://techcrunch.com/2018/10/16/gartner-picks-digital-ethics-and-privacy-as-a-strategic-trend-for-2019/

"Tim Berners-Lee on the huge sociotechnical design challenge" by Natasha Lomas. Published 10.24.2018 in TechCrunch: https://techcrunch.com/2018/10/24/tim-berners-lee-on-the-huge-sociotechnical-design-challenge/

"Fitness app that revealed military bases highlights bigger privacy issues" by Selena Larson. Published 1.29.2018 in CNN Business. http://money.cnn.com/2018/01/29/technology/strava-privacy-data-exposed/index.html

"What Does a Fair Algorithm Actually Look Like?" by Louise Matsakis. Published 10.8.2018 in WIRED: https://www.wired.com/story/what-does-a-fair-algorithm-look-like/

"We must not treat data like a natural resource" by Lisa Austin, University of Toronto Law and Technology. Published 07.9.2018 in The Globe and Mail: https://www.theglobeandmail.com/opinion/article-we-must-not-treat-data-like-a-natural-resource/

Government Publications

This report is published by the American National Science Foundation's Council for Big Data, Ethics, and Society. It addresses security, privacy, equality and access to help guard against repetition of known ethical problems and inadequate preparation for future issues: http://bdes.datasociety.net/wp-content/uploads/2016/05/Perspectives-on-Big-Data.pdf

A Government of Canada report to parliament on Personal Information Protection and Electronic Documents Act and the Privacy Act: https://www.priv.gc.ca/en/opc-actions-and-decisions/ar_index/201516/ar_201516/

Books

Weapons of Math Destruction by Cathy O'Neil. Published 2016 by Crown, an imprint of the Crown Publishing Group, a division of Penguin Random House LLC, New York. https://weaponsofmathdestructionbook.com/

About the Authors

Rochelle Fairfield

Rochelle is an Integral Practitioner and activist, dedicated to promoting a non-colonizing growth mind-set in tech and AI. She learned a great deal about governance and human's susceptibility to power during her career in commercial fishing and fisheries management. An ordained Zen monk and Executive Director of the Human Data Commons Foundation, Rochelle contributes to transdisciplinary thought leadership in AI ethics and data industry governance.

Heather Mann

Heather is a data strategy consultant and industry professional with an interest in human decision-making, and how data can be used to support better decisions. She hold three degrees in Psychology, and for her PhD she worked with Dan Ariely at Duke University to investigate how dishonesty varies across cultures. She currently works as a Behavioural Scientist for Symend, a technology startup focused on improving the customer experience of debt resolution.

About the Human Data Commons Foundation

The Human Data Commons Foundation (HDC) is a non-profit organization based in Vancouver, Canada working to make data collection more salient, ethical and beneficial to the well-being of humanity.

HDC does this by fostering and facilitating interdisciplinary, inclusive stakeholder discussion about how to shape ethics and protocols for big data.

Our actions include:

- creating awareness through education
- generating ideas for change
- advocating for higher industry standards
- supporting seed projects

We envision a world where the big data playing field is level for everyone.

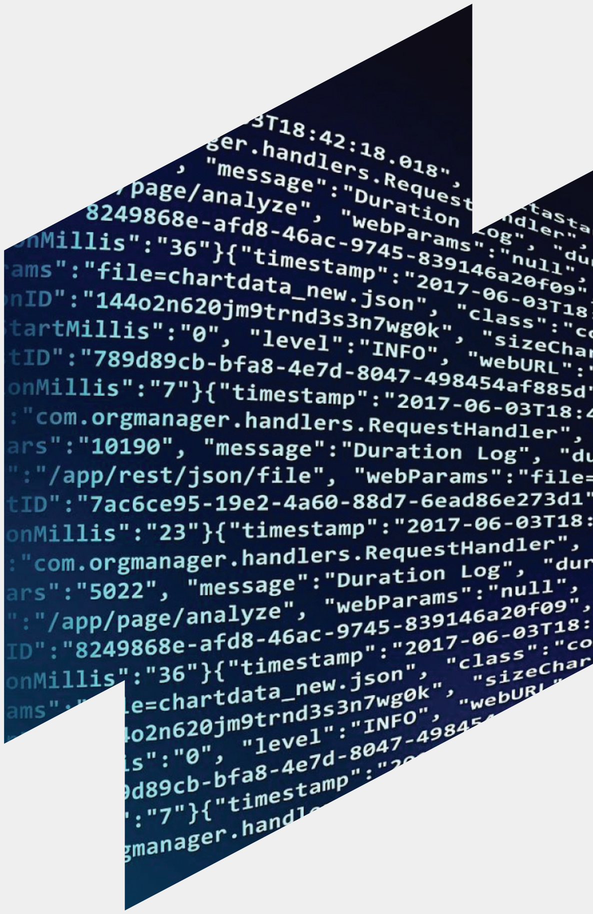
Human Data Commons Foundation
www.humandatacommons.org

hdc

DONATE NOW

Donate now to help us help you secure your data.100% of all funds will be used to support our research and education programs.

GO TO: **http://www.humandatacommons.org/donate**



```
3T18:42:18.018",
ger.handlers.Request
, "message":"Duration
/page/analyze", "webParams
8249868e-afd8-46ac-9745-839146a20f09"
nMillis":"36"}{"timestamp":"2017-06-03T18
ams":"file=chartdata_new.json", "class":"c
nID":"144o2n620jm9trnd3s3n7wg0k", "sizeChar
tartMillis":"0", "level":"INFO", "webURL":
tID":"789d89cb-bfa8-4e7d-8047-498454af885d"
onMillis":"7"}{"timestamp":"2017-06-03T18:4
:"com.orgmanager.handlers.RequestHandler",
ars":"10190", "message":"Duration Log", "du
":"/app/rest/json/file", "webParams":"file=
tID":"7ac6ce95-19e2-4a60-88d7-6ead86e273d1"
onMillis":"23"}{"timestamp":"2017-06-03T18:
:"com.orgmanager.handlers.RequestHandler",
ars":"5022", "message":"Duration Log", "dur
":"/app/page/analyze", "webParams":"null",
ID":"8249868e-afd8-46ac-9745-839146a20f09",
onMillis":"36"}{"timestamp":"2017-06-03T18:
le=chartdata_new.json", "class":"co
4o2n620jm9trnd3s3n7wg0k", "sizeChar
is":"0", "level":"INFO", "webURL
9d89cb-bfa8-4e7d-8047-49845
":"7"}{"timestamp":"20
gmanager.handl
```

Appendix A. Scorecard and Scoring

Category	Question	Response options	Points
User Rights	1. Are policies (ToS and Privacy Policy) easy to find?	Link to policy is evident on home page or product page	0.5 points per policy
		Link to policy is somewhat difficult to find	0 points per policy
		Policy not available on website	-0.5 points per policy
User Rights	2. Are users notified of changes to policies? (ToS and Privacy Policy)	Users are notified of material changes to policy	0.5 points per policy
		User has to check back for changes to policy	0 points per policy
		Not clear or no mention	-0.5 points per policy
		N/A	-0.5 points per policy
User Rights	3. Are policies (ToS and Privacy Policy) written in clear and readable language?	Language in policy is very user-friendly, and there are no conflicts between "user-friendly" portions of the contract and "legalese" portions of the contract	0.5 points per policy
		Language in policy is reasonably legible for an everyday user, may include some legal jargon.	0 points per policy
		Language in policy is difficult for an everyday user to read/understand, and/or "user-friendly" portions of the contract conflict with "legalese" portions of the contract	-0.5 points per policy
		N/A	-0.5 points per policy
User Rights	4. What are users' rights regarding data access and ownership?	Policy is explicit that user owns their data	1 point
		Policy implies that data belongs to the user	0 points
		User can access their data, but doesn't own it	-1 point
		User can't access their data or not clear	-1 point
User Rights	5. Are users given the right to be forgotten	At user's request, any data company holds on them will be erased	1 point
		At user's request, any data company holds on them will be anonymized but not erased	0 points
		At user's request, data company holds on them will be erased but may be retained on backup systems	0 points
		Not clear, no mention, or policy specifies that data will NOT be deleted	-1 point

Appendix A. Scorecard and Scoring

Category	Question	Response options	Points
Data Collection and Storage	6. What contact information does the company provide in case users have questions or concerns related to how their data is processed?	There is a link/email/phone for concerns related to privacy, that goes to a privacy department or officer	1 point
		There is a direct link/email/phone for concerns related to privacy, that goes to a GENERIC account	0 points
		There is no link/email/phone listed in the ToS or Privacy Policy	-1 point
Data Collection and Storage	7. Does the company use encryption when transfering data?	Policy states that data transfer between app/device and company is encrypted with specified method (e.g., TPS, SSL, HTTPS)	1 point
		Policy states that data transfer between app/device and company is encrypted, method not specified	0 points
		No mention of data encryption	-1 point
Data Collection and Storage	8. Is it clear which jurisdiction governs the contract?	It is clear which jurisdiction governs contract	1 point
		Vague / not clear which jurisdiction governs contract	-1 point

Appendix A. Scorecard and Scoring

Category	Question	Response options	Points
Third Party Sharing	9. On reading the policy, how well-contained does users' information seem?	There is a link/email/phone for concerns related to privacy, that goes to a privacy department or officer	1 point
		On reading the policy, it feels like user's information is somewhat contained	0 points
		On reading the policy, it feels unclear as to how far user's data will spread	-1 point
Third Party Sharing	10. On reading the policy, how clear is it on third party data sharing practices?	Policy states that user data is not shared with other companies	1 point
		Policy specifies which specific companies user data is shared with	1 point
		Policy indicates categories of companies user data is shared with	0 points
		Policy is very broad or unclear on third party sharing	-1 point

Appendix B. Excluded Companies

Company	Category	Reason Excluded
Apple (Apple Health)	Integration Platform	Terms of service for website, not for product. Online search for "Apple health Terms and Conditions" led to https://www.apple.com/ca/legal/. Not clear which if any were for Apple Health.
LG (Watch Sport™ - AT&T W280A)	Devices	Terms of service and privacy policy refer only to the website
Ketonix	Devices	Terms of service and privacy policy refer only to the website
SIDLYCare	Devices	Terms of service and privacy policy refer only to the website
Intheon (formerly Qusp)	Middleware Analytics	Terms of service and privacy policy refer only to the website
Timex	Devices	Terms of service and privacy policy refer only to the website
iMotions	Middleware Analytics	Terms of service and privacy policy refer only to the website
Sigma	Devices	Terms of service and privacy policy refer only to the website
Eyeris (EmoVu)	Middleware Analytics	Can't access policies (links to policies direct back to home page)
Vivametrica	Middleware Analytics	Can't access policies (unable to find links to policies)

Appendix C. Final Thoughts

It has been said that big data has replaced oil as the world's most valuable resource.[22] As with oil, we see big data as prone to colonizing attitudes. Industry players are tempted to extract as much of the resource as possible from wherever it can be found. When the resource is oil it is drawn from oil fields, whereas big data is culled from individuals like you, and not usually with an eye to fair compensation for it. Any time you interact with an app, gadget, platform, or software-based product you produce data that is being captured and increasingly used to train AI. In some cases, your data may be used to improve and fine tune the service you're using, but in other cases it may be used for purposes you don't know about and might not agree to if you did.

Maybe in the future, privacy does not exist – who we are is laid bare by our digital habits, searches, connections made visible. We would likely be horrified by some of what we discover about ourselves and one another as the veil of privacy is lifted. The possibility for increased judgement would certainly be there. On the other hand, the opportunity to recognize ourselves more as we actually are, warts and all, may move or motivate us to more acceptance, perhaps even compassion for our fellow humans.

Human Data Commons Foundation | www.humandatacommons.org