

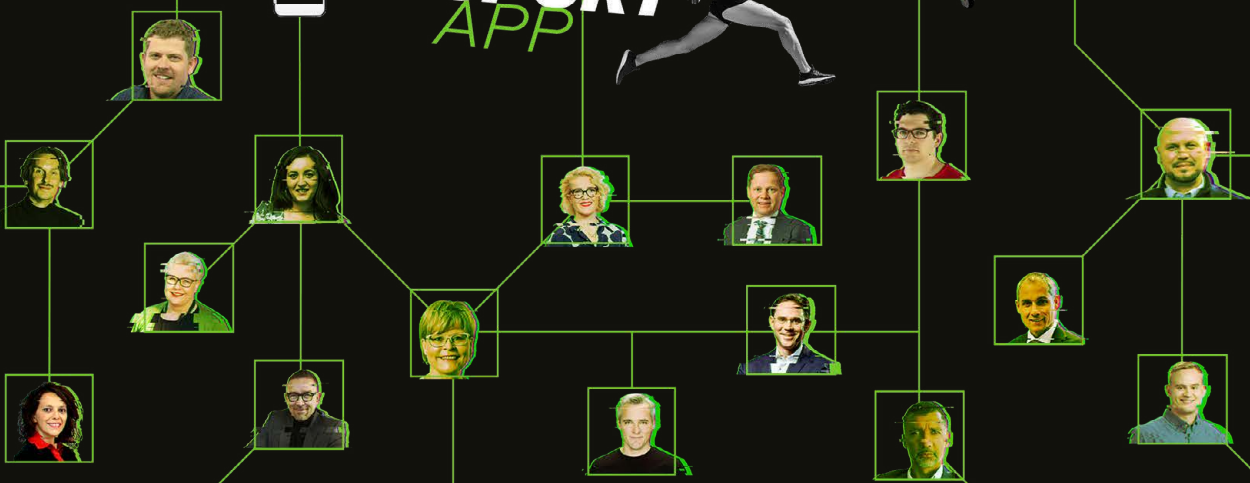
TRACKING DIGIPOWER

How data can be used for influencing decision-makers and steering the world

Tiina Härkönen, Riitta Vänskä, Jukka Vahti and Kristo Lehtonen



A network diagram with green lines connecting various icons and text. The icons include a group of people taking a selfie, a smartphone, a car, a runner, a shopping cart with sneakers, and a person's face. Text elements include 'SOME APP', 'SHARE', 'NAVIGATOR TRACK SPORT APP', 'ONLINE STORE', '-50% SALE!', and 'BUY'. The overall theme is digital connectivity and data-driven activities.



© Sitra 2022

Sitra Studies 215

Tracking Digipower – How data can be used for influencing decision-makers and steering the world

Working group at Sitra: Tiina Härkönen, Riitta Vänskä, Jukka Vahti, Kristo Lehtonen and Veera Heinonen.

Authors and mentors involved in the investigation:
The investigation was carried out in cooperation with Paul-Olivier Dehaye and Alex Bowyer, experts from Hestia.ai.

Visual design: Topias Dean
Sub-editor: Kirsi Suomalainen

Layout: PunaMusta Oy

ISBN 978-952-347-279-2 (PDF) www.sitra.fi
ISSN 1796-7112 (electronic publication)

THE SITRA STUDIES publication series features the results of Sitra's future-oriented work and experiments.

Contents

Preface	2
Summary	3
Tiivistelmä	4
Sammanfattning	5
1. Systemic change in the data economy can open the way to European success and democratic renewal	6
Digital oligarchy and the non-transparency of the data economy erode the foundations of democracy	6
The discussion on the importance of data is just beginning, and politicians are absent from the data economy debate	7
The Digipower investigation provides a unique insight into the flow of decision-makers' data	8
2. Test group of European policy makers and societal influencers	9
Test subjects included social influencers in different roles	9
Test subjects' Digipower journey	9
3. The investigation used three methods to collect data	27
4. Key findings: Digital power is self-reinforcing, individuals have no control over their data	30
Digital forms of power are self-reinforcing	30
The strongest understanding of the data economy comes from personal experience	31
Understanding your own data and how it is used is hard work and may not succeed at all	32
Does everyone get the same level of service for their subject access requests?	33
Gigantti provides by far the best responses to subject access requests	34
A subject access request is a way to measure the trustworthiness of a company	34
Advice from Finland's Data Protection Ombudsman Ms. Anu Talus	35
5. Conclusions: Time to build alternative futures	36
An example of data giants' surveillance methods: the Facebook Pixel	36
Downsides of concentrated power in the data economy	37
It is time to break the no-alternative narrative	38
6. Recommendations to address digital power imbalance	41
Sources	46
Appendix 1: A model of a subject access request	49

Preface

Data is the most important raw material and resource of our time. A historic technological and economic transformation is underway whereby digitalisation and the data economy are enabling a surge in prosperity and labour productivity.

But we have a problem on our hands. The current data-driven economy is unfair as digital power is concentrated in the hands of a few data giants. The distorted competitive landscape and the lack of ground rules for the data economy are detrimental to both people and societies.

The first step to remedy this situation is to increase our understanding of this new economy. As societies, we are only just waking up to the fact that key social media platforms and digital marketplaces are critical infrastructure for society. Platforms and their algorithms have a lot of power over the kind of information provided to individuals. This power can also be used to manipulate people, for example by transmitting false information. At worst, this can erode the foundations of the democratic system: the assumption of free and autonomous individuals capable of forming their own opinions and taking independent decisions.

As individuals, we have too little visibility over how much data is collected about us and how this data is used to profile us. It is unsustainable, both from a data economy development and democratic perspective, that individuals have no control over their own data.

Our Digipower investigation was carried out to shed light on the mechanisms of the current data economy and the use of digital power based on data. One of its main aims is to raise awareness of the nature of digital power in our daily lives. We also offer recommendations to improve the situation.

The problem boils down to the high concentration of digital power and the limited control by individuals. Therefore, the ultimate solution is to increase people's control and agency.

The investigation involved 15 decision-makers and other social influencers from European countries. We would like to express our warmest thanks to all the test subjects. Despite their busy routines, they gave their time and data to the investigation. Without them, this groundbreaking investigation would not have been possible.

Many thanks are also due to the whole talented team of our partner, Hestia.ai, and especially to **Paul-Olivier Dehaye** and **Alex Bowyer**.

24 May 2022

KRISTO LEHTONEN and VEERA HEINONEN

Kristo Lehtonen is the director of the fair data economy theme and Veera Heinonen is the director of the democracy and participation theme at Sitra.

Summary

The aim of the Digipower investigation was to explore the ecosystems and operating models of a data economy based on the collection and use of data. The information gathered in the investigation was used to profile the nature of digital power and its different dimensions from the perspectives of the economy, democracy and the everyday lives and rights of individuals, including privacy. The investigation made visible data economy mechanisms that until now have largely remained hidden. This is important because we need to better understand how the current data economy works so that we can envisage and build an alternative and fairer future for the data economy.

The investigation involved 15 decision-makers and other European social influencers as test subjects. They went through a “data mentoring programme” to help them understand the movement, quantity and quality of data about themselves using test phones, subject access requests and service providers' data portals. Based on the data obtained, experts at Hestia.ai, which provided the data mentoring, assessed what could be done with the data collected on the test subjects and how it might be used to influence them. The data received and the analyses based on it were discussed with the test subjects during mentoring.

Compared to Sitra's Digitrail survey conducted in 2019, which focused on the amount of data collection and the network of collectors, this investigation concentrated more on the flow of data between different services. Above all, it looked at how diverse a picture of the test subjects can be formed using data, and the power it gives to data collectors.

The investigation shows that data collection is not limited to digital environments, but that that detailed information on transactions in brick-and mortar shops may also be passed on to the most important power-holders in the data economy today: global platform companies. In many respects, the test subjects found the results of the investigation startling, which was also one of the points of the study. Raising awareness among decision-makers is a prerequisite for creating a level playing field for the data economy.

Using examples built from the data of individual test subjects, the Digipower investigation also sought to understand of the nature of digital power in general. The investigation stressed that in today's data economy, the winners are those organisations that are capable of collecting vast amounts of diverse data. Power is cumulative in nature, as new data can be used to refine and add value to data already held by an organisation. In digitalising societies, such power is infrastructural in nature and is exercised not only in relation to individuals, but also to businesses and society as a whole.

Based on the Digipower investigation and other work by Sitra on the data economy, the Sitra working group has developed a set of recommendations on how to build a fairer and more democratically sustainable data economy. Key measures include curbing the power of data giants, promoting the functioning of the data business market, increasing corporate responsibility and advancing the agency of individuals – meaning all of us – by supporting data economy skills and “digital literacy”.

Tiivistelmä

Digivalta-selvityksen tavoitteena oli läpivalaista datan keräämiseen ja hyödyntämiseen perustuvan datatalouden ekosysteemejä ja niiden toimintamalleja. Selvitystyön myötä saadun tiedon pohjalta pyrittiin hahmottamaan digitaalisen vallan luonnetta sekä sen erilaisia ulottuvuuksia niin talouden, demokratian kuin yksittäisen ihmisen arjen ja oikeuksien, esimerkiksi yksityisyyden, näkökulmista. Selvitystyössä tehtiin toistaiseksi pitkälti näkymättömiä datatalouden mekanismeja näkyviksi. Tämä on tärkeää, koska meidän tulee ymmärtää nykyisen datatalouden toimintaa paremmin, jotta voimme kuvitella ja rakentaa vaihtoehtoja ja nykyistä reilumpaa datatalouden tulevaisuutta.

Selvityksessä oli mukana testihenkilöinä 15 päättäjää ja muuta yhteiskunnallista, eurooppalaista vaikuttajaa. He kävivät läpi ”datavalmennusohjelman”, jossa heitä autettiin selvittämään itseään koskevan datan liikkeitä, määrää ja laatua testipuhelimien, tietopyyntöjen ja palveluntarjoajien dataportaalien avulla. Saadun datan perusteella datavalmennuksesta vastanneen Hestai.ai-yrityksen asiantuntijat arvioivat, mitä kaikkea kerätyllä datalla on mahdollista tehdä ja miten sitä voidaan käyttää testihenkilöihin vaikuttamiseen. Saadut tiedot ja niiden pohjalta tehdyt analyysit käytiin valmennuksen aikana läpi testihenkilöiden kanssa.

Verrattuna vuoden Sitran 2019 Digijälki-selvitykseen, joka keskittyi datan keräämisen määrään ja kerääjien verkostoon, tässä selvityksissä paneuduttiin syvällisemmin eri palveluiden väliseen datan liikkumiseen. Ennen kaikkea tarkasteltiin, kuinka moninaisen kuvan testihenkilöistä saa datan avulla muodostettua, ja millaista valtaa datan kerääjille sen kautta syntyy.

Datan kerääminen ei selvityksen perusteella rajaudu vain digitaalisiin ympäristöihin, vaan myös asioinnista kivijalkaliikkeessä voi mennä yksityiskohtaisia tietoja datatalouden nykyisille merkittävimmille vallankäyttäjille eli globaaleille alustayrityksille. Testihenkilöt kokivat selvityksen tulokset monella tapaa hätkähdyttävinä, ja tämä oli myös yksi selvityksen tavoitteista. Päättäjien tietoisuuden parantaminen on edellytys reilujen pelisääntöjen laatimiselle datataloudelle.

Digivalta-selvityksessä tavoiteltiin yksittäisten testihenkilöiden datasta rakennettujen esimerkkien kautta ymmärrystä myös digitaalisen vallan luonteesta yleisemmin. Selvitys alleviivasi sitä, että nykymuotoisessa datataloudessa voittajia ovat ne toimijat, jotka pystyvät keräämään suuria määriä monipuolista dataa. Valta on luonteeltaan kasautuvaa, sillä uudella datalla on mahdollista jalostaa ennestään toimijan hallussa olevaa dataa ja nostaa sen arvoa. Digitalisoituvissa yhteiskunnissa tällainen valta on luonteeltaan infrastruktuurista ja sitä käytetään paitsi suhteessa yksilöihin myös yrityksiin ja koko yhteiskuntaan.

Sitran työryhmä laati Digivalta-selvityksen ja muun Sitran datatalouteen liittyvän työn pohjalta joukon suosituksia siitä, miten nykyistä reilumpaa ja demokratian kannalta kestävämpää datataloutta tulisi rakentaa. Keskeisiä toimenpiteitä ovat esimerkiksi datajättien vallan suitsiminen, dataliiketoiminnan markkinoiden toimivuuden edistäminen, yritysvastuun lisääminen sekä yksilöiden, eli meidän jokaisen, oman toimijuuden tukeminen datatalousosaamista ja ”digitaalista sivistystä” tukemalla.

Sammanfattning

Målet med utredningen Digital makt var att belysa dataekonomins ekosystem och dess operativa modeller utifrån datainsamling och dataanvändning. Utifrån den information som erhållits från utredningen var syftet att förstå den digitala maktens natur och dess olika dimensioner utifrån såväl ekonomi och demokrati som den enskilda människans vardag och rättigheter, till exempel ur ett integritetsperspektiv. I utredningsarbetet synliggjordes mekanismer i dataekonomin som hittills i stort sett varit osynliga. Detta är viktigt, eftersom vi behöver förstå bättre hur den nuvarande dataekonomin fungerar, så att vi kan föreställa oss och bygga en alternativ och mer rättvis framtid för dataekonomin.

I utredningen deltog 15 beslutsfattare och andra samhälleliga, europeiska påverkare som testpersoner. De genomgick ett "datacoachningsprogram" där de fick hjälp med att utreda rörelserna, mängden och kvaliteten på den data som gällde dem själva via testmobiler, informationsbegäranden och serviceproducenters dataportaler. På basis av insamlad data bedömde experterna från företaget Hestia.ai, som ansvarade för datacoachningen, vad man kan göra med dessa insamlade uppgifter och hur de kan användas för att påverka testpersonerna. Tillsammans med testpersonerna gick man under coachningen igenom den mottagna informationen och de analyser som hade gjorts på basis av den.

Jämfört med Digispår-utredningen som genomfördes av Sitra år 2019 och fokuserade på mängden insamlad data och nätverket av insamlare fördjupade sig denna utredning i hur data rör sig mellan olika tjänster. Framför allt granskade man hur pass mångsidig den bild som man kan skapa av testpersonen är på grundval av data, och vilken typ av makt detta kan ge insamlaren av data.

Enligt utredningen är insamlingen av data inte enbart avgränsad till digitala miljöer. Även vid besök i fysiska butiker kan detaljerad information hamna hos de i nuläget största makthavarna inom dataekonomin, det vill säga de globala plattformsföretagen. Testpersonerna upplevde att utredningens resultat på många sätt var häpnadsväckande och detta var även ett av målen med utredningen. Att öka medvetenheten hos beslutsfattarna är en förutsättning för att upprätta rättvisa spelregler för dataekonomin.

Med utredningen Digital makt försökte också förstå den digitala maktens natur mer generellt genom exempel konstruerade utifrån data från enskilda testpersoner. Utredningen betonade att vinnarna i den nuvarande dataekonomin är de aktörer som kan samla stora mängder mångsidig data. Makten är kumulativ till sin karaktär, eftersom man med ny data kan förädla data som man redan har från tidigare och således öka dess värde. I ett digitalt samhälle är makt av detta slag infrastrukturell till sin natur och används förutom i förhållande till individer även i relation till företag och samhället i stort.

Med stöd av utredningen Digital makt och annat arbete som Sitra utfört med anknytning till dataekonomin utarbetade Sitras arbetsgrupp ett antal rekommendationer om hur en rättvisare och ur demokratins synvinkel mer hållbar dataekonomi kunde byggas. Centrala åtgärder är till exempel att kontrollera datajättarnas makt, främja dataaffärsmarknadens funktioner, öka företagsansvaret samt förbättra aktörskapet hos individer, det vill säga var och en av oss, genom att stödja kompetens inom dataekonomi och "digital bildning".

1. Systemic change in the data economy can open the way to European success and democratic renewal

Europe can harness the full potential of data by tackling the most glaring problems of the data economy. By identifying and describing those problems, solutions can be found that support people and businesses and defend democracy.

The companies that have advanced the most in the use of data utilise their own data and their partners' data in a variety of ways to improve the speed, quality and price competitiveness of their services or products. For example, retailers use sales data to place advance orders with their suppliers. Restaurant delivery platforms can anticipate orders during peak periods. Transport companies use data to predict traffic congestion. Social media platforms recommend interesting content to their users and online shops recommend products to their customers.

Currently, the structures, operating practices and incentives of the data economy, are not fair either for individuals or most businesses. Data about individuals, its exploitation and economic benefits are increasingly concentrated in the hands of a few global data giants such as Google, Meta, Amazon and Microsoft. The opacity of data use and algorithms built on it are eroding the basis of trust between people and towards society's institutions. This also poses a threat to the sustainability of democracy.

At the same time, our ability to understand the long-term impacts of harm to privacy is limited by lack of transparency,

misinformation and lack of knowledge.

Lack of privacy may only be perceived as a problem if you are operating on the wrong side of the law, meaning that its importance for an individual's autonomy, mental health and ability to develop as a human being is not deeply understood.

Curbing the collection and use of personal data through regulation, for example, is causing a backlash in some parts of the business world, especially in the management of large platform companies, as it is said to restrict their ability to innovate. In reality, the success of the data economy, fair rules for companies and individual rights are not in conflict with each other, but are in fact a prerequisite for competition and innovation.

Digital oligarchy and the non-transparency of the data economy erode the foundations of democracy

Surveying the mechanisms of the data economy and digital power is important not only from the perspectives of individuals and businesses, but also for democracy and

inclusion. In recent times, our daily lives have moved into digital environments faster than our values, structures and practices that have traditionally sustained our democratic social order. This in turn leads to a discrepancy between political rhetoric and approaches and our everyday experience.

In this context, everyday life means, for example, the way we spend time, interact and work with others, find information in a complex world, entertain ourselves and get things done. Democratic structures and practices, on the other hand, refer to the ways in which the surrounding society provides us with information and influences matters that concern us or that we otherwise consider important, such as through different initiatives, elections, positions of trust or civic activism.

In Sitra's view, we should both defend and reform democracy. Defending democracy means, for example, combating or circumventing existing barriers to participation. In digital environments, such barriers have been identified in various surveys and studies both in Finland and elsewhere as disinformation, stirring up confusion, online harassment and the polarising nature of algorithms in social debate. The importance of understanding these phenomena has become even more important since Russia's invasion of Ukraine and the increased pressure on the security policy situation.

The non-transparent nature of the complex networks of data collection and exploitation and the uselessness and poor usability of data from the individual's perspective reduce the preconditions for democracy. The fact that the public administration maintains the status quo and does not actively promote the use of data generated with public funds or encourage the public to exercise their data rights can also be seen as forms of exercising digital power. This is despite the fact that in a highly educated country like Finland, civic "data activism" could produce new models of

participation and collaboration that could also benefit and support representative decision-making.

The discussion on the importance of data is just beginning, and politicians are absent from the data economy debate

The amount of data, its impact on all areas of society and the value of data-driven business are growing rapidly. Nonetheless, we know little about how data about us is collected and used. There is still a lack of understanding of the importance of online, data-driven monitoring and profiling for society as a whole.

A 2022 survey by Sitra and the media monitoring company Meedius International found that the debate on the data economy is rather narrow, lacks an accountability angle, and was actively dominated by researchers while politicians were almost invisible. The Digipower investigation was primarily intended to pay attention to data collected about decision-makers, but it also wanted politicians themselves to pay attention to the collection of data and its social relevance.

The international debate on the data economy has highlighted worries to do with the fairness of digitalisation and platform economy development. Here, the role of politicians is absolutely crucial. We need more social debate on the data economy, especially with regard to personal data, and politicians must be actively involved in this debate, not only as decision-makers and legislators but also as custodians of the overall interests of society and pluralism.

The Digipower investigation provides a unique insight into the flow of decision-makers' data

The Digipower investigation sought to track power in several ways: to find out whether and how the data collected about the decision-makers involved as test subjects could be used as a means to influence them. At the same time, we tracked both the practices and the entities that collect data on individuals when using different online services.

The aim of the study was not only to identify new kinds of mechanisms of influence but also to define what digital power means. It also sought to explore how traditional and digital power intersect. The investigation showed that in today's data

economy, digital power is mainly exercised by data platforms, such as various “social media” services, which are key nodes in digital networks. The investigation also revealed that one of the key forms of digital power is the ability to use different data sources to build up a complete picture of how people, goods, money, ideas and opinions move in these networks.

The longer-term goal of the Digipower investigation is to build on its findings to develop solutions, in partnership with a range of actors, to make the data economy fairer and more sustainable, not only for business but for the defence and further development of democracy.

The Digipower investigation was carried out in cooperation with the Swiss consultancy Hestia.ai.

2. Test group of European policy makers and societal influencers

Decision-makers and influencers from five European countries wanted to learn about the data economy approaches, personal data flows and to understand digital power. Despite having a good understanding of the data economy, many of the test subjects were surprised at the scale and extent of personal data collection.

The investigation continues Sitra's 2019 survey on individuals' digital footprints, which tracked how much data is collected about the users of digital services and how many third parties their data flows to. Back then, one of the six test subjects was a Member of Parliament, which aroused interest on the flow of data from decision-makers.

Key findings of the survey:

- a vast amount of data is leaked to unknown actors
- the effectiveness of the General Data Protection Regulation (GDPR) is limited
- our privacy is constantly at risk
- the largest platform and data companies run their business in a way that is problematic for society and people's overall well-being

Test subjects included social influencers in different roles

The Digipower investigation involved 15 socially active and well-known test subjects from Finland, France, Ireland, Italy, and the United Kingdom. Most of them were political decision-makers from left to right, but other socially significant influencers, such as EU officials, were also involved.

For data collection and analysis, a secure "bubble" was formed with selected researchers, allowing the investigation to be carried out with respect for the data security and privacy of the test subjects. Each test subject was also able to decide what information was documented on websites, social media or, for example, by this publication.

Participation in the investigation was demanding and the schedules of the test subjects were exceptionally challenging, forcing some to drop out midway. And for some test subjects the data findings are not given in detail in the publication.

Test subjects' Digipower journey

“There's still work to be done on transparency in the data economy and the use of power.”

The test subjects had different motivations for taking part in the investigation, but all were united by a desire to understand the mechanisms of the data economy and their impact on individuals' lives and society. Knowledge of the data economy varied somewhat among the test subjects, depending on their background and role. Each test subject selected the companies of most interest to them for their data journey and particular attention was paid to the data collected by these companies.

Both local and international companies were among those selected by the test subjects. The flow of data was monitored and data requested from both traditional retailers and data giants.

Although the investigation researchers are experts in their field, dealing with companies was often difficult and time-consuming for the test subjects. The majority of the companies selected for the investigation either did not respond to requests for data or provided incomplete responses. Typically, responses omitted relevant information, such as about data bought externally or derived from personal data and profiling.

Test subjects in the investigation

- **Anders Adlercreutz**, Member of the Finnish Parliament
- **Leïla Chaibi**, Member of the European Parliament
- **Filomena Chirico**, European Commission official, Member of the cabinet of Commissioner Thierry Breton
- **Christian D'Cunha**, European Commission/ Directorate of Communications Networks, Content and Technology
- **Stéphane Duguin**, CEO of the CyberPeace Institute
- **Atte Harjanne**, Member of the Finnish Parliament, Chairperson of the Green Parliamentary Group
- **Jyrki Katainen**, President of Sitra
- **Dan Koivulaakso**, State Secretary
- **Miapetra Kumpula-Natri**, Member of the European Parliament
- **Markus Lohi**, Member of the Finnish Parliament
- **Tom Packalén**, Member of the Finnish Parliament
- **Sirpa Pietikäinen**, Member of the European Parliament
- **Mark Scott**, Chief Technology Correspondent, Politico
- **Niclas Storås**, journalist, HS Visio
- **Sari Tanus**, Member of the Finnish Parliament

Anders Adlercreutz

OCCUPATION:

Member of Parliament, Swedish People's Party in Finland

EXPERIENCE IN THE DIGITAL WORLD:

active since 1997

FAVOURITE SOCIAL MEDIA CHANNEL:

Twitter



“You make a deal and think you’re the buyer, but actually you’re the one being traded.”

Highlights from findings

Finland was in the process of acquiring new fighter jets through an international competitive tender worth billions of euros. Boeing, the American fighter manufacturer, targeted Adlercreutz several times in its extensive Twitter campaign. He became a target of advertising because he was of a certain age, used Twitter in Finland and belonged to the so-called look-a-like audience. Artificial intelligence used by social media platforms analyses user online behaviour, such as purchasing behaviour, and ads can be targeted to similar people who would normally buy the desired product. So, somewhat surprisingly, Twitter uses natural persons as targeting criteria in addition to other targeting methods.

The Washington Post collected information about Adlercreutz’s interests based on the stories he read and specific keywords and expressions used in their content, while generating its own digital advertising offering to compete with Google and Facebook. The data revealed an operator called Clavis, a product created by the Washington Post and inspired by Amazon’s successful recommendation engine. The Washington Post has compiled a list of Adlercreutz’s interests based on the articles he reads and specific

keywords and expressions used in their content. Clavis uses this to recommend new content with a high degree of accuracy. Readers are also segmented using third-party data.

But recommendations are not only done for editorial purposes, and the Washington Post’s Clavis tool is also used in sales to measure the effectiveness of advertisements and so to price them. Jeff Bezos bought the Washington Post in 2013, so the development seems natural. In just a few years, Amazon has managed to become the third largest corporation in digital advertising, just behind Google and Facebook. On top of this, Clavis is also able to easily combine data with Amazon’s cloud services.

How did your participation in the Digipower investigation change your understanding of the data economy and how it works today?

“I expected to be surprised by the amount of data collected about me – although of course you imagine that you see through the attempts to influence you digitally.

The investigation clarified my conception that service providers don’t just sell their services to you, but that data to an increasing extent travels in two directions. Sure, I pay, but someone else pays for the data. This earning logic involves much more than just a buyer and a seller.

The scale of activity surprised me, although I was aware of it as such. I was also surprised by the complexity, the number of players and the integration of data collection practices. Companies collecting data are well integrated into this world and into the service systems. They are specialised in packaging, pricing and selling my behaviour and preferences to others.”

How do you think the rules of the data economy should be changed?

“I got access to my data through being a Member of Parliament and being involved in this project – but the access shouldn’t depend on a process like this. People should be able to clearly see what their information is being used for, who is using it and who is benefiting from it – and this flow of information should be made visible. One way of doing this could be to price services differently. A newspaper subscription could have the option that data could go here and there, in which case the data flow would have to be opened up: when you hear news from us, data about your activity will be passed on.

Data collection will inevitably lead to each of us being fed the kind of content that we like. It leads to [filter] bubbles. The news

feed is driven by algorithms, which certainly have an impact on policy making. Everyone gets the messages that they like and thinks that everyone else can see the same content. How does that in turn affect our image of the world when there is only one kind of truth available?”

What do you think everyone should understand about the data economy?

“What everyone should understand about the data economy is that service providers do not provide services in one direction to consumer, but provide services to many others. The picture of the data economy is bigger than what you see when you enter into an agreement, and you are the subject of the transaction in that entity.”

Leïla Chaibi**OCCUPATION:**

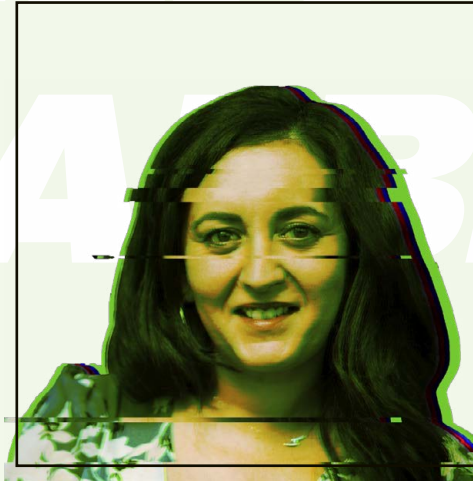
Member of the
European Parliament

**EXPERIENCE IN
THE DIGITAL
WORLD:**

13 years in social
media

**FAVOURITE
SOCIAL MEDIA**

CHANNEL:
Instagram



“Law alone will not bring about change – we also need resources for implementation.”

How did your participation in the Digipower investigation change your understanding of the data economy and how it works today?

“Before the investigation, my image of the data economy was vague. I understood that the topic was important, but now my understanding is concrete. Now when I’m online, I can see that there’s another hidden world in the background.

When I received the results of the investigation, I was taken aback and angry. I had not been previously aware of the scale of the issue. What surprised me most was the data collected by Uber. Uber had information that even my colleagues didn’t have – they know the location of my office and where my friends live. Even some websites, such as the newspaper Le Monde did the same. They rely on people not asking for their data.”

How do you think the rules of the data economy should be changed?

“I think this market needs to be regulated – we can’t let companies regulate themselves. They are more powerful than the law. Legislators need to regulate this activity and for example increase transparency. We need resources for international oversight of data collection. Such a watchdog already exists, but it lacks resources – they have very few employees.

People are not aware of how big the problem is and how many more resources would be needed to solve it. That is why companies can do what they want, because no one is holding them accountable. Law alone will not bring about change – we also need resources for implementation. We also need to inform the public. By using the internet, you agree to give up your right to your data.”

What do you think everyone should understand about the data economy?

“Everyone should care about the issue because the data economy affects everyone. It’s not just about people with computers, it concerns everyone. Many companies seem to be selling a service, but in reality their main source of income is data collection. Uber is not about selling car rides; their main business is getting data. There are things that I don’t remember but they do. They have a kind of a snapshot of my life.

People should be able to see their own data to understand what it really is. For instance, an instructor or an instructional video to show how to read their data. Without expert assistance, the data is very hard to understand.”

Filomena Chirico

OCCUPATION:
European
Commission
official, Member
of the cabinet of
Commissioner
Thierry Breton



Highlights from Findings

Filomena Chirico obtained her data from Twitter, Netflix, Uber, Politico and Google. Google and Facebook were poor at responding to subject access requests.

Filomena Chirico's work makes her a particular target of interest for data economy operators. IAB Europe, an organisation representing the interests of operators in digital marketing and advertising ecosystems, used data about her age, location and the Twitter users she followed to communicate their views to Chirico. Politico Europe also targeted her with advertising on Twitter and used the areas of her political interest for further targeting.

Christian D'Cunha

OCCUPATION:
Works at the
European
Commission,
Directorate-General
for Communications
Networks, Content
and Technology,
which was
responsible, inter
alia, for writing the
Data Act.



Highlights from Findings

Christian D'Cunha obtained his data from Uber, WhatsApp, Google, Facebook, Colruyt, Strava, Mobile Vikings and BBC. Uber and Facebook responded poorly to subject access requests.

Christian D'Cunha was informed by the UK Labour Party that the data handling company used by the party had been a target of cybercrime and his data had been compromised, even though he was no longer a member of the party at the time. There should not have been any information about him in third party databases. The case also revealed that D'Cunha's data had been exchanged between the party and Facebook, as well as between his telecommunications operator Mobile Vikings and Facebook.

Stéphane Duguin

OCCUPATION:
CEO of the
CyberPeace
Institute

**EXPERIENCE
IN THE DIGITAL
WORLD:**
Early adopters

**FAVOURITE
SOCIAL MEDIA
CHANNEL:** I don't
use social media.



“Everyone should understand that platforms benefit a lot from people.”

How did your participation in the Digipower investigation change your understanding of the data economy and how it works today?

“Participating in the Digipower investigation confirmed my previous understanding of the data economy. The results didn't surprise me. The platforms collect a lot of data, and it is difficult to understand why without the link to the data economy. All this data is collected and packaged without providing any value to users. This didn't come as a surprise – the situation is just as bad as we thought.

The investigation itself was a positive experience, but the results confirmed just how uncontrolled the situation is. I have seen data in my work, but it was sad to see how ambitious laws like the GDPR are still so inadequate. Accessing data is very difficult for ordinary citizens – without support, getting access to data is problematic. This complexity surprised me.”

How do you think the rules of the data economy should be changed?

“The rules of the data economy could be changed by investing in and supporting access to personal data. It would be important for people to understand the subject and seeing your own data would help strengthen it. Once people see their own data, they would start to ask why so much of it is collected, and the platforms would have to be responsible and held accountable for the data they collect.

Data collection shouldn't be regulated by habit, behaviour or the market. It should be regulated by law, in a way that takes into account the public interest and impact on people. There is the potential for capacity building and training for people. The public has a very important role to play in understanding why data processing needs to be regulated.

What do you think everyone should understand about the data economy?

“Everyone should understand that platforms benefit a lot from people. The idea of getting something cheap or for free is very misleading, as platforms benefit considerably from us. In EU democracies, people are well aware of data collection and actively inquire about their rights when public authorities collect data. This is a very good thing. But the same people provide personal data to platforms that do not comply with the data security laws at all. This is very worrying.”

Atte Harjanne

OCCUPATION:

Member of Parliament and Chairperson of the Green Parliamentary Group

EXPERIENCE

IN THE DIGITAL

WORLD: All my life, or 37 years, internet since the mid-1990s

FAVOURITE

SOCIAL MEDIA

CHANNEL:
Instagram



“What everyone should understand about the data economy is that data is power.”

Highlights from findings

Atte Harjanne obtained his data from Sanoma, Wolt, Voi, Bookbeat, Spotify and Google.

The data showed that the media house Sanoma has built a very accurate profile of Atte Harjanne's interests. The company also has information on his purchasing power, right down to the devices he uses. Gambling had been marked as one of his interests, as Harjanne has had to follow the gambling industry for his job.

How did your participation in the Digipower investigation change your understanding of the data economy and how it works today?

“The Digipower investigation did not change my understanding of the data economy but updated and refined it. It was quite enlightening to see how the GDPR is being complied with in practice.

As a consumer in today's data economy, I'm aware that everything leaves a trace and that companies accumulate a huge amount of data on what I do. But it's still thought-provoking to get your data into your hands and go through it. At the same time, I was left thinking that data in itself does not tell us

everything about how it is used and combined. There is still work to be done on transparency in the data economy and the use of power.”

How do you think the rules of the data economy should be changed?

“Everyone should be able to monitor and control their own data and how it is collected and used. It is important that legislation keeps pace with technological developments and that democracy is in the driver's seat in a changing world. Current regulation does not yet fully guarantee this, and the problem is that data policy does not receive the attention it deserves in decision-making or by political parties. The structures of public administration and decision-making do not sufficiently reflect the changing world in this respect.

Legislation needs to be strengthened, but that's not easy to do. Legislation should at the same time be up to date but still based on general principles so that it does not become immediately obsolete. It is definitely worth using the EU's muscle in this area, but this does not exclude smart national regulation.

Data collection is now a central part of the business logic of online services and there is probably no going back to the world before. It would be essential for the users to have genuine power over and perspective of how their data is collected and used. I doubt that everyone understands the huge amount of user data that is currently collected and stored. Personally, I don't necessarily realise it myself.”

What do you think everyone should understand about the data economy?

“What everyone should understand about the data economy is that data is power. A data economy that concentrates power is shaping the world at an incredible pace, and it would be important for democratic decision-making to be in the driver's seat.”

Jyrki Katainen

OCCUPATION:
President of Sitra

**EXPERIENCE
IN THE DIGITAL
WORLD:**

I don't feel like I'm a digital native, but a large share of my life revolves around digital applications.

**FAVOURITE
SOCIAL MEDIA
CHANNEL:**

Twitter – maybe not my favourite, but the only one I use.



“This strengthened my understanding of our roles as silent partners.”

Highlights from findings

Jyrki Katainen obtained his data from Sanoma, K-Group and Twitter. Apple, Twitter, Google and Zalando were poor at responding to subject access requests and, with the exception of Twitter, Katainen failed to obtain his data from them.

The data collected by the major retail chain K-Group amounted to a 172-page document of Katainen's data. Most of it comprised purchase and other data accumulated during his relationship with the retailer. Some of this data goes to Google via Google Analytics, among other things. When Katainen searched the store's app for a recipe for spaghetti carbonara, the data was sent to Google.

K-Group places its customers into five categories: enthusiasts, indulgers, woke, comfort-seekers, established/conservative. Each month, they analyse customer data and use the results to categorise each customer. This was what Katainen's categorisation looked like in December 2019 according to different probabilities:

- Enthusiast 17%
- Indulger 12%
- Woke 20%
- Comfort-seeker 38%
- Established/Conservative 13%

The data is used to recommend appropriate content. For example, a Christmas recipe for plum pastries is more likely to be recommended for customers who are weighted towards the categories enthusiasts, indulgers and comfort-seekers.

Boeing, the American fighter jet manufacturer, also targeted Katainen on numerous occasions in its extensive Twitter campaign.

How did your participation in the Digipower investigation change your understanding of the data economy and how it works today?

"My assumption was that data does move online. But what surprised me the most about the investigation was how widely data ultimately spreads. The data collected was traded between different actors, and the scale of that was also surprising. The investigation strengthened my understanding of our role as silent partners. There isn't much users can do, regardless of whether they give permission to the use of their own data or not.

The investigation did not reveal any major surprises or overwhelming risks. But it did confirm my perception that the system is not sustainable. We need fairer rules for the data economy so that data can be shared and used in a way that generates trust. Unfortunately, we cannot now rule out the possibility of data falling into the wrong hands or being used for the wrong purposes. At the same time, I don't want to demonise all companies. But it's worth remembering that data security is never totally foolproof."

How do you think the rules of the data economy should be changed?

"People should have more control over data management. Especially as data will be shared and used massively more in the future. On the other hand, leveraging data will also increase productivity in every sector, and for

that to happen people must be able to trust and be protagonists in the data economy. At present we are too dependent on gatekeepers.”

What do you think everyone should understand about the data economy?

“It’s good to understand that a lot of data about us is really being passed on and to a very wide range of parties. Information is also shared in situations we would not even think of. Even if you, say, visit a bricks-and-mortar shop, when you pay, information that

too sends data to many different parties. Few people even understand why this happens and who is getting our data.

We are also being targeted by attempts to influence us, which can happen either in a quite positive sense or, unfortunately, also in a manipulative way. All the information that is being forwarded about us will also be retraded. The trading of data can take place even much later. This information can be used for a wide range of purposes, such as influencing political opinions, advertising and other commercial purposes or for wielding influence in a much wider sense.”

Dan Koivulaakso

OCCUPATION:
State Secretary,
Left Alliance

**EXPERIENCE
IN THE DIGITAL
WORLD:**
about 14 years on
social media, longer
online

**FAVOURITE
SOCIAL MEDIA
CHANNEL:**
I guess it's a love-
hate relationship
– I probably like
Instagram the most,
although I don't use
it much.



“Everything you put online becomes a commodity and people become a resource for the data economy.”

Highlights from findings

Dan Koivulaakso obtained his data from Sanoma, Google, WhatsApp, HSL, MTV and Yle. He was able to confirm that Signal did not have any of his personal data. Google and Facebook responded poorly to subject access requests.

Dan Koivulaakso compared different media houses and their applications. As expected, while Yle used only a few tracking tools, the Helsingin Sanomat application was full of them. Somewhat surprisingly, the commercial Finnish television channel MTV used tracking moderately.

How did your participation in the Digipower investigation change your understanding of the data economy and how it works today?

“Participating in the Digipower investigation underlined my earlier views of the data economy and how it works. The internet remembers things that you don't.

I carried out the Digipower experiment in a secure way, meaning that I did not use my own normal accounts. I provided

information about the services that contain my data in connection with the subject access request. I'm sure that my own social media accounts would have revealed something even more dystopian. Now I only used accounts created for the experiment. Because of my preconceived fears, all the worst features still remained hidden from me.”

How do you think the rules of the data economy should be changed?

“One solution to how to regulate the data economy could be to break up companies that are too big and, for example, communicate more widely about the features of the GDPR. The service promise given to users based on legislation is strong, but people should be more informed about it. The rights to delete personal data should be strengthened.

In the data economy, there should be a broad discussion on what is delimited as private and how it is to be safeguarded. This would guarantee rights in reality, not just in theory. This experiment was an excellent example, and on this basis I would encourage everyone to make subject access requests to the services they use. You will be surprised at what data has been collected and how much is stored.”

What do you think everyone should understand about the data economy?

“Everyone should know that the Internet doesn't forget. Understanding how the data economy works has probably been easy for those of us for whom the web only became part of our lives in late teens or early adulthood. Adopting them was a choice then, not something we automatically grew into. We need a broad awareness of the impacts of the data economy.

People generally think that there's certainly no interesting data about them online

that would need to be protected. That's not the point. Everything put online becomes a commodity and people become a resource for the data economy. People should have the right to control it themselves. Making subject access requests must be made far simpler than it is. The GDPR gives us good opportunities for that. It will also force companies to handle data better.”

Miapetra Kumpula-Natri

OCCUPATION:

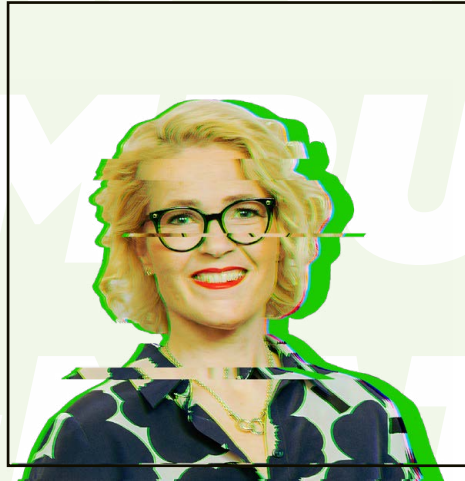
Member of
the European
Parliament, Social
Democratic Party

EXPERIENCE IN THE DIGITAL WORLD:

Twitter for about
10 years, Facebook
slightly longer

FAVOURITE SOCIAL MEDIA CHANNEL:

Twitter for
following news
items, Facebook
and Instagram for
sharing news



“People should have more control over their own data”.

Highlights from findings

Miapetra Kumpula-Natri received information from HSL, Gigantti, Uber and Google.

Data from the purchases made by Miapetra Kumpula-Natri at the home electronics chain Gigantti's bricks-and-mortar shop, including her personal data, was sent from the company to both Google and Facebook. Clicking on a link in the chain's digital advertising letter on the phone revealed Kumpula-Natri's location at her holiday home even though the location detection capability was not used. According to the investigation researchers, it is possible that Gigantti identified the IP address of the device's Internet connection and used it to determine her location. In this case, the test subject could not pinpoint a situation where she would have given permission to track her location. The subsequent response by Gigantti confirmed that they had used the IP address.

When Kumpula-Natri requested her data from the home electronics company, she also found out how much data about her the company had bought from other companies specialising in personal data collection. For example, data about her assumed financial situation and family status had been obtained elsewhere, although much of it was inaccurate or false.

The data and the responses to the subject access requests showed that by accepting a cookie on Gigantti's website, the customer also accepted 231 cookie partners, including a cookie of Russian service Yandex with a retention duration of 10 years. This cookie partner and cookie were removed by Gigantti in March 2022.

How did your participation in the Digipower investigation change your understanding of the data economy and how it works today?

“Before I took part in the study, I assumed that there would be a lot of data collected about me. With the introduction of stricter consent requests, I realise how variable the practices for collecting consent still are. Wherever you are, current data is collected about you even if you just have your Wi-Fi or different software open.

I knew quite a lot already, but the experiment gave me a more concrete analysis of what a company might know. It rendered every keystroke delay visible. The investigation showed me the amount of trading where data was sent to dozens or even hundreds of places. For example, my data was sent offshore where it is stored for 10 years. The user isn't told what data is being transferred and where.

The most surprising part of the investigation was the scale of the trading. The lack of transparency was also surprising, although many might think that it doesn't matter if the data is only for advertising purposes. But there is a fundamental issue of privacy in targeted advertising, and it's not just marketing. It's also important to consider whether everything remains at the level of marketing. On the other hand, it is also great that small-scale entrepreneurs can target their marketing properly. But is it necessary on such a large scale?”

How do you think the rules of the data economy should be changed?

“Transparency in the data economy is the next more challenging step. The end user should always be in control of their data. The user also needs to know what data has been collected about them and where it is going. Data must be under the user’s control – and not so that when you use a particular app everything about you becomes a commodity.

It would increase user understanding if it was made more transparent and apparent what each click means. It is also a prerequisite

for being able to manage your digital identity securely.”

What do you think everyone should understand about the data economy?

“People should have more control over their own data. Society should create a safe environment in the digital economy for everyone. The starting point should be such that, as a rule, the online environment would be safe and easy for everyone to use.”

Niclas Storås

OCCUPATION:
HS Visio/ journalist

**EXPERIENCE
IN THE DIGITAL
WORLD:**
Nearly 20 years

**FAVOURITE
SOCIAL MEDIA
CHANNEL:**
I don't really like
any of them. I spend
the most time on
Instagram.



“People tell Google things they don't even tell their best friends.”

How did your participation in the Digipower investigation change your understanding of the data economy and how it works today?

“I've been following the topic for years and participating in the investigation did not change my perception. But the amount of data obtained through the investigation was surprising. It was only when a company sent me the data they have about me that I saw how much there really is.

In terms of quantity, the material consisted of dozens of folders with several different files and, for example, all my YouTube viewing history over a ten-year period. I had no idea of the full amount until it was sent to me. Of course you know all the time that there will be a trail, but you take it more seriously when you see it.

We are quietly surrendering an enormous amount of information about our own lives. A single scrap of Google data is not significant as such, but when you accumulate material over ten years, you get a huge mass of data. Online searches show where you are in your life, or whether you have done some renovations or bought a bicycle. In other

words, everything you have been thinking about. This is how we help companies to profile both ourselves and other people like us.”

How do you think the rules of the data economy should be changed?

“In terms of regulation, it would be important for companies to open up the logic by which algorithms work. Companies should also be forced to include an option for users to turn off tracking. Apple, for example, already has one.

Regulation must come through legislation, but companies must also develop it themselves. At the moment, the EU is in the process of creating a different regulation to deal with algorithms and cracking down on data giants. If you compare the situation to traffic, there were no traffic rules at first. It was only when accidents started to happen that different traffic rules were developed and seat belts were introduced. Similarly, the digital economy should be made safer for all of us.

Everyone should have the right to ask to see all data concerning them. At the same time, we should remember that whenever we give personal information about ourselves, we work for the company concerned. People tell Google things they don't even tell their best friends.

At the same time, however, we should not make a complete about-turn that would destroy the basis for many businesses and make it difficult for many people to work. This is a balancing act for which legislation, for example, can provide a safe framework. We need to think about how to make it as easy as possible for small entrepreneurs to explore and gain benefits from data, but without power being concentrated in the hands of large corporations as has been the case.”

What do you think everyone should understand about the data economy?

“In the data economy people should be given the opportunity to share the things that interest them. Then they could see content related to those topics. Users should also be told why their behaviour is being tracked and what the data collected is being used for.”

Sari Tanus**OCCUPATION:**

Member of
Parliament,
Christian Democrats

**EXPERIENCE
IN THE DIGITAL
WORLD:**

Seven years on social
media, much longer
with online search
engines, since the
early 2000s.

**FAVOURITE
SOCIAL MEDIA
CHANNEL:**

Facebook



“The data economy is already a huge economic sector and will grow even more in the future.”

Highlights from findings

Sari Tanus obtained her data from Instagram, S-Group, Sanoma and Gigantti. She requested her data from Aamulehti but was surprised to receive data from Helsingin Sanomat as well, as the companies both belong to the Sanoma Group.

She was profiled by the use of, among other things, data bought from a third party, Bisnode (owned by the American company Dun & Bradstreet). Both her and her entire household were assessed according to wealth, purchasing power, risks and level of education.

How did your participation in the Digipower investigation change your understanding of the data economy and how it works today?

“The Digipower investigation clarified how extensively data is collected, grouped, set to a certain format and disseminated. Data can be a tool for power, but it also has economic value. It can be re-sold and sold at different levels. Every click and article leave a trace.

The level of detail about our purchases and clicks on articles was astonishing. There

are thousands of pages of material detailing these things. I had never heard of many of these companies, and they were all over the world. When companies were asked for more detailed information, some provided it, while others you had to ask again, or didn't send it at all.

All this surprised me, and I felt kind of exasperated – not just for me, but all of us being used as a resource by the data economy on such a large scale.”

How do you think the rules of the data economy should be changed?

“Regulation should be clear, transparent and enabling. As clear as possible so that ordinary folk can understand it. Enabling so that as the technology develops, the measurements and data needed for research are there. Freedom of speech and opinion must be remembered, and it must be possible to decide and determine your own affairs. This is a very big challenge because the sector is developing by leaps and bounds. We don't even know what the challenges will be in a few years' time. Legislation should be proactive.

From an international perspective, it should be forward-looking in the wider context, but this is a big, complex and difficult task. There may be certain EU policies, but it would be important to keep decision-making power in our own hands.”

What do you think everyone should understand about the data economy?

The data economy is already a huge economic sector and will be even bigger in the future. There are huge numbers of international companies, huge amounts of power and huge amounts of money moving around in the sector. Everyone should have a realistic view of it. On the economic and business side, I believe that Finland does not have

a realistic overview of the issue. Important things will be lost if we don't consciously hold on to them.”

3. The investigation used three methods to collect data

The Digipower investigation traced data and companies in three different ways. The test subjects used their rights under the GDPR, downloaded their data from data portals and also monitored the movements of their data with a monitoring app.

Three complementary methods were used in the investigation:

1. Where available, personal data was downloaded through a download portal provided by an individual service (the largest services generally have portals for downloading data).
2. A subject access request (SAR) was sent to the service provider in accordance with the GDPR.
3. The data flows to/from the mobile app of each service provider was monitored through the Tracker Control application downloaded onto the test phone.

Data download portal

In response to the growing number of data requests, many companies have sought to facilitate the time-consuming and labour-intensive process by providing self-service tools that enable users to access their own data.

Even if only a small percentage of users of companies' digital services actually submit successful subject access requests, this percentage could still mean an unmanageable number of requests for companies with millions of users.

For this reason, companies such as Google and Facebook have introduced automatic "Mydata download" on their websites or applications which any user can

log in to and download zip files of their own data.

Subject access request

A subject access request is a mechanism set out in the GDPR for users to ask for information about, and copies of, what data is collected on them and how it is used. The request is usually sent by email, although in some cases companies require a paper form or an electronic form to be filled in to start the process.

Once a subject access request has been submitted and the service provider has successfully identified the user's account and/or verified their identity, the organisation has 30 days to respond.

The GDPR currently applies to all people living in the European Union and all customers of companies operating in the European Union.

Tracker Control monitoring application

For about two weeks, the test subjects used an Android test phone with the Tracker Control app developed by Oxford University researcher **Konrad Kollnig**.

- Tracker Control monitors the outbound connections of the mobile services of the companies investigated

and the domains to which these applications connect.

- Tracker Control acts as a local virtual private network (VPN), making it possible to see the connections.
- Tracker Control does not see the content of messages or what data is transmitted and in which direction. The iOS operating system includes a similar functionality since version 14.0.

The responses data download portals and separate subject access requests provided on an individual's data may be incomplete or even non-existent. Tracker Control provides an overview of connections between applications and trackers in a standardised format covering applications and participants.

The coaching that the researchers provided for the test subjects was an

integral part of the Digipower investigation. The test subjects were guided in how to interpret their own monitoring results data in the HestiaLabs user interface, which displayed the CSV data produced by the Tracker Control application in a visual format. The HestiaLabs user interface also allowed data returns from subject access requests or data download portals to be visualised and explored.

This user interface does not send data to any internet server, but all data remains local to the browser to ensure privacy. At a basic level, the UI provides the ability to view CSV files as tables and JSON files as directory trees. In general, responses to subject access requests are usually received in a form that is difficult to interpret, so Hestia.ai created easy-to-understand interfaces for viewing files in order to clarify and harmonise the responses provided by all companies.

Figure 1. Example of Tracker Control data visualised using HestiaLabs' interface



The Digitrail survey conducted by Sitra in 2019 examined how much data is transferred to third parties and how many third parties there are. The same methods were used in part in the Digitrail survey and the Digipower investigation. In the Digitrail survey, the movements of the data were monitored for selected services and third parties were searched and identified using

the WebXRay database and manual check.

Although the two studies used similar methodological tools, there are differences between them. The Digipower investigation focused in more depth on the flow of data between different services, but more importantly examined the diversity of the picture of the test subjects that data provides and the power it gives to data collectors.

4. Key findings: Digital power is self-reinforcing, individuals have no control over their data

The collection and use of personal data on decision-makers and influencers is at least as extensive as for ordinary people. Digital power is used to influence decision-makers in a variety of ways and policymakers are vulnerable to online influencing.

At the time of writing, the EU's GDPR has been in force for four years, but data giants from outside Europe still do not comply with it. There is a deliberate indifference and evasiveness towards individuals' requests for data – and thus toward European legislation. At the same time, European companies have invested enormous sums on developing their skills, processes and systems to comply with the GDPR.

An individual person's ability to control the collection and use of their own data is almost as poor as was seen during Sitra's Digitrail survey three years ago, and the right to privacy, among other things, is still not fully respected. The same can be said about the situation of politicians and other social influencers. The results of the Digipower investigation show that decision-makers have as much difficulty as ordinary individuals in obtaining information on the movements of their data, and that data is collected on them just as much as on others.

Digital forms of power are self-reinforcing

The world's most successful data giants, such as Google and Meta, have been creating an "ecosystem of ecosystems" – or perhaps more accurately a super-ecosystem – of data

collection with numerous data collection technologies and methods. It would be difficult for them to credibly justify the serious problems in replying to the subject access requests in this investigation on the grounds of lack of resources or technology. The companies developing metaverses, artificial intelligence and other state-of-the-art technologies must therefore have other reasons for not complying.

There have been calls for years for transparency in digital advertising ecosystems. There has been hardly any change, as it simply does not serve the majority of organisations in the super-ecosystem: the giant platforms, the large social media, the major technology companies, marketing technology companies, data traders and data processors each get their share of the business as it currently is.

Hestia.ai calls the super-ecosystem a "digiscape", a digital landscape. A digiscape includes the idea of both the physical and a completely digital environment and the combination of the two, as the physical and the digital world are inexorably intertwined in numerous ways. A digiscape consists of both the context and the contents provided in different contexts.

The large companies operating in digiscapes exercise enormous power on both the context and the content, as the most popular

platforms and the largest services have become the infrastructure that society needs. We should understand that if the warped practices of the data economy remain unchanged, the power of large companies will continue to grow as different aspects of life – moments, connections and events – are copied and digitalised. They will also become continuous, observable and further analysable data streams.

In today's data economy, the winners are those entities that can collect vast amounts of diverse data. At the same time, they benefit from the “feedback loop” of the system, where new data can refine the data already in possession of the organisation and increase its value. In this way, power is further increased and concentrated. The large amount of data enables them to make a large number of accurate analyses, predict what will happen and thus to forecast, to make choices that are optimal for their own activities. The end result is that those with the greatest digital power are able to partly shape the future in the direction they want.

At present, international data giants sit on top of large and diverse data flows, defining and shaping both context and the content. They have infrastructural power not only on individuals but also on companies and society as a whole. A digital oligarchy has emerged.

The strongest understanding of the data economy comes from personal experience

The Digipower investigation also aimed to raise people's awareness of the collection/use of data and thus to increase their activism for a fairer and more just data economy. We examined the effectiveness of the tools currently available to individuals to access data and understand corporate data practices.

Sitra's 2019 Digitrail survey also examined the performance of the GDPR slightly over a year after its entry into force. At the time, no responses to the subject access requests were received. The conclusion drawn was that the GDPR did not sufficiently protect individual rights, there was a lack of transparency, the provisions on data protection were inadequate and people could not see how their data is circulated.

The means available to people to obtain their own data are to access it through the service provider's data portal and to submit a subject access request. The self-service tools offered by companies, data download portals, only provide some of the data that individuals have the right to access under the GDPR. The data obtained through a portal represents data that the company has voluntarily made available, but there is no guarantee that all the legally required or user requested data has been provided. In addition to using the portal, it is still necessary to request additional data by email, as a subject access request under the GDPR, if one wants to obtain a complete picture.

A subject access request asking in detail for information about your data is a more effective way to find out what data has been collected and how that data has been processed than relying on data download portals (which in any case are often not available). For this process, we created an email template (Appendix 1) based on a template developed by Personaldata.io. The subject access request template goes through all the provisions of the GDPR related to data use, data provision and explanations to spell out exactly what a company should return.

We identified specific categories of data that have to be returned and clearly communicated that all data from all sources must be given from applications, websites, devices, personal visits and external sources, in a machine-readable format and with explanations.

The email message requesting the data was sent to the data protection officer of the

service provider or to the specific email address provided in the provider's privacy policy. In the Digipower investigation, the test subjects collectively requested information on the data collected on them from a total of 40 different service providers in ten different countries.

The UK has legislation identical to the EU's General Data Protection Regulation, so the country applies the provisions in the same way as in the EU. In the case of the test subject in the UK, we adapted the wording of the request template accordingly, and in the case of the participant in Switzerland (outside the EU), the model was adapted to refer to the corresponding Swiss act (Loi fédérale sur la protection des données et ordonnance).

Many companies respond to subject access requests even from outside the EU and UK, meaning your own data may be available to you worldwide, even where no legal right exists.

Understanding your own data and how it is used is hard work and may not succeed at all

Accessing your own data is a time-consuming and difficult task. In the Digipower investigation, the coaches did everything they could to make the process easier for the test subjects but submitting subject access requests still required a lot of work from the participants. Some test subjects could not continue because of their workload and had to cancel their participation in the part dealing with the subject access request.

Several of the test subjects experienced delays in receiving responses despite the fact that the GDPR requires a response within 30 days. Some companies invoked the three-month delay allowed by law immediately upon receipt of the request due to the complexity of the request (including Booking.com and KLM, and Google in the case of

Chirico), but in some other cases the companies did not respond at all within 30 days.

In some instances, the companies started counting the 30-day response period from a date other than that of the original message sent by the test subject. It is probably not the intention that a company can first spend three weeks approving the subject access request and only then start the 30-day processing period.

There were clear shortcomings with the data returned. The responses received were often disappointing, resulting in information that was difficult to understand and use, or that did not answer in the questions posed. Even when the subject access requests were responded to, not all of the information requested was provided. Previous studies also show that information such as profiling data and the information on data sharing, exchange, and the acquisition of data through third parties is often missing or poorly provided.

Four test subjects (Anders Adlercreutz, Filomena Chirico, Christian D'Cunha and Niclas Storås) submitted a subject access request to Facebook, but in violation of the GDPR, no data was provided. Facebook was the clearest example of how the provided data did not answer the questions asked at all. The company also systematically avoided responding to individual, tailored subject access requests, thus refusing to provide data that clearly falls within the rights under the GDPR. Facebook views its data download portal as being sufficient.

Specialist Riitta Vänskä, a member of Sitra's working group carrying out the investigation, made a request under the GDPR to Facebook at the same time as the test subjects. Facebook actively answered with extremely long email messages for about a month, initially directing Vänskä to retrieve her data from Facebook's automated portal. Vänskä repeatedly justified her request for additional information by explaining that she had already retrieved data from the portal but that the service did not provide all

of the data that an individual is entitled to receive under the GDPR. Time and again, Facebook's responses were copies of their privacy policy and terms of use.

After a month of exchanging messages, Facebook stated that they had complied with the GDPR and the company would not give any more information.

Vänskä submitted a complaint about Facebook's actions to the Finnish Office of the Data Protection Ombudsman on 24 January 2022 and was still waiting for a response in April. In March 2022, she received a notification from the office explaining that the service was backlogged and the complaint had yet to be processed, and that it would most likely be dealt with using a one-stop-shop principle together with the member states.

Because the Meta Platforms (Facebook) is in an EEA territory, Ireland, Ireland is the lead data protection authority. The message

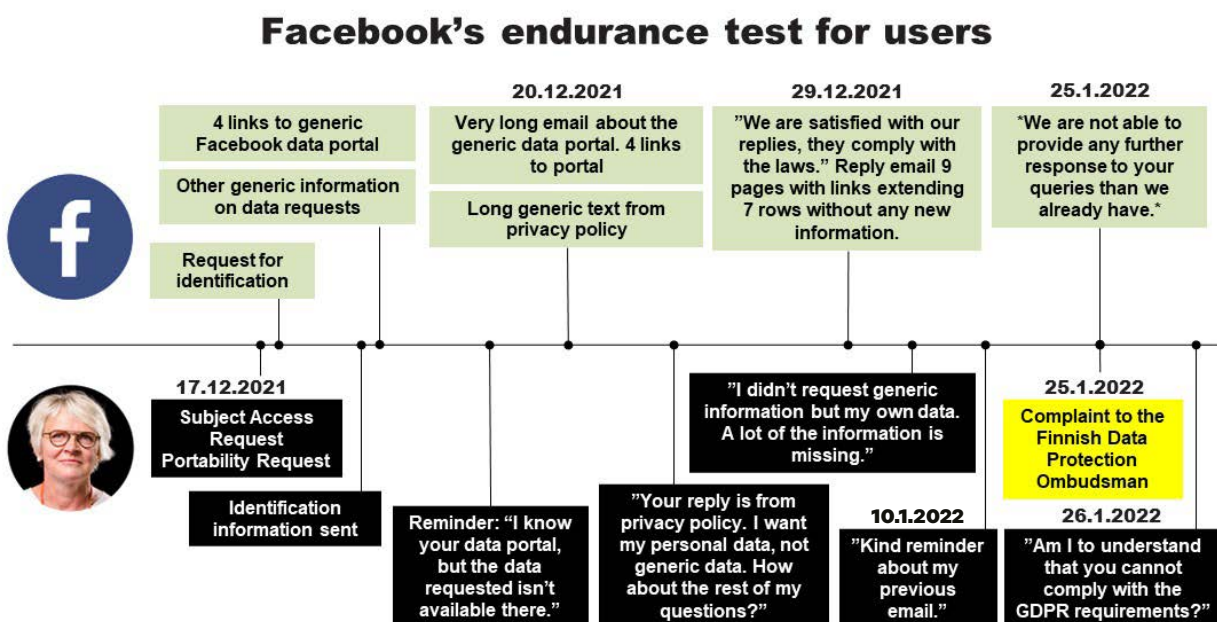
also said that the process would last more than a year. Vänskä later discovered that the overall process was slow because of the different processes of the member states, but that it would speed up in future.

Does everyone get the same level of service for their subject access requests?

There was much variation in the responses received by the participants, both in terms of quality and in the level of detail and comprehensibility of the data. For example, there were great differences in the quality of data provided to different test subjects by Gigantti, Apple, Google and HSL (Helsinki Region Transport).

In several cases, the test subjects had to rely on their own contacts within the

Figure 2. Meta Platforms is testing the endurance of subject access requesters. The example of Riitta Vänskä's attempt to obtain information on data in possession of Facebook, especially data that cannot be retrieved through the Facebook portal.



organisations to get the companies to respond. It is noteworthy that journalists had more leverage on this issue than politicians. Both Mark Scott and Niclas Storås were able to speed up the responses from companies (FullContact and Stockmann) by contacting their communications or media departments. The priority given to journalists has been observed in other similar efforts and may be due to the desire of companies to avoid bad publicity.

It seems that the test subject's decision-making role in society also influenced the responses to the subject access requests. The response received by Filomena Chirico from Google differed from the responses received by the others. In the experience of the researchers conducting the investigation, one difference was that Google directed the others to obtain the data through the data download portal, but in Chirico's case, it first invoked the three-month delay allowed by law and communicated the difficulty of obtaining data.

Three months later, Google sent Chirico a tailored file that provided more detailed information on her data than others have been able to obtain. The response was closer to the official subject access request than the information received through the portal. Did Google's data protection office recognise her role in the cabinet of EU Internal Market Commissioner Thierry Breton, who is also responsible for platform regulation? And did the company decide to offer better quality answers to Chirico than to the ordinary citizen?

Gigantti provides by far the best responses to subject access requests

Of the hundreds of subject access requests seen by the researchers – including outside the scope of this investigation – the Finnish home electronics retail chain Gigantti was by

far the best respondent. The responses provided to Sari Tanus and Miapetra Kum-pula-Natri consisted of a five-page PDF file with a detailed, transparent and understandable explanation of what data was included in the response, why the data was kept and how the data was used or shared. Another positive exception was that the response used the terminology and the framework used by the person in the email request, and was clearly written in a way that was tailored to the person making the request.

In the case of Gigantti, we found out about their data exchange relationships with its partners through their transparent response, whereas most of the companies in the investigation did not disclose any such information through the subject access request. Instead, for these companies we had to rely entirely on the data produced by the Tracker Control for our interpretations.

A subject access request is a way to measure the trustworthiness of a company

But is it worth submitting subject access requests in the first place and do the responses to the requests affect people's trust in companies? Sari Tanus's experience with a request for information under the General Data Protection Regulation is an example of a positive and empowering effect on individuals.

Tanus found out through a data request sent to Aamulehti that the newspaper was buying Bisnode data on her (the company selling the data is Dun & Bradstreet). She then submitted a new data request to Bisnode. She found out what data the company had, where the data had been received from and what Bisnode was doing with it. This gave Tanus a richer picture than usual of the data ecosystem in which her data was moving.

Most target companies were less transparent than the test subjects originally assumed. It was clear that the trust of the test subjects in these companies decreased when they were not transparent about the data collected on the user and how it was used.

Advice from Finland's Data Protection Ombudsman Ms. Anu Talus

Why should you ask for your own data?

By requesting your own data, you can find out what data on you is being processed and whether the data is up-to-date, accurate and necessary. This is the only way to get a realistic idea of all the data that has been collected on you.

This will enable you to also exercise your other rights under the GDPR. For example, you can ask for data to be erased or for inaccurate data to be corrected.

When should you contact the Data Protection Ombudsman?

If the company does not provide the data. Before contacting the Ombudsman, you should also contact the company's data protection officer.

For example, the Data Protection Ombudsman can order the company to provide the data. In some situations, an administrative fine may be imposed on the company.

What are the possible consequences for an organisation of non-compliance with the GDPR?

The GDPR contains a comprehensive toolkit. The sanctions for businesses may be a warning or an administrative fine, which can be very high.

We have issued decisions in which we have taken the view that a controller cannot make it more difficult for data subjects to

exercise their rights. For example, a company cannot impose its own formal requirements on the exercise of rights.

For larger actors, such as Facebook and Google, cases are settled in the context of the cooperation procedure with the data protection authorities of the EU.

According to the Digipower investigation, the process to obtain data from Facebook may take more than a year. Is this likely to improve?

During an administrative procedure, it must be ensured that everyone, including the data controller, has legal protection. In the case of Facebook, this is usually a cross-border matter. When the authorities of the 27 + 3 states cooperate in a procedure that involves translating documents and consulting with different parties, the procedure takes time.

Currently, there are several thousand cases pending, which also adds to the processing times.

The European Data Protection Board (EDPB) actively seeks to develop internal procedures to improve the efficiency of case handling.

Rules of the game for the collection and use of data are being developed through EU-level "regulatory tsunamis". Do you think this will improve the position and rights of individuals?

In some respects, the reforms can clarify the current situation and thereby improve the position of the individual. However, I consider the proposed reforms to be problematic to the extent that they contain overlapping provisions with the GDPR.

It is therefore essential that the widening of the regulatory framework does not obscure who is the competent authority. If the division of powers between the authorities becomes blurred, the position of the individual will also weaken. People will then not know where to turn and, in the worst-case, European practice will be fragmented.

5. Conclusions: Time to build alternative futures

The current imbalance of the data economy and digital power is maintained by the narrative of no alternative. However, history is full of examples of market-changing policies, actors and regulations that can help regain a balance.

The world's most valuable companies are global data giants whose main raw material is the data collected from individuals. A clear conclusion from the Digipower investigation is that individuals have far too little control or visibility over the use of their own data. In addition, data misuse reduces trust in the data economy.

The Digipower investigation reveals the large amount of data collected, but above all the ability of data giants to combine data from different sources to profile individuals. Even experts find it difficult to understand of the organisations in the “digiscape”, their roles and the long-term impacts of their policies on society.

With free products, individuals are not customers, but, through data, treated as commodities that could even be called the modern data proletariat. Individuals generate data, the most valuable raw material in the economy, almost for free and it is handed over to or taken by data giants. People are like the families who worked in coal mines 150 years ago, whose underage children, their noses black with coal dust, are also mining coal. The rights of children, let alone the six-day working week, are unknown and people do not know any better.

The first step in changing the current situation is to raise awareness, which is why the Digipower investigation was carried out.

An example of data giants' surveillance methods: the Facebook Pixel

The Facebook Pixel is one of the methods used by Meta. It is a piece of code that is installed on the website of a company advertising on Facebook, such as an electronics store. The user does not notice the code and has no way of knowing that it has been installed. The Facebook Pixel allows the company to better track the effectiveness of Facebook advertising and also enables retargeting. For example, if you search for electronic products online but don't yet buy anything, the next thing you see on your Facebook or Instagram feed might be advertising for similar electronic products.

An online retailer operating in an EU country passes on the purchase data from their bricks-and-mortar store to the data giant to be able to optimise the effectiveness of the advertising campaign it has bought from the data giant and to monitor the impact of online advertising all the way back to the brick-and-mortar store. Despite its size, the local retailer is much smaller than a global data giant. At the same time, it is part of the problem: by passing on the data of its users, it enables surveillance capitalism and unintentionally further increases the market power of the data giants.

In negotiations between companies, whoever has the more market power can dictate the terms and conditions. Because advertising on Facebook can be effective and is therefore popular, Facebook is able to collect more and more information on the users of Facebook, even when they visit other websites.

Downsides of concentrated power in the data economy

The negative effects of infrastructural power appear stronger when they are scaled up from the individual to the population, but this also makes them harder to understand. It is therefore important to unravel complex mechanisms with the help of simple everyday examples.

When ordinary people are asked whether they find privacy important, their answers may vary widely depending on their digital and media literacy. Sitra's findings show that young people, especially those who are otherwise knowledgeable have a short-sighted attitude to privacy protection, and so it is therefore important to raise the importance of privacy in the social debate. The Digipower investigation focuses on systemic harms unintentionally caused to groups of people, but individuals may also aim to deliberately cause harm to other people.

In the preprint (Kröger et al.), data collected from individuals has been classified into 11 harm categories, or ways of causing harming to individuals from data collected about them (be it data collected from digital services or the physical world).

1. Consuming data for personal gratification
2. Generating coercive incentives
3. Compliance monitoring
4. Discrediting
5. Assessment and discrimination
6. Identification of personal weak spots
7. Personalised persuasion
8. Locating and physically accessing the data
9. Contacting the data subject
10. Accessing protected domains or assets
11. Reacting strategically to actions or plans of the data subject

When using conventional digital services, there is no time to think about all the potential drawbacks of the current data economy system. However, because digital services are absolutely necessary for society to function and everyday life to run smoothly, we need to minimise the harmful effects on individuals and the risks related to their use. Changing the harmful patterns of the data economy will free up more space and expertise to deliver significant benefits to individuals rather than potential harm.

A significant change in the market was the decision taken by Apple in 2020 to enable the users of iPhone and iPad to opt out of tracking. While the organisations lobbying for digital advertising and major social media are signalling that data collection for third parties cannot be a significant problem because people themselves accept the cookies or terms of service, the majority of iPhone and iPad users are opting out of having their behaviour tracked by digital services. In international figures published in 2021, the opt-out rate was between 80 and 96 per cent of users.

The most successful business model in the current personal data economy, the unlimited collection of data outside the individual's control, cannot withstand close scrutiny. The models created over two decades and the old and new organisations that have developed through them have

failed to self-regulate in a way that supports individual rights. We therefore cannot count on them to make the change.

It is time to break the no-alternative narrative

Rationalising the current situation justifies a general passivity and creates a narrative that perpetuates the illusion that there is no alternative. We should examine and challenge the beliefs that this narrative is based on.

A typical part of the no-alternative narrative is the idea that mass data collection is a necessary price for free services that work well and that this cannot be changed because then services may become chargeable. Similarly, it has been argued in the past that you cannot ban child labour because products would become more expensive. Today, children's rights – as part of other fundamental and human rights – naturally delimit economic activity. At the same time, the overall well-being of society has increased and has been distributed more evenly.

Another part of the no-alternative narrative is the claim that “people are not interested in data being collected on them”. According to Sitra's research, lack of trust among individuals is the main stumbling block in the data economy, discouraging the use of digital services. It is a cognitive bias that wrongly generalises one's own opinion to apply to the entire population. On the other hand, it may also be a case of deliberate influence and a way of protecting the current winners in the data economy.

Linked to the narrative of no alternative is the myth of technological innovation by data giants as the main source of market power. But their market power is not primarily based on the most innovative technology, but on the amount of data they collect. Rather, monopolies and oligopolies have tended to turn against competition and, as

a result, their innovative power is weakened relative to their competitors. However, their market power allows them to focus more on preventing competition to safeguard their position.

In turn, monopoly profits from monopolies lead to welfare losses for the rest of society. Companies with alternatives that are technologically more advanced or meet the preferences of consumers better cannot compete if the digital giants buy them out or fair competition in the market is prevented by such things as supplier or customer traps to prevent data moving between competing services. Consumers therefore have no alternative but to submit to the conduct of the data giants.

Individual submission to the status quo can be seen as a psychological coping mechanism. Draper and Turow have argued (2019) that digital resignation is a rational response to constant monitoring. Consumers, in their view, are persuaded to surrender their data and this creates unequal power relations between companies and individuals.

Some companies use dubious strategies, such as misnaming, which describes “attempts to hide practices by using misleading expressions”. One of these is the claim of better services to justify the collection of individual data, when in reality the data is being used only for more accurate profiling of the individual to maximise advertising revenue, even at the expense of consumers' own preferences.

In order to move away from the false no-alternative narrative the current reality and its arrangements must be identified and described.

The power of the data giants must be curbed

Sitra sees a fair data economy as one of the three themes that will most change society in the future. The current data economy is not

fair enough because the interests of data giants are overemphasised at the expense of individuals and society.

Data is collected without people's understanding or control. Companies also buy data from other companies specialising in data collection, and this data may include information on such things as assumed wealth, family situation or political orientation.

Data giants such as Facebook are able to combine information about individuals from across the internet. The average person or business have little choice but to use the services of data giants. It is also difficult for business to operate without digital advertising or the main online trading platforms. The conditions are not negotiable for businesses or individuals, as market power is strongly on the side of data giants – and they also use the situation to strengthen their own power.

The practical examples highlighted in the investigation showed that the digital power in the hands of the world's largest data giants can be used effectively to profile and segment individuals into different groups of people with the aim of anticipating and influencing future behaviour. The more data an organisation can collect and combine from different sources, the more sophisticated the profiling and analysis of our movements, thoughts, purchasing behaviour and interests. As data accumulates among a few organisations, so does market power.

It is important to understand that business models based on data collection are not inherently a problem. For example, a two-way business model of making revenue by selling advertising is a viable approach for many services and products, such as free games, and is based on people's consent. The problem is the business model combined with the massive volume of collected data, lack of transparency and control by individuals over their own data use, the accumulation of that data by a small number of organisations, and the tendency of these organisations

to operate on the borderline of legality or in an ethically questionable way. This combination has created the current “surveillance capitalism” and its attendant side-effects.

The surveillance society and understanding the data economy

In the former East Germany, people lived in a surveillance society controlled by the secret police, Stasi. At the time, someone wrote down what texts a person read, where they went, what conversations they had, what they liked and disliked, and also recorded what they bought. No one would tolerate such surveillance now and yet there are similar features in today's societies.

According to **Shoshana Zuboff** (2019), the current data economy can be called surveillance capitalism. She defines surveillance capitalism as an economic system that treats the human experience and activity as raw material to be freely exploited for commercial purposes, based on undermining people's right to self-determination. Surveillance capitalism includes today's main “free services”, such as online search engines, marketplaces, operating systems and app stores, where the logic of profit is based on the exploitation of data collected from users.

Gillian Tett (2021) wrote that the current data economy could be viewed through the lens of cultural anthropology: today's data economy could be an imbalanced barter economy. **Omaar Hodan** (2021) on the other hand argues that the barter economy is not a suitable point of comparison, as the same data can be used simultaneously by a large number of companies if they so wish and is unlimited in relation to the traditional products of the barter economy.

Both authors are right in saying that the framework for assessing the data economy is important. It is particularly important

because decision-makers need to be able to understand what they are regulating.

From a surveillance society to MyData thinking

Surveillance capitalism includes the idea of user data as a freely available resource and the presumption of the right of that the commercial organisation to decide how to use data on behaviour.

The MyData approach, in which Finland has for a long time aimed to be a pioneer, offers an excellent starting point for rectifying the current situation. The starting point is to enable the widest possible use and sharing of data, with the individual having the right to decide on the re-use of data, which takes place with the user's permission. A strong advocate of MyData thinking is MyData Global, of which Sitra is a founding member.

6. Recommendations to address digital power imbalance

The lack of alternatives and the tacit acceptance of the current model must be translated into action at all levels of society. Systemic change can be achieved through cooperation between individuals, businesses, public actors and civil society and by investing in the development of concrete tools and policies.

Recommendations

- The basic right to one's own data must be strengthened.
- Market manipulation should be prohibited and the manipulation of people avoided.
- Companies must make ethical data use part of their corporate responsibility and subject to internal and external auditing.
- Venture capital investors should demonstrate accountability when looking for investment targets.
- It must be possible to prevent anti-competitive behaviour in advance.
- The largest data giants are immune to fines and thus sanctions need to be strengthened.
- Supplier and customer traps must be reduced.
- The basics of data economy should be included in the curriculum.
- Promotion of information literacy must be made a national project.
- Data collectives, MyData operators and DAOs should be used to strengthen the bargaining power of individuals.

Basic right to own data must be strengthened

The right to one's own data should be seen as a human right and it should be highlighted more in the public debate. The aim of the GDPR has been to provide people with the means to manage their personal data. The Commission's 2022 proposal for the Data Act would extend these rights from personal data to all data.

Under the Data Act, if an individual has participated in the creation of the data, they will have rights in relation to it, such as the right to access the data collected by the smart fridge or a newish car and share it with third parties such as repair shops. The proposals for the Data Act are examples of initiatives to give individuals more control over their own data.

Market manipulation is prohibited, but the manipulation of people must be avoided as well

Data combined with artificial intelligence contains particularly high potential to increase the productiveness of work and resolve problems in society. But it also creates new risks. The mass collection of data combined with artificial intelligence not only allows for the prediction of human behaviour, but also for the manipulation of behaviour.

The principle of self-determination and the prohibition of manipulation are included in the ethical guidelines of the high-level expert group set up by the European Commission to investigate the ethics of artificial intelligence (European Commission 8 April 2019). Our right to self-determination and respect for basic and human rights must be taken into account during the entire lifecycle of the design and use of a service or product.

Manipulation through data collection and profiling may involve subjugation, coercion, deception, conditioning or condescension, and should be prohibited in the EU. For example, it should be possible to conduct reviews and audits of the largest organisations to identify systemic risks.

Companies must make data use part of their corporate responsibility

The Cambridge Analytica scandal, in which Facebook breached not only the law but also its own rules, would not have been possible without a corporate culture that did not value privacy. The motto of Facebook's founder **Mark Zuckerberg** describes the corporate culture: move fast and break things.

At junctures in history, individual entrepreneurs and companies can play a role in determining the direction of an entire sector. Their motivation may be to produce results by any means, but also to turn the world into a better place.

In almost all cases, companies that want to market their products digitally, often as widely as possible, will need to use the services of the data giants for marketing and advertising. Instead of an unlimited supply, advertising channels are a predictable handful of organisations, and companies are increasingly dependent on them.

Every European company, even a small one, should now assess its own digital presence in different channels and its marketing, communication and customer relationship

processes from the viewpoint of protecting privacy and increasing the value of data. Processes and service providers should be audited and scrutinised to see where valuable data is flowing for free outside the company, to increase the value of the world's largest companies rather than the company's own intelligence capital. Alternative new practices should be required from both in-house marketing and customer-facing functions and from suppliers of digital services.

The Digipower investigation found that data was sent to Russian data giants Yandex and VKontakte via a retail chain website. Companies' knowledge of the data economy and their awareness of its mechanisms should be increased, as they may also be surrendering data without realising it.

Venture capital investors can demonstrate responsibility when looking for investment targets

Individual influential investors can have a major impact on how the internet economy develops. For example, Finnish Silicon Valley investor Jyri Engeström has spoken (Futucast 2021) about how in the early days of Web 2.0 individual influential business founders and owners rejected the open standards and decentralisation model so as to concentrate market power.

The EU could consider creating a classification system or taxonomy for sustainable financing in the data economy, in the same way that taxonomy for environmental classification is being created as part of the EU's environmental legislation. Different data economy certificates and reporting obligations could also be developed through collaboration in the sector, in the same way that responsibility reporting has been developed in the last few decades.

Preventing anti-competitive behaviour must be possible

Markets dominated by a few large actors typically need to be regulated. For example, Finland has a very competitive mobile access market, even though there are only a few large operators on the market. This would not have been possible without the right kind of regulation.

This has been studied in a Sitra working paper comparing the telecoms and financial markets with the current regulation of the data economy over recent decades (Sitra 1 October 2020). More than a decade ago, the European Commission required telecoms operators to enable their customers to keep their telephone number when changing their subscriptions. Business opposed the change. However, it made it easier for people to shop around for mobile subscriptions.

This type of regulation complements competition law by prohibiting market-damaging behaviour in advance. Examples of such regulation include the Commission's proposals for the Digital Markets Act (DMA), the Digital Services Act (DSA) and the Data Act. New regulation is required and it needs repeated iterations. At the same time, over regulation must be avoided, as it may in itself constitute a barrier to market access.

Competition law must take into account the wider well-being of consumers

Competition law has traditionally safeguarded the interests of consumers by examining pricing. A free service may not be an optimal solution for the overall well-being of the consumer. Competition law and its application should be developed on the basis of a wider concept of consumer welfare, which takes into account not only the price but also qualitative elements such as the protection of privacy and the actual choice of consumers.

The largest data giants are immune from fines – sanctions need to be strengthened

Competition law needs a global revamp, as the rules built before the internet age no longer work fast or efficiently enough.

Competition law was created at the end of the 19th century in the US in response to the excesses of companies that had gained a dominant market position. The best-known case was the petroleum company Standard Oil. The company sought to prevent competition by buying its competitors or driving them out of the market with predatory pricing. Standard Oil was split up into 34 separate organisations in 1911.

In the US, the Federal Trade Commission (2021) has been investigating Meta (Facebook) for abuse of market power. The company bought its most threatening competitors Instagram and WhatsApp and has allegedly bought smaller competitors to bury their innovations when its own innovation was not enough to keep up with the competition.

Acquisition controls in the EU also need to be improved and the use of takeovers to eliminate the worst competitors should be more sensitively prevented.

The European Commission fined Google EUR 2.42 billion in 2017 for favouring its own products in a search engine (European Commission 27 June 2017), EUR 4.34 billion in 2018 for illegal practices regarding its dominant position in relation to the Android operating system (European Commission 18 July 2018) and EUR 1.49 billion in 2019 for abuses in online advertising. According to the Commission, some such abuses had continued for more than ten years and prevented innovation in other companies and competition (European Commission 20 March 2019). Even after this, similar charges have been brought against Google for abusing its dominant market position (Ilta-Sanomat 14 May 2021).

In the Cambridge Analytica scandal, Facebook illegally handed over data collected on its users to third parties. The data was used among other things to target fake information in elections to confuse voters. Facebook was fined about five billion dollars (YLE 13 July 2019). According to the information disclosed by whistle-blower **Frances Haugen** in 2021, Facebook still continues to allow the spread of false information for commercial reasons.

Current competition legislation is too slow to act and has insufficient deterrent effect. It can take up to ten years to go through the various courts, during which time the competitors will already have lost out and individuals and society as a whole will not have benefitted from competition and new innovations.

The DMA and the Commission's 2021 updated interpretative guidelines on merger control will improve the situation. Their effectiveness needs to be assessed and existing and new tools need to be created as appropriate. It is particularly important to prevent so-called 'killer acquisitions', whereby large companies buy promising competitors out of the market at an early stage. Fines must be raised to a level that has a genuinely deterrent effect on unlawful activities.

Supplier and customer traps must be reduced

The majority of the data in the world is unstructured and not in digital format. Data does not move enough and is siloed, which are deliberately built supplier traps.

Data must be made to move safely and transparently across sectoral and company boundaries within Europe, as envisaged in the EU Data Strategy. This will be enabled not only by legislation but also by so-called soft digital infrastructure, which refers to standards, ground rules, interfaces and tools, such as digital identity and real-time economy solutions.

National programmes are needed to speed up the creation of data economy standards, interfaces, rules and tools. There are already many examples of these in Finland and in the EU, and various experiments have been conducted widely (Sitra IHAN 2022; Sitra Gaia-X Finland 2022). In Finland, the architecture work of the Virtual Finland project in the public sector uses the rulebook developed by Sitra and the fair data economy architecture model, which requires reusability of data. (Ministry for Foreign Affairs 2022)

Business skills of smaller companies in the data economy must be improved

Improving SME competence particularly about the business logic of the data economy would not only help SMEs take advantage of the opportunities provided by the data economy but would also improve their negotiating position with larger companies.

Sitra has collaborated with the Finnish Ministry of Economic Affairs and Employment to develop the Growth from Data business programme, which is aimed at sharing data and open to all (Sitra 18 June 2021). National, similar business programmes to support SMEs should be implemented in other EU countries.

Basics of data economy to be made part of national syllabuses

A hundred years ago, agrarian society started to use electric power. Although few of us know everything about electric power, people still know how to turn on a light switch. The light would not come on without new skills. People have learnt to use

electric appliances and businesses have learnt to use electric power. The data economy is at the beginning of this century-long effort.

The basics of the data economy should be made a new European civic skill. Germany, for example, published a national data strategy in 2021, with the aim of creating a “data culture” and “improving data literacy” in Germany. (Ahvonen et al. 2022)

Sitra is creating the first sets of teaching material related to the basics of the data economy together with the Finnish National Agency for Education. The Digiprofile test is already available, allowing individuals to test their knowledge, attitudes and activity online. More is needed in a broad collaboration across all school subjects and degree programmes. (Helminen 2021)

Promoting information literacy should be made a national project

The information literacy and more widely the ‘digital education’ must meet the requirements of the changing information environment. Traditional media literacy is no longer enough. This requires the ability to identify disinformation and a deeper understanding of the digital use of power and one’s own role in it.

In practice, it should be ensured that information literacy is included in teaching at different levels of education and in the democratic work of municipal education services (schools, libraries, folk high schools). In addition, the resilience of democracy should be strengthened by supporting independent media and NGOs.

Data collectives, MyData operators and DAOs to strengthen the bargaining power of the individual

When we log on to a new online service, we have to accept terms and conditions of use that can often be described as “take it or leave it” terms and conditions, where the individual has no bargaining power. In the future, bargaining power could be balanced by data collectives, for which various forms of organisation exist and should be explored as possible solutions. The collectives could be groups of people who convene to manage data together or they may be based on voluntary disclosure of one’s own data for research use. MyData operators, for their part, are envisaged in the international MyData movement.

The next stage of the internet, commonly known as Web 3.0, is based on a decentralised model and the principle of individuals controlling their own data. One way to realise collective data control is through the decentralised autonomous organisation, DAO. Because DAOs make it possible to create organisations without centralised power or management, they could even be called the Cooperative 2.0 of the digital age.

Decentralisation is expected to reduce the power of centralised platforms and gatekeeper companies. In the best-case, digital power will be distributed more evenly to support the democracy of tomorrow. The result may also be more innovation and more effective competition, but the future direction and ground rules depend on companies and entrepreneurs, legislators, researchers, public officials, NGOs and the public. All of us.

It is also essential to find new tools for people, alone and with others, to use data about themselves to learn, participate and improve their daily lives and society.

Sources

The Digipower investigation reports

Bowyer, A. et al. 2022. [Digipower Technical Reports: Auditing the Data Economy through Personal Data Access](#) (The case studies report). (accessed on 23 May 2022)

Pidoux, J. et al. 2022. [Digipower Technical Reports: Understanding Influence and Power in the Data Economy](#) (The narrative report). (accessed on 23 May 2022)

Other sources

Ahvonen, K., Bremer, O., Djakonoff, V., Koponen, J., Mikkonen, J., Toivanen, M. (toim.). 2022. [Finland's strengths, challenges and opportunities in building a data economy](#) (in Finnish). Sitra memorandum.

Draper, N.A., Turow, J. [The corporate cultivation of digital resignation](#). Sage journals 8.3.2019. (accessed on 21 April 2022) or Memo 27 June 2017. (accessed on 21 April 2022)

European Commission. [Antitrust: Commission fines Google €4.34 billion for illegal practices regarding Android mobile devices to strengthen dominance of Google's search engine](#). (PDF). Press release 18 July 2018. (accessed on 21 April 2022)

European Commission. [Antitrust: Commission fines Google €1.49 billion for abusive practices in online advertising](#). (PDF). Press release 20 March 2019. (accessed on 21 April 2022)

European Commission. (2021/C 113/01). [Antitrust: Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service](#) – Factsheet. Memo 27 June 2017. (accessed on 21 April 2022)

European Commission. [Antitrust: Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service](#) – Factsheet. Memo 27.6.2017. (haettu 21.4.2022)

European Commission. [Mergers](#). (accessed on 21 April 2022)

European Commission. High-level expert group on artificial intelligence. [Ethics guidelines for trustworthy AI](#). 8 April 2019 (accessed on 21 April 2022)

Federal Trade Commission. [FTC Alleges Facebook Resorted to Illegal Buy-or-Bury Scheme to Crush Competition After String of Failed Attempts to Innovate](#). 19 August 2021. (accessed on 21 April 2022)

Futucast. Jyri Engeström. [NFT, decentralisation and the future of the internet](#) (in Finnish). Video 15 November 2021. (accessed on 21 April 2022)

Helminen, J. 2021. The growing importance of the data economy challenges the education sector to update competence. (in Finnish). Sitra's guest blog. (accessed on 21 April 2022)

Helsingin Sanomat. Facebook whistle-blower Frances Haugen lost a friend to misinformation but denies hating the social media giant: "I love Facebook. I want to save it" (in Finnish). Article on 5 October 2021. (accessed on 21 April 2022)

Ilta-Sanomat 14.5.2021; Digitoday. Huge fines imposed on Google for abuse of market position. (in Finnish). Article on 14 May 2021. (accessed on 21 April 2022)

Jalavisto, M. 2021. Individuals' lack of trust is a stumbling block for the data economy. Sitra's article. (accessed on 21 April 2022)

Kippo, J., Kemppainen, E-L. 2022. The potential offered by data is overshadowed by the behaviour of data giants on online platforms – this is how Finland discussed the data economy in 2021. Sitra's article from a report by Meedius International on online media discussions. (accessed on 4 May 2022)

Kröger, J., Miceli, M., Müller, F. 2021. How Data Can Be Used Against People: A Classification of Personal Data Misuses. (accessed on 20 April 2022)

Lehtonen, K., Halenius, L. The EU defines ground rules for the data economy – this is how new regulations will affect you and all of us. (in Finnish). Sitra's blog. (accessed on 4 May 2022)

Lehtonen, K., Halenius, L., Rastas, T. The Data Act promises the same keys to success to everyone participating in the data economy. Sitra's article. (accessed on 4 May 2022)

Meta Business Help Centre. About Meta pixel (in Finnish). (accessed on 21 April 2022)

Ministry for Foreign Affairs. Developing a Virtual Finland platform. (accessed on 20 April 2022)

Ministry of Transport and Communications. Finland to lead the way in MyData. Press release 4 July 2018 (accessed on 20 April 2022)

MyData Global. MyData operators. (accessed on 20 February 2022)

Official Journal of the European Union. (2021/C 113/01). Guidance on the application of the referral mechanism set out in Article 22 of the Merger Regulation to certain categories of cases. Communication from the Commission. (accessed on 3 May 2022)

Omaar, H. 2021. No, The Data Economy Is Not A Barter Economy. Center for Data Innovation. (accessed on 21 April 2022)

Paasikivi, O., Tuohino, J., Mansnérus, J., Lång, J. 2022. Possibilities to use artificial intelligence in the public sector. (in Finnish). Sitra Studies 206.

Raunio, E. 2020. The future of the European data economy at stake – Lessons from telecom and financial markets. Sitra Working Paper.

Sitra. 2022. Principles-based frameworks and tools for fair data economy (PDF).

Sitra's Gaia-X Finland project. (accessed on 20 April 2022)

Sitra's Growth from data programme. (accessed on 20 April 2022)

Sitra's IHAN® project. (accessed on 21 April 2022)

Tett, G. 2021. The Data Economy Is a Barter Economy. Harvard Business Review. (accessed on 21 April 2022)

Vänskä, R., Härkönen, T. 2020. On the trail of personal data – The flow and use of data collected from individuals using digital services (Digitrail survey). Sitra Studies 169.

World Economic Forum. Advancing Digital Agency: The power of data Intermediaries. Insight report, 15.2.2022.

Yle News. Facebook faces huge fines as a result of the Cambridge Analytica scandal – will have to pay five billion dollars for leaking user data (in Finnish). 13 July 2019. (accessed on 21 April 2022)

Zuboff, S. 2019. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. London: Profile Books.

Appendix 1: A model of a subject access request

This message is a transparency request under the General Data Protection Regulation, including a subject access request, a portability request, and other specific provisions. Please note that it is not legal to require data subjects to use an in-house form[1].

I would like to request a copy of all my personal data held and/or undergoing processing. This is both a subject access request and a portability request. This message is not in any way to be considered as a complaint.

I am aware of your automated data download service. This request goes beyond the data that your automated download service makes available, therefore please process this request manually rather than referring me to the download service.

COPIES OF MY PERSONAL DATA

This request covers all my personal data, from usage of both your app and your website, including specifically all data belonging to each of the following categories:

1. Volunteered data - data that I have explicitly shared with you
2. Observed data - data that you have collected about me and my activity through my use of the service or through interactions with your staff
3. Derived data - data you have created and stored about me as a result of analysis, processing or inspection of my data or service activity.
4. Acquired data - data you have acquired from any external sources including e.g. credit checks, other users or advertiser's lists.
5. Sharing & Handling data - all metadata and handling information (as detailed below) about who you have shared my data with, as well as copies of the specific data that has been shared, and details of how you have stored and handled and processed my specific data.
6. Consents - where you rely on my consent to process my personal data, please provide details of those consents including when and by what means I gave consent, and the scope of that consent, and how I might change that consent if I wanted to.

In particular, please make sure that all personal data you state in your privacy policy you collect, create, store or share is provided. Please also make sure to include data for which you remain controller that is held by third parties.

For any of these categories where you do not hold any data about me, please explicitly confirm that you do not hold any data of that type about me

ADDITIONAL SPECIFIC DATAPPOINTS

In addition, I have recently learned about the following specific datapoints you hold. Please provide any data you have for me for each of the following. Where you have no data about me for a given one of these fields, please confirm that fact to me explicitly.

- Xcheck data
- Civic amplification score
- Close-friendness data
- meaningful people data;
- world2vec vector
- "Feed Unified Scoring System" data
- Five user interest segments

ARTICLE 20

For data falling within the right to data portability (GDPR article 20), which includes all data I have provided and which have been indirectly observed about me [2] and where lawful bases for processing include consent or contract, I wish to have that data:

- sent to me in commonly used, structured, machine-readable format, such as a CSV file. A PDF is not a machine-readable format [3].
- accompanied with an intelligible description of all variables or abbreviations.

ARTICLE 15

For all personal data not falling within portability, I would like to request, under the right to access (GDPR, article 15):

- a copy sent to me in electronic format. This includes any data derived about me, such as opinions, inferences, settings and preferences. Note that opinions, inferences and the like are considered personal data [4]. For data that is available to the controller in machine readable format, it must be provided to me in that form in accordance with the principle of fairness and provision of data protection by design.

If your organisation considers me a controller for whom you process

Furthermore, if your business considers me the controller of any personal data for which your business acts as processor, please provide me with all the data you process on my behalf in machine readable format in accordance with your obligation to respect my determination of the means and purposes of processing.

METADATA ON PROCESSING

This request also includes the metadata I am entitled to under the GDPR, including details as follows:

INFORMATION ON CONTROLLERS, PROCESSORS, SOURCE AND TRANSFERS

- The identity of all joint controllers of my personal data, as well as the essence of you contracts with them (GDPR Article 26).
- Any third parties to whom data has been disclosed, named with contact details in accordance with Article 15(1)(c). Please note that the European data protection regulators have stated that by default, controllers should name precise recipients and not “categories” of recipients. If they do choose to name categories, they must justify why this is fair, and be specific, naming “the type of recipient (i.e. by reference to the activities it carries out), the industry, sector and sub-sector and the location of the recipients [3]. Please note that in the case of any transferred data processed on the basis of consent, there is no option to just name categories of recipients without invalidating that legal basis [5].
- If any data was not collected, observed or inferred from me directly, please provide precise information about the source of that data, including the name and contact email of the data controller(s) in question (“from which source the personal data originate”, Article 14(2)(f)/15(1)(g)).
- Please confirm where my personal data is physically stored (including backups) and at the very least whether it has exited the EU at any stage (if so, please also detail the legal grounds and safeguards for such data transfers).

INFORMATION ON PURPOSES AND LEGAL BASIS

- All processing purposes and the lawful basis for those purposes by category of personal data. This list must be broken down by purpose, lawful basis aligned to purposes, and categories of data concerned aligned to purposes and lawful bases.

Separate lists where these three factors do not correspond are not acceptable (Article 29 Working Party [6]). A table may be the best way to display this information.

- The specified legitimate interest where legitimate interest is relied upon (Article 14(2)(b)).

INFORMATION ON AUTOMATED DECISION-MAKING

- Please confirm whether or not you make any automated decisions (within the meaning of Article 22, GDPR). If the answer is yes, please provide meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for me. (Article 15(1)(h))

INFORMATION ON STORAGE

- Please confirm for how long each category of personal data is stored, or the criteria used to make this decision, in accordance with the storage limitation principle and Article 15(1)(d).

IDENTIFYING INFORMATION

I understand that according to Article 11 GDPR, and particularly Art 11.2, you might need additional information to identify me for the purpose of this request. I have provided this data in the form of my Facebook profile URL within your contact form. If you need any further information please contact me.

If you do not normally deal with these requests, please pass this email to your Data Protection Officer. If you need advice on dealing with this request, any European Data Protection Authority should be able to provide you with assistance.

In accordance with the law, I look forward to hearing from you within one month of receipt.

References:

[1] UK Information Commissioner's Office, Subject Access Code of Practice (9 June 2017) p13; Information Commissioner's Office, 'Guide to the GDPR: Right to access' (22 May 2019): 'Even if you have a form, you should note that a subject access request is valid if it is submitted by any means, so you will still need to comply with any requests you receive in a letter, a standard email or verbally [...] although you may invite individuals to use a form, you must make it clear that it is not compulsory'.

[2] Article 29 Working Party, Guidelines on the Right to Data Portability (WP 242), 13 December 2016, 8.

[3] Article 29 Working Party, Guidelines on Transparency under Regulation 2016/679 WP260 rev.01, 11 April 2018.

[4] See Case C 434/16 Peter Nowak v Data Protection Commissioner [2017] ECLI:EU:C:2017:994, 34.

[5] Article 29 Working Party, Guidelines on Consent under Regulation 2016/679 WP259 rev.01, 10 April 2018, 13.

[6] Article 29 Working Party, Guidelines on Transparency under Regulation 2016/679 WP260 rev.01, 11 April 2018, page 35.

SITRA

SITRA STUDIES 215

The Sitra Studies publication series features the results of Sitra's future-oriented work and experiments.

ISBN 978-952-347-279-2 (PDF)
www.sitra.fi

SITRA.FI

Itämerenkatu 11–13
PO Box 160
FI-00181 Helsinki,
Finland
Tel. +358 294 618 991
 @SitraFund