Turun yliopisto
University of Turku

ETHICAL FRAMEWORK FOR A FAIR, HUMAN-CENTRIC DATA ECONOMY

WP 1: Citizens' values
report for IHAN

Authors:

Minna M. Rantanen
Jani Koskinen

28.10.2019

Turun kauppakorkeakoulu • Turku School of Economics
Future Ethics Research Group

Abstract

European Union (EU) has been the forerunner and visionary of data protection and privacy. However, the data economy is global business where the values of EU drives are not yet fulfilled or understood. Sitra's IHAN project is aiming to make a difference by creating the foundation for a fair data economy by treating the individuals as key stakeholders.

The goal of this report is to seek the values of citizens considering the fair data economy. The research is based on the previous value research by Schwartz and Hofstede. This is complemented with a survey on citizens (n=8004) – focusing on fair data use – that was conducted for Sitra by Kantar TNS Ltd in the four member states of EU: Finland (n=2000), France (n=2000), Germany (n=2004) and the Netherlands (n=2000). The answers of two open questions, that were part of the survey, were used as empirical material for our analysis. Thematic analysis was used to find values that should be considered when developing fair human-centric data economy.

The main finding – based on previous value research of Hofstede – is that the values differ in different European countries and it seems that we do not share the similar "European values". Instead, people in EU have diverse values based on their cultural background. However, this is not the simply answer. Even though we do not have similar values basis, we do share same common values and people are able to create new ones, which should be the goal of when creating a people-centric data economy. If we cannot find commonly acceptable values, the data economy continues to be field filled with doubts, exploitations and scandals, where mere economical needs are at the centre of ecosystem.

In thematic analyse we found eighth different themes: User's control over data and data sharing, Transparency and being informed, Security, Trust and fairness, Compensation or benefits for users, Supervision and rules, Attitudes towards data collection and data economy and Unable to answer/does not know. The founded values were also located to Schwartz's model and main finding was that Transparency seems to be connecting value for other values. It is understandable as transparency is only way to ensure the fulfilment of other values. Without transparency one may not see how convincing security or trust could be achieved.

These themes show the issues that should be taken account when developing fair data economy that meets the justified values of people. Therefore, it seems that key factors for achieving fair human centric data economy –that is in line with values of humans– are transparency and open communication with people all the time. We need open and rational discourse were people are truly seen as equal stakeholders and their interests are no bypassed by other stakeholders.

Keywords: Values, Human-centric, Data economy

Tiivistelmä

Euroopan Unioni on toiminut visionäärisenä edelläkävijänä tietosuojaan ja yksityisyyteen liittyvissä asioissa. Niiden perustana olevat arvot eivät kuitenkaan ole vielä täysin hyväksyttyjä datatalouden globaalissa liiketoimintaympäristössä, saati ymmärrettyjä. Nyt käynnissä olevalla Sitran IHAN-hankkeella pyritään luomaan perusta reilulle datataloudelle missä yksilöt ovat keskeinen toimija ja sidosryhmän jäsen.

Tässä Työpaketti 1 raportissa käsitellään kansalaisten arvoja. Arvojen tarkastelu perustuu pääosin IHAN-hankkeessa Kantar TNS:n Sitralle toteuttamaan kansalaiskyselyyn (n=8004) neljässä EU-maassa: Suomessa (n=2000), Ranskassa (n=2000), Saksassa (n=2004) ja Hollannissa (n=2000). Pyrkimyksenä on selvittää, millaisiin arvoihin datatalouden tulisi kansalaisten mielestä perustua ollakseen reilua. Tämän lisäksi tutkimuksessa käytetään kahta aiemmin tehtyä tutkimusta: Schwartzin arvoteoriaa ja Hofsteden toteuttamaa tutkimusta arvoista.

Yksi päälöydöksistä on se, että ei ole olemassa Euroopan yhteistä arvopohjaa, vaan arvot eri puolella Eurooppaa ovat vaihtelevia perustuen kulttuurien eroihin eri maissa. Tämä ei kuitenkaan ole ainoa totuus, vaan on mahdollista löytää yhteisiä arvoja, joiden avulla voidaan luoda reilun datatalouden yhteinen arvopohja. Tämä arvopohja on hahmotettavissa kansalaiskyselyn analyysin perusteella.

Empiirisenä aineistona analyysille käytettiin Kantar TNS:n kyselyn kahteen avoimeen kysymykseen saatuja vastauksia, jotka analysoitiin temaattisesti. Tutkimuksen keskeisin tulos on, että tutkittujen maiden arvot eivät ole keskenään yhtenäisiä. Aineistosta erottui kahdeksan teemaa, nämä olivat: (1) Käyttäjän kontrolli dataan ja sen jakamiseen, (2) Läpinäkyvyys ja tiedottaminen, (3) Turvallisuus, (4) Luottamus ja reiluus, (5) Käyttäjien saamat korvaukset tai hyödyt, (6) Valvonta ja säännöt, (7) Tietojen keräämiseen ja tietotalouteen liittyvät asenteet ja (8) Ei osaa sanoa / ei tiedä.

Tulokseksi saadut teemat kuvaavat kansalaisten arvostamia asioita, joihin datatalouden tulisi perustua ollakseen reilua. Kun löydetyt arvot järjestettiin käyttäen Schwartzin arvomallia, läpinäkyvyys nousi keskeiseksi muita arvoja yhdistäväksi tekijäksi – ilman läpinäkyvyyttä ei ole mahdollista tarkastella miten arvot toteutuvat datataloudessa.
Tutkimusten löydösten perusteella voidaan luoda datatalouden arvopohja. Arvopohjan luomisessa reilulle ihmiskeskeiselle datataloudelle keskeisenä edellytyksenä on läpinäkyvyys ja jatkuva avoin kommunikaatio ihmisten kanssa. Aiheesta tarvitaan kuitenkin lisätutkimusta syvällisemmän ja moniulotteisemman kuvan muodostamiseksi.
Avainsanat: Arvot, Ihmislähtöisyys, Datatalous,

Table of Contents

List of figures

List of tables

1 INTRODUCTION

Data has become the basis of modern knowledge economy. With the development of new data gathering methods, analysis tools and processes, new business and industrial domains are expected to be born while old ones are expected to be renewed. Thus, data have become more valuable and will continue to do so in the future. However, data are not valuable if they are not used. A mere hoarding of data without analysis to form information does not create value. In the best-case scenario, data are used to form information that is then transformed to knowledge. Hopefully, this knowledge will then lead to wisdom about the topic, generating long-lasting value in the data economy.

The use of personal data has risen during the digital age with the help of social media. Personal data have mainly been used to target advertising, although the use of personal data could hold potential for greater benefits, such as personalised medicine and services in general. However, current big tech companies and their failures in data protection have brought privacy issues into public awareness, with the result that individuals became more concerned and reserved about sharing their personal data. It is clear that people are unhappy with the current practices in the global data economy, and we have to find a way to re-create trust to realise the potential benefits of the data use.

Europe has long been falling behind the global development of the data economy, but the current situation offers us an opportunity to develop a trustworthy data economy. The EU has shown that it has the power to govern the European data economy with regulations based on values with the General Data Protection Regulation (GDPR), which emphasises individuals' right to privacy and mastery over their data. Although the implementation of the GDPR has not been as straightforward as was expected, it is a demonstration of the EU's role as a value leader and referee of the European data economy.

This report has been produced for Sitra's IHAN® project. IHAN aims to build the foundation for a fair and functioning data economy. Its main objectives are to create a method for data exchange and set up European-level rules and guidelines for the ethical use of data. See more:
www.sitra.fi/ihan

The creation of a trustworthy data economy is not a simple issue, and many attempts to develop it have and are being made. The GDPR has inspired initiatives to create a new paradigms for the digital economy, such as in the case of the Mesinfos project in France, while the DECODE project aims to provide tools for individuals to own their data. Although the goals of these existing projects are admirable, and their results needed, there is still a clear lack of a holistic view on the data

economy. Solving challenges one by one does not necessarily lead to a better data economy, and thus, we need to take a more holistic view on the data economy and acknowledge its ethical aspects.

We view the data economy as a complex ecosystem that includes both social and technological aspects. However, to create knowledge from personal data with technical tools, individuals are the key to a viable data ecosystem, both due to the data they provide and their ability to create technological applications to serve the data economy and its different stakeholders. Thus, to create a viable data economy, we should consider the individuals and their needs, as well as the nature of technology.

In information systems science, socio-technical development has a long history. Socio-technical theory emerged in the 1950s. It challenged the idea that systems could be optimised just by the standardisation of the task and implementing technology. Socio-technical design is based on the idea that, to create an efficient system, both technology and the humans using and affected by it need to be considered. Since the possibilities to modify human beings are limited, many socio-technical design methods set human needs as a starting point of technological design and endorse user participation in the design process.

User participation can be seen as a democratic and human-centric approach in systems design. However, it has been developed for organisational information systems that are much more limited in size and complexity. In the case of data economy ecosystems, applying a socio-technical design is rather impossible due to the vast amount of people involved. Nevertheless, the basic idea of socio-technical theory – the need to consider both the social and technical aspects of a system – is a good starting point for the development of human-centric data economy.

The goal of the IHAN project is developing a fair data economy. However, this raises a question: What kind of data economy could be considered a fair one? Fairness is a normative claim that includes the idea of treating people[1] equally, without favouritism or discrimination. However, fairness is also a subjective issue. What constitutes fairness depends on how one perceives an action that is being evaluated. Thus, our evaluations stem from our concepts of fairness and prior knowledge.

As we do not live in isolation, norms also affect the evaluations of our actions. Norms are guidelines, rules or prohibitions that guide our actions in a particular group or social unit. They are also often formalised and made into laws that state what is appropriate and what is not from an authoritarian perspective. Norms stem from the values of the society and the desire to safeguard them. Values refer to what is held to be important by an individual or group. These vary between individuals, but there is often some common ground on what is held to be important in a particular context. Thus, values offer a way to pivot the participatory design from small-scale systems to larger ecosystems as they can be

[1] Also other stakeholders should be considered. However, in context of fairness there always are people behind other stakeholders. If not, there would not be point for fairness either.

used as the substitutes of needs. In our research, we focus on the values of individuals that should be made into the norms of the fair data economy. A focus on individuals does not mean that the other stakeholders are forgotten, since a functioning fair data economy needs to be fair to all or it perishes. Individuals are the key to the fair data economy, since their participation in the data economy is pivotal – without individuals, there are no personal data to be shared or utilised.

It is notable that values affect everything we do, and thus, seemingly value-free technology also has inbuilt values. These can be derived from the values of developers or the people initiating the development. For example, in an organisational setting, efficiency is and has been one of the leading values in information systems development. Efficiency is not a bad value, but it may not be something that people using the system value or need. People using the system could, for example, value ease of use. If ease of use is to be the leading value in the development of a system, efficiency could be a by-product of this.

Values that are implemented in the system should be carefully considered. An important reason for careful consideration stems from the subjective nature of values – what one considers as a priority, others may see as a trivial issue. Thus, we need to have a way of distinguishing core values from secondary ones. In other words, we need a way to justify the values that should be the basis of the human-centric fair data economy. In this ethics is a valuable approach due to its proximity to values and long tradition in consideration of what should be done and what should be considered good.

Our research is theoretically grounded in socio-technical design, values and ethics. The aim of this research project is developing an ethical guideline for the human-centric fair data economy (IHAN). This research is divided into three parts. First, we analyse the values of individuals in the context of the fair data economy. In other words, we aim to find an answer to the following question: *What is valued in the context of the data economy*? This research forms the value basis of the fair data economy from the perspective of individuals.

As values are personal, they must be further analysed to find the most justified basis for the fair data economy. Thus, the second research question is as follows: *What values should be implemented into the norms of the fair data economy?* To answer this question, we use ethics as a tool to find the intrinsic values from the values of individuals. In other words, we aim to find the values that are valuable in themselves, not as a means of achieving some other value. This research results in an ethical framework of the fair data economy, which will be presented in the second report of the project.

In the third phase of the research, the ethical framework is used as a basis for the ethical guidelines of the fair data economy. This is crucial, since without practical guidelines, the ethical framework is not likely to be taken into use by relevant stakeholders. Thus, the third and final question of this research is: *How can we implement the ethical values in*

practice? The results of this phase stem from the earlier phases and will give concrete guidelines for incorporating values in a fair data economy.

This research introduces a human-centric, value-sensitive methodology for developing data economies and a concrete way to implement its results – ethical guidelines – in practice. This report includes the findings of the first phase and serves as an introduction to the background of the whole research.

This report is constructed as follows. The next section introduces three different views on values that should be considered in relation to the data economy. Because cultural differences can affect values, in section 3, cultural differences are briefly considered. These considerations are limited to the EU due to the overall focus of the IHAN project. In section 4, a qualitative analysis of the open questions on the citizen survey is presented with its related results. In section 5, we discuss the results. Finally, we conclude in section 6.

2 VALUES IN THE HUMAN-CENTRIC DATA ECONOMY

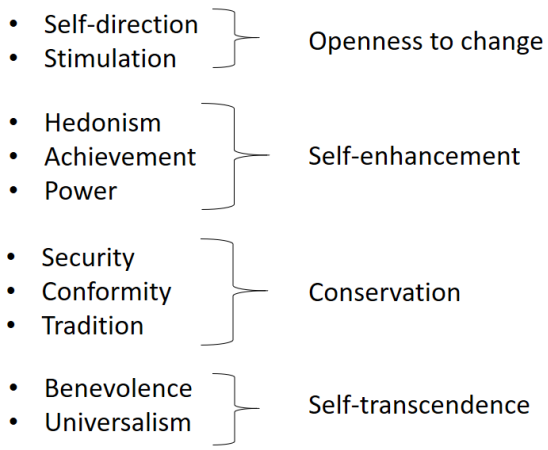
Values can be defined as elements that are considered important by an individual or group; they are often discussed in general or in a particular context. Thus, when values are discussed, there is a need to clarify between different types of values and how they should be considered in developing and governing data economy ecosystems. In this research, we focus on three types of values – personal, organisational and ethical values.

2.1 Personal values

Figure 1. Schwartz's basic values.

Personal values refer to the values of an individual person. Each person holds a multitude of values that create the personal value system. Values help to guide our behaviour and reflect what we think is valuable. Values and value systems vary between persons. For example, one person may value security, whereas another may value excitement. However, some values – often referred to as *basic values* – can be recognised across cultures. These are often broad in the sense that they are not linked to any specific situations. Thus, many cultures have some constancy in value sets, although on an individual level, values can vary.

Schwartz (2012) defines ten basic values that are likely to be universal since they are based on the universal requirements of human existence; these are as follows: the needs of individuals as biological organisms, means of coordination in social interaction and welfare and survival of groups. These values and the four categories they can be divided into are presented in Figure 1.



Values have unique characteristics that differentiate them from, for example, social norms. Schwartz (1992) presents six main features of values. First, values are *beliefs that are linked inextricably to affect*. This means that values manifest through emotions. For example, people that value independence feel irritated and aroused when their independence is threatened and secure when they can protect their independence. The notion that values are beliefs means that they are subjective mental representations that are learned rather than inherited. This means that values are affected by social norms and expectations – in other words, they are derived from social and cultural mores. However, although values are learned, they are still relatively permanent attributes of a person and are not easily changed.

Second, values refer to *desirable goals that motivate action*. Thus, values influence the ways we act and the goals we assert. Because of this characteristic, values also affect cooperative action and the motivation behind it. However, values and the motivations that drive people to work towards common goals can vary. To one person, cooperation can be valuable, whereas other may see it as a way to achieve something else that they hold valuable, such as self-enhancement.

Third, values *transcend specific actions and situations*. This means that values are not tied only to specific actions or situations, but instead, are held as important in a multitude of actions and situations. For example, honesty is often valued in the workplace, school, politics and with both friends and strangers. This separates values from norms and attitudes, which usually refer to certain actions, situations or objects. Values, norms and attitudes are still somewhat intertwined, for example, in social situations. Norms refer to expected behaviour, whereas attitudes are positive or negative emotions, beliefs and intentions that affect behaviour. Attitudes are affected by values, but they are less permanent and tied to the situation or object.

Fourth, values *serve as standards or criteria*. This means that the values that one holds guide selection or evaluation of actions, policies, people and events. Thus, values serve as a basis for decisions about, for example, the goodness, justification and worthiness of an action, its outcome or other people's actions. Although these evaluations are constantly made, people rarely make them consciously. Awareness of value considerations often arises when held values have conflicting implications.

Fifth, values are *ordered by importance relative to one another.* This means that personal values form an ordered system of priorities that characterises a person as an individual. This hierarchical feature affects decision making and makes it possible to compare values with conflicting implications. Although the set of values that a person holds is relatively small, the hierarchy between them can be hard to distinguish because people are rarely conscious of it.

Finally*, the relative importance of a multitude of values guides action*. As suggested above, actions or considerations are guided by multiple values, and thus, making a decision could require a trade-off between competing values. Values affect actions when they are relevant in the context and important to the actor; therefore, not all values are seen as important in all situations. For example, in a context of studying, competing values can be hedonism and achievement. Although one can select studying one day over Netflix, it does not mean that the selection will be the same on the next day.

Since we are rarely conscious of our values, the topic is not easy to study despite its fundamental role in human-related issues. What makes researching values even more challenging is that values are personal, but in the research context, they will be vulnerable to biases. For example, sustainability has gained a normative status as a value, and people often state that they value it even if they do not act on it. In other words, people tend to articulate what they should value instead of what they truly do value.

Since values are something that guide our actions and considerations, they play an important role in our daily lives; they also affect the ways that we design, develop and use information technology. Thus, technological systems are never value free, since they serve certain purposes of the people that have ordered and developed them. On the user side, values affect the ways that technological systems are taken into use and employed. If a technological system is considered to be in conflict with one's values, people are likely not to use it or misuse it. Thus, it is parallel that personal values of the stakeholders – users, developers and so forth – are considered.

2.2 Organisational values

Aligning stakeholder values with the values embedded in a technological system is not only a burden – it can have major benefits for all stakeholders. This alignment of values is often called *value congruence*, which has been an interest of organisational studies because achieving it can offer competitive and financial advantages.

Organisational values differ from personal values and are closer to norms since they are meant to guide individuals' actions in the organisation. Organisational core values are defined as '[t]he organization's essential and enduring tenets – a small set of general guiding principles; not to be confused with specific cultural or operating practices; not to be compromised for financial gain or short-term expediency' (Collins & Porras, 1994, p. 73). Thus, organisational core values are seen as relatively permanent features of an organisation like personal values are for a person. However, it must be noted that organisational values are selected sets of values; thus, they are not learned and subjective to a similar sense as personal values. Moreover, the organisational values are highly contextual, since they are often effectual in only that organisation. Thus, organisational values

do not fulfil the characteristics of personal values as presented in the previous sub-section. Hence, organisational values are more norms of a certain organisation than individual values, although it is possible that employees will share some of the organisational values.

Organisational core values often reflect the personal values of the founder or founders of a company. Organisational values are often modified to fit the circumstances and era to guarantee that they are understood by the employees. Employees' clarity on organisational values is an important factor in value congruence; it can lead to many positive effects, such as better communication, higher predictability, increased trust towards the organisation and greater job satisfaction. Aligning organisational values with personal values has also been linked to a more committed workforce, since explicit organisational values can help people find meaningfulness in their work and help employees to be more motivated towards the goals of the business organisation.

It has been indicated that companies that are explicitly value led outperform others and are also perceived as better companies by potential employees that share those values. Thus, clear and meaningful organisational values can distinguish an organisation and help it to achieve its goals. However, not all companies have explicit core values, and in some case, the values are hollow, which can do more harm than good. For example, Lencioni (2002, p. 5) states:

> *Most value statements are bland, toothless, or just plain dishonest. And far from being harmless, as some executives assume, they're often highly destructive. Empty value statements create cynical and dispirited employees, alienate customers, and undermine managerial credibility.*

He further argues that building authentic organisational values takes a lot of work, but fixing the problems created by the poorly understood role and nature of organisational values is even harder. This demonstrates the complexity of values and their effects in an organisation well.

It must be noted that, in value congruence, clarity on personal values is a significant aspect. It is more likely that, for example, a positive attitude towards work and ethical practices are gained if employees have clarity over both personal and organisational values. To gain potential benefits from value congruence and its factors, one should provide an explicit, carefully considered set of organisational values that are understood by employees, as well as supporting an atmosphere where clarity on values in general is cherished.

In the development of the human-centric fair data economy, we should aim for explicit, carefully considered sets of values that constitute the organisational values and norms. By acknowledging the existing values, we should be able to make the values of data economy congruent with the values of individuals and thus, gain the potential benefits.

2.3 Moral values and ethical values

Ethical consideration is highly important when designing any technology. The concept of fairness has ethical aspects – as do some values. Before proceeding to values, which should be endorsed in a fair data economy, we need to consider the concept of morality and ethics, since ethics can be considered the philosophical study of morality (Feldman, 1978). Morals and ethics are not synonymous, although in many cases, they are misleadingly used as if they were. Moral values are characterised by an 'oughtness' factor that originates within individual, social group or society, and such demand concerns members that share that moral belief. However, the 'oughtness' is often more like an intuitive feeling of what is good or bad, and thus, hard to explicitly explain. Therefore, like other values, the moral values or codes can be and usually are different between individuals, groups and societies. Because of this, people usually live under three different levels of moral codes, namely, their personal code, the codes shared by people they interact with and the codes of the society they are living in. Similarly, like values in general, values and moral values can be conflicting. People may value overwhelming wealth, and at the same time, understand that greed is not a good thing. The problem with morals is that, in many cases, moral codes are not rational or well argued; thus, analysing the conflict between values and morals seems to be unfruitful. If values and morals are both learned and grown in, we need another approach for analysing what the justified values are.

As an example, if we look to the history of mankind, we can see different values and moral codes in different societies, and they often seem to be contradictory. These contradictions are now easier to see because of the development of the internet. In contrast, the development of technology in the 20th century and new millennium has shown how radically the values in society can change. For example, the emergence of social media has radically changed the moral codes of communication. Even 20 years ago, it would have been considered rude, for example, to be aggressive towards somebody in public, whereas now, aggression has become a somewhat acceptable way of communicating online—at least by some. These kinds of changes have also influenced the offline environment.

Discussions about moral norms and ethics can be seen as an integral part of modern technological development. People actively discuss the 'wrongdoings' and endorse the 'rightness' of service providers and technology developers. Moral judgements have become mainstream, and it is not uncommon that, once some behaviour is deemed immoral, these discussions affect businesses negatively. For example, in the financial sector, revelations about banks that use tax havens affected the general trust towards banks and the share prices in stocks plunged.

In the field of technology, the effects have been similar. For example, due to the Cambridge Analytica scandal, Facebook lost millions of dollars via decreasing stock prices, and 3 million users left the platform. It is notable that the company gained the lost revenue

back in just a few months and the amount of customers is not that big in proportion to the total number of Facebook users. However, this scandal affected the general trust towards Facebook. Nevertheless, it did not affect user behaviour in general: People still see this social media platform as an important part of their social life. Thus, one could conclude that if a platform is able to become important enough to users and there are no real alternatives in the market, people will use the platform even if they consider the company's actions immoral.

The example of Facebook shows that the lack of alternative services puts users in a hard position. Even if someone's acts would be considered immoral – which would normally lead to a lack of trust and then avoidance – in situations like this, avoidance is not possible because there are no viable alternatives. Thus, tech giants have been able to build monopolistic/oligopolistic positions. This limits the moral agency of the users and creates the possibility of moral disengagement. In other words, situations like this can lead to individuals being powerless to act on their morals and convinced that ethical standards do not apply in a particular context.

In the context of the fair data ecosystem, we should acknowledge the moral agency of individuals and create a market that allows alternative options for users. These alternative technical services should be designed to be ethical and governed ethically. This would allow the service providers to gain and maintain trust. To decide on what is considered ethical, we need guidelines of ethical conduct in the data economy. The need for ethical guidelines is essential so that judgements do not rely merely on subjective moral values but objective interpretations of commonly accepted rules. These rules should be designed from individuals' perspective so that data ecosystems will truly safeguard them and their role as active moral participants.

2.4 Towards the values of the data economy

Three types of values – personal, organisational and ethical – form a good basis for understanding how values can affect the data economy and why they should be considered when designing any socio-technical system. It is noteworthy that these types of values are not separate from each other but overlapping (see Figure 2).

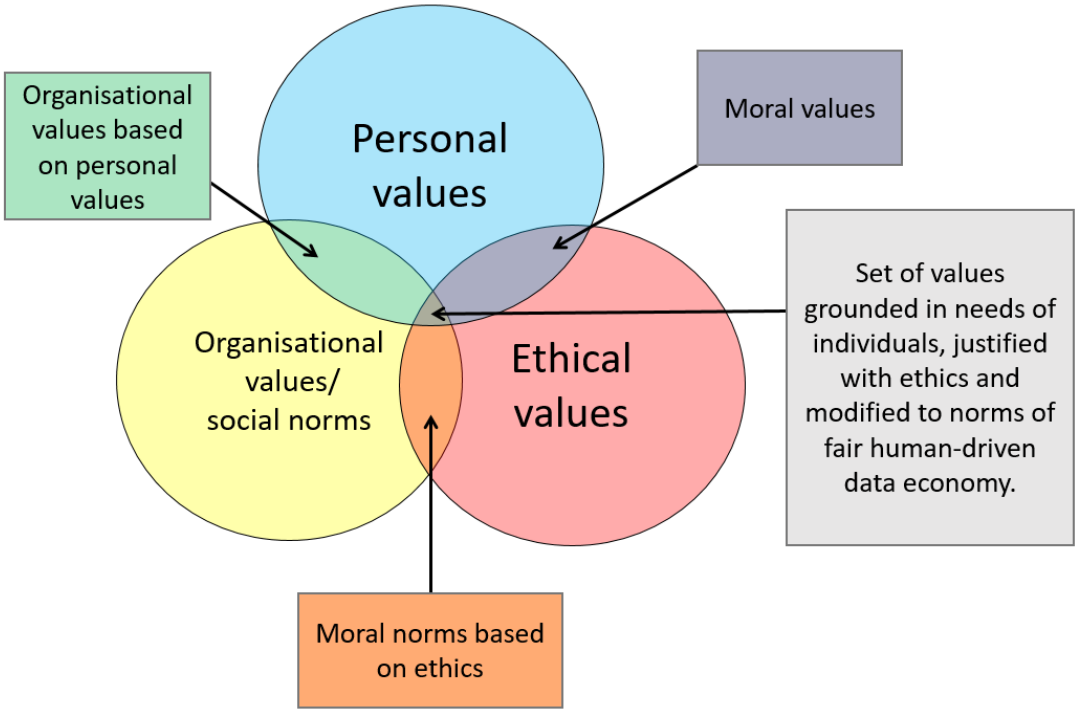


Figure 2. Values in the data economy.

When personal values overlap with ethical ones, we can talk about a person's moral values. Similarly, when organisational values/social norms overlap with personal values, we can talk about organisational values that are based on personal values. The overlap between organisational and ethical values can be described as moral norms that are based on ethics. It is noteworthy that, in a sense, moral values are personal values, whereas organisational values can reflect personal values. All three main categories and their intersection exhibit small differences in their characteristics. The scope of this study, and thus, the values of the data economy, are in the middle of the diagram. This refers to a set of values that are grounded in the needs of individuals, justified with ethics and modified to the norms of a fair human-centric data economy.

Since it is impossible to consider all the values of all individuals, we need a way to justify the selected ones. Due to ethics' relation to morals and morals' relation to values,

ethics is a suitable tool for analysing the justification of values. As ethics is a brand of philosophy – based on rational and scientific argumentation – that aims to find what is good and bad in a normative sense, it seems a promising basis for analysing which values are worth cherishing and which should be altered or even rejected.

Habermas' (1996) rational discourse is based on the view that all stakeholders have the possibility to participate in discourse. All arguments in rational discourse are evaluated in terms of how convincing and plausible they are. Those arguments can vary depending on the issue at hand, and they can be based on logic, ethics or another justified basis. What is notable is that no strategic games are allowed in rational discourse. A strategic game is a way of influencing others where some participant is trying to bargain for some outcome by using something other than a better argument, and this is not allowed.

We can test our values by analysing them with ethical theories, and in philosophy, the strongest rational argument should win instead of relying on strategic games where people with more negotiation power win (see e.g. Habermas, 1996). The philosophical approach in rational discourse is more suitable compared with approaches of natural sciences that focus only on the 'natural' rationality, and thus, bypasses the humanistic and holistic view – which is needed in the context of values. Thus, when analysing values through ethics, we should be able to avoid the major pitfalls that are based on strategic games instead of trying to make deals between different stakeholders, which usually ends up with the strongest winning. As in science, better evidence – better philosophical argument – should always win.

It has already been noted[2] that Habermas' rational discourse is a promising approach for gaining insight into and understanding social aspects of information systems. It has also been noted[3] that Habermas' work has played an important role in IS development and has great potential to develop it further. These insights demonstrate the potential of Habermasian discourse in the context of the data ecosystem. In comparison with participatory approaches, it does require extensive representativeness, but it still endorses participation of all willing parties, and most of all, asserts rules for how the discussion should proceed through negotiation. This makes the Habermasian discourse a good approach to the democratic development of norms, although it is an ideal.

Especially important is the idea that norms should be formed through a discourse in which all stakeholders – here, individuals, service providers, government (national and EU level), third sector – have the opportunity to be involved. Such a participatory approach to norm formation is needed, because otherwise, norms would fail to be legitimate in the eyes of those who ought to follow them.

---

[2] Lyytinen, K. & Hirschheim, R. (1988) Information systems as rational discourse: an application of Habermas's theory of communicative action. Scandinavian Journal of Management, 4(1), 19–30.

[3] Ross, A. & Chiasson, M. (2011) Habermas and information systems research: New directions. Information and Organization, 21(3), 123-141.

Rational discourse is a way of gaining common ground where agreement or a shared background cannot be seen as granted. Discourse has four criteria that must be met for it to be described as rational. These are **clarity**, **truthfulness**, **correctness** and **appropriateness**. In addition to these criteria, there are certain ground rules. First, actors have the **possibility to participate in discourse** and **express their arguments**. Second, all actors **need to accept the better and more rational argument over a lesser one**.

In the IHAN project, it is important to support communication procedures between all stakeholders during the project. Likewise, after the project ends, there is a need for a clear governance model where all stakeholders are included in developing the basis (such as a rulebook, ethical guidelines and blueprint) of the IHAN ecosystem together and in a fair manner.

Summary

- Values are things held to be important by an individual or a group
- Personal values refer to the values that an individual holds and that guide their actions
- Organisational values are more like norms that guide actions in an organisation
- Achieving value congruence between personal and organisational values can create great benefits
- Moral values are personal values that reflect 'oughtness', and they are used when making moral judgements
- Ethics is a way to make objective moral judgements and justify actions
- Creation of ethical guidelines should be based on the values of individuals, clearly presented and well justified
- Negotiating ethical guidelines of the data economy is required, and it should be rational discourse: clear, truthful, correct and approriate discourse between all willing parties, where the most rational argument wins

3 CULTURAL DIFFERENCES

Although culture and values are two different things, that values are learned makes the two elements inherently intertwined. Since cultures affect values, and we analyse the values of different cultures, it is important to acknowledge the slight differences between these concepts. In this study, we aim to determine the European value basis for the data economy, and as we are using data from only four countries, we need to understand the difference among the cultures that we are examining. To highlight the existing cultural aspects and their potential effects on the survey results, we now give a brief perspective on the culture of the European countries in general, then take a closer look at the cultures of Finland, France, Germany and the Netherlands.

3.1 Cultural dimensions

Culture is a concept that was described as 'coding of the mind' by Geert Hofstede. This metaphor highlights the effect that culture has on the human mind – the way we perceive the world and act in it. Hofstede researched cultures and realised that there are six dimensions that can be used to describe all cultures, which are as follows:

1) Power distance index (PDI);
2) Individualism versus collectivism (IDV);
3) Masculinity (MAS);
4) Uncertainty avoidance index (UAI);
5) Long-term orientation versus short-term normative orientation (LTO); and
6) Indulgence versus restraint (IND).

More detailed descriptions of these dimensions are presented in the 'Dimensions of national cultures' box below. The data describing the current cultural dimensions are kept up by the Hofstede Insights, an organisation focussing on operationalising academic culture research in the business area based on Hofstede's work[4].

It must be noted that there is no clear 'European culture'. National cultures in the EU vary a lot (see Table 1). The biggest differences in the dimensions of power distance are masculinity. In the power distance, Austria has a culture where no inequality is tolerated, whereas in Slovakia, it is accepted that some people have more power than others do and they use it. Similarly, Slovakia is a masculine culture, whereas Sweden is a feminine one.

It must be acknowledged that these dimensions and how cultures score on them do not represent 'good' or 'bad' societies. Cultural dimensions only characterise cultures in a

[4] See https://www.hofstede-insights.com/ for information on the Hoftede insights and https://www.hofstede-insights.com/models/national-culture/ for description of dimensions.

quantitative manner so that comparison is possible. This quantitative approach also encapsulates the culture in a few numbers. Thus, it must be acknowledged that, in reality, cultures are not this unified, and individuals can value different things despite the scorings. Overall, these cultural dimensions allow us to review cultural differences, but making too far-reaching conclusions based on these alone is not appropriate.

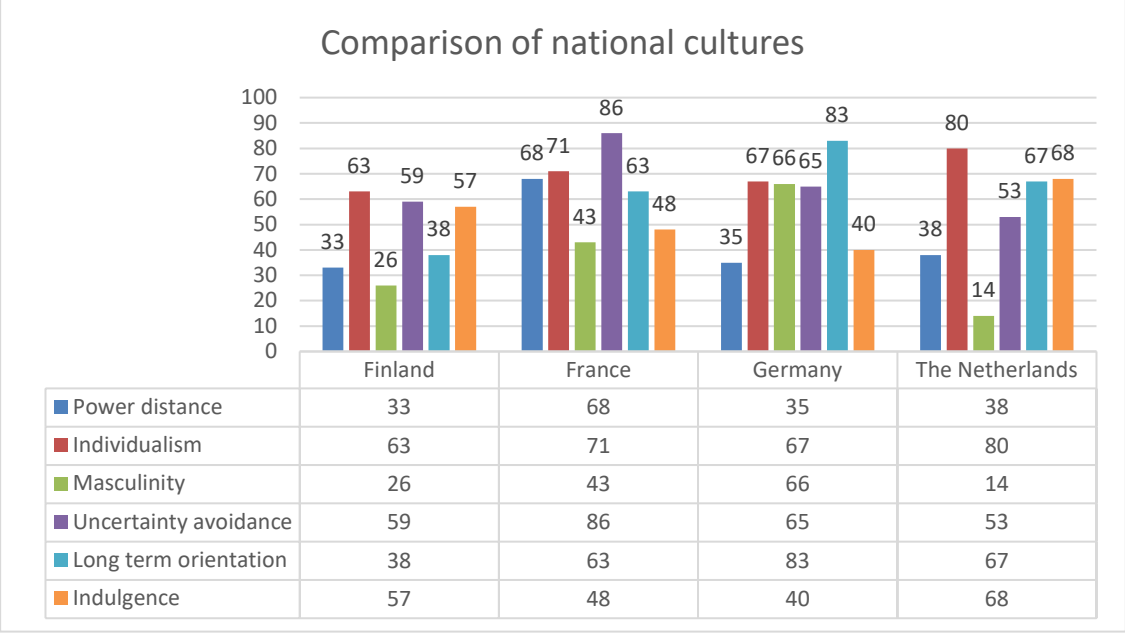
Table 1 Comparison of Culture Scores (Source: Hofstede Insights)

		PDI	IDV	MAS	UAI	LTO	IND
Selected countries	AVG.	43.5	70.3	37.3	65.8	62.8	53.3
EU	AVG.	51.4	58.6	46.0	70.6	57.5	43.4
	Min	11	27	5	23	24	13
	Max	100	89	100	100	83	78

* Note: Data from Cyprus are unavailable.

3.2 Cultures of the studied societies

It is notable that Finland, France, Germany and the Netherlands have some differences in their national cultures (see Table 1 and Figure 3), but they represent a more modest variation of scores than all the EU countries together.



	Finland	France	Germany	The Netherlands
Power distance	33	68	35	38
Individualism	63	71	67	80
Masculinity	26	43	66	14
Uncertainty avoidance	59	86	65	53
Long term orientation	38	63	83	67
Indulgence	57	48	40	68

Figure 3 Comparison of national cultures (source: Hofstede Insights).

Finland, France, Germany and the Netherlands all have unique cultural characteristics that may affect their citizens' values and attitudes. Thus, it is important to understand the differences and similarities of these cultures before analysing citizens' values. Their scores are fairly close to the average scores of all the EU countries. However, there are some slight differences: The selected countries represent more individualistic and indulgent cultures and score slightly higher on long-term orientation. Thus, it cannot be stated that these cultures fully represent the plurality of European cultures, although they are fairly close to average.

In general, the four countries present fairly low PDI cultures. However, in comparison to the other countries, France has a high PDI. This means that French people accept a fair degree of inequality and centralisation of power, whereas the other three countries, prefer equal rights and decentralisation of power. In terms of PDI, the studied countries represent an average EU culture fairly well, although it must be noted that the PDI varies widely, and thus, an average does not guarantee generalisability.

In terms of individualism (IDV), all the countries are fairly similar, scoring moderately high or high. This means that the cultures in all four countries are individualistic. This dimension is about the degree of interdependence a society maintains amongst its members. In individualist societies, people are supposed to look after themselves and their direct family only. In collectivist societies, people belong to 'ingroups' that take care of them in exchange for loyalty. Although most European countries are individualistic, it must be remembered that not all of them are. For example, Portugal, Slovenia, Bulgaria, Romania, Croatia and Greece are collectivist societies, and thus, differ significantly from the four countries studied in this research.

For MAS, our four countries represent both masculine and feminine cultures. The fundamental issue in the MAS dimensions is what motivates people. In masculine societies, people want to be the best, whereas in feminine cultures, people want to like what they do. Germany is a masculine society, with the highest score amongst the four countries. This means that, in Germany, performance is highly valued, and this value is embedded, for example, in the schooling system and work life. In contrast, the Dutch culture is the most feminine culture, with the lowest score on MAS. Thus, in the Netherlands and cultures alike, it is important to keep the life/work balance and make sure that everybody is included. On average, the four countries represent somewhat feminine cultures, whereas the EU countries in general represent slightly more masculine societies.

The dimension of the UAI refers to the extent to which the members of a culture feel threatened by ambiguous or unknown situations and have created beliefs and institutions that try to avoid them. In general, all four cultures prefer avoiding uncertainty. This means that all of these cultures have an emotional need for rules, and precision and punctuality are the norm. In contrast, cultures scoring high on the UAI value security over change, which may affect attitudes towards innovations. Although these four countries have high

UAI scores, they still score below the EU average. In other words, it can be stated that culture in the EU is not open for uncertainty, which can create a challenge for the development and creation of the fair data economy. Thus, it is of the utmost importance that people are actively given the possibility to be part of the development process and the whole ecosystem is based on transparency.

LTO can be a challenge. This dimension describes the tension between maintaining links with the past and dealing with the challenges of the present and future. Societies with a low LTO are normative societies that, for example, prefer to maintain traditions and norms, while viewing societal change with suspicion. For example, Finland can be classified as a normative society, where traditions are respected and the focus is on achieving quick results. In contrast, France, Germany and the Netherlands are more pragmatic, with higher scores on LTO. In these kinds of societies, thrift and effort are encouraged, and people adapt their traditions easily. In comparison to all member nations of the EU, on average, these four countries represent slightly more pragmatic societies. Germany is the most pragmatic culture, whereas Finland is the fourth most normative. Thus, in terms of LTO, these four countries represent the cultures of the EU fairly well.

Finally, the IND dimension refers to the extent to which people try to control their desires and impulses based on how they are raised. Societies with relatively low IND scores are restrained, with stronger control of desires and impulses. People in these kinds of societies are restrained by social norms and feel that self-indulgence is somewhat wrong. They also have a tendency towards cynicism and pessimism. In contrast, in indulgent societies with high IND, people have a positive attitude and tendency towards optimism. In indulgent societies, leisure time is appreciated, as well as acting as one pleases and spending money as one wills. In this dimension, Finland and the Netherlands represent indulgent cultures, whereas Germany and France are more restrained.

In comparison with the EU countries, these four states are more indulgent, since an average culture is slightly more restrained than indulgent. It must be acknowledged that, although these dimensions are often presented as separate ones, the real culture is a result of their synergism. For example, the French have a slightly restrained culture and score high on uncertainty avoidance. This implies that the French are not as relaxed as is commonly assumed.

Since the focus of this study is on personal values, not cultural ones, the interplay of cultural dimensions and the search for the cultural value base of the EU is beyond the scope of this research. This small insight, however, does highlight the importance of acknowledging the different cultures affecting the values of European individuals.

Summary

- Personal values are affected by cultures and cultural norms
- There is no clear ‘European culture’, but cultures in the area vary
- The four countries selected as subjects of the citizen survey all have their distict cultures that can affect the results
- In comparison with other countries in the European Union, these four cultures are fairly representative, although it must be noted that they
    - Expect slightly more equal power distribution
    - Are more individualistic
    - Are more feminine
    - Are little more open to change
    - Are sligtly more long term oriented
    - Are more indulgent

Dimensions of national culture

1) **Power distance index (PDI):** This dimension expresses the degree to which the less powerful members of a society accept and expect that power is distributed unequally. The fundamental issue here is how a society handles inequalities among people. People in societies exhibiting a large degree of power distance accept a hierarchical order in which everybody has a place and that needs no further justification. In societies with a low power distance, people strive to equalise the distribution of power and demand justification for inequalities of power.
2) **Individualism versus collectivism (IDV):** The high side of this dimension, called individualism, can be defined as a preference for a loosely knit social framework in which individuals are expected to take care of only themselves and their immediate families.
   Its opposite, collectivism, represents a preference for a tightly knit framework in society in which individuals can expect their relatives or members of a particular ingroup to look after them in exchange for unquestioning loyalty. A society's position on this dimension is reflected in whether people's self-image is defined in terms of 'I' or 'we'.
3) **Masculinity versus femininity (MAS):** The masculinity side of this dimension represents a preference in society for achievement, heroism, assertiveness and material rewards for success. Society at large is more competitive. Its opposite, femininity, stands for a preference for cooperation, modesty, caring for the weak and quality of life. Society at large is more consensus oriented.
4) **Uncertainty avoidance index (UAI):** The uncertainty avoidance dimension expresses the degree to which the members of a society feel uncomfortable with uncertainty and ambiguity. The fundamental issue here is how a society deals with the fact that the future can never be known: Should we try to control the future or just let it happen? Countries exhibiting strong uncertainty avoidance maintain rigid codes of belief and behaviour, and they are intolerant of unorthodox behaviour and ideas. Societies with weak uncertainty avoidance maintain a more relaxed attitude, in which practice counts more than principles.
5) **Long-term orientation versus short-term normative orientation (LTO):** Every society has to maintain some links with its past while dealing with the challenges of the present and future. Societies prioritize these two existential goals differently. Societies scoring low on this dimension, for example, prefer to maintain time-honoured traditions and norms while viewing societal change with suspicion. In contrast, cultures that score high on this dimension take a more pragmatic approach: They encourage thrift and efforts in modern education as a way to prepare for the future.
6) **Indulgence versus restraint (IND):** Indulgence stands for a society that allows relatively free gratification of basic and natural human drives related to enjoying life and having fun. Restraint stands for a society that suppresses gratification of needs and regulates it by means of strict social norms.

Source: Hofstede Insights, https://www.hofstede-insights.com/models/national-culture/

4 VALUES OF THE FAIR DATA ECONOMY: AN ANALYSIS OF THE SURVEY

4.1 The data

To determine the values of citizens in relation to the fair data economy, we conducted a thematic analysis of the citizen survey results. The survey was conducted by Kantar TNS Ltd for Sitra. The questionnaire was distributed through online panels in four countries in 2018. In total, 8004 individuals answered the questions (N = 8004). The numbers of answers are presented in Table 2.

Table 2 Answer Rates by Country and Questions

Country	Number of respondents	Number of open question respondents
Finland	*n* = 2000	Q20 *n* = 15 (0.8%) Q21 *n* = 705 (35.3%) Q23 *n* = 762 (38.1%)
Germany	*n* = 2004	Q20 *n* = 7 (0.4%) Q21 *n* = 615 (30.7%) Q23 *n* = 581 (29.0%)
Netherland	*n* = 2000	Q20 *n* = 6 (0.3%) Q21 *n* = 568 (28.4%) Q23 *n* = 537 (26.9%)
France	*n* = 2000	Q20 *n* = 8 (0.4%) Q21 *n* = 497 (24.9%) Q23 *n* = 527 (26.4%)
Total	*n* = 8004	Q20 *n* = 36 (0.4%) Q21 *n* = 2385(29.8%) Q23 *n* = 2407(30.1%)

The survey included questions that aimed to clarify the use of digital services and attitudes towards personal data collected during use. The aim was to clarify how individuals feel about the potential use of their data from the perspective of data protection and privacy. The survey included four sections: 1) background information, 2) rights in relation to data and attitudes towards terms of use and privacy settings, 3) trust towards service providers and increasing trust and 4) disclosure of information and its management and the concept of fair data service.

In the survey, there were 23 questions, including 3 open questions. In this qualitative analysis, we focus only on analysing the answers to the open questions:

Q20: Which features would you expect/require from the above-mentioned centralised service in which you could manage all of the data about you accumulated in different applications/services? – Other, please specify.

Q21: Service providers collect a lot of data about you. What should the use of your personal data be like to be fair?

Q23: If applications/services that use data fairly had 'a fair digital service label', what should it guarantee at the least?

The numbers of answers to these questions were more modest than the answers to the quantitative questions. In total, 2385 respondents answered question 21 and 2407 answered question 23. However, question 20 was answered only 35 times, representing a mere 0.4% of respondents; in contrast, questions 21 and 23 were answered by 30% of the respondents. Thus, no analysis has been done for question 20 because of its low answer rate.

4.2 The research process

The aim of this research was to determine the value basis of European citizens in relation to the fair data economy. To accomplish this, open questions were analysed. All the questions concerned attitudes or expectations towards the fair data economy. Thus, it is fair to assume that these answers also reflect what citizens value in the context of the data economy.

First, after acquiring the data, the researchers gained familiarity with the material. Open answers were read and notes taken about the initial ideas of values that could be seen in them. After familiarisation, the data were coded with NVivo Pro software. In practice, all answers were read through and a code or codes were attached to them if they reflected any particular value. Some of the codes were based on the preliminary analysis of the data, but some were added during the coding process.

Since coding of the answers is an interpretative process, it is prone to subjectivity. To ensure objectivity, the coding was discussed and decided on among the authors. Despite this, it must be acknowledged that thematic analysis still is an interpretative approach, and as such, open to other interpretations as well.

Some answers were too vague for making interpretations about values. Thus,it is impossible to interpret what the values behind this answer are, and thus, we refrained from coding answers such as these to avoid interpretations that were too farfetched. However, most answers did reflect some values, and thus, the number of analysed answers remained high enough. In total, 3831 value instances were identified from the open questions

clearly enough to be included. One answer could possess one or several values, and some did not have any that could be clearly interpreted as value statements.

After coding all the answers, similar themes were combined. For example, the answers, 'openness, trustworthiness, protection, possibility to choose what it is used for and by whom, **ability to erase** and/or correct data (which is currently possible in theory only)' (Q23, FI25) and 'I can view them quickly and **delete them if I want to**. Information collected for a single study' (Q23, FR130) clearly reflect the **possibility to delete or remove information** as a value, although the responses differ in the choice of words. Combining the coding resulted in eight distinctive themes that red throughout the data; these are presented in the results section in further detail.

4.3 Research limitations and future research

This study can be used to consider which values are the most relevant ones when developing the fair data economy ecosystem. The research can be used to analyse the values that should at least be considered when developing such an ecosystem. However, the analysis is based on only two open questions in a larger, mainly quantitative based, survey. Moreover, the questions were not developed by researchers, and they were not designed for this type of research. Thus, the results cannot be seen as evidence of values, but instead, as providing insight into relevant issues for future research.

There is a need for empirical interpretative research to gain deeper understanding of citizens' values and motivations in the context of the data economy and personal information. Likewise, the answers were analysed and categorised by the researchers, and thus, involved a subjective view of values. However, this study indicates that people have values they want to be respected in the fair data economy. These values should be evaluated more deeply in the future, with a view to implementing them in practice.

5 RESULTS

When open answers of the questionnaire were analysed by theme, distinctive and repeated values were found. The thematic analysis resulted in eight main themes, which are presented in Table 3. All these themes, except theme 8, consist of several subthemes. For example, theme 1 includes all subthemes in relation to individuals' desire to control their data and data sharing, such as asking for consent, limitations of secondary use and possibility to remove information. Examples of answers implying subthemes are handled in more detail in the subsections below.

Table 3 Themes Found from the Open Answers

	Major themes	References
1	User's control over data and data sharing	1829
2	Transparency and being informed	1396
3	Security	1346
4	Trust and fairness	796
5	Compensation or benefits for users	251
6	Supervision and rules	226
7	Attitudes towards data collection and the data economy	173
8	Unable to answer/does not know	61
		6017

The themes included all the coded answers. Most of the answers included between one and three distinctive themes. Some subthemes could have been included in a multitude of themes due to some overlap in issues. For example, notions like 'no selling to third parties' could have been handled under theme 7 instead of theme 1. These selections were discussed by the authors, but it is still noteworthy that the boundaries of these themes are not clear cut, and the presented themes are based on the authors' interpretation about their similarities.

It must also be noted that not all the answers were coded: Some answers were too brief to draw any conclusions from. For example, for both questions, some respondents replied with the name of an organisation without context or explanation. Since it is impossible to know what exactly was being argued with these answers, they were not considered as a theme. The amount of uncoded answers was fairly low, and on average, 98% of all the data were coded.

In the following subsections, themes 1–7 are discussed in further detail. Theme 8 is not discussed further, since none of the 61 respondents who stated that they were unable

to answer provided any explanations. In addition, these answers were relatively rare, and there was no apparent link between the question and nation of the respondent.

5.1 User's control over data and data sharing

Answers that included some notion about control over data and data sharing formed the largest distinguished theme, with six different subthemes. The subthemes are presented in Figure 4.

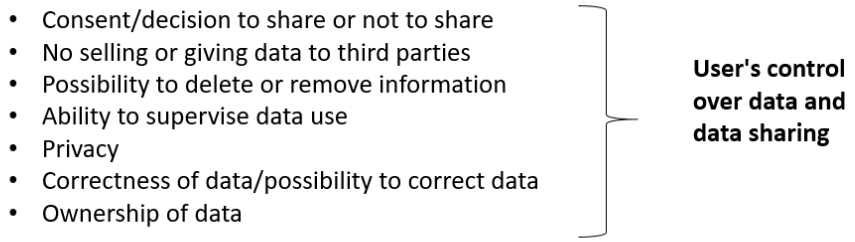


Figure 4 Themes related to control over data and data sharing.

The respondents clearly highlights consent as a crucial part of the fair data economy. The themes under this main theme are represented equally in all countries, although word choices vary.

Ways of stating the willingness to control one's data vary a lot, but many state that they should be the ones deciding whether data are shared and whom they are shared with. Thus, many of the answers reflect **privacy** as a value, since privacy refers to people's ability to determine independently when, how and to what extent information about them is communicated to others. Privacy is largely seen as self-explanatory or implied with other choices of words:

> *Not disclose any very personal data due to personal safety and privacy. (GER Q23 ID264)*

> *With the utmost care and privacy. (NE Q21 ID562)*

Many of the respondents express that **consent** should also be asked for when data is passed to third parties:

> *Asking me for permission; my data would only be disclosed onwards with my permission and I would be paid for it. I could easily withdraw my data*

> *at any time. I don't want a long list of my data but an easy way to withdraw it, if necessary. (FI Q23 ID355)*

This answer also highlights the **dynamic nature of consent**. The respondent expresses that he/she would like to 'easily withdraw data', which could refer to revoking of consent or **deletion of data**. Some respondents are more explicit about their wishes to delete data:

> *It is reported in advance what is collected and why. The data could be deleted or its collection prevented if you want. You can personally limit what data is collected. (FI Q21 ID186)*

Informed consent also links to transparency and being informed (theme 3). Some of the respondents emphasise **the clarity of terms and conditions** as part of consent. For example,

> *To ask the user beforehand for their informed consent (and every time it is useful), to clearly explain how they are used and what for, to clearly lay out which actions to take to amend or delete them, to mention how long they are stored for. (FR Q23 ID365)*

It is noteworthy that acceptance of terms and conditions can prevent service use; for example, one resbondent states:

> *I should be told which information is collected and what it is used for. When I am aware of it, I will have the possibility of not using the service if I do not accept the collection of data. I am not on Facebook, for example, because I don't accept its terms and conditions. (FI Q21 ID384)*

The answers also explicitly and implicitly highlight the **ability to supervise the data and their use** in addition to consent:

> *The minimum is to have insight into all the data. What is being collected for what and that you can change that yourself. (NE Q23 ID260)*

> *Data security and appropriate use of data is primary. Similarly, information about who is collecting and what data about it should be easily available, what it is used for and how I can influence the volume and nature of the data collected with my own choices and have it erased if I want. (FI Q21 ID610)*

Ability to supervise the data and its sharing is also seen as important in terms of **passing data to third parties**:

> *I want to know myself who else receives my data. I believe a lot of the time, third parties simply DO receive the data. I have just had an experience like that which very much surprised me. (GER Q21 ID199)*

In general, the respondents seem to have a **negative attitude towards secondary use of information**, especially selling information to third parties. Answers like '*no forwarding of data of any kind*', '*Not shared to everyone*' and '*No sale to third parties*' are common. However, some do see secondary use of data as a possibility if they can control the data use:

> *It should happen in a transparent way and I should have the opportunity at any time to undertake changes or deletions. And the provider must always specifically get my permission if he wants to pass on the data to third parties. (GER Q21 ID251)*

Many respondents, especially in France, wish for **access** to see what kinds of data are collected about them. Ability to supervise data and data use also link to **ability to correct data**. Without seeing the data, it is not possible to correct it:

> *It should guarantee that the consumer has access to editing their own data, erasure, etc. It should also guarantee who it discloses data to, and to my mind, it might also be good to have automatic approval as an option, as well as manual approval, so that you can decide which third parties you want your data to be disclosed to. (FI Q23 ID 474)*

Although people clearly have interest in gaining control over their personal data and its use, only a few people directly mention the **ownership of data** and suggest that they should have it. However, seeing data as personal property is a reoccurring implicit theme that is also visible in the demands for compensation (theme 5). Thus, control over the data and data sharing also relates to the idea that individuals see personal data as a valuable asset that should not be exploited.

Overall, it becomes clear that respondents feel strongly that they should have control over their data and its usage in a fair data economy. The answers seem to highlight the values of **power and self-direction**, more specifically, values of **privacy** and **autonomy**—an ability to make decisions without being controlled by anyone else. However, the responses also give practical requirements for privacy and autonomy, namely, informed consent that can be withdrawn and the ability to have control over who receives the datasets. To control consents, the individuals should also be able to supervise the status quo of the data collections and their use. This is also a requirement from those who wish to supervise the correctness of their data. In return, people could be willing to edit the data, which could enhance the quality of the datasets.

For a fair data economy, needs for control over data and data sharing not only pose challenges but also potential solutions. Although the respondents seem reluctant to share their data with third parties, giving the individuals autonomy over their data and means to control it could lead to quality datasets. However, current practices are not seen as sufficient, and thus, serving the needs of respondents means redesigning the ways in which data are collected and distributed. Individuals should be treated as active participants, and their autonomy should be respected in all practices.

5.2 Transparency and being informed

Demands for transparency and being informed formed the second largest theme. The subthemes are presented in Figure 5.

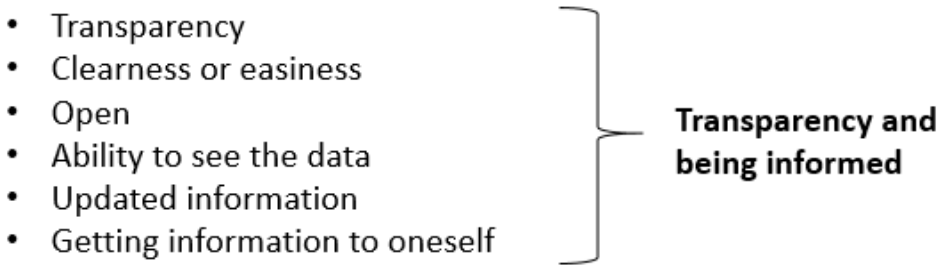


Figure 5 Themes related to transparency and being informed.

Transparency and being informed overlap with the ability to control data and their usage. A small difference is that people wanting to be informed in this theme do not suggest they need some power over the issue; they just want to obtain information about the data usage. Thus, this theme could also be described as **being aware** of what data are collected and how they are used within and outside an organisation. Being aware of the data and data usage are crucial in terms of control over data; this theme can be seen as an enabler of the power over data.

In this context, **transparency** refers to organisations and its activities being open to public scrutiny. It is a precondition for the ability to be aware of an organisation's actions. Transparency is often presented without further explanation, suggesting that people have a strong intrinsic conception of it. Often, the respondents view transparency as a clear and simple manner of explaining data usage:

> *That each provider collecting our data asks for our opinion and explains clearly and simply what they are being used for. (FR Q21 ID253)*

> *Open and clear, that is, saying in plain language which data and for what purpose it will be used. Very many data protection approval policies are complicated legal jargon that you have to read many times to really get the essence. These should be expressed more clearly and concisely. (FI Q21 ID60)*

In addition to being seen as attributes of transparency, **clearness and easiness** are seen as attributes of data, as well as the issue of **usability**:

> *Present all data in a clear way, an overview of how this data should be used, be able to forbid passing on the data to third parties without having to go without the service. (GER Q21 ID 382)*

> *First, currently, the statutory information about the use of your own data is often ultimately difficult and bureaucratic to obtain. This should be made considerably easier. Second, the fact that I have to review long lists of different purposes of data use and amount of advertising on each site I visit is far from easy and user friendly. It feels like it was intentionally made as difficult as possible so that as many as possible would just automatically accept everything. A central, neutral and non-commercial service in which I can specify at least broader guidelines on what I want to allow in terms of the use of my data would make it a lot easier. (FI Q21 ID166)*

Transparency is often linked to or used synonymously with **openness**:

> *Complete openness of what the data provided by a person is used for, and the person is able to change their permission settings at any time, easily and conveniently. The person is given an opportunity to protect themselves. (FI Q23 ID214)*

Informing users about changes is also seen as part of transparency. Thus, transparency is seen as a continuous action, not just static statements that are occasionally approved:

> *Open and transparent towards the clients, that at all times you can gain insight into your data and can remove them, that they never share your information with third parties without your personal consent, that they don't just change the conditions without informing you about it. (NE Q23 ID 269)*

A**bility to see the data** in addition to monitoring data use is seen as an informative action:

I should receive a summary of the data used at certain intervals and whenever I request it. (FI Q21 ID400)

To have access to my data so that I have insight into what is being collected. (NE Q21 ID483)

Thus, transparency is inherently linked with the information that is given to the people that the data are about. This information is crucial in making decisions about the data use and giving informed consent. In contrast, transparency is seen as a means to gain trust (theme 4), and thus, it can also be seen as instrumental by nature. From the perspective of values, transparency and being informed connect to self-direction, universalism, conformity, security and power. Thus, they have a wide value basis. Links between different themes and value categories emphasise the complex nature and effects of transparency and being informed. Although it seems clear that transparency is an instrumental value, it is an instrumental value that enables more values.

5.3 Security

Security-related themes were the third most common theme in the answers. The subthemes of this theme are presented in Figure 6.

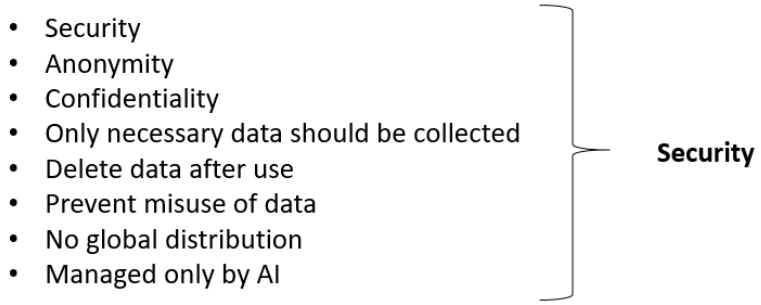


Figure 6 Themes related to security.

Respondents from all the countries highlighted security and/or privacy as a feature of the fair data economy. It is clear that **data security –** protecting data from unauthorised access and data corruption – is and should be sufficient in any technological solution. Many of the respondents just stated that the requirement for the fair data economy or fair data label is that the system is secure. Often, security is mentioned along with trustworthiness and reliability, linking this theme to theme 5, 'Trust and fairness'.

The respondent from Finland manages to summarise multiple subthemes in his/her answer:

> *The data must not be identifiable and possible to link directly to me. Data must not be disclosed without permission. Intellectual property rights belong to me, so I must be compensated for their use. The compensation could also be, for example, the right to use the services instead of money. My data may only be stored within the EU, for example. Security must be a priority, and I only accept two-tier identification. My data may only be disclosed for non-profit or scientific purposes. (FI Q21 ID568)*

Anonymity reoccurs in the dataset without further explanations. The respondents often state that the data should be anonymous, which generally means that it is not possible to identify the person the data are about:

> *Anonymous usage, personal identification is not possible. (GER Q23 ID37)*

However, it seems that, in some cases, people's perceptions about what data should be anonymised or what constitutes sensitive information vary:

> *To commit to store them securely, to not share any sensitive information (address, phone number, medical and financial data for instance) and to really delete any data deleted by the user. (FR Q23 ID381)*

> *I should be informed transparently about usage. The data is anonymously passed on (without names). (GER Q21 ID190)*

> *To not share phone number and email. (FR Q23 ID437)*

Confidentiality is highlighted in many answers and often used as a synonym for data security:

> *Confidential data is not disclosed to other parties without permission, and no access for inappropriate people. (FI Q21 ID318)*

> *Strictly confidential, access possible for me at any time. (GER Q21 ID499)*

It is notable that **disclosing information or data without** permission is also seen as a violation of security. Thus, there is a contradiction between the general notion of data security and what people see as security:

> *Guarantee 100% data security, that is, the data is not leaked anywhere. Absolutely guarantee that the data is ONLY used by the service provider in question and not disclosed to anyone. (FI Q23 ID489)*

Although many of the respondents demand total data security, some accept that leaks are possible and note that, in the case of such events, honestly **informing the users** is seen as crucial:

> *That data security is kept, and if there is a leak, it is honestly reported. When requested, the data is actually erased and it must be possible to tell who has processed the data. (FI Q23 ID183)*

The respondents also emphasise the nature of the data collected, especially that o**nly necessary data should be collected** and that **data should be deleted after use**. Only necessary data for the necessary time are permissible in relation to both privacy and security:

> *The terms and conditions of use should be clear. Use of personal data should be restricted to the absolutely necessary minimum for the operation of the service in question. (FI Q21 ID283)*

> *[D]ata should be kept for a limited time, a minimum period of time for specific reasons, then regularly deleted. (FR Q21 ID171)*

Some of the respondents also mention area-based restrictions in their answers. Most want to **restrict global distribution** of data and keep data on national servers. Saving the data to foreign servers is identified as a risk:

> *No sharing, copying or saving outside the national territory. (FR Q21 ID195)*

Preventing misuse of data is also frequently mentioned, although it seems that perceptions of what misuse of data truly means vary. Misuse seems to entail anything from advertising to criminal activities and profiting, and thus, it remains unclear what constitutes misuse according to the respondents.

Some of the respondents note the possibility of **human errors** as part of security, which in some cases, also means that people would only want their **data to be handled via technical solutions**, such as artificial intelligence:

> *An external auditor should periodically (often) check that the data is reliably and securely stored, no 'whoops' human error situations or hacking should possibly happen. (FI Q23 ID721)*

Not surprisingly, it seems that security is a highly important value in relation to the fair data economy. What is somewhat surprising is the general nature of answers in this theme. It seems that the content and nature of security are self-explanatory concepts, although it is clear that people's conceptions of them vary remarkably. Nevertheless, it is

apparent that security is something that should be taken as a value of the fair data economy to safeguard, for example, privacy.

5.4 Trust and fairness

Trust and fairness are the fourth biggest reoccurring theme in the answers. Subthemes of trust and fairness are presented in Figure 7.

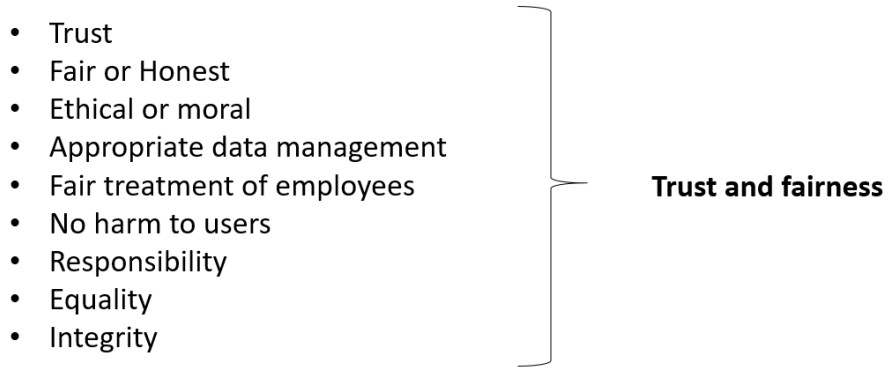


Figure 7 Themes related to trust and fairness.

Trust is a strong belief in the reliability, truth or ability of someone or somebody. Thus, trust is a subjective feeling that someone or something is trustworthy. Building trust is complex due to the subjective nature of trust. As stated above, transparency has a strong connection to trust. In addition, security is often seen as a measurement that allows trust. Thus, trust is interconnected with different themes in this study. Fairness, as explained above, is also a subjective issue, and thus, similar to trust. Fairness is a naturally cross-cutting theme in this study. How people perceive actualisations of trust and fairness is an important issue in the development of the fair data economy.

In the answers to the questions about the fair data economy, trust and trustworthiness are often mentioned, as well as other similar attributes describing the benevolence of the data economy. Another notable point is that answers reflecting distrust are not uncommon, but these answers are presented in more detail in theme 7, since they clearly form a separate theme.

Typically, **trust or trustworthiness** are just mentioned as the preferred attributes of the fair data economy or fair data label and not discussed further. However, some explain how they perceive trustworthiness, that is,

> *[Being] trustworthy is keeping promises and not amending the terms and conditions in the middle of everything. (FI Q23 ID271)*

Similar answers that highlight the keeping of promises often describe **reliable** action, although this term is frequently used without further context. In addition, **fair** and **honest** are often used as adjectives, but only rarely explained:

> *Knowing what exactly is being collected about you and how you can access it, then it's fair for me. (NE Q21 ID545)*

> *Obtained in an honest manner, meaning with my consent, and to be shared honestly, meaning with my consent and to pay me for it. (NE Q23 ID262)*

Ethical or moral data collection and use are mentioned multiple times. Ethical actions are seen both as a duty and a goal in the fair data economy:

> *To sell information ethically and not first and foremost to make money. (FR Q23 ID457)*

> *[F]or causes that are relevant to the users' moral values, causes that are ethical. (FR Q23 ID354)*

Appropriate data management is suggested multiple times, but this idea is rarely explained further. It seems that appropriate data management is seen as **respectful** use of data, **integrity** and **responsibility**:

> *With the greatest respect. Not for the services, apart from the usually obligatory need to give my data. (NE Q21 ID560)*

> *Rightful use with respect to the wishes of the person. (NE Q23 ID340)*

Fair treatment of employees is a surprisingly significant theme in the results, although it seems that some of these instances result from mixed-up perceptions about the fair data label and fair-trade label. Nevertheless, fair treatment of all parties, including the employees of the organisations, is an issue that should be considered:

> *No child labour, no wages that are abnormally low, employees being treated well. (NE Q23 ID223)*

> *Fair working relationships for staff regarding the usage/service. Involvement of the customer in profits due to data usage. (GER Q23 ID124)*

Some of the respondents note that collection and use of data should not harm them or those close to them. **No harm** is linked to exploitation, using the data against the people and non-intentional causing of harm. Some also link avoidance of harm to others and

respond that data or data use should not create inequality. **Equality** and equal rights between people and people and organisations are also mentioned a few times.

Trust and fairness, with all the subthemes, reflect the qualities that are seen as positive attributes of a fair data economy. However, as can be noted, conceptions of these subthemes are vague and demand more profound research. Nevertheless, **trust** is undoubtedly an enabler of a functioning data economy, and thus, valued in itself, as is **fairness**. These are values of universalism, benevolence and conformity. They could also be described as moral values, since they include a factor of 'shouldness' of behaviour. Thus, they could also be described as **justice –** a just behaviour.

5.5 Compensation or benefits for users

Compensation or benefits for users was the fifth most common theme, although it was considerably rarer than previous themes. The subthemes of this theme are presented in Figure 8.

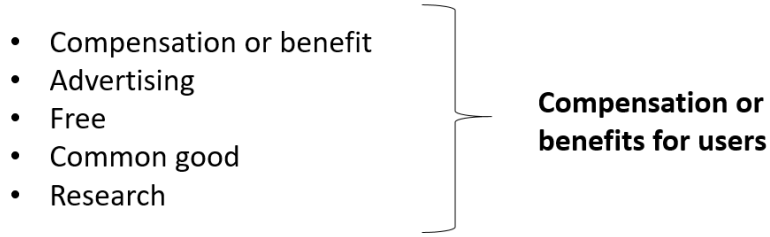


Figure 8 Themes related to compensation or benefits for users.

As can be noted from the subthemes, they vary from general compensation to more specific benefits. Respondents demanding compensation or benefits can roughly be divided into those demanding some compensation or benefit, those who see the current use of personal data as insufficient and those willing to give their data for use in service of the common good and benefits.

Respondents demanding **compensation and benefit** most often ask for some sort of benefit in general:

> *By compensating me, by offering me benefits to enjoy… (FR Q21 ID38)*

> *I am willing to release data if I am rewarded appropriately. (GER Q21 ID172)*

> *Data should be managed to know me better, and therefore, help improve my quality of life. (FR Q21 ID87)*

This general requirement for some sort of benefit is accompanied by demands for **sharing money** that is made with the data or **monetary compensation** when collecting the data:

> *I want to be informed every time MY data earns money, and additionally, I would like 50% of the money earned with my data transferred to my account! (GER Q21 ID196)*

> *The user should be offered compensation for the potential use of their data. (FR Q21 ID261)*

Targeted advertising is mentioned multiple times, often as a negative issue. It is seen as something that should not be done with personal data. Only one respondent wishes that the personal data would be used to make the advertisement more personalised, although he/she also feels that current practices are insufficient:

> *I would not accept use for advertising, although unfortunately, it is probably the most common use. (FI Q21 ID535)*

> *I benefit personally, for example, personalised offers and advertisements. While the targeting of advertisements is possible, I still get car and nappy advertisements even though I don't have a driver's licence or children. (FI Q21 ID233)*

It was equally observed that services considered to be fair are seen as **free of charge** and the fair data economy should not demand monetary investments from the users. However, using freeness as a way to justify collecting personal data and benefiting from it is not seen as sufficient practice:

> *That these service providers do not misuse my or my family's data. They would be reliable and free of charge and easy to use. (FI Q23 ID136)*

> *It is fine if money is earned with the data. It is not okay to make out you are offering a free service. (GER Q21 ID239)*

Some respondents link the fair data economy directly to the **common good** and state that it would be fair to use their data if they could benefit society and/or humankind. Some see the common good as the only fair way for benefiting from their data and have negative attitudes towards the commercial use of the data:

> *[S]uch that the society would benefit from it and I too would benefit from it even if in terms of knowledge. (FI Q21 ID 527)*
>
> *To make the world a better place but not for your own interests. (NE Q21 ID498)*
>
> *No commercial sales of our data. Only for general interest. (FR Q23 ID189)*

Although scientific **research** can be seen as one form of common good, there are multiple answers that conceptualise it as fair use of data:

> *Scientific and non-profit studies. (FI Q21 ID576)*
>
> *So that scientists have free access to it. (GER Q21 ID495)*

Demanding compensation or benefits clearly reflects values of **hedonism**, but at the same time, **benevolence** and **universalism**. There is also a strong connection to **self-direction** and the claim for **power** over data use, which links this theme to theme 1 (user's control over data and data sharing). Compensation or benefit to oneself is clearly an issue of what is seen as **just** – it is not permissible to collect and use data without users profiting from it in some manner as well. It is notable that this theme also connects to transparency. Organisations should communicate their intentions with data clearly and point out what the benefits are of sharing data for individuals and society.

5.6 Supervision and rules

Answers including demands for supervision and rules for the fair data economy formed the sixth biggest theme in the dataset. Similarly to the previous theme, these issues were more rarely mentioned than the first four themes were. The subthemes of this theme are presented in Figure 9.

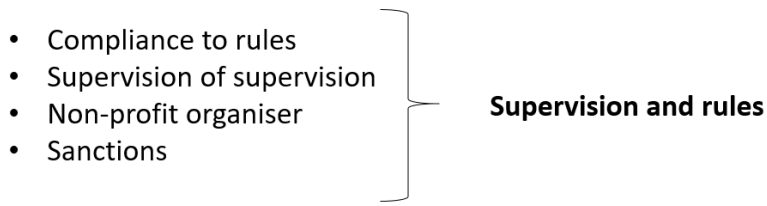


Figure 9 Themes related to supervision and rules.

Rules and the need for supervision as a theme includes subthemes that highlight the role of laws, policies and their supervision in the data economy. **Compliance with rules** mainly highlights the lawfulness with current laws and regulations, but it also relates to the need for new rules. However, most of the answers do not mention specific laws or what the rules should be. The most noted regulation is the GDPR:

> *Accessible for me, according to legal directives. (GER Q21 ID21)*

> *We have a law regarding personal data, and I assume there's an organisation responsible for checking it as well. They could perform this task. (NE Q23 ID526)*

As can be noted, people are not quite sure who supervises even current laws and regulations in practice. **Supervision** of the fair data economy has a twofold significance. On the one hand, people demand supervision of it, but on the other, they want supervision of the supervision:

> *They should be controlled regularly and seriously (not a label they can buy). (FR Q23 ID350)*

> *That it is being checked – whether they actually operate in the way the people whose personal data they are using want it to be done. (NE Q23 ID416)*

> *To have a central supervising organisation appointed by the government. It has branches that are specialised in data collection for healthcare, terrorism and other security, as well as statistics about education. These should supervise the private service providers that have to work according to strict guidelines. (NE Q21 ID482)*

Some also impose restrictions on the nature of the organiser and demand that the institute or organisation responsible for supervision and the data label should be a **non-profit organisation**. From the suggested institutions, governments are most often mentioned:

> *Checks by a government agency that is accountable to the politicians, for a customers' council, and in the public through the media. (NE Q21 ID68)*

> *Approved by an independent organisation that is being checked by impartial people: double-checking. (NE Q23 ID41)*

> *It should be controlled by governmental services. (FR Q23 ID171)*

Sanctions for misbehaviour in the fair data economy or for misuse of the fair data label are mentioned by only a few people. Most highlight compensation, but making the misbehaviour public is also seen as a form of sanction:

> *Hefty sanctions and compensation to users if it is found that misuse has taken place. (FI Q23 ID92)*

> *[A]nyone who messes up is out, and it is published. (GER Q23 ID7)*

Demands for rules, compliance and sanctions most clearly reflect values of **conformity** and **security**, although there is a link to **benevolence** through **responsibility** and **power** through **authority**. In addition, **justice** can be seen as an even stronger value in this theme than in the previous themes. It is notable how this theme overlaps with **transparency** and **trust –** the rules, their supervision by a trusted party and noted misbehaviours should be understandable and available to the public.

5.7 Negative attitudes towards data collection and the data economy

The last large theme is negative attitudes towards data collection and the data economy. The subthemes of this theme are presented in Figure 10.

Figure 10 Themes related to negative attitudes.

Negative attitudes in the respondent's answers are not as common as solution-centred answers, but they still form a clear theme. As stated above, many of the negative attitudes can be seen as distrust towards the current system.

Cynical attitudes are the most common subtheme of negative attitudes. These respondents are doubtful about the fairness of the data economy and/or about the effect of the fair data label. Most of these cynical answers are sceptical or directly pessimistic. Unfortunately, the respondents do not explain where their attitudes come from:

> *It's all empty words. Empty words. It's about nothing. It all doesn't happen so fairly and it will not become so in the future either. Strategic considerations combined with wanting to make a profit, and that's how it will likely remain. (NE Q23 ID203)*

> *I'm very much against this. Another organisation that will decide for me what is fair and good. It's like ISO 9001; it's about nothing. (NE Q23 ID 175)*

> *This is a sham anyway. Nobody is prevented from having lucrative trade with data if there is a profit in it for him. (GER Q23 ID479)*

Distrust and negativity are also targeted to data collection in general. Some of the respondents state that there **should not be any data collection or use**. Again, this attitude is rarely explained:

> *The best would be to not collect any data. (FI Q21 ID462)*

> *It would be fairest if you were to not collect and use any data at all. (GER Q21 ID262)*

The suggestion of a fair data label provokes some resistance. Some strongly express that there **should not be any label**, mostly because they perceive that labels are not efficient safeguards:

> *Labels are not worth anything. Either they are fake or they are simply bought by companies who can afford to. (FR Q23 ID177)*

These negative attitudes make visible some of the challenges of the fair data ecosystem. It is clear that these respondents do not trust organisations collecting data nor perceive that they could benefit from a mechanism like a fair data label. Such attitudes could reflect values of security and tradition, as well as self-direction and power, but with the limited amount of explicit reasoning in the answers, it is hard to say for sure. However, it should be noted that strong negative attitudes exist, and they challenge the idea of the fair data economy.

5.8 Summary of thematic analysis of values

To visualise values and their connections, we have located them in Schwartz's (1992, 2012) model (see Figure 11). The main finding is that transparency can be seen as a con-

necting value between different kinds of basic value categories. Thus, even if transparency is an instrumental value, it is necessary to ensure fulfilment of other values. An instrumental value is a value that is necessary for upholding other values, and thus, it gains its meaning through them. As an example, if we want to ensure that we achieve informed consent or trust, we need to make the data economy ecosystem visible and transparent or individuals cannot verify what we have done. If there is a lack of transparency, it can also increase the cynical attitude and lessen the autonomy of people, as they cannot know, and therefore cannot control, the information collected from them. Ultimately, it seems that the key factors for achieving a human-centric data economy – in line with human values – are transparency and open communication with people when developing it.

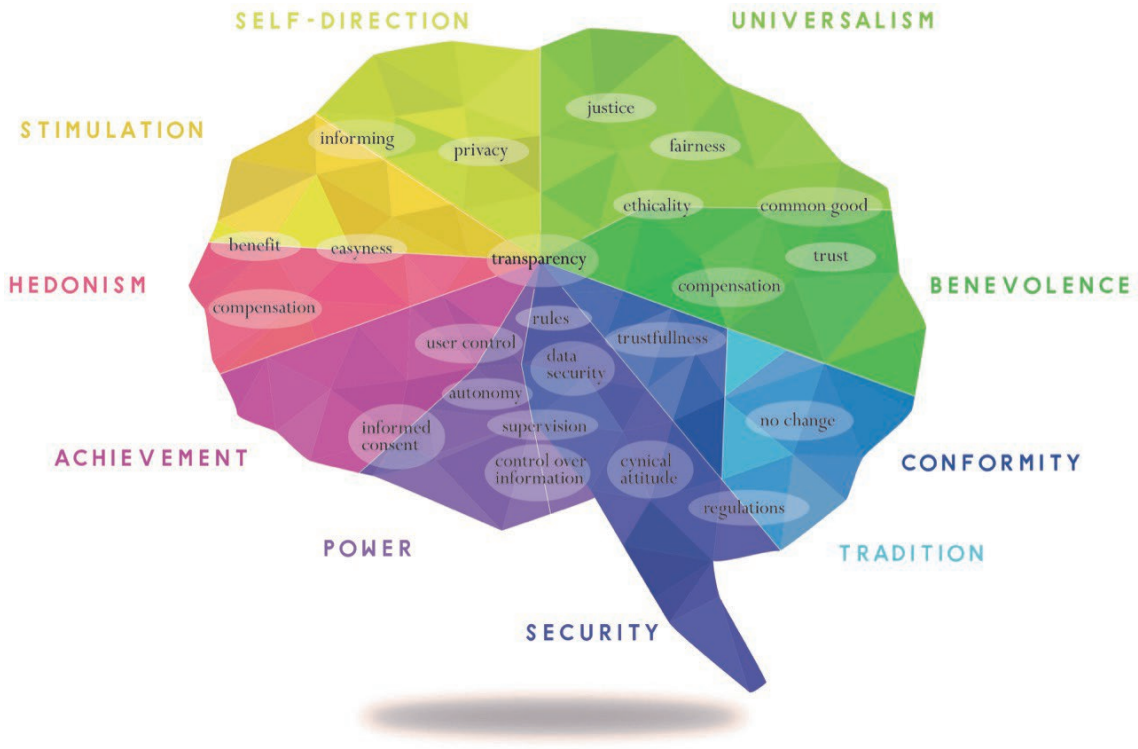


Figure 11 Values organised according to Schwartz's model of basic values.

6 DISCUSSION AND CONCLUSION

The four most commonly expressed values in the citizen survey (Q21 and Q23) were security, consent, transparency and anonymity. These can all be seen as preconditions/parts for having autonomy in the digital society. By autonomy, we mean the possibility and capacity to decide for oneself and pursue a course of action in one's life without outside pressure from factors/actors forcing one to take an action that would not be taken otherwise. Security is needed in the data economy ecosystem so that people can trust the ecosystem to be safe to use. The reason for security being mentioned could be that it has been visible in public discussion in different areas, so it has gained a status of de facto issue to be noted and guaranteed. Consent gives the right to make decisions on what people want to do with their information or to whom they will allow the right to use it. In some cases, anonymity is the only way to give people the opportunity to engage in self-expression, and thus, real autonomy is needed.

Transparency is an instrumental value that is required to create trust between all actors, and it did rise as the central issue for achieving and verifying the realisation of all other values. Together with open communication, it gives the possibility to ensure that people's values can be made visible. Hence, to ensure that people's values and the ethicalness of the fair data economy are reached, there is a need for rational discourse based on transparency and the possibility to affect how the data economy works and how it is governed. The rational discourse and consensus between all stakeholders need to form the basis of ecosystem governance if we truly want it to be fair. All stakeholders need to be given the opportunity to participate (in reality, there should be representation of all stakeholders) in the discourse, thereby ensuring that the data economy ecosystem would be just.

We also need different kinds of research instead of relying only on surveys, which lack the needed depth. A promising approach would be relying on the concept of discourse ethics and Yetim's (2006) metacommunication model to make the values, aims and views of different stakeholders visible. This could be done with a series of workshops where different stakeholders gather together to co-create the rules and guidelines of the fair data economy.

The fair data economy is an issue that has far-reaching consequences, and thus, it should have a strong basis that is ethically justified. Without this, it is possible that people will lose their trust of the data economy – which has already happened in some cases – and this could be a real barrier to gaining the advantages that the data economy can offer us.

REFERENCES

Collins, J. C., & Porras, J. I. (1994). *Built to last: Successful habits of visionary companies*. New York, NY: Harper Business.

Feldman, F. (1978). *Introductory ethics*. Upper Saddle River, NJ: Prentice-Hall.

Habermas, J. (1996). *Between facts and norms*. Cambridge, United Kingdom: Polity Press.

Lencioni, P.M. (2002, July). Make your values mean something. *Harvard Business Review*, 5–9.

Lyytinen, K. & Hirschheim, R. (1988). Information systems as rational discourse: An application of Habermas's theory of communicative action. *Scandinavian Journal of Management*, *4*(1), 19–30.

Ross, A. & Chiasson, M. (2011). Habermas and information systems research: New directions. *Information and Organization*, *21*(3), 123–141.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theory and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1–65). New York, NY: Academic Press.

Schwartz, S. H. (2012). An overview of the Schwartz theory of basic values. *Online Readings in Psychology and Culture*, *2*(1).

Yetim, F. (2006). Acting with genres: Discursive-ethical concepts for reflecting on and legitimating genres. *European Journal of Information Systems*, *15*(1), 54–69.