The Internet, Big Data & Platforms

Digital Transformation in Learning for Active Citizenship

By Nils-Eyk Zimmermann

dare
Democracy and Human Rights Education in Europe
BLUE LINES

The Internet, Big Data and Platforms.
Part of the reader
"Smart City, Smart Teaching: Understanding Digital Transformation in Teaching and Learning."

Authors: Nils-Eyk Zimmermann
With guest contributions of Viktor Mayer-Schönberger, Manuela Lenzen, Irights.Lab and José van Dijck and contributions of Elisa Rapetti and Marco Oberosler
Copy-editing: Katja Greeson
Design: Katharina Scholkmann (layout), Felix Kumpfe, Atelier Hurra (illustration)
Publisher: DARE – Democracy and Human Rights Education in Europe vzw., Brussels 2020
Editors of the series: Sulev Valdmaa, Nils-Eyk Zimmermann

The project DIGIT-AL – Digital Transformation in Adult Learning for Active Citizenship – is a European cooperation,
coordinated by Association of German Educational Organizations (AdB)
with
DARE – Democracy and Human Rights Education in Europe vzw. (BE)
Centre for International Cooperation CCI (IT)
Education Development Center (LV)
Jaan Tõnisson Institute (EE)
Partners Bulgaria Foundation (BG)
Rede Inducar (PT)

Supported by:

Federal Ministry for Family Affairs, Senior Citizens, Women and Youth

Co-funded by the Erasmus + Programme of the European Union

The project is supported in the framework of the Erasmus+ program of the European Commission (Strategic Partnership in the field of Adult Education). Project Number: 2019-1-DE02-KA204-006421
The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Preface:

Into Digital Transformation

The social, economic, cultural and political impact of digital change in education and learning

Digitalisation is an essential part of our lives across all dimensions. Many people think that it is a technological process, i.e. it is mainly about computer servers, algorithms, Internet and the like. But that is only half of the truth. For example, it is difficult to separate digitalisation from almost all activities in our lives. When we shop online – are we online or are we shopping? When we play computer games – are we playing or are we at the computer? And when we are active in social media, we are both social and active in an electronic medium. Moreover, our health system is already digitised, the pollution of the planet is, to a growing extent, caused by digital technology, and activities such as navigating a car or collaboration in civil society are increasingly facilitated by digital technology.

This example seeks to point out that what we ultimately understand by ”digitalisation” depends very much on how we look at the topic. It is after all possible to engage in all the aforementioned activities without information and communication technology (ICT). In this sense, we prefer the term *digital transformation*, because it explains a social, cultural or economic process in which things are done seemingly differently – made possible by information and communication technology. In this sense, education for digital transformation is learning about social, economic and cultural processes and about understanding the differences caused by technology. As such, in further exploring the topic, it is important to:

1. Look at both the technology and the nature of economic, social and cultural activities, for example, what we do in different social roles as digital customers, digital activists, digital workers and digital citizens.

2. Take an interest in the difference that digitalisation brings to such activities. What is changing thanks to new technology? What impact does it have on society?

There is No Overly Complex Issue for Education

A lot of curiosity and increasing concerns regarding digitalisation today have to do with its ‘engine room’ - the fascinating global infrastructure of the Internet, its enormous costs and hunger for energy, Big Data, AI, and the increasing economic value of digital platforms.

In particular, the growth of new kinds of platforms, fuelled by digital business models successfully capitalizing on users, is a widely visible phenomenon of this new technological and economic configuration. Consequently, their users are at the same time subjects and objects of digital change. They experience the opportunities made available through new, platform-mediated forms of interaction, but also feel uncomfortable since they are also symmetrically affected in their role as autonomous subjects. The right to independent information, privacy and security are, from this perspective, not yet sufficiently respected in the digital sphere.

The migration of substantial parts of working and communication processes to the digital sphere during the last decades is also simultaneously a benefit and a challenge. One aspect is technical mastery – access to current technology and the ability to use it in a competent way. A more fundamental aspect is that the “digital self” is completing people’s analogue identity. Their digital traces are accompanying people’s lives with related consequences for their various social roles as private subjects, employees and citizens.

Feeling overtaxed by all the associated challenges and concerns is a bad prerequisite for learning and a bad basis for considering future personal and social decisions. It is high time for adult education and youth work to do something about this double-edged sword.

In particular, adult citizenship education has a lot of experience teaching complex social issues and could transfer its methodology and approach to the topic of digital transformation. We know, for example, that nobody needs to be an economist to be able to co-decide on political decisions affecting the economy. We also are capable of understanding the social impact of cars, despite very limited knowledge of automotive engineering. Considering that it is possible to acquire knowledge about digital transformation, could we not even enjoy learning about Big Data, robotics, algorithms or the Internet of tomorrow similar to the way we passionately discuss political issues such as transport, ecology, or democracy? We should not, however, be blinded by the technical complexity of the digital transformation. It is important that we pay more attention to the social dimension, the intentions behind a technology, exploring its effects and regulations.

Although not familiar with all technical or legal details, most people intuit that it is ill-advised to give out personal information without consent. We suppose what the right to privacy should entail and what distinguishes conscious decisions from uninformed ones, and in our analogue world, we discourage the ”used car salesmen” of our society from taking unsuspecting customers for a ride. After all, most of us have experienced the discomfort of having been deceived as a result of not understanding the fine print.

If we transfer this insight to a pedagogy of digital transformation, we must admit that we should also be willing to explore new aspects of the technical dimension such as data processing or the nudging mechanisms in online platforms. But that is not the only priority! The most important thing is that we know what our *rights* and *ethical foundations* are and how they relate to the new digital contexts and are able to act accordingly. These questions are not solely related to privacy and safety, as seemingly no aspect of social life is unaffected by digital transformation.

Using this foundation, we might further explore the potentials and risks of digitalisation in context, assessing its impact. Personal rights, for instance, entail privacy issues, but digital transformation has also led to new opportunities for co-creating, better information, or involvement of citizens in decision-making processes. On this basis, we are then able to define the conditions and rules under which certain digital practices should be rolled-out or restricted.

E*lectronic* communication has changed the character of *human communication* as a whole. There are fewer impermanent ideas or assertions that go undocumented, to later be searched and rehashed. This change is both positive and negative, for example from the perspective of an employee who may be judged based on past decisions which live forever online. Pedagogy might help people to better understand the risks and benefits associated with electronic communication.

In addition, it will be a creative challenge to imagine the technology we want to develop as a society and what will help us to initiate social, economic and cultural changes in the future. In this regard, it is also important to develop a view towards the so-called ‘skill gaps’ and ‘digital gaps’ people may face when mastering digitalisation. What is the purpose of defining a gap; for whom is the gap relevant; in whose interest is it to argue the risk of gaps as opposed to their benefits?

Why Democracy and Rights-based Learning Makes the Difference

The essence of a definition of democracy and rights-based education can be found in the Council of Europe's Declaration regarding Education for Democratic Citizenship (EDC), which is "education, training, awareness-raising, information, practices, and activities which aim, by equipping learners with knowledge, skills and understanding and developing their attitudes and behaviour, to empower them to exercise and defend their democratic rights and responsibilities in society, to value diversity and to play an active part in democratic life, with a view to the promotion and protection of democracy and the rule of law" (CoE CM/Rec(2010)7).

Transferred to the context of learning about digital transformation, we extract three core questions from this:

1. *What digital transformation competence* – knowledge, skills, values and attitudes – do citizens need to understand the digital transformation in their society and how it affects them in their different social roles?

2. How are *fundamental rights and ethical foundations* related to the transformation? Where do they shift their nature, what weakens them and what kind of development strengthens their enforcement?

3. What *active civic competences* do citizens need to contribute to the transformation, including participation in relevant public discourses and decisions, self-organisation and social engagement, and the development of social innovations?

Stakeholders from many different sectors have high expectations in education. In particular, they demand from earning for active citizenship a better preparation of Europeans for big societal changes. Only if we implement ideals of democracy "by design" into digital progress will we create a *democratic* digital society.

Enjoy and Explore

This reader series aims to introduce selected key aspects of digital transformation to educators and teachers in formal, non-formal or informal education. Our perspective is *Education for Democratic Citizenship* and our main goal is to motivate you as educators in adult education and in youthwork or other education fields to dive into the topics connected to digital transformation with curiosity and critical thinking as well as ideas for educational action. In other words: Nobody has to adore technology, but it is definitely worthwhile to become more comfortable with it. Digital transformation is a reality and as such, in principle, relevant for any specific field of education, any subject, or pedagogy.

Together we might work on a broader understanding of what digital literacy is and explore as educators and learners in lifelong learning processes how it affects our lives. With a strong aspect of democracy and human rights in lifelong learning, we should lay the foundations for a democratic digital transformation and empower learners to find a constructive and active position in this transformation.

We aim to provide basic insights into some of the various aspects of digital transformation as a basis for further exploration. They tackle the digital-self, participation, the e-state, digital culture, media and journalism and the future of work and education. In each of the publications we also present our ideas as to how education might take up this specific topic.

You may access, read, copy, reassemble and distribute our information free of charge. Also, thanks to digital transformation (and the Erasmus+ program of the European Commission) we are able to publish it as an "Open Educational Resource" (OER) under a "Creative Commons License" (CC-BY-SA 4.0 International).

The Internet, Big Data, Platforms

Our digital transformation today is rooted in earlier digitalisation in different parts of society. In particular, the emergence of the non-centralised internet, globalisation, networked technology, technical advancement, new ways of networked collaboration and the vision of ubiquitous computing have abetted the transformation toward the dominant topics in discourse around digital transformation today. Topics like the platform economy, big data and artificial intelligence. But the Internet has also helped other ideas break through, in particular, new open and non-centralised models of creation, communication and collaboration. As a global infrastructure, there is also an environmental impact associated with the physical network of cables, satellites, data centres, and antennas. In this publication, we introduce some of these topics. In this context, we would like to thank the guest contributors. Viktor Mayer-Schönberger explains the concept of big data, Manuela Lenzen describes the emergence of AI and José von Dijck uses the metaphor of a tree to explore the concept of the platform economy, in order to make it more comprehensive to broader audiences.

1. From the Microchip Revolution to the Internet

From its very beginning, the narrative of digital transformation is that we are close to entering a new historical configuration. Progress is a leitmotiv in digital discourses. However, if we look back at the prospects and developments of the past, we can also learn about the developments leading us into a digital future. In his article on big data in this publication, Viktor Mayer-Schönberger writes, “in the context of big data, it is also possible to forecast the future based on analyses of past or present behaviour”. Therefore, let’s start with the evolution of digital transformation.

After the pioneering work on computers in the 1940s and with the microchip revolution of the 1960s, the binary technology found its way into different domains of social life. First, computers appeared in offices, moving beyond the military and science sectors. The invention of the integrated circuits and transistors resulted in plentiful affordable and available electronic devices enabling the use of information and computing technology (ICT) on a broad scale. Already then, people reflected on the opportunities (and threats) of technological progress for humanity. However, the ICT revolution seemed to be a manageable process. Digitalisation and mechanisation were perceived as a mainly positive development serving humanity to create social, economic and cultural progress. BBC presented “Tomorrow‘s World: Home Computer Terminal” in September 1967 to the audience:

“Industrial consultant Rex Malik feels the business world‘s pulse from his bedside. Stock prices and market trends are available to him through Europe‘s first home computer terminal. This terminal is linked to a giant brain ten miles away in the heart of London. It‘s one of two Malik has installed for experimental purposes, because he wants to know if they could run his life and his home. For him they‘re simple to operate and experts predict that in 20 years time all new houses will be built with special computer points and the terminals will be cheaper to rent than today’s telephones. There‘s no complicated language to master. Before he can understand what the computer is saying the unseen brain sends its messages in good old fashioned English. [...] “

The vision of connecting people and things in networks sparked the imagination of system designers and software developers. The first electronic communication networks were developed as the development of the computer as a device not limited to research institutions - state and big corporations were on the horizon. ALOHAnet was the first wireless network (1971). Robert Metcalfe invented the ethernet in 1973, connecting different devices by a standardized cable, for instance, computer terminals, servers, printers and other devices. Clear-sighted he postulated an exponential increase of the value of networks through their enlargement (which was later called the Metcalfe Law). Better technology plus more connection, the "knowledge society" or "information society" was waiting to be built up under such premises.

Having originated in a military network project in 1968 called ARPANET, the Internet has emerged step-by-step into what we know today. The first newsgroups, mailbox networks and emails connected people from their home computers via telephone line in the 1980s. The World Wide Web (WWW) was established in 1991, and over time powerful hardware and finally mobile devices and broadband connections became more widespread and available starting in the late 1990s and early 2000s. Around 1993/94, the Mosaic and Netscape Browsers launched amid a WWW boom. Suddenly, people were able to share their content and information about themselves, visible for all on webpages. In 1995, this was remarkable progress which needed to be explained to offline audiences: "Those able to process their documents with HTML might simply publish them worldwide. These opportunities make up the fascination of the WWW" (Die ZEIT, 1995). Users found already in the early Web a broad diversity of content, from funny things like a coffee pot camera (first webcam) to the first news sites and also information of some non-governmental organisations.

The Web was an emerging field for experiments - a borderless, shared and low-hierarchy communication space thanks to the openness of the technology, which was the mission of the World Wide Web Consortium (W3C) from the beginning. This new WWW age seemed to also form a new digital culture. De digitale Stad in Amsterdam or the Internationale Stadt Berlin were early innovative projects aiming to connect the vision of global citizenship with arts and local networking – McLuhan's "global village" found its expression here.

Also, the first e-commerce (eBay in 1995/96, Book Stacks Unlimited already in 1992) and first media websites went online during 1994/95. Later, with Web 2.0, the whole space became increasingly user-friendly and interactive. Peer platforms involved many people in sharing, communicating and downloading (the older readers will remember their first file download).

And where are we now? On the one hand, the progressive leitmotiv is still intact. The technical and economic opportunities are growing. The knowledge society is advancing. Ordinary people have access to networking in ways even the most privileged could not have imagined fifty years ago. We enjoy more accessible, more intuitive, cheaper and also more beautifully designed technology. We have gotten used to digital work and private communication, often across borders. Granny uses video calls and message services. The digital progress paradigm is still intact.

The bright picture of digital transformation can be symbolized by the presentation of new products. Microsoft set the standard for the celebration of soft- and hardware as "pop stars" with the presentation of Windows 95 by Jay Leno for a 2,500-person audience. Apple brought a new narrative into these celebrations and created the image of digital transformation decisively – clearly designed, simple, intuitive and even fashionable.

In your view, what impact do the most recent digital technologies currently have?

	Very positive	Somewhat positive	Negative overall	Very negative
Economy	23%	52%	13%	3%
Quality of life	17%	50%	18%	4%
Impact on society	15%	49%	25%	5%

Source: EUC-EB, 2017

However, over the longer run, this clean and positive surface has gotten its first scratches. While digitalisation was perceived as a project driven by heterogeneous visions and milieus in the initial decades, where big business, creatives and bottom-up visionaries co-created a new global culture in synergy, we discuss the Internet differently today. While in the early years, discussions about the Internet centred more around its global governance model and led to the establishment of the organisation ICANN (governing the internet domains), today, platforms and their power – big data and the digital economy – are currently dominating the narrative of digitalisation. Soon the internet also became a space for economic phantasy. Venture capital caused a first dot-com bubble that burst around 2003. Soon the digital market grew further, surviving the economic crisis in 2008. Today, the most profitable public corporations are technology conglomerates. "In effect, digital platforms have become systemically important in the digital economy, similar to the financial sector itself" (Nogared & Støstad, 2020, p.7).

Most valuable companies

Market capitalisation in 2020 (second half year)

Apple Inc.	$ 1.576.000.000.000
Microsoft	$ 1.551.000.000.000
Amazon.com	$ 1.432.590.000.000
Alphabet Inc.	$ 979.700.000.000
Facebook, Inc.	$ 675.690.000.000
Tencent	$ 620.920.000.000
Alibaba Group	$ 579.740.000.000
Berkshire Hathaway	$ 432.570.000.000
Visa	$ 412.710.000.000
Johnson & Johnson	$ 370.590.000.000

Source: Wikipedia: List of public corporations by market capitalization

Today, the digital transformation feels ambiguous for many people, based on the present expectations of Artificial Intelligence, but already prompted earlier by discussions on job rationalisation, "illegal" file sharing/music industry (early 2000s), data breaches (for example, the AOL leak 2006, Google street view 2007), WikiLeaks (in particular, the Afghanistan and Iraq files in 2010), "filter bubbles" (2011), the NSA leak (2013), and "fake news" (coined in 2016). It is a big dragon that is neither understood nor one we try to learn to ride. However, we seem fascinated by the creature, since it is an interesting, entertaining and helpful beast.

In contrast to the confidence of earlier decades more people feel that the speed of developments would make things complex and confusing. Many fear the domination of a technology-centred perspective in society and the power behind platforms and especially big data, raising privacy issues and problems regarding their autonomy.
Also, the growing desires of states and authoritarian regimes for control and surveillance give reasons for concern.

Whom do we trust? Whom do we see as competent, as guiding us through the transformation? Digital transformation, although a Europe-wide and global process, is

Big Data

Persons that see more disadvantages	Persons that see more advantages
51%	**31%**
Persons that would rather pay for a service, compared to paying nothing but giving their data in return:	Persons that would not pay:
55%	**39%**

8000 respondents in 8 EU countries

Source: Vodafone Institute for Society and Communications (2016)

embedded in very different civic cultures and governance contexts, which hinders our ability to give universal answers to these questions. Although in some countries, trust in government and the state seems to be relatively high, in others, this is lower. For instance, in Estonia electronic voting is accepted by many citizens, but less so in Germany and Italy – and for very different reasons. While in some states, the remembrance of negative experience with state surveillance is very present in the debate, in others the discourse is more dominated by the fear of the power and surveillance capacity of private platforms. Also in the judiciary, different traditions and perspectives exist, which are certainly not homogenized by the EU or CoE courts. The strength of the voice of critical civil society varies from country to country and relates to its ability to reach out to media and politics. In this sense, these examples illustrate that what might be important and relevant for some contexts might play a less important role in the discourses and decisions about digitalisation in other countries. Digital transformation is always embedded in a specific civic culture. Education needs to deal with specific contexts and visualize their relation to other European and non-European situations.

The Evolution Described with a Notebook

What has happened technologically since the microchip revolution? The developments can be illustrated through the notebook or address book. Some decades ago, everybody used a small book with their friends' telephone numbers. These books were very valued, which was reflected in their material quality. Gradually these were replaced by digital address books, for instance, in an email program or in contact lists in cell phones. Soon, less people bought address books.

In a next step, digital notebooks became connected, or "smart". Zuboff introduced this term in 1989 describing that computed information "renders events, objects, and processes that become visible, knowable, and shareable in a new way" (Zuboff, 2015, p. 76).

We were able to copy and paste mass entries from databases, automatically collecting numbers and email addresses of people we were in contact with and sending around emails to hundreds of recipients. Over time we forgot how to remember a telephone number, because it was saved automatically in our mobile phone (from today's perspective, we have to say "non-smart").

As the Internet started connecting all our devices and the smartphone became our central communication management tool, new opportunities were opened. Our notebook might now be migrated into a cloud which means, technically, from a client computer to a server, and can now be accessed through many different devices. It has become independent from the material place of storage which relieves us from fearing its physical loss. If your mobile phone is broken or stolen, access your notes or addresses from your cloud space simply by using a new device.

The coexistence of more and more apps and of more and more devices around us is putting the vision of ubiquitous computing into reality. Digitalisation pioneer Mark Weiser postulated in 1991, that a lot of our devices today would be more or less "invisible in fact as well as in metaphor" (Weiser, 1991). Our devices are small and intuitive, and we don't even recognize them as computers.

Their value lies on the one hand in their small size and intuitiveness, but their impact is their connection to servers, to other systems or to data processing. In an Internet of Everything, the machine is embedded in our social context and so are intelligent plug sockets, fridges, automotive body computer modules, factory robots, or home media centres. Also wearables (and even some implants) have "become social actors in a networked environment" (Spiekermann, 2010, p. 2). Out of these observations, one can draw a general pattern.

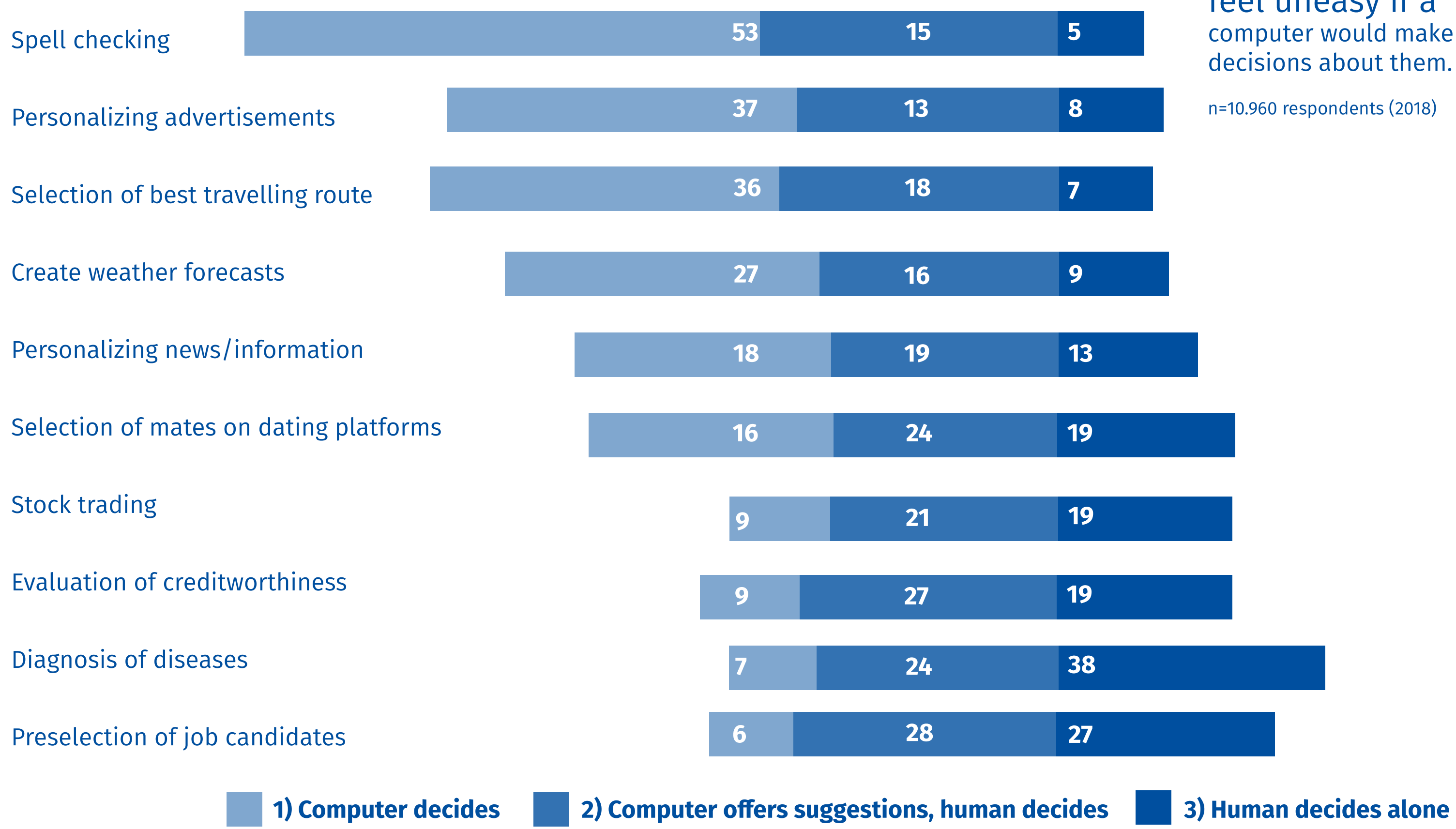
Ubiquitous Computing: A technological vision of many, often small, and very differently connected computing devices, deeply embedded in our daily routines, interacting intuitively with us and with each other.

Internet of Everything: Computed devices for different purposes, of different sizes and with different abilities interact with other devices (Internet of Things) and with the surrounding space through facility-installed technology (Smart Home) and their social environment.

Digitisation: Conversion or reinvention of analogue contents, products, functions or processes in order to process them with computers.

Acceptance of Algorithms in the EU

For which of the following tasks....
1)... would you find it acceptable for a computer to make decisions on its own?
2) ... would you find it acceptable for a computer to make suggestions, but only if a human makes the final descision?
3) ... do you think a human should decide alone without any suggestion from a computer?



64%
feel uneasy if a computer would make decisions about them.

n=10.960 respondents (2018)

Source: Grzymek, Puntschuh, Bertelsmann Foundation (2019)

The Evolutionary Pattern of Digitization

- Analogue practices undergo a digitisation.
- Digitised devices and services link to others. New devices beyond the server or desktop computer appear.
- Linkage and networking of devices, services and data in digital processes (recording, extracting, comparing, monitoring or analysis of data) enable new forms and business models.

Digitalisation draws its dynamics in particular out of the opportunities of the latter two.

Towards Datafication

When we say these devices are becoming active, this means they are actively generating and processing data. “Smartness” relies on a combination of

- many (different) ICT devices,
- mediated through networks or servers,
- which have not only storage but also data processing capacity.

Datafication: extracting personal data from user interaction, processing it digitally and turning it into (added) value.

With growing “smartness” the amount of data and the server capacity grows, which is needed to process all the data and manage cloud spaces. It is also now possible to merge different kinds of data. Although formerly, shopping data and other household expenses like rent, gas/water/electricity and bank transfers would have been documented in a household book, now they can be merged digitally, making a digital notebook more meaningful: Users not only gain a better overview but also a clearer picture thanks to built-in analysis and evaluation functions.

Further, these devices also create new data through their usage (such as metrics, location data, and metadata), allowing better analysis (for example, from where and how often somebody accessed their notes) or by attaching data to specific persons. For instance, a digital picture stores the copyright holder, the date and location it was taken and the camera information.

But how to make sense out of all the data? The more distinct information we have, the less an individual is able to process it. Now "big data" comes into play. This term describes an automated method of gaining insight on the basis of quantitative data by building statistical correlations and relations between a variety of data types using a massive amount of data. Big data could help users of a notebook to draw new conclusions – and also the owner of the big data servers and algorithms, the platforms, to draw conclusions about their customer. Even the analysis of different data which seems not to be in logical relation to each other might lead to valuable insights. If data from other persons were also to be available, this would be even better. For instance, could an analysis software conclude: "Bulgarian males between 30 and 40 searching for 'conspiracies' on the internet usually spend more time around later evening in a social network and also more frequently buy cook books." Maybe, this is not very interesting or valuable for the concrete user person, but it definitely is for the marketing of cooking books.

By processing different data, like information on nationality, gender, book sales and individual social network time, big data is modelling social reality through statistical approximation. This leads to the ability to forecast human behaviour (the user will be a cook book buyer) or to understand societal processes (if a lot of people from a certain location share critical remarks regarding governments, listen to death metal and send emails with links to critical news sites, this might lead to demonstrations), or even to intervene in these (when people that tend to listen to death metal receive dessert cook books for free, they share fewer critical articles).

An analogous forerunner of this way of thinking is perhaps the scoring of a person's creditworthiness, which is often consulted when credit decisions are made, or for rentals. Here, too, very different data are brought together. Information that is generally available – such as place of residence, gender or age – is combined with experiential data – such as payment discipline in certain neighbourhoods, in age groups or among genders. In

Big Data:
Method of gaining insight on the basis of quantitative data by building statistical correlations and relations (between a variety of data types and a massive amount of data). Facilitated by algorithmic computing. Via modelling social reality through statistical approximation, the fundamental aim of big data is to forecast human behaviour, to understand societal processes, or to influence human activities.

Algorithm:
A set of computational rules and steps proceeding data with the purpose of extracting information out of it or triggering action.

Platforms:
Digital infrastructures that facilitate and shape personalised interactions among endusers and complementers, organised through the systematic collection, algorithmic processing, monetisation, and circulation of data.
(Poell et al., 2019, p. 3).

addition, personal data helps to narrow down this general picture more precisely, such as one‘s own payment behaviour, family situation or profession.

The example from online commerce also makes us think about how attractive this type of data processing can be and already is in very different application areas. Data of many people or many data of one individual can help insurance companies to calculate or even control their risks or retailers to tailor their offers and customer service. Human activity can be analysed and controlled more precisely, for example at work, in traffic jams, in social media, for monitoring places or in many other areas of society.

A public scenario for the use of such technology is the delivery of public services or the maintenance and management of public infrastructure. The European strategy for data of the EU explains some of the examples of use: “Data is created by society and can serve to combat emergencies, such as floods and wildfires, to ensure that people can live longer and healthier lives, to improve public services, and to tackle environmental degradation and climate change, and, where necessary and proportionate, to ensure more efficient fight against crime” (EU COM 2020/66 final).

Ubiquitous computing led to an inflation of data, often very personal data like fitness or other body data. Now one might also use devices for more intimately, documenting personal moods and thoughts. The body data collected by a Smart Watch could complement the information. The collected information about temperature, pulse, heart frequency or movement would give a person a better overview of when and under what circumstances they were extraordinarily active. It would be possible that a medical app is not only documenting, but also monitoring. For instance, it could nudge the user when it is time to have a break from sitting or send a signal when people need medication. Comparing this individual data with the data of others would set the individual experience in a social context, informing people about social standards, aberrations or norms.

These situations bring us to digital transformations’ most critical point – the autonomy and privacy of individuals which is potentially affected when external parties such as a platform or the state, process personal data and analyse behaviour. Users and platforms create not only personal data traces or data shadows, but also digital selves, the presence of individuals in the digital sphere which goes far beyond a mere extension of their analogue appearance. The question for individual users is how they will manage it. The brochure “The Digital Self” dives deeper into these aspects. Moreover, how might they make a claim for their human rights, which are connected to this appearance or identity – such as privacy, integrity, free expression, and others? The idea of ubiquitous computing is risky by default, since it is necessarily connecting first, second and third parties, and also sharing, storing and processing data vividly and, for their users, chaotically. In addition, individuals feel at a disadvantage before the authority of algorithmic systems. This is because they often neither know how a decision or evaluation has come about nor can they have it revised, similar to the credit scoring mentioned above.

Degree of concern about third parties accessing personal information shared online

Criminals/fraudsters	55%
Advertisers/ businesses	31%
Foreign governments	30%
Your country's secret services/intelligence services	26%
Government	20%
Law enforcement agencies	17%
Your (or any potential) employer	17%

Survey in the EU 27

Source: FRA 2020

Value-Centred Development and Control

Intuitiveness of usage is too often connected with lacking overview or control by the people affected by datafication. Conversely, those who have the technical possibilities and the algorithms are gaining influence.

Security, overview and transparency are from this perspective very important aspects for human rights-sensitive regulations. With every new development of the digital sphere, these rights need again to be made relevant and tangible – creating *strong digital human rights* and their enforcement.

Algorithmic computing and AI as the underlying technology of big data are opening new opportunities to communication, collaboration, insight and work, but also are potential dangers and may harm people's ability to communicate, collaborate and work (FRA, 2018). All the benefits or harm of technological development is dependent on how technology is implemented and regulated.

Since algorithmic models are human constructs, it is evident that they are following human assumptions. In this sense, they are not neutral. "A model's blind spots reflect the judgments and priorities of its creators" (O'Neil, 2017, p. 33).

In particular, minority communities complain about biased design of technology and being discriminated against by unfair algorithms. O'Neil is giving examples for such biased or even partial algorithms, for instance in university rankings or in police work. The AI white paper of the EU Commission mentions in particular biometry and AI, "the use of AI applications for recruitment processes as well as in situations impacting workers' rights" (for example, performance tracking) as very risky technology (EU COM 2020/65 final). It discusses its governance under strict regulatory limitation.

Hardware is similar. For example, camera sensors had problems with dark tones which was an issue in the past for portrait photography of people with a darker skin colour. Today the issue is perpetuated in some facial recognition systems, which less reliably recognize a variety of skin tones. The challenge is also to design processes, software and hardware according to democratic values and also to invest in the value-related education of ICT specialists. But according to what criteria? One answer gives the model of Value Sensitive Design (Friedman, Kahn, Borning, 2006): Human welfare, ownership and property, privacy, freedom from bias, universal usability, trust, fairness, autonomy, informed consent, accountability, courtesy, identity, calmness, and environmental sustainability. Since 2006, the approach has been broadened and further developed.

Also on a political level, regulation and policy creation with stronger ethical perspectives has gained importance in recent years. The General Data Protection Regulation (GDPR) from 2016 is the central element of the EU data protection law (EP, EC Regulation 2016/679). As a directive, it is superordinate to the national legislation. When it comes into force, the new Digital Service Act will provide the rules of the game on the digital market. Also, in regard to AI, an Independent High-Level Expert Group on Artificial Intelligence was set up by the EU commission exploring Ethics Guidelines for Trustworthy Artificial Intelligence (IHLEG 2019).

A Vision for Europe

In 2020, digital transformation arrived at a stage, where platform economy, AI and big data have been mainstreamed and become the central pillar of the digital economy. National governments and the EU have big expectations toward this development. Compared to 2018, it aims to double the number of data professionals to 10.9 million people by 2025 and nearly triple the value of the EU 27 data economy to €829 billion, which amounts to 5.8% of the EU's GDP (EUC-2020-02-Factsheet).

The EU is aiming to become a global leader in an ethical AI and big data approach, as expressed in the EU Commission's Artificial Intelligence for Europe (EUC COM(2018) 237 final), the Europan data strategy (EU COM 2020/66 final) and the 2020 White Paper on Artificial Intelligence (in continuation of the 2014 Digital Agenda). The white paper is a non-legal but important document of the EU Commission in order to present and discuss its strategy: "The enormous volume of new data yet to be generated constitutes an opportunity for Europe to position itself at the forefront of the data and AI transformation. Promoting responsible data management practices and compliance of data with the FAIR principles will contribute to build trust and ensure re-usability of data" (EU COM 2020/65 final, p. 8).

FAIR Data

Findable	(i. e. identifiable, including metadata, searchable)
Accessible	(i. e. retrievable metadata, open and free protocols)
Interoperable	(i. e. possible data exchange and re-use, open and established standards)
Re-usable	(i. e. clear and enabling licences)

Source: Wilkinson et al. (2016)

The European data strategy is led by the vision of a balanced "European way": "In order to release Europe's potential we have to find our European way, balancing the flow and wide use of data, while preserving high privacy, security, safety and ethical standards". In particular this vision builds on a "single European data space" (EU COM 2020/66 final).

The position of the network European Digital Rights (EDRi) in regard to an update of the European Digital Service Act lets an alternative vision shine through, the re-introduction of the old ideas of a decentralised and diverse internet ecosystem: "What is more, the DSA can stimulate the plurality and diversity of the online ecosystem with the emergence of new providers and real alternative services and business models by lowering barriers to enter the market and regulating some of the most toxic activities of the currently dominant platforms" (EDRi 2020).

Their counterparts, the industry lobby organisation DIGITALEUROPE advocates for a market-friendly governance: "Creating Common European data spaces would support the objective of making more data available for AI applications to thrive. It is however important to ensure that the development of such data space schemes is based on a robust and market-friendly governance framework, ensuring voluntary participation to the schemes" (DIGITALEUROPE 2020).

Looking back on the visions of the past while keeping an eye on the present demands and requests of the various interest groups and their proposals is opening the opportunity to explore our visions of the digital future tomorrow.

It is, in the end, the citizens deciding on what idea of digital transformation they would like to follow. A basic condition is foundational understanding. Therefore, the following chapters introduce some of the key concepts behind the digital transformation which were already mentioned in this short introduction. The infrastructure of the internet, big data, platforms, standards and openness, and artificial intelligence.

Conclusions for Education

One access point to discuss digital transformation is the vision behind the transformation process. We introduced already some possible visions: Overcoming the border between technology and real life, global connectedness, unlimited access to culture (including media, movie, texts), co-creating a peer-to-peer culture, extracting (economic or social) growth and value out of data, individualised society, automated (robotized) society, and control.

These are very different **cultural, social, economic and political ideas** and therefore are also inscribed in different political programs, personal attitudes, advocacy agendas and business models. Since digitalisation or digital transformation are umbrella terms for this diversity, learners might explore their vision for transformation.

Also the **evolutionary perspective** might help learners to reflect developments and to explore their picture of future digitalisation. For instance, along their individual internet or technology biography: When did they get into touch with what kind of technology? What did they do then? What were key events in their lives? What has changed for them personally? Where were hopes and hypes, and also threats and disappointments?

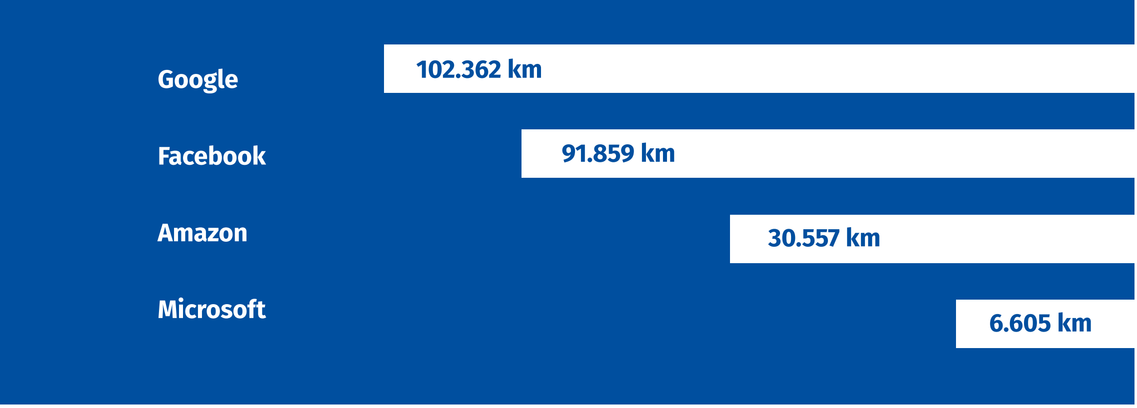
The factual digitalisation does not always embody a consciousness about it. Learners might raise awareness about their **individual connectedness** within the digital sphere. What kind of devices do they use? How do these interact, with whom? How do they work, or what do they know about them?

Since algorithmic computing and platforms rely on prod-users, users of platforms and producers of content or interaction in one person, nearly all adult learners have had **experience** with big data, artificial intelligence, scoring/rating or algorithmic selection or filtering. Where, and what kind of experience? A concrete individual reflection can be a motivational driver to enter learning about these abstract concepts and the technology.

2. The Machine Room behind the Internet

Digital transformation relies on the necessary infrastructure in form of cables, networks, data centres and electricity. The ambitions in regard to the Internet of Everything, autonomous vehicles or smart infrastructure require bandwidth and fast data transport. Governments, researchers, and ICT companies are currently working feverishly on the development and expansion of the new mobile phone standard 5G, and on the necessary networks that meet these requirements. This is connected with huge investments. According to a report to the European Parliament, for the EU “it will cost €500 billion to meet its 2025 connectivity targets, which includes 5G coverage in all urban areas” (European Parliament 2019). Google alone already owns a lot of servers and machines, seemingly running 2.5 million machines globally (Strickland & Donovan, 2020). The tech giants are accelerating their efforts toward the material backbone of the Internet, for instance by investing in submarine cables (BroadbandNow) or in low-orbital satellites like the project Starlink, by the company Space X.

Submarine Cable Kilometres owned By...



Source: BroadbandNow

In times of climate change, another global aspect is *energy hunger*. The efficiency gains by technological advancement are insufficient in compensating the growing need for electricity. Although the global platforms are moving toward renewable energy, the rebound effect lets us consider, where to twiddle the necessary knobs (Greenpeace, 2017). According to Greenpeace, by 2030, 13% of the global electricity will go to data centres.

IT Sector Electricity Consumption

7% of global electricity in 2017

13% of global electricity for data centres in 2030

Source: Greenpeace, 2017

In line with efforts to create a carbon free energy future, digital practices might also be better reflected. For instance, video streaming is clearly identified as an activity with a huge potential for reduction. Will we experience again a way back to "old fashioned" downloads, and will there be incentives to reduce consumptive internet traffic? The current developments show us moving in the opposite direction, rather mainstreaming the model of streaming digital entertainment by investing in the necessary server power.

Today mobile internet is a crucial condition for digitalisation. 75% of the EU population was connected to mobile internet in 2019, which is a clear growth compared to 2012 (36%). In Norway and Sweden, with 93%, most people are connected (Eurostat, TIN00083). If one includes non-mobile devices in this calculation, an even higher 85% of Europeans were online (Eurostat, TIN00028). 394 million Europeans currently use smartphones and 83% of all mobile connections are via smartphone. (GSMA, 2018). But also many other types of electronic and digital devices are continuously attached to us, such as smart watches, hearing aids and pacemakers. The trend is moving toward many devices per capita; According to Cisco, we can expect there will be 9.4 devices per capita in Western Europe and 4 in Eastern Europe by 2023 (Cisco, 2020).

Also the *affordability* of hardware is a core condition for digital transformation. For us in Europe it is constantly increasing, similar to North America. For instance, for the price of an Apple II in 1977, we can buy more than 3

Video

Video streaming is a tremendous driver of data demand, with 63% of global internet traffic in 2015, and is projected to reach

80% of global internet traffic by 2020.

It is necessary to spend

5h writing and sending emails without stopping

(i.e. 100 short emails and an attached document of 1 Megabytes) to generate an electricity consumption analogous to that generated by watching a

10-minute video.

Source The Shift Project, 2019, p.33

solid business laptops from Lenovo (T490 s) or 6 to 8 simple consumer notebooks today (USA Today, 2018). However, in other regions of the world, smart and mobile technology is still a luxury, leaving many people behind in digital participation. Although it is an extreme example, in Sierra Leone, one has to work an average of half a year in order to buy the cheapest locally available smartphone. In India, it would take two months (Alliance for Affordable Internet, 2020). Internet costs are similar. In African countries, 1GB data is 7.12% of the average monthly salary. In order to put it into relation: "If the average US earner paid 7.12% of their income for access, 1GB data would cost USD $373 per month!" (Alliance for Affordable Internet, 2019).

Since digital access, which also implies access to digital education, is unequally distributed globally, we need to consider the global dimension more consciously in our reasoning on the social, political, cultural and economic impact of digital transformation. The Alliance for Affordable Internet advocates for a mixture of stimulating competition among broadband providers in these markets, more state-investment in network infrastructure, and also facilitating complementary public internet access points.

Global Digital Divide: Access to Internet

	2018	2023
Central/Eastern Europe	65%	78%
Western Europe	82%	87%
Middle /East Africa	24%	35%

Source: Cisco Annual Internet Report 2018-2023

Individuals using mobile devices to access the internet on the move - % of individuals aged 16 to 74.

	2012	2015	2017	2019
EU 28	**36**	**57**	**65**	**75**
Belgium	44	69	75	86
Bulgaria	13	38	56	64
Czechia		45	60	73
Denmark	61	78	83	92
Germany	31	63	75	77
Estonia	37	61	68	78
Ireland	51	69	75	84
Greece	23	44	53	63
Spain	38	67	78	87
France	43	61	68	81
Croatia	38	50	51	72
Italy	16	26	32	50
Cyprus	25	59	70	79
Latvia	25	44	57	67
Lithuania	17	38	55	70
Luxembourg	63	80	82	86
Hungary	18	52	62	72
Malta	40	64	72	76
Netherlands	55	76	87	89
Austria	45	64	74	82
Poland	22	44	40	59
Portugal	21	45	58	63
Romania	7	38	53	70
Slovenia	30	51	63	76
Slovakia	38	54	64	71
Finland	56	73	79	
Sweden	70	77	87	**93**
UK	63	79	84	88
Norway	75	83	87	**93**

Source: Eurostat, TIN00083

Decreasing Prices of Computers

1977 **Apple II - $5,389** original price: $1,298

1985 **Commodore Amiga 1000 - $3,028** original price $1,295

1999 **Compaq ProSignia 330 - $4,076** original price: $2,699

2020 **Lenovo Thinkpad T490 - 1.500,00 €**

Source: USA Today (2018/06/22)

Others Work Harder for their Device

Cheapest available smartphone, share of monthly income

Sierra Leone	636%
Burundi	221%
India	205%
Niger	189%
Central African Republic	122%

Source: USA Today (2018/06/22)

Furthermore, the method of producing devices (including ever cheaper smartphones, tablets, TVs or digital notebooks) is leading to reduced life cycles and less willingness to repair or reuse devices. Prices which do not factor in the social and ecological costs of increasing ICT consumption, obsolescence by design, complicated repairability or lacking software support are pushing consumers to buy new products more often.

Modular solutions, like old desktop computers which allowed owners to replace or renew parts, are extinct. The website Ifixit is empowering consumers to repair their devices with published guides and advocacy for a "right to repair", according to the provocative question: "Would you buy a car if it was illegal to replace the tires?" Beyond repair, refurbishing is also still a niche. Some resellers are offering checked and repaired hardware (often the longer lasting business hardware). However, there is a global demand for used and refurbished mobile phones of the top brands and models, going against the intentions of some producers like Apple, which is lobbying against a right to repair. Still, they are not able to deplete the small independent repair shops completely.

Others try to unlock devices from their outdated, no longer maintained operating systems (rooting) to install free operating systems. The Free Software Foundation for instance is giving hints as to how to install a free Android operating system on smartphones and tablets (for instance LineageOS).

The socio-political conception leading toward more sustainability and conscious use of resources is the circular economy. The EU is pushing it currently forward in the framework of its Green New Deal. In particular the EU Commission aims to come up with regulatory measures "for electronics and ICT including mobile phones, tablets and laptops under the Ecodesign Directive so that devices are designed for energy efficiency and durability, repairability, upgradability, maintenance, reuse and recycling." It proposes also "to work toward establishing a new 'right to repair'" and a Circular Electronics Initiative (EUC COM(2020) 98 final).

This is also possible because most devices in 2020 are produced in China and Vietnam, countries with low wages, and because the necessary raw materials (rare earth) often come from conflict regions. A circular economy approach softens the negative environmental and social effects of raw material exploitation. The demand for raw materials is constantly increasing and globally they are unequally distributed. In their report,

Circular Economy:
The value of products and materials is maintained for as long as possible. Waste and resource use are minimised, and when a product reaches the end of its life, it is used again to create further value.
(EUC DG GROW).

Smartphone Production

Producing a smartphone of 140 g demands about 700 MJ of primary energy.
Producing a smartphone generates in France

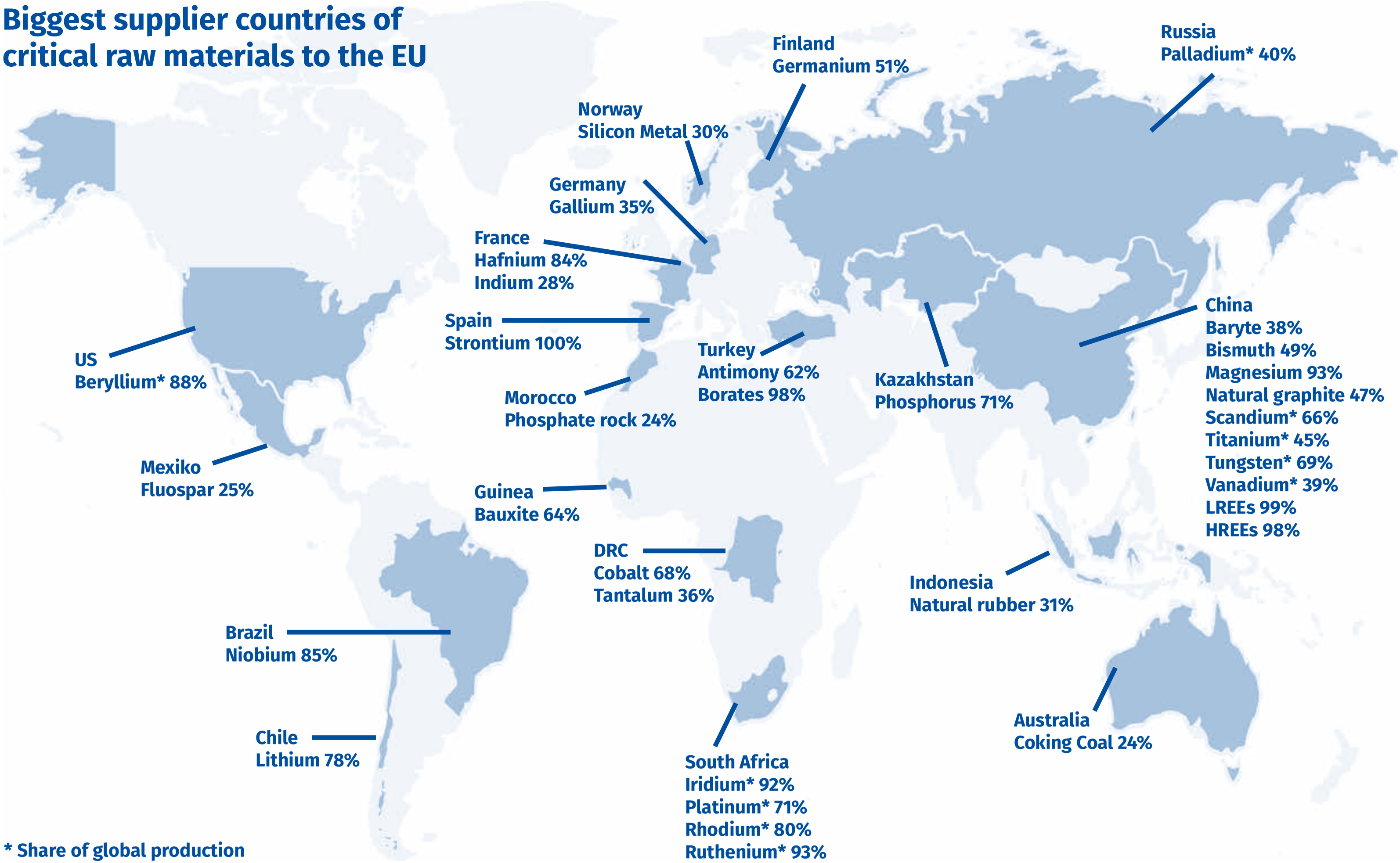
400 times more emissions than its utilisation.

If a person uses a smartphone from the age of 10 to the age of 80 and it is replaced every two years, the result is the equivalent of

200,000 km travelled by train.

The Shift Project, 2019, p.30

Biggest supplier countries of critical raw materials to the EU



* Share of global production

Source: EU-COM 2020/474 final

"Critical Raw Materials Resilience: Charting a Path towards greater Security and Sustainability", the EU explores the European raw material dependence, with a view on global demand, concluding that "despite improvements in materials intensity and resource efficiency" still 110% more raw materials need to be exploited in 2060 compared to 2011 and by a total of 167 billion tons (EU-COM 2020/474 final, p. 5).

Today, initiatives for fair trade or conflict-free IT aiming to strengthen the position of workers involved in the manufacturing process and also the position of production societies in world trade have not yet made a significant impact, although some initiatives like the project, Make ICT Fair (engaging for more fair public procurement policies) or Fairphone, are raising awareness about the production conditions of hardware. But in general, a fair European approach to "Critical Raw Materials Resilience" would need to prove that ethical words and fair global cooperation are a priority of European policies and economic practices.

The Shift Project, 2019, p. 60

Consuming 1 € of digital technology induces direct and indirect energy consumption 37% higher than what it was in 2010.

This trend is the exact opposite of what is generally attributed to digital technology and runs counter to the objectives of energy and climatic decoupling set by the Paris Agreement.

The Shift Project, 2019, p. 32

Smartphone Consumption

During the lifecycle of a smartphone, it is factually 33 times more energy consuming than its direct annual electricity consumption.

Conclusions for Education

The Internet and the digital transformation as a whole effect the whole world, but in different ways. Digital transformation can be explored as a phenomenon of globalisation, also included in global competence learning, for example in line with the global competence framework of OECD PISA:

"Global competence is the capacity to examine local, global and intercultural issues, to understand and appreciate the perspectives and world views of others, to engage in open, appropriate and effective interactions with people from different cultures, and to act for collective well-being and sustainable development" (OECD PISA, 2018).

Among others, accessibility, affordability and ownership of the infrastructure necessary for the digital transformation must be reflected. While we usually address the topic of inequality between influential global platforms and their users, the global imbalance also needs to be considered. Manyfold dependencies - economic, cultural and political - exist and are increasing (for instance, foreign investment in digital infrastructure, limited digital sovereignty, and cultural biases).

Since Europe is part of the global internet and if Europe wants to create a "European way" to digitalisation, Europeans also need to ask what kind of global vision they share and what global responsibilities evolve from this ambition – what would the European way be abroad?

Other topics also confront relevant aspects of digital transformation. Education for Sustainable Development (ESD) tackles relevant aspects in its various ESD goals (UNESCO 2017). Environmental education also covers topics like energy consumption, circular economy, sustainable production, repairability or recycling, and needs to be extended to the context of digital transformation.

In 2020, Europe's educational sector, from schools to non-governmental organisations offering non-formal learning to adult audiences, has experienced acutely the difficulties connected with lacking broadband (Wi-Fi) access, unequal affordability of hardware for students, a lack of reliable privacy-sensitive servers and software, and lacking digital competence referring to infrastructural aspects. Despite facilitating knowledge about the backbone and the material foundation of the internet, the sector has been called to action and investment.

3. What is Big Data? Accelerating the Human Cognitive Process

By Viktor Mayer-Schönberger, professor of internet governance and regulation at the Oxford Internet Institute

Using Internet searches to predict the spread of flu; predicting damage to aircraft engine components; determining inflation rates in real-time; catching potential criminals before they even commit the crime: The promises of big data are as astounding as they are complex. Already, an army of service providers have specialized in providing us with big data's „benefits" - or competently protecting us from them. A lot of money will be made based on this advice, but what big data is exactly remains largely unclear.

Many may intuitively equate the term „big data" with huge amounts of data to be analysed. It is undoubtedly true that the absolute amount of data in the world has increased dramatically over the past decades. The best available estimate assumes that the total amount of data has increased a hundredfold in the two decades from 1987 to 2007. [1] By way of comparison, historian Elisabeth Eisenstein writes that in the first five decades after Johannes Gutenberg invented a movable-type printing system, the amount of books in the world roughly doubled. [2] And the increase in data is not letting up; at present, the amount of data in the world is supposed to double at least every two years. [3] A common idea is that the increase in the quantity of data will at some point lead to improved quality. However, it seems doubtful that an increase in quantity of data alone will lead to the big data phenomenon that is expected to profoundly change our economy and society. [...]

The fundamental characteristics of big data may become clearer if we understand that it allows us to gain new insights into reality. Big data is therefore less a new technology than a new, or at least significantly improved, method of gaining knowledge. Big data is associated with the hope that we will understand the world better – and make better decisions based on this understanding. By extrapolating the past and present, we expect to be able to make better predictions about the future. But why does big data improve human insight?

Relatively More Data

In the future, we will collect and evaluate considerably more data relative to the phenomenon we want to understand and the questions we want to answer. It is not a question of the absolute volume of data, but of its relative size. People have always tried to explain the world by observing it, and as a result, the collection and evaluation of data is deeply connected with human knowledge. But this work of collecting and analysing data has always involved a great deal of time and expense. Consequently, we have developed methods and procedures, structures and institutions that were designed to get by with as little data as possible.

In principle, this makes sense when few data points available, but it has also led to terrible mistakes in some cases. Random sampling as a proven method for drawing conclusions with relatively few data points has been available to us for less than a century. Its use has brought about great progress, from quality control in industrial production to robust opinion polls on social issues, but random sampling remains a Band-Aid solution, lacking the density of detail needed to comprehensively depict the underlying phenomenon. Thus, our knowledge based on samples inevitably lacks detail. Typically, using random samples only allows us to answer questions that we had in mind from the very beginning, so knowledge generated from samples is at best a confirmation or refutation of a previously formulated hypothesis. However, if handling data becomes drastically easier with time, we will more often be able to collect and evaluate a full set of data related to the phenomenon we want to study. Moreover, because we will have an almost complete set of data, we will be able to analyse it at any level of detail desired. Most importantly, we will be able to use the data as inspiration for new hypotheses that can be evaluated more often and without having to collect new data.

The following example makes this idea clear: Google can predict the spread of flu using queries entered into its search engine. The idea being that people usually seek information about the flu when they themselves or people close to them are affected by it. A corresponding analysis of search queries and historical flu data over five years did indeed find a correlation [4]. This involved the automated evaluation of 50 million different search terms and 450 million combinations of terms; in other words, almost half a billion concrete hypotheses were generated and evaluated on the basis of the data in order to select not just one, but the most appropriate hypothesis. And because Google stored not only the search queries and their date but also where the query came from, it was ultimately possible to derive geographically differentiated predictions about the probable spread of the flu [5].

In a much-discussed article from several years ago, the then editor-in-chief of Wired, Chris Anderson, argued that the automated development of hypotheses made human theory-building superfluous [6]. He soon revised his opinion, because as much as big data is able to accelerate the process of cognition in the parametric generation of hypotheses, abstract theories are not very successful. Humans therefore remain at the centre of knowledge creation. Consequently, the results of every big data analysis are interwoven with human theories and thus, also with their corresponding weaknesses and shortcomings. So even the best big-data analysis cannot free us from resulting possible distortions [7]. In summary, big data not only confirms preconceived hypotheses, but also automatically generates and evaluates new hypotheses, accelerating the cognitive process.

On Quantity and Quality

When little data is available, special care must be taken to ensure that the data points collected accurately reflect reality, because any measurement error can falsify the result. This is particularly serious if all data come from a single instrument that is measuring falsely. With big data, on the other hand, there are large collections of data that can be technically combined relatively easily. With so many more data points, measurement errors for one or a handful of data points are much less significant. And if the data come from different sources, the probability of a systematic error decreases.

At the same time, more data from very different sources leads to new potential problems. For example, different data sets may measure reality with different error rates or even depict different aspects of reality, making them not directly comparable. If we were to disregard that and subject them to a joint analysis anyway, we would be comparing apples with oranges. This makes it clear that neither a highly accurate, small amount of data points nor a diversely-sourced, large amount of data are superior to the other. Instead, in the context of big data, we are much more often faced with with a trade-off when selecting data. Until now, this goal conflict has rarely arisen as the high cost of collection and evaluation mean we typically collect little data. Over time, this has led to a general focus on data quality.

To illustrate this, in the late 1980s, researchers at IBM experimented with a new approach to automated machine translation of texts from one language to another. The idea was to statistically determine which word of one language is translated into a specific word of another language. This required a training text that was available to researchers in the form of the official minutes of the Canadian Parliament in the two official languages, English and French. The result was astonishingly good, but could hardly be improved upon subsequently. A decade later, Google did something similar using all the multilingual texts from the Internet that could be found, regardless of the quality of the translations. Despite the very different — and on average probably lower — quality of the translations, the huge amount of data produced a much better result than IBM had achieved with less but higher quality data.

The End of Causal Monopolies

Common big data analyses identify statistical correlations in the data sets that indicate relationships. At best, they explain what is happening, but not why. This is often unsatisfactory for us, as humans typically understand the world as a chain of causes and effects.

Daniel Kahneman, Nobel Prize winner for economics, has impressively demonstrated that quick causal conclusions by humans are often incorrect [8]. They may give us the feeling of understanding the world, but they do not sufficiently reflect reality and its causes. The real search for causation, on the other hand, is usually extraordinarily difficult and time-consuming and, especially in complex contexts, is only completely successful in select cases. Despite a considerable investment of resources, this difficulty in identifying causation has led us to only sufficiently understand causality when analysing relatively less complex phenomena. Moreover, considerable errors creep in simply because researchers identify their own hypotheses and only set out to prove their ideas. [...]

Big data analysis based on correlations could offer advantages here. For example, in the data on the vital functions of premature babies, the health informatics specialist Carolyn McGregor and her team at the University of Toronto have identified patterns that indicate a probable future infection many hours before the first symptoms appear. McGregor may not know the cause of the infection, but the probabilistic findings are sufficient to administer appropriate medication to the affected infants. Although perhaps not necessary in some individual cases, in the majority it saves the life of the infant and is therefore the pragmatic response to the data analysis, especially because of the relatively few side effects.

On the other hand, we have to be careful not to assume that every statistical correlation has a deeper meaning, as they also may be spurious correlations that do not reflect a causal connection.

Findings about the state of reality can also be of significant benefit for research into causal relationships. Instead of merely exploring a certain context on the basis of intuition, a big data analysis based on correlations allows the evaluation of a large number of slightly different hypotheses. The most promising hypotheses can then be used to investigate the causes. In other words, big data can help to find the needle of knowledge in the haystack of data for causal research.

This alone makes it clear that big data will not stop people from searching for causal explanations. However, the almost monopolistic position of causal analysis in the knowledge process is diminishing as the what before the why is more often prioritized.

Approximation of Reality

In 2014, science magazines around the world reported an error in Google‘s flu prediction. In December 2012 in particular, the company had massively miscalculated its forecast for winter flu in the U.S., and far too many cases had been predicted [9]. What happened? After a thorough error analysis, Google admitted that the statistical model used for the flu forecast had been left unchanged since its introduction in 2009. However, because people‘s search habits on the Internet have changed over the years, the forecast was misleading.

Google should have known that. After all, the Internet company regularly updates many other big data analyses of its various services using new data. An updated version of the forecast, based on data up to 2011 resulted in a much more accurate forecast for December 2012 and the following months.

This somewhat embarrassing mistake by Google highlights another special feature of big data. Until now, we have tried to make generalizations about reality, which should be simple and always valid, but in doing so, we have often had to idealize reality. In most cases this was sufficient. However, by trying to understand reality in all its detail, we are now reaching the limits of idealized conceptions of the world. With big data it becomes clear that with idealized simplifications we can no longer grasp reality in all its diversity and complexity, but must understand each result of an analysis as only provisional.

Accordingly, we gratefully accept each new data point, hoping that with its help, we will come a little closer to reality. We also accept that complete knowledge is escaping us, not least because the data is always merely a reflection of reality and thus incomplete.

(Economics) Primacy of Data

The premise of big data is that data can be used to gain insights into reality. Therefore, it is primarily the data, not the algorithm, that is constitutive for gaining knowledge. This is also a difference to the „data poor“ past. When little data is available, the model or algorithm holds greater weight, as it must work to compensate for the lack of data. This also has consequences for the distribution of informational power in the context of big data. In the future, less power will be given to those who merely analyse data than to those who also have access to the data itself. This development will ground in fact the unease of many people towards organizations and companies that collect and evaluate ever larger amounts of data.

Because knowledge can be drawn from data, there are massive incentives to capture more and more aspects of our reality in data. In other words – to coin a phrase – to increasingly „datify“ reality. [. . .] If the costs of evaluation and storage decrease, then it suddenly makes sense to keep previously collected data on-hand and to reuse it for new purposes in the future. As a result, from an economic point of view, there are also massive incentives to collect, store and use as much data as possible, without apparent reason, since data recycling increases the efficiency of data management.

Big data is a powerful tool for understanding the reality in which we live, and those who use this tool effectively benefit from it. Of course, this also means the redistribution of informational power in our society – which brings us to the dark side of big data.

Permanence of the Past, Predicted Future

Since Edward Snowden‘s revelations about the NSA‘s machinations, much has been written about the dangers of big data. The first thing usually mentioned is comprehensive monitoring and data collection, but the threat scenario goes beyond the NSA.

If simple availability and inexpensive storage encourage unlimited data collection, then the danger exists that our own past will catch up with us again and again [10]. On the one hand, it empowers those who know more about our past actions than we ourselves can perhaps remember. If we were then regularly reproached for what we said or did in earlier years, we might be tempted to censor ourselves, hoping that we would not run the risk of being confronted with an unpleasant past in the future. Students, trade unionists and activists might feel compelled to remain silent because they might fear being punished for their actions in the future or at least treated worse. According to psychologists, holding on to the past also prevents us from living and acting in the present. This is how literature describes the case of a woman who cannot forget and whose memory of every day of the past decades blocks her in her decisions in the present.[11]

In the context of big data, it is also possible to forecast the future based on analyses of past or present behaviour. This can have a positive impact on social planning, for example when it comes to predicting future public transportation flows. However, it becomes highly problematic if we start to hold people accountable on the basis of big data predictions about future behaviour alone. That would be like the Hollywood film „Minority Report“ and would call into question our established sense of justice. What is more, if punishment is no longer linked to actual but merely predicted behaviour, then this is essentially also the end of social respect for free will.

Although this scenario has not yet become reality, numerous experiments around the world already point in this direction. For example, in thirty states in the United States, big data is used to predict how likely it is that a criminal in prison will re-offend in the future, and thus, to decide whether or not they will be released on parole. In many cities in the Western world, the decision of which police patrols operate and where and when they do is based on a big data prediction of the next likely crime. The latter is not an immediate individual punishment, but it may feel like it for people in high-crime areas when the police knock on the door every evening, even if just to ask nicely whether everything is alright.

What if big data analysis could predict whether someone would be a good driver before they even pass their driving test? Would we then deny such predicted bad drivers their licences even if they could successfully pass the test? And would insurance companies still offer these people a policy if the risk was predicted to be higher? At what conditions?

All these cases confront us as a society with the choice between security and predictability on the one hand and freedom and risk on the other. But these cases are also the result of the misuse of big data correlations for causal purposes — the allocation of individual responsibility. However, it is precisely this necessary answer to the why that the analysis of the what cannot provide. Forging ahead anyway means no less than surrendering to the dictatorship of data and attributing more insight to big data analysis than is actually inherent in it.

[...]

This text is a shortened, translated and author-approved version of the article, “Was ist Big Data? Zur Beschleunigung des menschlichen Erkenntnisprozesses”, by Viktor Mayer-Schönberger.
It was published originally in: Aus Politik und Zeitgeschichte/bpb.de 6.3.2015 in German, under a

Endnotes:

1. Martin Hilbert/Priscilla López, The World’s Technological Capacity to Store, Communicate, and Compute Information, in: Science, 332 (2011) 6025, S. 60–65.

2. Elizabeth L. Eisenstein, The Printing Revolution in Early Modern Europe, Cambridge 1993, S. 13f.

3. John Gantz/David Reinsel, Extracting Value from Chaos, 2011, http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf« (24.2.2015).

4. Jeremy Ginsburg et al., Detecting Influenza Epidemics Using Search Engine Query Data, in: Nature, 457 (2009), S. 1012ff

5. Andrea Freyer Dugas et al., Google Flu Trends: Correlation With Emergency Department Influenza Rates and Crowding Metrics, in: Clinical Infectious Diseases, 54 (2012) 4, S. 463–469.

6. Chris Anderson, The End of Theory, in: Wired, 16 (2008) 7, http://www.wired.com/science/discoveries/magazine/16-07/pb_theory« (24.2.2015).

7. danah boyd/Kate Crawford, Six Provocations for Big Data, Research Paper, 21.9.2011, ssrn.com/abstract=1926431 (24.2.2015).

8. Daniel Kahneman, Schnelles Denken, langsames Denken, München 2012.

9. David Lazer/Ryan Kennedy/Gary King, The Parable of Google Flu: Traps in Big Data Analysis, in: Science, 343 (2014) 6176, S. 1203ff.

10. More in detail: Viktor Mayer-Schönberger, Delete – Die Tugend des Vergessens in digitalen Zeiten, Berlin 2010.

11. Elizabeth S. Parker/Larry Cahill/James L. McGaugh, A Case of Unusual Autobiographical Remembering, in: Neurocase, 12 (2006), S. 35–49.

I ♥ A.I.
OPEN SOURCE
BOTTOM-UP
BLOCKCHAIN
EMBASSY OF DECENTRALIZATION
SHOP
TILL YOU DROP!
WORK
CONNECT WITH YOUR PEERS
BUY!
PLATFORMIZATION
5G
coming soon

4. Platforms and the Decentralised Internet

With the increasing importance of algorithms and big data, the programming aspects and the software on computers and servers gain importance. The term *platform* describes the different kind of digital services that organize "personalised interactions" which are "organised through the systematic collection, algorithmic processing, monetisation and circulation of data" (Poell et al, 2019, p. 3). Facebook, Twitter, LinkedIn and Instagram are social media platforms, aiming to connect people and facilitate exchange. Platforms can be a space for two parties matching and exchanging goods, for instance accommodation (AirBnB), a car ride (Uber), work (Amazon Mechanical Turk) or a product (Amazon Marketplace or eBay). Other platforms enable their users to share content (Flickr), to develop content together (like maps via OpenStreetMap or 3D models on Thingiverse) or provide other spaces for collaboration (like learning platforms, the different collaborative Google services or project management softwares). Crowdfunding is a new way to finance projects. In the public sector, platforms are enabling and organizing social and public services, for instance in public administration or in the health system (see also the publication on E-Governance). Many more examples could be added to the list of different platforms, and it would still be incomplete.

In the employment sector, platforms have challenged traditional working relations with the platform worker, a new kind of employment between self-employment and labor contract. AirBnB and Booking.com are disrupting the accommodation sector. Social media platforms are challenging old media: "Over four in ten Europeans now say they use online social networks every day" (EUC-EB, 2018, p. 17). Also in the educational field, platforms are gaining importance, for instance in education to learn analytics or credentialing.

Although their aims are different – networking, exchanging, sharing, collaborating, co-creating, earning or learning – what all platforms have in common is, they are providing a digital infrastructure enabling people to interact with others. A shared feature is also that they are processing data, which can be personal data, process data, statistical data or product data. This was not always at the forefront. Platforms are a natural way of human self-organization and as such, witness to the first collaborative steps of the internet.

Platform Economy

Platformisation:
Penetration of the infrastructures, economic processes, and governmental frameworks of platforms in different spheres of life. Reorganisation of cultural practices and imaginations around platforms (Poell et al., 2019, p. 6).

Together with the diffusion of platforms across all sector of the society, the *platform economy* has emerged. Its triumph is strongly connected with the technological capacities and the processing approach of big data. It went far beyond the simple digitisation of former analogue relations and services. With datafication, a shift toward accelerating new data and its processing is taking place. Information about the users and their interactions become essential for functioning, for the relation between people on the platform, and also for the value of the platforms, which often coincides with a shift in the platforms' mode of value creation. The huge necessary investments are stimulated by an investor driven venture capitalism and also by massive direct or indirect state investments, for instance in security or surveillance technology (Zuboff, 2018, p. 113 ff.).

The common phrase, "data would be the raw material of digitalisation", hints at an intensified process of datafication driving the development of platforms. "Datafication combines two processes: the transformation of human life into data through processes of quantification, and the generation of different kinds of value from data" (Mejias & Couldry, 2019, p. 3). It is not only statistical process data used by platform designers to, for example, measure whether transactions are proceeding successfully. They are increasingly trying to gain insight into the users.

What are the different kind of values generated out of platform users? Obviously, extracting information out of these interactions gives hints for improvements on how the offering can work better or more comfortably and intuitively for specific users or user types. Also, the insight into a variety of different interactions enables a platform to understand their product better and how it serves users. With growing market shares, platforms are also playing out the knowledge asymmetry between them and the users, for instance, binding them or setting the rules unilaterally.

Another opportunity goes beyond the triangle user, platform and other users. Information might be used to create income and value through the involvement of third parties. The dictum "free is not for free" highlights that the income model of platforms - for instance a social media platform like Facebook or Twitter - is often not the users low fee or zero fee, but the value of the information about them relevant for others, their "behavioural surplus" (Zuboff, 2018). A third party can be a platform's business client (for example from the advertising industry), a research institution (for example, some learning platforms give insights to researchers), or the state (for example, analysis of health data from public health platforms or censorship on social media platforms). A thief could also be seen as a third party, exploiting the multiple security gaps which are inherent to such a level of intensive communication.

More active users lead to more user data and to new users. Consequently, the interest of third parties in the platform is increasing. This conditionality suggests a strategy of unconditional growth. One related effect is a growing imbalance between platforms, a division between those benefiting from the growth effect and the competitors that are far behind. The effect is not limited to the digital sphere and also affects the real world. AirBnB offers cheaper, individual accommodation and at the same time, drives gentrification and unemployment in the hotel sector. Uber is destroying local cab services and driving taxi drivers into precarious working conditions.

Growing numbers of users, their datafication and following extraction of behavioural surplus are key drivers for the global platform economy. As a result, enterprises with an already huge immaterial asset of technology and data, for instance the big global conglomerates - Google, Facebook, Amazon, Alíbaba, Apple, etc. - might easier set up new platforms thanks to the peculiarity that unlike other raw material data is not depleting, only outdating. In order to keep its value, it needs to be updated on a regular basis or be recontextualized. Behind Google classrooms, stands the database of the world's biggest platform enterprise. It gains intimate insights in behaviour of the most vulnerable group in society, youth, when schools and families are not aware of privacy issues (Landesdatenschutzbeauftragter Rheinland-Pfalz, 2020). Without the Google data and ability to draw information and extract value out of these, the tool would not be worth more than the software of a talented mid-sized enterprise.

These companies are consequently extending their services, for instance by offering platform computing to others. Similar to a consumer listening to music on Spotify (renting the right to access the music catalogue and to stream), companies are also renting software access (software as a service) or might even book Artificial Intelligence as a Service. Amazon Web Services, Microsoft Azure, Google Cloud Platform (GCP), and Alibaba Cloud are the global top and dominating the cloud market with a market share of around 60%.

In consequence, this means also that they own a recent part of the internet's infrastructure. Mozilla Foundation warns: "It's a new development for online platforms to also be the owners (or co-owners) of the delivery infrastructure. At a time when there is already significant concern about the consolidation of power by the biggest technology companies in multiple realms, and telcos are merging with traditional media companies, it raises questions about who (literally) controls the internet, and how we wish to see it develop in the future" (Mozilla Foundation, 2019).

While these examples paint a rather dystopian picture for many of us, there is also reason for optimism. Remembering what once made the internet strong: de-centralised development, trust in the power of the many convinced by the idea of diversity, and the important ideals of non-profit and civil engagement. And in fact, these types of platforms also find their niche. Couchsurfing, for instance, is a platform still characterized by reciprocal hospitality, and user data is still not being monetised. In the field of crowdfunding, many different platforms are also able to coexist. TED disrupted education but is still no venture capital driven enterprise. Although the funding for non-profits is relatively small in comparison to the investments in startups, many platforms in Europe support specifically non-profit or cultural projects. Local or independent shopping platforms also try to compete. In response to Amazon, 700 German bookshops formed Genialokal, a common online shop that allows customers to order books for pick-up at the nearest bookshop or directly to the purchaser's home. Last but not least, public administration has huge potential, for example by providing open data and building the ground for local platform ecosystems.

As the EU's Next Generation Internet Initiative puts it, the challenge is "to shape the future internet as an interoperable platform ecosystem that embodies the values that Europe holds dear: openness, inclusivity, transparency, privacy, cooperation, and protection of data"(EU NGI, 2020).

Interoperability

Regarding monopolist tendencies through enforced growth or aggressive merges and acquisitions, an effective counteragent is to open interfaces and standards. Imagine, you could choose voluntarily your favourite messenger, independent of the communication partners. You use Signal, your mother uses WhatsApp, and a friend something specific like Conversations. Similar to the early days of telecommunication, only the protocols for data transfer are important, not the design of the device or its software. Some participants use telephones with dial plates, some have replaced them with those with keyboards, and others use smartphones. The open line connects everybody, which is opposite of the current lock-in to specific platforms and apps. Emails are also set up under this premise or even the Web. Its inventor, Tim Berners-Lee, wrote already in his initial concept: "Information systems start small and grow. They also start isolated and then merge. A new system must allow existing systems to be linked together without requiring any central control or coordination" (Berners-Lee, 1989/90).

What, if we instead had a Google, Apple, Russian and EU version of the World Wide Web? Technically this concept of enforcing non-centralisation on the basis of shared standards is called i*nteroperability.* It is the easiest way to dissuade hardware and software producers and also platforms to exclude competitors from the game. Also, if nations and platforms try to monopolise or separate "their" internet from the "big" internet, the open standard is a big barrier. It enables citizens to outsmart these gatekeepers, for instance by using a TOR browser (obscuring ones IP address and location to browse anonymously) or a VPN connection (tunnelling the censorship walls between a user and a server in the "free" internet). Interoperability also enables smaller platforms to cooperate and to gain size. The Fediverse network is an alliance of smaller free messengers and protocols aiming to increase interoperability on a free and open basis

However, since these tools are less intuitive than the tools provided by the carefree closed shop platforms,

Interoperability:
Ability of a system to exchange with another system and use data provided by the other system on the basis of a shared standard and in absence of central control.

you need additional competence, no matter how easy these offers are to use. First of all, you have to spend more time, realising that in the end this extra effort will be rewarded in the form of more freedom and privacy.

Interoperability is as a part of the before mentioned "FAIR principles" also strategically relevant for the future of the European Digital Single Market. For instance, it allows improved communication between public administrations (EUC-DIGIT, 2017). Also, in numbers-based telecommunication, television and radio, the EU prioritises interoperability (EU Directive 2018/1972). However, the EU is still reluctant in regard to a regulation of private markets and aware of limiting the public determination to control: "Standardisation should remain primarily a market-driven process". In particular, it is carefully trying to exclude "number-independent interpersonal communications services" from the interoperability regulations.

The Mozilla Foundation, however, is advocating for more engaged steps in this direction: "A healthy balance of power in our global internet ecosystem depends on a delicate interplay between governments, companies and civil society. We need effective competition standards and technical interoperability – between the products of different companies – to ensure that the internet grows and evolves in ways that accommodate the diverse needs of people around the world" (Mozilla Foundation, 2019, p. 98).

Decentralisation

When interoperability is a condition for more competition and a more diverse technological ecosystem, decentralised software or platforms would bring in this diversity. A lot of *Open Source* products and communities are coming into play. Let's take the example of video conferencing and clouds, which were the most striking examples during the COVID-19 pandemic in Europe for users. The most dominant platform was Zoom Video Communications (300 billion daily users). Also Microsoft's Teams and Skype (which will be shut down in 2021), GoToMeeting, Cisco's Webex or Google Hangouts were very popular. They all have in common that they are operated as a central platform. The advantage for clients is that they don't need to take care of technical aspects like enough computing space, updates, and installations. But for sure such all-in-one solutions require a certain size and financial capacity and a critical mass of users in order to make them competitive and to convince investors to invest.

Decentralised software works differently. It's installed on many different servers, not controlled by the original developers. The installing server-provider or institution (i.e. a university or school) is responsible for the installation. The advantage is that the provider or institution is able to control security and privacy and often is also able to decide, what kind of features (such as plugins or add-ons) will be installed. BigBlueButton and Jitsi are examples of de-centralised video conferencing software. Everybody can download and install it on their personal web server or rented webspace. Decentralised infrastructure requires an ecosystem of trustworthy providers and maintenance and also the willingness on the part of the consumer to invest in development and maintenance. While the software is often free, the service providers and local IT companies earn with installation and maintenance, something that might be seen as a negative. However, the need to constantly manage updates, security problems and user satisfaction, might also be a positive thing, because software installed under their control allows people and institutions to decide what is going to be installed, processed and stored by the software or the provider.

Together with improved interoperability, decentralisation incites competition and gives users more autonomy and opportunity for choice. Also the governance or co-governance of decentralised platforms by states and users is easier in comparison to the governance of multinational monopolists. Moreover, a bigger share of the value creation remains locally, which makes the national financial authorities and the local economy happier.

Furthermore, *open* interfaces and code give decentralised developers opportunities to roll out specific add-ons and to improve the software according to the needs of users. One example is the cloud software Nextcloud. Thanks to decentralised contributions and a diversity of add-ons, it developed from an Open Source alternative to Dropbox to an increasingly individually adjustable collaboration platform. "Vibrant communities of innovators are working to create alternatives to centralised systems by upscaling local connectivity, spinning up decentralised products and protocols and even creating independent alternatives to publishing on the big platforms" (Mozilla Foundation, 2019, p. 98).

COVID-19 has also shown that decentralised providers and software have not been adaptable enough to compete with the big players on the market. These solutions often cannot be developed quickly, problems might be rooted in the wrong (decentralised) configurations, testing and distribution of hardware cannot take place as fast and comprehensively as a global corporation and they experience a lack of manpower to further develop software and iron out weaknesses.

Giorgio Comai paints a realistic picture when mentioning also the challenges connected with a transformation of the internet toward more decentralisation: "In these years, as a society, we have delegated to the tech giants so many choices, including the responsibility to decide what can be legitimately published in a shared space like social networks: In a decentralised system, each entity could reasonably set up different rules, for example allowing alternative approaches to managing the flow of contents that are shown to users, benefiting pluralism and freedom of expression, but also creating new problems that the technology giants currently solve for us, including security, moderation, and control of access to data" (Comai, 2019).

On the other hand, investment in the improvement of Open Source and alternatives to central infrastructures has a long-term effect. A communication tool developed for one city can also be used in another city; a learning platform add-on developed for one university can be used by an unlimited number of schools. Furthermore, they might even be connected in a *federated* way. A stable server for schools could more easily also be made accessible to local non-profit organizations. The examples make clear that in particular public bodies would be important catalysts for alternatives to centralised platforms and for the development of the necessary open software. They would gain most through a more resilient and independent technological infrastructure.

The Platformisation Tree

By José van Dijck. professor of Comparative Media Studies at the University of Amsterdam and president of the Royal Netherlands Academy of Arts and Sciences

To envisage the platform ecosystem's hierarchical and interdependent nature, we imagine a tree that consists of three interconnected layers: the roots of digital infrastructures all leading to the trunk of intermediary platforms which branches out into industrial and societal sectors that all grow their own twigs and leaves. The tree metaphor emphasizes how platforms constitute "living" dynamic systems, always morphing and hence co-shaping its species. Like air and water can be absorbed by leaves, branches, and roots to make the tree grow, platformisation is a process in which data are continuously collected and absorbed. Data (knowingly) provided and (unknowingly) exhaled by users form the oxygen and carbon dioxide feeding the platform ecosystem. Due to the ubiquitous distribution of APIs, the process of absorbing data and turning them into nutrients—a metaphorical kind of photosynthesis—stimulates growth, upward, downward, and sideways. Each tree is part of a larger ecosystem—a global connective network driven by organic and inorganic forces. Resisting the temptation to build on this metaphor, we instead concentrate on the three layers that constitute its basic shape: roots, trunk, and branches (Figure 1).

The roots of the tree refer to the layers of digital infrastructure which penetrate into the soil; roots can run deep underground and spread widely, connecting trees to one another. Roots signify the infrastructural systems on which the Internet is built—cables, satellites, microchips, data centres, semi-conductors, speed links, wireless access points, caches and more. Material infrastructures enable telecommunications and networks like the Internet and intranets to send data packages. Online traffic is organized through coded protocols, such as the TCP/IP protocol that helps identify every location with an IP-address, and a domain name system (DNS) for proper routing and delivering of messages. The World Wide Web is one such protocol system which

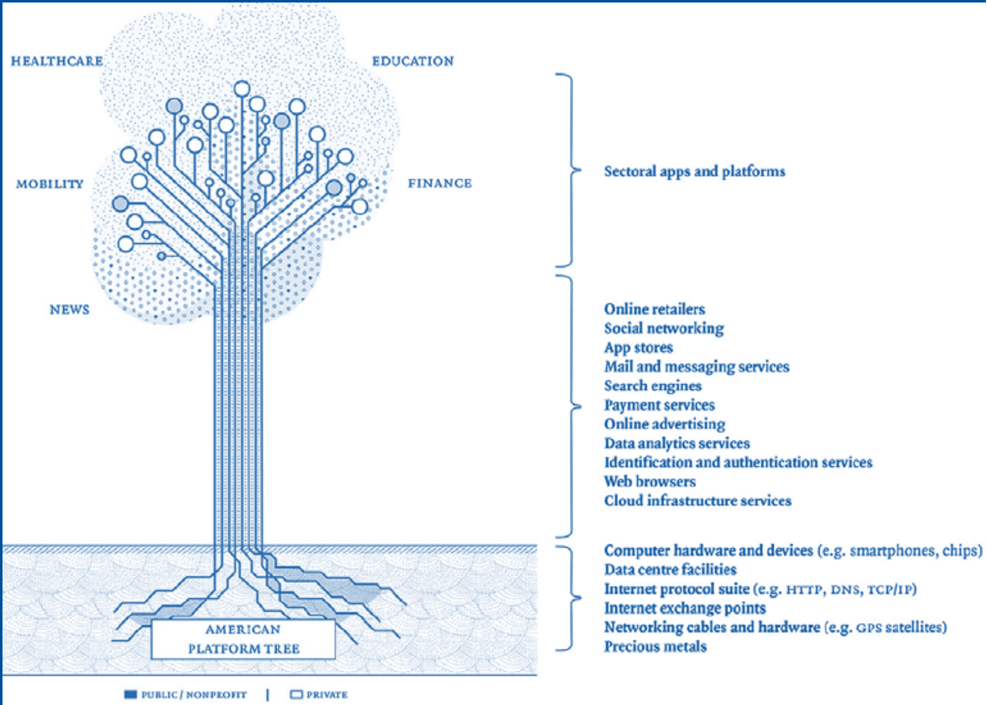


Figure 1. American Platform Tree (Giant Sequoia). Designed by Fernando van der Vlist.

helps routing data seamlessly across the net. Internet service providers (ISPs) can provide the infrastructure on which clients can build applications, such as browsers.

All separate root elements contribute to a global digital infrastructure—a structure on which many companies and states depend to build their platforms and online services. The Internet itself was originally meant to serve as a "utility," independently organized and managed, indifferent to various geopolitical and corporate interests, to guarantee the global fluidity of Internet traffic. For instance, the Internet Corporation for Assigned Names and Numbers (ICANN) represents the ideal of multi-stakeholder governance, an ideal that has come under pressure as companies and states are extending their powers to appropriate the "deep" architecture of the Internet. On one hand, tech firms privatise vital parts of the infrastructure (Malcick, 2018; Plantin et al., 2018). Google, for instance, invested billions of dollars in data centres across the globe and underwater cables for data distribution. On the other hand, states and governments increasingly seek control over digital infrastructures, illustrated by American government interventions in Huawei's efforts to develop 5G networks in Europe.

While control over the "deeper" infrastructural layers has privatised and politicised, we can see similar struggles in the layers situated in the gradual changeover between the roots and the trunk of the tree, for example consumer hardware and cloud services.

Hardware devices such as mobile phones, laptops, tablets, digital assistants (Siri, Echo, Alexa) and navigation boxes allow for Internet activity to spread among users. Inside these devices, hardware components—including hubs, switches, network interface cards, modems, and routers—are tied to proprietary software components such as operating systems (iOS, Android) and browsers (Chrome, Explorer, Safari). The architecture of cloud services forms a blueprint for data storage, analytics and distribution; control over cloud architecture increasingly informs the governance of societal functions and sectors. Amazon Web Services, Google Cloud, and Microsoft Azure dominate this layer, and while states and civil society actors become increasingly dependent on them, public control over their governance is dwindling. Blurring the boundaries between "digital infrastructure" and "intermediary services" allows for further incorporation.

The intermediary platforms in the trunk of the tree constitute the core of platform power, as they mediate between infrastructures and individual users, as well as between infrastructures and societal sectors. The stack at this level includes identification or login services (FB ID, Google ID, Amazon ID, Apple ID), pay systems (Apple Pay, Google Pay), mail and messaging services (FB Messenger, Google Mail, MS Mail, Skype, FaceTime), social networks (Facebook, Instagram, WhatsApp, YouTube), search engines (Google Search, Bing), advertising services (FB Ads, Google), retail networks (Amazon Marketplace, Prime), and app stores (Google Play, Apple). This list is neither exhaustive nor static. None of these intermediary platforms is essential for all Internet activities, but together they derive their power from being central information gateways in the middle, where they dominate one or more layers in the trunk, allowing them to channel data flows upward and downward. What characterises intermediary services is that (1) GAFAM platforms strategically dominate this space while there is hardly any non-market or state presence and (2) these super-platforms are highly interdependent, governing the platform ecosystem through competition and coordination. [...]

When we move to the branches that sprout out of the trunk of the tree, we may see their volume expanding and diversifying into smaller arms and twigs, allowing for foliage to sprawl infinitely toward the sky. The branches represent the sectoral applications which are built on platform services in the intermediary layer (trunk) and enabled by the digital infrastructure (roots). The numerous branches of the tree represent the many societal sectors where platformisation is taking shape. Some sectors are mainly private, serving markets as well as individual consumers; others are mainly public, serving citizens and guarding the common good. In principle, sectoral platforms can be operated by companies—including the Big Five, incumbent (legacy) companies, and (digital native) startups—but also by governmental, non-governmental, or public actors (Van Dijck et al., 2018). In practice, we have seen an increasing number of corporate players taking the lead in sectoral data-based services, even if these sectors are predominantly public (e.g. health, education).

The platformisation tree exemplifies a complex system that comprises a variety of human and non-human actors, which all intermingle to define private and public space. Unlike the “stack” metaphor, the platformisation tree shows the order and accumulation of platforms is not random but the result of invisible forces shaping the tree into its current form: from the circulation of its resources via its root structure and intermediary trunk all the way to feeding its twigs and foliage. As the tree grows bigger and taller, the influence of private actors’ operating platforms across all levels and layer of the tree is mounting. There is more diversity of players in the branches than there is in the trunk, just as there is (still) more diversity in the infrastructural roots than there is in the trunk. In the next section, we will focus on the dynamics of platformisation by scrutinizing the privileged position of intermediary platforms as “orchestrators in the digital ecology value chain” (Mansell quoted in Lynskey, 2017: 9).

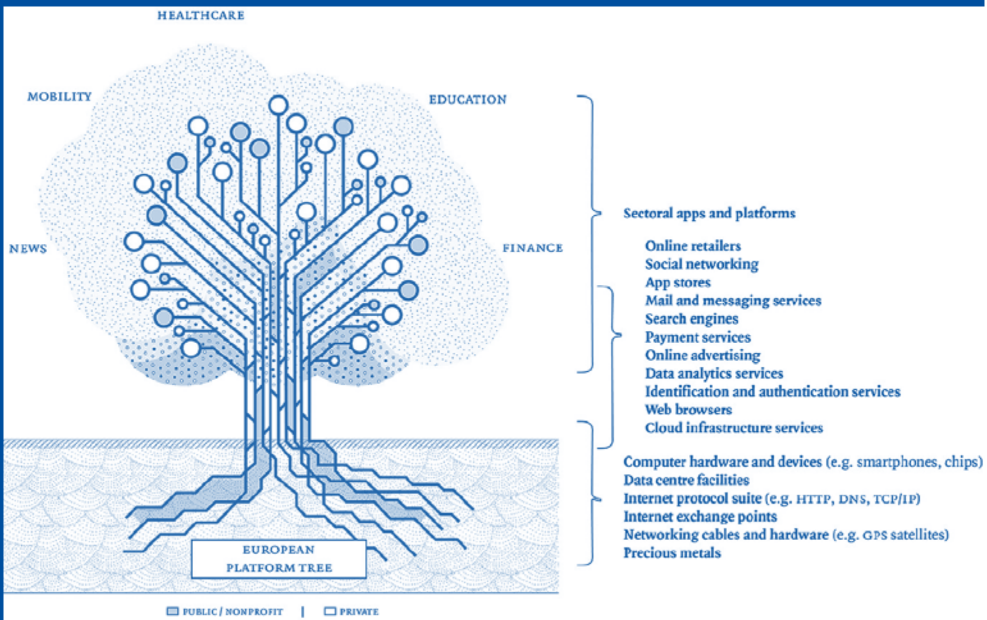


Figure 2. European platform tree

(...) The European tree does not have a trunk that grows taller and thicker fed by proprietary data flows, but it has a „federated,“ decentralized shape. It features switching nodes between and across all levels and layers, allowing users to change between platforms and define at each point how their data may be deployed. Such tree may help grow a different kind of ecosystem — one that allows for more variety, openness, and interoperability at all levels (Figure 2). (...)

Growing a diverse and sustainable platform ecosystem requires a comprehensive vision; the tree allows us to visualize a platform constellation that comprises multiple levels, visible and invisible, underground and above surface. By allowing a handful of tech companies to define the principles of a market-driven ecosystem, they are afforded all rule-setting and governing power over the world‘s information ecosystems. Focusing on single firms, markets, or individual platforms will not lead to profound, systemic changes. We need to see the forest for the trees in order to understand how to effectively govern their connective structures hidden in layers of code. The tree, although merely a metaphor, expresses the urgency to diversify the platform ecosystem in order to keep it sustainable. Without diversity, we can‘t grow a rich, nutritious forest; without a variety of actors with distinct and respected societal roles, we cannot control its unbridled growth; and without a set of principles, we cannot govern its dynamics. Changing a system starts with vision and visualization.

José van Dijck is author of the book “The Culture of Connectivity. A Critical History of Social Media” and co-author of “The Platform Society”. The text is an excerpt from the article “Seeing the forest for the trees: Visualizing platformization and its governance”, originally published by SAGE New Media & Society (Online First). The full version is available under: https://doi.org/10.1177/1461444820940293.

References:

Lynskey O (2017) Regulating “platform power..” LSE legal studies working papers 1. SSRN. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2921021

Malcick S (2018) Proteus Online: digital identity and the Internet governance industry. Convergence 24(2): 205–225.

Plantin J-C, Lagoze C, Edwards PN, et al. (2018) Infrastructure studies meet platform studies in the age of Google and Facebook. New Media & Society 20(1): 293–310.

Van Dijck J, Poell T and De Waal M (2018) The Platform Society: Public Values in an Online World. New York: Oxford University Press.

5. Openness

The Internet is organized in a decentralised and open way. It was created as a network, and as such, under the premise that each part of the network, author and user, is an equal participant. Looking from this perspective at the platform economy, the impression might be that its proprietary conception competes with this generally non-centralised vision of networks. From this perspective, openness is an important feature for alternative innovation to the growth models of proprietary platforms. Competitors strengthening and reinforcing the idea of openness are also a condition for balancing these two paths of digital transformation.

One striking effect of electronic platforms is that the difference between users and co-creators is blurring. In this sense, we all are producers. If you, as an educational institution, are publishing a Google-map with all locations of your experts network to present regional contact persons, it's not the educational institution producing the technology behind the map, but Google. Still though, you are going to create a kind of small platform, according to the definition above.

Described using a metaphor from the animal kingdom - we are symbiotic with the bigger and smaller fishes in the marine world. Still, however, the bigger fish too often controls your contributions. The other approach is to collaborate with many small fishes in a more symmetric way. OpenStreetMap and Wikipedia are such examples. The OpenStreetMap Foundation provides a map on a central server, the community fills this "aquarium". The rules for using and sharing are defined in an Open Database License provided by the Open Knowledge Foundation. A similar platform is Wikipedia, controlled by the Wikimedia Foundation. Also the software behind Wikipedia, the MediaWiki, is published under a free license (GNU General Public License). In consequence, many Wiki projects around the world might download the software and benefit from the development of the Wikimedia Foundation and the community around MediaWiki. The CC license logo might be found under many current publications. These explain under what conditions people might use and share creative works. These licenses are published by the international nonprofit organisation Creative Commons. Shared standards for messengers also exist. The XMPP protocol, which is the basis for many interoperable messenger software, is developed by a foundation. And the Open Document Format (.odt), is a format for text files published by the International Organization for Standardization (ISO).

Open Source is software which makes its code base transparent, allowing anyone to check what is programmed and use the software. Their users are encouraged to change and co-create the software within the limitations and opportunities described in the open license models. For instance this text was produced with the help of the open source office program LibreOffice, developed by the non-profit Open Document Foundation. Thingiverse, one of different platforms for 3D print templates, is not run by a nonprofit organisation, but from the 3D printer company MakerBot. It also publishes the contributions of the 3D printing community under an open CC Creative Commons license. This example shows that commercial actors may also have an interest in sticking and promoting open standards. In fact, a lot of open software projects are co-financed and otherwise supported by enterprises. In 2018, the five most active contributors to open source software were Microsoft, Google, Red Hat, IBM and Intel (Asay 2018/02/7).

It is a little bit like with football, a game played by many amateurs and professionals, people from all regions of the world. They work together for the sport's popularity and development. A common standard also exists: "The Laws of the Game are the same for all football throughout the world from the FIFA World Cup final through to a game between young children in a remote village". The International Football Association Board (IFAB) in Zurich is the custodian of the standard. Surely it might be that some influential platforms like the UEFA, Real Madrid or Manchester City would like to change the rules – a shorter game, bigger goals, smaller field or greener grass. But the price to leave the community might be high. Their players would be excluded from the international football scene. Maybe one wants to still participate in tournaments, or sell players? It can also be easier to get new players who already learned the rules of the game somewhere else. Real or City are the Googles and Amazons, compared to an amateur club in a Spanish or English village. They are playing the game under very different conditions. But all of them require the common standard or the joint commitment to football.

Therefore, the open source model gains more and more importance for the technology behind the surface, like for databases or operating systems of servers (Apache or nginx). Microsoft also runs its cloud, Azure, with an open operating system on Linux basis, although the Linux operating system on desktops/notebooks (with around 3% market share) is marginal in comparison to Windows (around 87%), or Mac OS (around 9%). The share of Open Source on mobile devices is different: Google's Open Source software, Android, dominates the market with around 68% (Mac OS has a market share of around 29%; NetMarketShare).

Openness in regard to data and systems is also becoming apparent. The concept of the sharing economy, for example, depends heavily on open and accessible data. Although open access and open usage of data is not explicitly part of the earlier mentioned FAIR concept, open data is key for innovation and alternative data-economic models in the EU Single Market, a condition for a rights-sensitive digitalisation of public infrastructure, or for publicly funded research: "FAIR principles should be implemented in combination with a policy requirement that research data should be Open by default” (EUC-RTD 2018, p. 21). Beyond research also the public sector has a crucial role to play, as a provider and producer of many different public data. Access and usage rights for such open data would enable the society, including diverse actors like entrepreneurs, civil society or state bodies, to develop innovative products, to fulfil their role as critical public, or to come up with evidence-based management and policies. The idea of open data is not limited to the central provision of infrastructural, environmental, planning, or public performance data in a public website or database. Open AI would also enable these to make use of algorithms and AI for public, not-for-profit and also for-profit purposes. The question here is, who has access to data, and how might affected persons and groups inform and control the systems and “their” data. In its report, “Steering AI and Advanced ICTs for Knowledge Societies”, UNESCO advocates decisively for openness and transparent systems: “Openness is an important attribute for publication of research and for ensuring transparency and accountability, as well as fair competition in the development and use of AI.” (Hu et. al., 2019, p. 86). Connected to this, is the necessity of free and open access to research knowledge, computing power and data for “bridging new digital divides that we are witnessing between and within countries” (Hu et al., 2019, p. 106).

Open Source:
Software with source code that anyone can inspect, modify, and enhance.
(OpenSource.com)

Open Access:
providing online access to scientific information that is free of charge to the user and that is re-usable. It includes peer-reviewed scientific publications and scientific research data
(EU Commision; EUC-RTD, 2017)

Open Data:
Free and accessible sets of (public) data, often provided through a database or a website.

Open Educational Resources:
Learning, teaching and research materials in any format or medium that reside in the public domain or are under copyright that have been released under an open license that permits no-cost access, reuse, re-purpose, adaptation and redistribution by others (UNESCO, 2019).

Conclusions for Education

In 2020, we experienced the power and advantages of global platforms in the educational sector. Although proprietary solutions were often faster to implement in online teaching and worked relatively reliable, the COVID-19 pandemic also showed their disadvantages: opaque user contracts, privacy concerns and data and security breeches. However, decentralised software was not able to compete, sometimes due to lacking availability, technical support or digital competence. The consequence is to learn from the crisis and to invest in decentralised software. Education for Democratic Citizenship/Human Rights Education **requires rights-sensitive tools and infrastructures**.

"We call for the promotion of decentralisation and a broad ecosystem of digital infrastructure operators in order to achieve digital sovereignty and dissolve dependencies on individual providers, through the dismantling of operator monopolies and the consistent use of open standards, free and open source software technologies" (Alliance Learning from the crisis, 2020).

Furthermore, the idea of open software and creative commons addresses the **proactive** aspects of civic education. Sharing and co-creating is an attitude and a skill. When using open educational resources or materials published as creative commons, the motivation is too often their cheap availability. But why do people share? Appreciation starts with using open materials or software, but finds its expression in giving feedback, co-creating and in self-publishing and sharing efforts. Using and providing open data, open access, the UNESCO-promoted **Open Educational Resources** or content under the aforementioned "Creative Commons License" are well-recognised opportunities. And joining coalitions and networks aiming to promote open (re)sources is a clear signal and a necessary step on their way to greater recognition.

Education can also become a role model in the **choice of digital methods** or smaller tools such as boards, messengers, Etherpads or surveys, bringing learners into contact with non-proprietary and more privacy-aware alternatives. This might be embedded in lectures about the idea of **digital openness and a decentralised internet**.

Last but not least, open science intends "to foster all practices and processes that enable the creation, contribution, discovery and reuse of research knowledge more reliably, effectively and equitably" (Mendez et al., 2020).

6. Algorithms and Artificial Intelligence

Algorithms in digital contexts are sets of instructions for computers. They make programming of machines possible, instructing computers to conduct various tasks, in contrast to only processing limited calculations.

The more complex these algorithms are, the more data they are able to process. The availability of better hardware allows algorithms to model complex situations. For instance, climate change models map our climate in a way that we understand better what measure out of a set of options would lead to our intended goal of reducing global temperatures.

Another vision is that computing might help us to understand or even predict and direct human behaviour, which would allow municipalities, mobility providers, energy suppliers and insurance companies to efficiently build and manage systems.

While up until this point, systems have depended on human decisions and programs written by humans, artificial intelligence (AI) is opening further opportunities. If machines were to improve by themselves or solve problems independent of human advice, automatisation could then enter a new stage. In particular, the progress in neuronal computing in connection with big data has helped the AI technology to gain new attention.

At the moment, AI systems are not really intelligent by definition of the word. Such a strong AI system "functions just like a human mind, which we would characterize as 'strong' AI" (Wrobel, 2017). Still they are, as Wrobel puts it, "exhibiting intelligent behavior", which can be seen as the key feature of weak AI.

However, the technical term weak is mislading, since AI is becoming stronger in its influence on society. Nearly every European citizen interacts with systems using AI technology. The vision that systems could support or replace human decision-making in a specific context is more tangible then ever – from car rides to decision-making in courts to automated communication with customers.

AI is a key technology in digital transformation, like the EU Commission concludes in its White Paper on AI: "AI is a strategic technology that offers many benefits for citizens, companies and society as a whole, provided it is human-centric, ethical, sustainable and respects fundamental rights and values" (EU COM 2020/65 final). The strategic importance given to AI is reflected also in the monetary ambitions of the EU. It aims to reach €20 billion in investments per year in AI by private and public sources. The commitment of the public sector in member states and the European Commission to the development of AI is an annual investment of €7 billion (EU COM(2018)795).

The condition for AI based computing is access to many and very different data. As Mayer-Schönberger put it in this publication, "if the data come from different sources, the probability of a systemic error decreases". The underlying social question is, if this is positive or negative for citizens. Many would say, under the premises of platform power that this is a concerning development. The Vodafone Institute for Society and Communications summarises based on a survey of big data: "Less than one-third of all respondents say that they think there are advantages associated with the big data phenomenon – over half of the participants say they see more disadvantages" (2016). Citizens particularly have doubts that their data is treated in a confident and responsible way by governments and companies. However, also democratically governed, non-proprietary AI systems and those intended for the common good are based on big data. In this sense, sceptical citizens could also have an interest in feeding AI systems with (their) data. Therefore, it is not enough for education to criticize big data and datafication simply as such, but also to go deeper into questions of ethical and rights-sensitive "crash barriers" and of effective democratic governance.

Think, Machine!

By Manuela Lenzen

Intelligent machines are an old dream of mankind. In recent years, they have brought us a step forward to machine learning processes. But human intelligence is still unrivalled. In 1955, the Rockefeller Foundation received an ambitious grant application: Ten researchers led by the young mathematician John McCarthy planned to make „significant progress" in just two months in a field that was given its name in this application: artificial intelligence. Their optimism was convincing, and the hand-picked group spent the summer of 1956 at Dartmouth College in Hanover, New Hampshire, finding out „how to make machines use language, form abstractions and concepts, solve the kinds of problems now reserved for humans and improve themselves." To date, there is no binding definition of artificial intelligence, but the capabilities mentioned in McCarthy's proposal form the core of what machines should do to deserve this title.

The Dartmouth conference is now considered the starting point for AI research, and the researchers were already in the midst of it at the time, the only thing the company needed was a catchy name. The neurophysiologist Warren McCulloch and the logician Walter Pitts had already designed the first artificial neural networks in 1943, the computer scientist Allen Newell and the social scientist Herbert Simon presented their program „Logical Theorist" at the conference, which was able to prove logical theorems. Noam Chomsky worked on his generative grammar, according to which our ability to form ever new sentences is based on an unconsciously remaining system of rules. If one spelled this out, one should not be able to bring machines to use language?

In 1959, Herbert Simon, John Clifford Shaw and Allen Newell presented their General Problem Solver 1, which could play chess, and Towers of Hanoi. In 1966, Joseph Weizenbaum made a name for himself with ELIZA, a dialogue program that mimed a psychologist. He himself was surprised by the success of the rather simple system that reacted to signal words.

Setbacks and New Approaches

Intelligent machines seemed to be within reach of the new discipline in this optimistic phase of departure. But setbacks were not to be expected either. A translation program for English and Russian, which the U.S. Army had wanted during the Cold War, could not be realized, and autonomous tanks could not be developed as quickly as the researchers had promised. At the end of the 1970s and again ten years later, military and government donors concluded that the researchers had promised too much and cut funding massively. These phases went down in history as the AI winter.

In retrospect, we can see more clearly today why the early AI researchers underestimated their project: „The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it", says the very second sentence of the funding application cited above. Such a precise description is still illusory today. After more than 60 years of AI research, we now see much more clearly how little human intelligence has been understood so far. While the first generation of AI researchers had focused on universal problem solvers, the first, more modest expert systems were created in the 1970s: Dialog programs that specialized in a specific field, such as the diagnosis of infections or the analysis of data from mass spectrometers. For these systems, experts were asked about their approach and tried to reproduce it in a program. But this type of programming, called „symbolic", covers only that part of human cognition that humans are aware of, that they can spell out. Everything that happens more or less unconsciously is lost in the process. For example, how do you recognize a familiar face in a crowd? And what exactly distinguishes a dog from a cat? This is where the machine learning methods score, which we owe the current boom in AI to: You shake up your fine structure yourself, you don't have to spell the world out for yourself.

Machine Learning and a New Boom

The field of machine learning comprises numerous different procedures, the most popular of which is currently deep learning, based on artificial neural networks (ANN). Such ANN are roughly modelled on the neural networks of the brain. Artificial neurons are arranged in layers to form a network. They pick up activation signals and calculate them into an output signal. This process is executed on conventional computers with processors optimised for this purpose. ANN have an input layer, which receives the data – for example the pixel values of an image – followed by a different number of hidden layers in which the calculation takes place, and an output layer which presents the result. The connections between the neurons are weighted, so they can amplify or weaken the signals. ANN are not programmed, but rather trained: They start with random weighting and produce a random result at first, which is then corrected again and again in thousands of training runs until it works reliably. Unlike humans, these systems do not need prior knowledge about possible solutions.

Computing with artificial neural networks also has early precursors: Frank Rosenblatt presented the Perceptron as early as 1958, a system that was able to recognize simple patterns with the help of photocells and neurons simulated with cable connections. It seemed clear to Rosenblatt at the time that the future of information processing would lie in such statistical rather than logical procedures. But the Perceptron often did not work very well. When Marvin Minsky and Seymour Papert explained the limits of this method in book length in 1969, ANN became quiet again. That this method is now experiencing such a boom is due to the fact that better algorithms are now available, such as procedures for multi-layer networks, that there is enough data to train these systems, and computers with sufficient capacity to realize these processes. In addition, they are proving their usefulness in daily use.

One Technology, Many Applications

Systems that work with machine learning now not only play chess and “go”, they also analyse x-rays or images of skin changes for cancer, translate texts and calculate the placement of advertising on the Internet. One of the most promising areas of application is called predictive maintenance: appropriately trained systems recognize when, for example, the operating noise of a machine changes. In this way, they can be maintained before they fail and paralyse production.

The Weak Points

Learning systems find structures in large amounts of data that we would otherwise overlook. However, their hunger for data is also a weakness of these procedures. They can only be used where there is enough current data in the right format to train them. Another problem is the opacity of the learning process: the system provides results but no justification for them. This is problematic when algorithms decide, for example, whether someone gets a credit. In addition, they use data from the past to build models that classify new data – and thus tend to preserve or reinforce existing structures.

A New Winter?

In view of these problems, there are more and more voices prophesying that the current hype will be followed by a phase of disappointment, a new AI winter. Indeed, debates about super-intelligence are likely to raise unrealistic expectations. But AI winters have come about because researchers have had their funding cut. Currently, we are seeing the opposite: national AI funding strategies are springing up, and more and more research centres and chairs are being established. Above all, however, today‘s machine learning methods are already delivering ready-to-use products for industry, commerce, science and the military. All this speaks against a new AI winter break.

But we should take a more realistic view of what is feasible: the current AI systems are specialists. In the complex world in which we move, they will by no means be able to do without human knowledge. Perhaps the future of AI systems lies in hybrid procedures that combine both approaches, symbolic programming and learning.

Manuela Lenzen holds a PhD in philosophy and is a German science journalist.

This article was originally published in German at: https://www.dasgehirn.info/, a project of the Gemeinnützige Hertie-Stiftung, Neurowissenschaftlichen Gesellschaft e.V. in cooperation with ZKM | Zentrum für Kunst und Medien Karlsruhe .Scientific support from Dr. Marc-Oliver Gewaltig

Control and Transparency of AI

In order to develop AI technology that leads to social benefit, human control and sustainability that is in line with human rights and democratic principles, these investments need also include the creation of strong framing conditions. Therefore, the EU invited a High-Level Expert Group on Artificial Intelligence to elaborate Ethics Guidelines for Trustworthy AI. In 2019, it presented criteria for trustworthy AI.

Criteria for Trustworthy AI

Trustworthy AI has three components: (1) it should be lawful, ensuring compliance with all applicable laws and regulations; (2) it should be ethical, demonstrating respect for, and ensuring adherence to ethical principles and values and; (3) it should be robust, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm. Trustworthy AI concerns not only the trustworthiness of the AI system itself but also comprises the trustworthiness of all processes and actors that are part of the system's life cycle.

The seven key requirements are:

1. **Human agency and oversight** Including fundamental rights, human agency and human oversight
2. **Technical robustness and safety** Including resilience to attack and security, fall back plan and general safety, accuracy, reliability and reproducibility
3. **Privacy and data governance** Including respect for privacy, quality and integrity of data, and access to data
4. **Transparency** Including traceability, explainability and communication
5. **Diversity, non-discrimination and fairness** Including the avoidance of unfair bias, accessibility and universal design, and stakeholder participation
6. **Societal and environmental wellbeing** Including sustainability and environmental friendliness, social impact, society and democracy
7. **Accountability** Including auditability, minimisation and reporting of negative impact, trade-offs and redress

Source: Independent High-Level Expert Group on Artificial Intelligence (IHLEG, 2019)

Another key issue is the reflection of interests behind a technology and of basic human assumptions. Both are influencing how just and fair the output of an algorithm will become. Tijmen Schep is using the term “mathwashing” for hiding the intentional or unintentional use of power and bias behind a technical fassade. “People design algorithms. They make important choices like which data to use and how to weight it. Data is not automatically objective either. Algorithms work based on the data we provide. Anyone that has worked with data knows that data is political, messy, often incomplete, sometimes fake and full of complex human meanings. Even if you have ‘good’ and ‘clean’ data, it will still reflect societal biases” (Schep).

The project, Algo.Rules, developed criteria for the design of algorithmic systems. As such, they might become an obligatory part of an ICT education, but also help political decision makers, citizens, or learners and providers of Education for Democratic Citizenship/Human Rights Education to understand better what kind of technology they are aiming to implement in their context – if in a municipality, a school, or an educational centre.

Algo.Rules

By Irights.Lab and Bertelsmann Foundation

Algorithmic systems are being implemented in a growing number of areas and are being used to make decisions that have a profound impact on our lives. They involve opportunities as well as risks. It is up to us to ensure that algorithmic systems are designed for the benefit of society. The individual and collective freedoms and rights that comprise human rights should be strengthened, not undermined, by algorithmic systems. Regulations designed to protect these norms must remain enforceable. To achieve this objective, we’ve developed the following Algo.Rules together with a variety of experts and the interested public.

The Algo.Rules are a catalogue of formal criteria for enabling the socially beneficial design and oversight of algorithmic systems. They provide the basis for ethical considerations as well as the implementation and enforcement of legal frameworks. These criteria should be integrated from the start in the development of any system and therefore be implemented by design. Given their interdependence on each other, the Algo.Rules should be treated as a composite unit. Interested stakeholders and experts are invited to join us in developing the Algo.Rules further and to adopt them, adapt them, expand them and, above all, explore opportunities to apply them in practice. Dynamic by design, the Algo. Rules should be fine-tuned, particularly in terms of their practical implementation.

1. **Strengthen competency:** The function and potential effects of an algorithmic system must be understood.

2. **Define responsibilities**: A natural or legal person must always be held responsible for the effects involved with the use of an algorithmic system.

3. **Document goals and anticipated impact:** The objectives and expected impact of the use of an algorithmic system must be documented and assessed prior to implementation.

4. **Guarantee security:** The security of an algorithmic system must be tested before and during its implementation.

5. **Provide labelling:** The use of an algorithmic system must be identified as such.

6. **Ensure intelligibility:** The decision-making processes within an algorithmic system must always be comprehensible.

7. **Safeguard manageability:** An algorithmic system must be manageable throughout the lifetime of its use.

8. **Monitor impact:** The effects of an algorithmic system must be reviewed on a regular basis.

9. **Establish complaint mechanisms:** If an algorithmic system results in a questionable decision or a decision that affects an individual's rights, it must be possible to request an explanation and file a complaint.

Conclusions for Education

The European education sector is still starting to accept the challenge in regard to AI education. However, the dynamic growth of this technology needs an equivalent provision of knowledge and competences among Europeans for how to deal with it and **find their position toward AI**.

The Council of Europe, with its focus on human rights, particularly emphasises adding a kind of **AI literacy** as a necessary prerequisite for the more distinctively promoted digital literacy. For instance, the authors of the study "Algorithms and Human Rights" make a claim for a broader "empowerment of the public to critically understand and deal with the logic and operation of algorithms" (CoE 2018, p. 43). From this perspective, education and information needs to also include the creation of new "additional institutions, networks and spaces where different forms of algorithmic decision making are analysed and accessed," and also a better empowerment of decision makers.

In line with this intent, the Council of Europe's Commissioner on Human Rights asks in "Unboxing Artificial Intelligence" for more investment in a more profound and **citizenship- and human rights-related education**: "Member states should invest in the level of literacy on AI with the general public through robust awareness raising, training, and education efforts, including (in particular) in schools. This should not be limited to education on the workings of AI, but also its potential impact – positive and negative – on human rights. Particular efforts should be made to reach out to marginalised groups, and those that are disadvantaged as regards IT literacy in general" (CoE 2019).

Beyond that, the education sector is not only expected to involve AI as a learning topic, but also as a technology. The EU's Digital Education Action Plan (2018, under revision by 2020) sets the scope:

- Making better use of digital technology for teaching and learning
- Developing relevant digital competences and skills for the digital transformation
- Improving education through better data analysis and foresight

(EU-COM/2018/022 final)

Here the Commission specifically expects that this should lead to "better use of data and AI-based technologies such as learning and predictive analytics with the aim to improve education and training systems" (EU COM 2020/65 final, p. 6).

Civil society organisations, researchers and think tanks have already started to think about the necessary frames and conditions for a **democratic- and human rights-sensitive development of AI** technology. Their findings can be a useful starting point for the creation of new AI literacy concepts, in particular in the different parts of adult learning. Cooperation between non-formal and formal education with these researchers and advocates might create synergies and help education to catch up.

Literature

Alliance for Affordable Internet (2019). The 2019 Affordability Report; Washington DC; Web Foundation; https://a4ai.org/affordability-report/

Alliance for Affordable Internet (2020). From luxury to lifeline: Reducing the cost of mobile devices to reach universal internet access. Washington DC, Web Foundation. https://a4ai.org/research/from-luxury-to-lifeline-reducing-the-cost-of-mobile-devices-to-reach-universal-internet-access/

Alliance Learning from the crisis. Learning from the crisis: empower civil society organisations! Retrieved from https://digitalezivilgesellschaft.org/en/ (2020/08/28).

Asay, M. (2018/02/7). Who really contributes to open source? In: InfoWorld. Retrieved from: https://www.infoworld.com/article/3253948/who-really-contributes-to-open-source.html

Andersson Schwarz, J. (2017). Platform Logic: An Interdisciplinary Approach to the Platform-Based Economy. Policy & Internet, 9: 374-394. https://doi.org/10.1002/poi3.159

Berners-Lee, T. (1989/90). Information Management: A Proposal. World Wide Web Consortium (W3C). https://www.w3.org/History/1989/proposal.html

BroadbandNow. Google Owns 63,605 Miles and 8.5% of Submarine Cables Worldwide. Retrieved from: https://broadbandnow.com/report/google-content-providers-submarine-cable-ownership/ (2020/08/25).

Comai, G (2019). The conditions for a pluralistic digital future: interoperability, transparency, and control over data. 13/05/2019 in Osservatore Balcani e Caucaso Transeuropa. Retrieved from: https://www.balcanicaucaso.org/eng/Projects2/ESVEI/News-Esvei/The-conditions-for-a-pluralistic-digital-future-interoperability-transparency-and-control-over-data-194436(2020/07/22).

Cisco (2020). Cisco Annual Internet Report 2018-2023. https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf

Council of Europe (CoE CM/Rec(2010)7). Recommendation CM/Rec(2010)7 of the Committee of Ministers to member states on the Council of Europe Charter on Education for Democratic Citizenship and Human Rights Education (Adopted by the Committee of Ministers on 11 May 2010 at the 120th Session). https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805cf01f

Council of Europe (CoE 2018). Algorithms and human rights – Study on the human rights dimensions of automated data processing techniques and possible regulatory implications. Committee of experts on internet intermediaries (MSI-NET), Strasbourg. https://edoc.coe.int/fr/internet/7589-algorithms-and-human-rights-study-on-the-human-rights-dimensions-of-automated-data-processing-techniques-and-possible-regulatory-implications.html

Council of Europe (CoE 2019). Unboxing Artificial Intelligence: 10 steps to protect Human Rights; Recommendations By the Council of EuropeCommissioner for Human Rights, May 2019, 50 Strasbourg. https://edoc.coe.int/fr/intelligence-artificielle/7967-unboxing-artificial-intelligence-10-steps-to-protect-human-rights.html

Die ZEIT (1995/03/03). Wie ein Weltbürger wandert. In Die ZEIT 10/1995.

DIGITALEUROPE (2020). DIGITALEUROPE's response to the EU Data strategy consultation. Brussels, 2020/05/29. https://www.digitaleurope.org/wp/wp-content/uploads/2020/06/DIGITALEUROPE-response-to-Data-strategy-consultation.pdf

EU Directive 2018/1972 of the European Parliament and of the Council of 11 December 2018 establishing the European Electronic Communications Code (Recast)Text with EEA relevance. http://data.europa.eu/eli/dir/2018/1972/oj

European Parliament, European Council (EU Regulation 2016/679). Regulation of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/ EC. General Data Protection Regulation. https://data.europa.eu/eli/reg/2016/679/oj

European Commission (EU-COM 2020/474 final). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – Critical Raw Materials Resilience: Charting a Path towards greater Security and Sustainability. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0474

European Commission (COM/2018/237 final). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – Artificial Intelligence for Europe.
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN

European Commission (EU-COM/2018/022 final). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on the Digital Education Action Plan.
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:22:FIN

European Commission (EU COM(2018)795). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – Coordinated Plan on Artificial Intelligence.
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:795:FIN

European Parliament (2019). 5G Deployment: State of play in Europe, USA and Asia. In-Depth Analysis Requested by the ITRE committee. Blackman, C./Camford Associates Ltd; Forge, S./SCF Associates Ltd. Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, Luxembourg.
https://www.europarl.europa.eu/thinktank/en/document.html?reference=IPOL_IDA(2019)631060

European Commission (EUC-EB 2017). Special Eurobarometer Report 460. Attitudes towards the impact of digitisation and automation on daily life. Survey conducted by TNS opinion & social at 51 the request of the European Commission, Directorate-General for Communications Networks, Content and Technology. Survey co-ordinated by the European Commission, Directorate-General for Communication (DG COMM “Strategic Communication” Unit). https://doi.org/10.2759/835661

European Commission (EUC-EB 2018). Standard Eurobarometer 88: “Media Use in the European Union” Report. https://doi.org/10.2775/116707

European Commission (EUC-DG GROW). Sustainability. Retrived at 2020/1030 from:
https://ec.europa.eu/growth/industry/sustainability_en

EU Commission (EUC-DIGIT (2017). Directorate-General for Informatics (DIGIT).New European interoperability framework - Promoting seamless services and data flows for European public administrations. https://doi.org/10.2799/78681

European Commission (EUC-RTD 2017). Directorate-General for Research & Innovation. H2020 Programme Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020.
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

European Commission (EUC-RTD 2018). Directorate General for Research and Innovation. Directorate B – Open Innovation and Open Science Unit B2 – Open Science. Turning FAIR into reality (2018). Luxembourg: Publications Office of the European Union. https://doi.org/10.2777/54599

European Commission (EU-COM 2020/65 final). Directorate-General for Communications Networks, Content and Technology: White Paper Artificial Intelligence.
A European approach to excellence and trust. https://op.europa.eu/s/oaNu

European Commission (EU COM 2020/66 final). Communication from the Commission to the European Parliament, the Council, the European Ecopnomic and Social Commmittee and the Committee of the Regions. A European strategy for data.
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0066

European Commission (EU-COM-2020-02 Factsheet). The European data strategy – Shaping Europe's digital future. Directorate-General for Communication.
https://ec.europa.eu/commission/presscorner/api/files/attachment/862109/European_data_strategy_en.pdf

European Commission (EU-COM (2020) 98 final). Circular Economy Action Plan; For a cleaner and more competitive Europe; Brussels, 11.3.2020;
https://ec.europa.eu/environment/circular-economy/pdf/new_circular_economy_action_plan.pdf

EU Court of Justice- Advocate General's Opinion in Case C-434/15 Asociación Profesional Elite Taxiv Uber Systems Spain SL, PRESS RELEASE No50/17, Luxembourg, 11 May 2017.
https://curia.europa.eu/jcms/upload/docs/application/pdf/2017-05/cp170050en.pdf

European Digital Rights (EDRi 2020). Platform Regulation Done Right. EDRi Position Paper on the EU Digital Services Act. Brussels, 9 April 2020.
https://edri.org/wp-content/uploads/2020/04/DSA_EDRiPositionPaper.pdf

European Union Agency for Fundamental Rights (FRA 2018). #BigData. Discrimination in data-supported decision making. Luxembourg, Publications Office of the European Union, 2020.
https://doi.org/10.2811/343905

European Union Agency for Fundamental Rights (FRA 2020). Your rights matter: Data protection and privacy. Fundamental Rights Survey. Luxembourg, Publications Office of the European Union, 2020.
https://doi.org/10.2811/292617

Grzymek, V., Puntschuh, M., Bertelsmann Foundation (2019). What Europe Knows and Thinks About Algorithms. Results of a Representative Survey. Bertelsmann Foundation, Gütersloh 2019.
https://doi.org/10.11586/2019008

Independent Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission (IHLEG 2019). Building trust in human-centric AI. Document made public on 8 April 2019.
https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines

European Union (EU NGI 2020). Next Generation Internet Initiative. Retrieved from https://www.ngi.eu/.

Eurostat (Eurostat TIN00083). Individuals using mobile devices to access the internet on the move % of individuals aged 16 to 74. Online data code: TIN00083, last update: 15/04/2020 23:00;
https://ec.europa.eu/eurostat/databrowser/view/tin00083/default/table?lang=en

Eurostat (Eurostat TIN00028). Internet use by individuals; % of individuals aged 16 to 74. Online data code: TIN00028, last update: 15/04/2020 23:00.
https://ec.europa.eu/eurostat/databrowser/view/tin00028/default/table?lang=en

Friedman, B., Kahn, P. H., Jr., and Borning, A. (2006). Value Sensitive Design and information systems. In P. Zhang and D. Galletta (eds.), Human-computer interaction in management information systems: Foundations, 348-372. Armonk, New York; London, England: M.E. Sharpe. Reprinted (2008) in K.E. Himma and H.T. Tavani (Eds.), The handbook of information and computer ethics, 69-101. Hoboken, NJ: John Wiley and Sons, Inc. Reprinted (2013) in N. Doorn, D. Schuurbiers, I. van de Poel, and M. E. Gorman (Eds.), Early engagement and new technologies: Opening up the laboratory. Dordrecht, Germany: Springer. https://doi.org/10.1007/978-94-007-7844-3_4

Greenpeace (2017). Clicking Clean: Who is Winning the Race to Build a Green Internet? Cook, G. et al., Greenpeace Inc.: Washington. https://www.clickclean.org

GSMA Association (2018). The Economy 2018. London 2018. https://www.gsma.com/mobileeconomy/wp-content/uploads/2018/05/The-Mobile-Economy-2018.pdf

Hu, X; Neupane, B; Flores Echaiz, L.; Sibal, P.; Rivera Lam, M. (2019). Steering AI and Advanced ICTs for Knowledge Societies. A Rights, Openness, Access, and Multi-stakeholder Perspective. UNESCO Series on Internet Freedom. UNESCO Publications Office, Paris. https://unesdoc.unesco.org/ark:/48223/pf0000372132.locale=en

Landesdatenschutzbeauftragter Rheinland-Pfalz (2020). Anforderungen für den schulischen Einsatz von Google-Classroom. Retrieved from: https://www.datenschutz.rlp.de/fileadmin/lfdi/Dokumente/Orientierungshilfen/anforderungen-google-classroom.pdf (2020/04/27).

Mendez, E; Lawrence, R.; MacCallum, C. J.; Moar, E. et. Al (2020). Progress on Open Science: Towards a Shared Research Knowledge System. Final Report of the Open Science Policy Platform. European Commission Directorate-General for Research and Innovation Directorate G — Research & Innovation Outreach. https://doi.org/10.2777/00139

Mozilla Foundation (2019). Internet Health Report. Transcript Verlag, Bielefeld. https://www.transcript-verlag.de/978-3-8376-4946-8/internet-health-report-2019/

NetMarketShare. Operating System Market Share´. Retrieved from: https://netmarketshare.com/operating-system-market-share.aspx (2020/08/03).

Nogared, J; Støstad, J-E (2020). A Progressive Approach to Digital Tech; Taking Charge of Europe's Digital Future. FEPS – Foundation for European Progressive Studies, SAMAK – The Cooperation Committee of the Nordic Labour Movement. Brussels: 2020. Retrieved from: https://www.feps-europe.eu/attachments/publications/feps_samak%20a%20progressive%20approach%20to%20digital%20tech.pdf

O' Neil, C. (2017). Angriff der Algorithmen: Wie sie Wahlen manipulieren, Berufschancen zerstören und unsere Gesudnheit gefährden. Hanser Verlag: Munich. Translation of: Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy. Crown: New York 2016.

OECD PISA (2018). Preparing our Youth for an Inclusive and Sustainable World. The OECD PISA global competence framework. Directorate for Education and Skills, Paris. https://www.oecd.org/pisa/Handbook-PISA-2018-Global-Competence.pdf

Pekka Raatikainen, M.J.; Arnar, D. O.; Zeppenfeld, K.; Merino, J. L.; Levya, F.; Hindriks, G.; Kuck, K-H (EUROPEACE 2015). Comparative analysis of EHRA White Book data 2009-2013: Statistics on the use of cardiac electronic devices and electrophysiological procedures in the ESC countries: 2014 report from the European Heart Rhythm Association (EHRA) In: Europace – 2015/01/23 17 Suppl 1. https://doi.org/10.1093/europace/euu300

Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. New Media & Society, 20(1), 293–310. https://doi.org/10.1177/1461444816661553

Poell, T. & Nieborg, D. & van Dijck, J. (2019). Platformisation. Internet Policy Review, 8(4). https://doi.org/10.14763/2019.4.1425

Schep, T. Mathwashing. Retrieved from: https://www.mathwashing.com/

The Shift Project (2019). Lean ICT: Towards digital sobriety – Report of the Working Group directed by Hugues Ferreboeuf for the Think Tank The Shift Project. March 2019. Retrieved from: https://theshiftproject.org/wp-content/uploads/2019/03/Lean-ICT-Report_The-Shift-Project_2019.pdf

Spiekermann, S. (2010). About the “Idea of Man” in System Design – An enlightened version of the Internet of Things? In Architecting The Internet of Things, edited by D. Uckelmann, M, Harrison, F. Michahelles, Springer Verlag, 2010, p. 25-34. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2046497

Strickland, J.; Donovan, J. How Google Works. HowStuffWorks. Retrieved from: https://computer.howstuffworks.com/internet/basics/google5.htm (2020/09/16).

Mejias, U. A. & Couldry, N. (2019). Datafication. Internet Policy Review, 8(4). https://doi.org/10.14763/2019.4.1428

UNESCO (2019). Recommendation on Open Educational Resources (OER). Retrieved from: https://en.unesco.org/themes/building-knowledge-societies/oer/recommendation

UNESCO (2017). Education for Sustainable Development Goals: learning objectives. Paris. https://unesdoc.unesco.org/ark:/48223/pf0000247444

USA Today (2018/10/03). Comen, E.: Check out how much a computer cost the year you were born. Retrieved from: https://eu.usatoday.com/story/tech/2018/06/22/cost-of-a-computer-the-year-you-were-born/36156373/ (2020/08/06).

van Dijck, J. (2020). Seeing the forest for the trees: Visualizing platformization and its governance”, SAGE New Media & Society (Online First), First published July 8, 2020. https://doi.org/10.1177/1461444820940293.

Vodafone Institute for Society and Communications (2016). Big Data. A European Survey on the Opportunities and Risks of Data Analytics. January2016. TNS Infratest, Munich. Retrieved from: https://www.vodafone-institut.de/wp-content/uploads/2016/01/VodafoneInstitute-Survey-BigData-en.pdf

Weiser, M. (1991). The Computer for the 21st Century in: Scientific American 09/1991; 94-104.

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

Wrobel, S. (2017). “We’re far from the end of this progress”. An interview with Prof. Dr. Stefan Wrobel, Director of the Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS and Professor of Computer Science at the University of Bonn. In: Berkler, K.; Köhler, H.; Möhlmann, R.. Trends in Artificial Intelligence. Fraunhofer Gesellschaft e. V., Munich. Retrieved from: https://www.bigdata.fraunhofer.de/ki (2020/10/01)

Zuboff, S.(2018). The Age of Surveillance Capitalism. The Fight for a Human Future at the New Frontier of Power. Profile Books, London 2019.

Zuboff, S (2015). Big Other: Surveillance Capitalism and the Prospects of an Information Civilization (April 4, 2015). Journal of Information Technology (2015) 30, 75–89. https://doi.org/1 0.1057/jit.2015.556

Sharing is Caring

This series "Smart City, Smart Teaching: Understanding Digital Transformation in Teaching and Learning" is an Open Educational Resource (OER) supported by the European Commission.