



encore

MAGAZINE ON INTERNET
AND SOCIETY

VOLUME 2022/2023

CONVERSATIONS

**Artificial intelligence
and the human**

**Beate Rössler on being
autonomous in the uncanny valley**

TOOLKIT

Future(s) of the digital world

ARTICLES

Designing digital democracy

**Why explainable AI needs
such a thing as society**

A lion for sustainable AI

**Workers' rights and
platforms' responsibility**



hiig.de/encore2022

**SELECTED CONTENT
ON OUR SIX RESEARCH
TOPICS IN FOCUS**

Platform regulation and data governance

Artificial intelligence and society

Digitalisation and sustainability

Autonomy, autonomous systems and robots

Digital organising and the future of work

Digital infrastructures of the knowledge society

EDITORIAL

Dear ChatGPT, please produce a short creative editorial for the relaunch of the research magazine of the Alexander von Humboldt Institute for Internet and Society *encore*, using the symbol of a butterfly unfolding out of its cocoon after 10 years switching from a paper magazine to an online-only version, including topics such as sustainability, autonomy, public interest AI, future of work, and data sovereignty.

editor's comment: remarks from a human and academic perspective added

Welcome to the relaunch of *encore*, the research magazine of the Alexander von Humboldt Institute for Internet and Society. Like a butterfly emerging from its cocoon, we have shed our print-based past and have emerged as a new online-only publication, ready to explore the complexities of the digital world.

As we embark on this new chapter of *encore*, we are excited to bring you a diverse range of voices and perspectives from a team of esteemed researchers hailing from multiple disciplines. From experts in computer science, economics and law to **political scientists**, sociologists and philosophers, we will offer fresh insights and new perspectives on the most pressing issues facing the digital world.

One of the challenges is, for instance, how to work with emerging and ever powerful AI tools like ChatGPT in a transparent and responsible manner.

This edition tackles the major societal issues of our time, focusing on digital infrastructure and inclusiveness, the future of work, digital sustainability, and the impact of autonomous systems and AI. We'll also explore the kind of future society we want and what it takes to achieve it. For instance, we'll examine the environmental benefits of AI and the role of digital tools in protecting rainforests. Additionally, we'll delve into the concept of public interest AI and what digital democracy might entail.

We invite you to join us on this journey of discovery as we explore the ever-evolving digital landscape and uncover the ~~true impact of technology on our lives~~ **societal impact of technological change**. At encore, we will be at the forefront of the conversation, providing thought-provoking and cutting-edge insights on internet and society research.

nice wording, but sometimes AI makes bold statements without proper citations; we need a language that is more precise and scientific

chatGPT

ChatGPT,
Chatbot by OpenAI



Björn Scheuermann,
Director at HIIG



Stephan Bohn,
Project leader at HIIG

MATTHIAS C. KETTEMANN

Designing digital democracy

FURTHER ARTICLES

Civil society and AI: Striving for ethical governance

Content moderation on digital platforms: A more intensive horizontal effect of freedom of expression

Data governance for a human-centred research centre and health learning hub

LECTURES

What is content moderation?

Freedom of expression in Central and Eastern Europe

Platform regulation and data governance

PI

AI

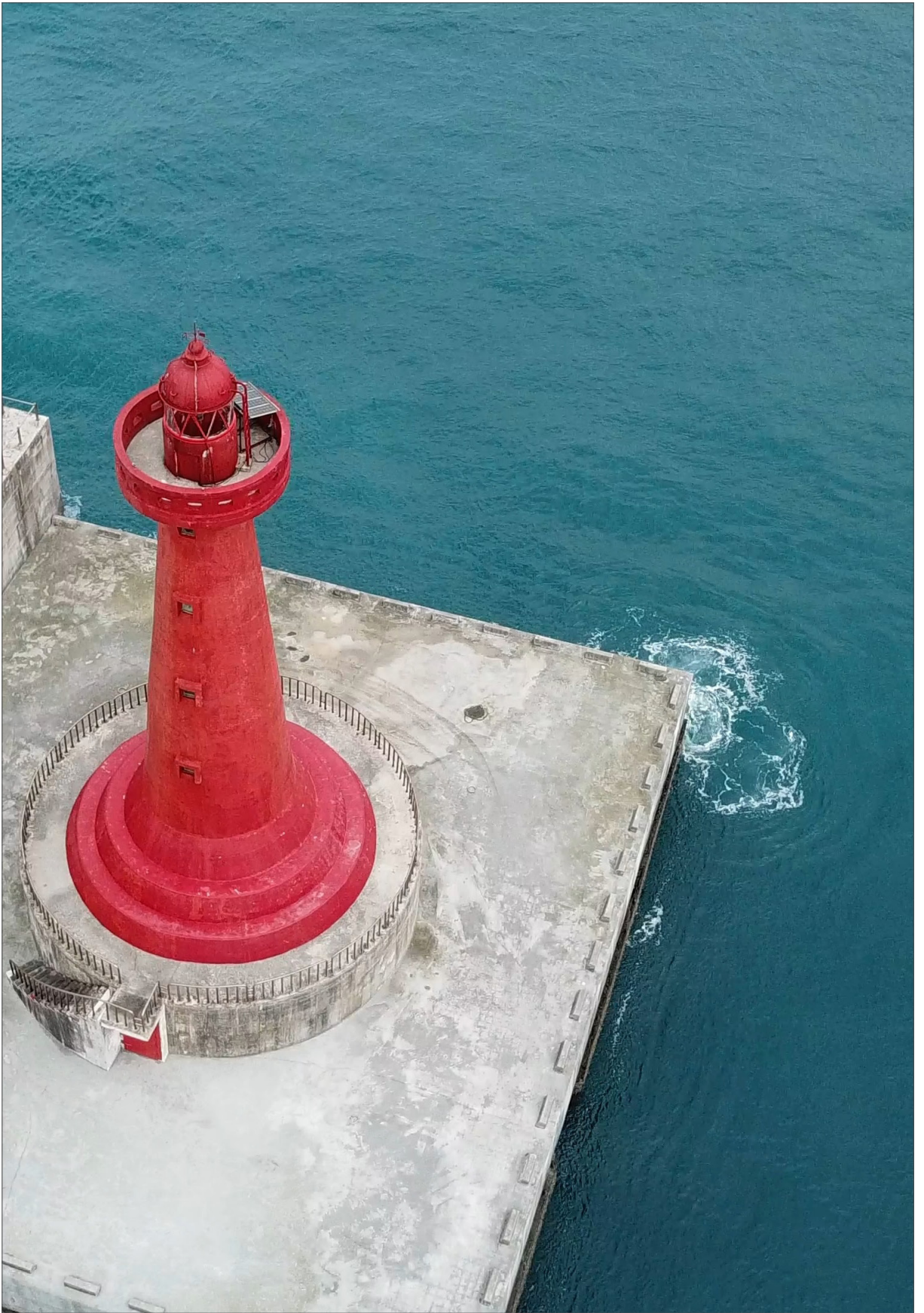
Artificial intelligence and society

FOCUS

Platform regulation and data governance

Digital platforms play a crucial role in social discourse. Besides providing access to products, entertainment and information, they also enable the politicisation of online spaces. In these, people discuss with each other, but also spread hate speech or disinformation. At HIIG, we are investigating how these new digital public spheres can be regulated by law. In addition to safeguarding human rights, the rule of law and democratic values, we are also focusing on the huge amounts of data managed by platform companies. How can this data be “shared” between state, private and public actors for new innovations for the benefit of society?

[VISIT TOPIC OVERVIEW ONLINE](#)



MATTHIAS C. KETTEMANN

Designing digital democracy

Germany's most important philosopher of democracy and democratic discourses is unhappy. In his latest book, Jürgen Habermas argues that "half-publics" are taking the place of public spaces and that democratic discourse is being challenged by heated online debates. But the character of the debates itself is not the central issue. The bigger challenge is making sure that the private rules and practices of platforms that shape online debates are aligned with public values. A number of platforms and NGOs have started to develop deliberative approaches to platform rules. But designing digital democracy is challenging.

Jürgen Habermas is unhappy. When his *Structural Change of the Public Sphere* appeared 60 years ago, he saw individual communication and participatory culture in danger from mass media, film, radio and television. Passive listeners and viewers would no longer engage democratically; instead, they would only consume. Fast forward to the year 2020, and he is unhappy again, as he eloquently explains in his *New Structural Change of the Public Sphere and Deliberative Politics*. Now, the culprits are not passive listeners and overbearing mass media organisations. Instead, it is the excessive number of (overly active) speakers online and the platforms that allow them to post cat memes, hate speech and COVID disinformation.

A lot has changed in 60 years. The platforms themselves have become rule-makers, rule-enforcers, and judges of their own decisions. They have created communication spaces where discourse, which necessarily impacts democratic values, is subjected to the demands of the attention economy. Is it time for a reset? Should we include more societal groups in the process of developing rules on what can be said online? The German Academies of Sciences and Humanities certainly think so. They recently called for the participation of “representatives of governmental and civil society bodies as well as (...) users (...) in decisions about principles and procedures of content curation”.

A DEMOCRATIC RESET

It is therefore not very surprising when the current German government commits to “advancing the establishment of platform councils” (i.e.,

institutions that oversee the rules and practices of platforms) in their Coalition Agreement. But how should those councils be constructed? As mini-parliaments, supreme courts, or councils of wise persons? Half a year later, not much moved on the political side. In response to a minor parliamentary question (Kleine Anfrage) from the CDU/CSU parliamentary group in June, the federal government replied that it was “actively involved in the development of concepts for setting up platform councils” and that platform councils could “represent a sensible addition to the legal framework”.

The signs pointing to a new trend towards new bodies that seek to improve the legitimacy of platform rules, practices and decisions are there. A major social network, Meta, has created an Oversight Board to help with content decisions and algorithmic recommendations. The same social network is experimenting with deliberative processes at scale. A gaming label is experimenting with player councils to help programmers make exciting choices. German public television’s advisory council wants to create a people’s panel to ensure more input into programming decisions. The world’s largest online knowledge platform has, since its inception, let users (and user-editors) decide upon content-related conflicts. All of these examples share one fundamental goal: ensuring that decisions on communication rules, for people and/or mediated through algorithms, are better, more nuanced, and considered more legitimate through broader involvement.

SOLVING THE FREE SPEECH CHALLENGE?

Scholars and NGOs have become increasingly involved in the debate as well. In 2021 I co-authored an introductory study on social media councils, exploring the concept and their origins in media councils. Tech journalist Casey Newton has suggested that to build trust, platforms should try a little democracy. David Kaye, together with ARTICLE 19 and Stanford's GDPI, published a detailed study on the potential of social media councils, which ARTICLE 19 then followed up with a report on their social media councils experiment in Ireland. At Harvard, Aviv Ovadya has suggested that citizens assemblies can help policymaking beyond corporate CEOs and partisan pressure.

What can democratic approaches to platform governance achieve? Will they “solve” the challenge of ensuring democratic discourse spaces while at the same time leaving platforms enough space to innovate and set internal rules? First of all, securing free speech is a regulatory challenge *that cannot be solved*; it is a so-called wicked problem. Nor can *public health* or *climate change* be solved. In order to ensure freedom of expression and a lively political discourse (because the institutional dimension of *free speech* is often forgotten), we do not need just less regulation and more freedom. Elon Musk's decision to allow Donald Trump and Kanye West (whose content on Instagram was reduced or removed because of anti-Semitic statements) back on Twitter (where his content was removed because of anti-Semitic statements within one day of his return) is only formally a gain in freedom of expression.

IMPLEMENTATION DIFFICULTIES

There are basically two choices and neither is easy to implement. SciencesPo's Rachel Griffin recently reminded us that to address the legitimacy deficits of platform decision-making, platforms can choose a “multistakeholderist response to increase civil society’s influence in platform governance through transparency, consultation and participation” or a “rule of law response” that extends “the platform/state analogy to argue that platform governance should follow the same rule of law principles as public institutions”. Or, of course, a mixture of the two is possible – like, arguably, the Meta Oversight Board. Though it has been cited approvingly by the Special Rapporteur on Freedom of Expression in her most recent report (“Many other companies provide little or no information on their operations, much less a public channel of appeal and review”), scholars like Riku Neuvonen (Helsinki and Tampere Universities) and Esa Sirkkunen (Tampere University) show why the board does not (yet) meet the democratic promise of social media councils (and thus cannot be counted as a true “Supreme Court”).

DELIVERING DELIBERATIVE DEMOCRACY

Let’s return to the unhappy philosopher. Jürgen Habermas is concerned with the prospect of a society that is shattered into “semi-publics” and loses its common points of reference. The spaces in which communication takes place seem to have gained a peculiar “anonymous intimacy: By previous standards, they can be understood neither as public nor as private, but most likely as a sphere inflated to the public sphere of a communication

hitherto reserved for private correspondence.” We call these “hybrid spaces” because *private* rules and *private* algorithmic recommendation regimes shape and influence communications that are relevant for *public* values and interests. And it is precisely in these spaces that the future of digitally mediated democracy is being negotiated – and designs for digital democracy are being piloted.

Yet, who is supposed to help implement new models of democratic decision-making in the digital age? For Jürgen Habermas the answer is clear: the state. In an essay also printed in *Neuer Strukturwandel*, he concludes with a reminder of the responsibility of constitutional law for the stabilisation of a society’s order of truth: “It is not just a political decision, but a constitutional imperative to maintain a media order that ensures the inclusive character of the public sphere and a deliberative character of the formation of public opinion and will”.

STATES MATTER

States matter. In today’s complex society, democratic states are not primarily seen as a threat to freedom but also as its *guarantor*. Indeed, states have human-rights-based obligations to *respect*, *protect* and *fulfil/enable/ensure* human rights. For a state to fulfil its obligations, it is not sufficient for it to just avoid censoring opinions. States have to actively design media orders to enable democratic discourses. Democracies are based on communicative interactions between their citizens. This requires – constitutionally – a communication order that is institutionally protected. Freedom of

communication and media freedoms should thus be located within a system of various institutional guarantees. As the media law experts Keno Potthast and Wolfgang Schulz wrote in an [expert opinion](#) for the Berlin-Brandenburg Academy of Sciences and Humanities, democracy as understood by the Basic Law presupposes that the state will ensure the functioning of a free and open, individual and public formation of opinion.

ACTIVATING ALL STAKEHOLDERS

As the Academies of Arts and Sciences [note](#), designing digital democracy is a project for all stakeholders: “academia and the providers of digital infrastructures and services (... platforms and public service media), but also NGOs and start-ups. Science can develop and provide innovative concepts (...) for a democracy-friendly design [of online communication spaces].” As far as science is concerned, this is what HIIG’s [Platform://Democracy](#) project, funded by [Stiftung Mercator](#) and implemented at HIIG, the [Leibniz Institute for Media Research | Hans-Bredow-Institut](#) and at the [University of Innsbruck](#), has set out to do.

A PLATFORM FOR DIGITAL DEMOCRACY

This blog post, which draws from a longer one published on [te.ma](#), a platform for open science and civil discourse, is part of the [Platform://Democracy](#) project. It is funded by [Stiftung Mercator](#) and led by [Matthias C. Kettemann](#) with [Josefa Francke](#) and [Christina Dinar](#). The project explores whether and how platform councils have the potential to align public values and private

orders. In four regional research clinics, the project sheds light on how to provide the normative infrastructure for better rule-making, rule-enforcing, and rule-adjudication structures in (primarily online) hybrid communication spaces. In these clinics in Europe, Africa, Asia/Pacific/Australia and the Americas, participants will exchange experiences on models that aim to increase the quality of deliberative democracy in online settings.

FURTHER READING

Interested in digital democracy? Do you understand the added societal value, and potential drawbacks, of digitalisation? Rules, and the normative order of the Internet, are changing rapidly. But rules continue to matter in designing social innovation: Do you know who rules the internet? And how you can contribute? In order to understand how democracy and the public sphere can be governed in the 21st Century, a solid understanding of the role of social media in harnessing and projecting opinion power while fighting disinformation helps.

MAURICE JONES

Civil society and AI: Striving for ethical governance

The involvement of civil society has been identified as key in ensuring ethical and equitable approaches towards the governance of AI by a variety of state and non-state actors. Civil society has the potential to hold organisations and institutions accountable, to advocate for marginalised voices to be heard, to spearhead ethically sound applications of AI, and to mediate between a variety of different perspectives. But despite the proclaimed ambitions and visible potential, civil society actors face great challenges in actively engaging in the governance of AI.

[READ FULL ARTICLE](#)

AMELIE HELDT

Content moderation on digital platforms: A more intensive horizontal effect of freedom of expression

To what extent can and should digital platforms be bound by freedom of expression even though they are private companies? This question is central to the issue of content moderation. One solution could be a more intensive horizontal effect of fundamental rights. This article summarises a doctoral thesis on this particular question from a constitutional law perspective.

[READ FULL ARTICLE](#)

JÖRG POHLE, ANNIKA ULICH, AND
MAX VON GRAFENSTEIN

Data governance for a human-centred research centre and health learning hub

Even if actors are bound by common values and goals, they may have different ideas, interests and risk perceptions regarding the handling of data worth protecting. These ideas need to be brought together and the conflicting interests reconciled. This requires good data governance. We have developed a cooperative and inclusive data governance process that is pragmatic, goal-oriented and at the same time scalable.

[READ FULL ARTICLE](#)

LECTURE BY KRISZTINA ROZGONYI
AND MARIUS DRAGOMIR

Freedom of expression in Central and Eastern Europe

Media systems in Central and Eastern Europe (CEE) have been subject to significant transformation processes over the past two decades. In the face of populist tendencies and with large swathes of the media being captured by governments and oligarchs, the space for independent journalism has dramatically shrunk in most of the region's nations. How can economically vulnerable media regain editorial independence and stand up against the powerful propaganda channels? To find answers to this question, this edition of the lecture series features two scholars who specialise in the digital transformation of media systems in democratic societies and the particular challenges in CEE.

Marius Dragomir talks about the changes experienced by media systems in Central and Eastern Europe and presents the findings of his research into the impact of media capture on independent journalism. Krisztina Rozgonyi analyses how digital societies in Central and Eastern Europe are embedded in a politically manipulated communicative context and sheds light on its historical roots.

[WATCH FULL LECTURE](#)



JEANETTE HOFMANN

What is content moderation?

With our new explainer format *Digital & Disziplinlos*, our HIIG experts take you on interdisciplinary journeys to the digital spheres. In its first issue, our research director Jeanette Hofmann explores the concept of content moderation: platforms are increasingly under public pressure to fight hate speech and other problematic content on their social media services. How do they manage the huge amounts of text posts, videos or photos we post daily?

[WATCH FULL VIDEO \(GERMAN WITH ENGLISH SUBTITLES\)](#)

INSIGHTS AND POWER CONVERSATION SERIES

Great expectations: What research expects from platforms and platforms from research

At the beginning of 2022, HIIG started the distinguished conversation series “Insights and Power” in cooperation with the Leibniz Institute for Media Research | Hans-Bredow-Institut. For the inaugural event, we invited YouTube CEO Susan Wojcicki to engage in a discussion with HIIG director Wolfgang Schulz. In their conversation, both emphasised the importance of collaboration between science and platforms. “Academics play a key role to give a transparent understanding of platforms”, said Susan Wojcicki.

[WATCH FULL CONVERSATION](#)

FREQUENTLY ASKED QUESTIONS

Platform regulation and data governance

What are digital platforms?

From social networks to video platforms to messenger apps: digital platforms and their services shape our everyday lives. As central players in our networked world, they have great economic power. For example, they decide under what conditions we can consume music, films or goods online. They also decide what information is displayed on our feeds and when. Many larger platforms have created their own markets and rules in recent decades, with which they exercise powerful influence on our social discourse.

Why do we need socio-political regulation of digital platforms (governance)?

Platforms have become powerful actors that set their own rules and enforce them with their algorithms and the design of their services. They do not answer to society nearly enough; their bosses are not elective and in no way representative (vastly more white and male than average), nor has anyone legitimised their decisions about what we can see and buy. States and the political order, however, are responsible for respecting, protecting and guaranteeing the rights of their citizens. This also applies to private actors, and especially the platforms.

FREQUENTLY ASKED QUESTIONS

What is data governance and why is it important?

Data governance refers to the structures and processes that are established based on different ideas, interests and risk perceptions to deal with data that are worthy of protection. They can be personal data, business secrets or even open data. Often, data governance is about “sharing” data or using artificial intelligence (AI). On the one hand, good data governance aims to achieve the desired effects of data processing. This means generating new insights, driving innovation or simply connecting people with each other. On the other hand, undesired effects, ranging from surveillance and censorship to exploitation, discrimination or political repression, should be prevented or minimised. Good data governance therefore tries to implement in technical systems the best solutions that are acceptable to all stakeholders and that protect the interests of all participants as well as the society as a whole.

JUDITH FASSBENDER

**Why explainable AI needs
such a thing as society**

INTERFACE

Public Interest AI Interface

FURTHER ARTICLES

**AI as a flying blue brain?
How metaphors influence
our visions of AI**

**Content moderation: What
can stay, what must go?**

**AI made in X: the German
AI policy landscape**

LECTURES

**How the world sees intelligent
machinesInterface**

Artificial intelligence and society

AI

PI

Platform regulation
and data governance

Su

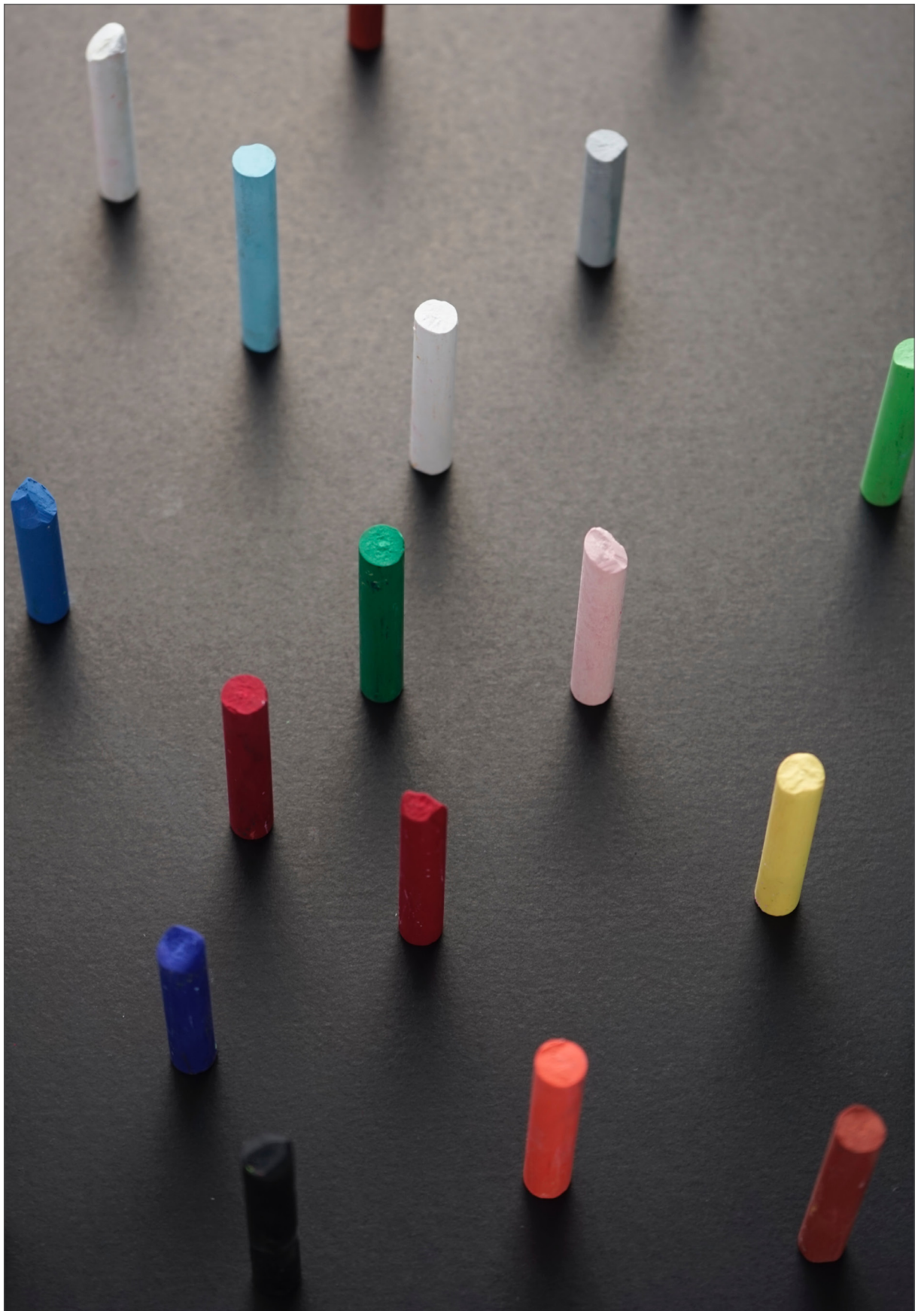
Digitalisation and
sustainability

FOCUS

Artificial intelligence and society

Artificial intelligence (AI) often refers to capabilities of technical systems that are similar to those of human intelligence. As such, the term AI is often associated with the automation of social processes. These AI systems are operating in diverse socio-technical contexts: from social media platforms to medicine to warfare. At HIIG, we investigate all the political, social and cultural facets of these discourses. Some of the discourses are connected, and some have their own dynamics. What lies behind the dazzling concept of AI?

[VISIT TOPIC OVERVIEW ONLINE](#)



JUDITH FASSBENDER

Why explainable AI needs such a thing as society

Have you ever asked yourself what the basis of your search-engine autocompletions is? For example, why did you get a suggestion to search for what it feels like to have heartburn, even though you weren't planning to search for anything like that at all? There is no standard for explaining such automated decisions as of yet. Moreover, today's explainable AI (XAI) frameworks focus strongly on individual interests but fall short in offering a societal perspective. This article will offer an introduction to communication in XAI and introduce the figure of the public advocate as a way of including collective interests in XAI frameworks.

The article is based on the findings of Theresa Züger, Hadi Asghari, Johannes Baeck and Judith Faßbender from the XAI Clinic in Autumn 2021.

NO OR POOR EXPLAINABILITY FOR LAY PEOPLE AND SOCIETY

Have you ever asked yourself why you get the search engine auto-completions you get? For example when you typed “how does” and your search engine suggested “how does ... it feel to die”, “how does ... it feel to love”, or “how does ... it feel to have heartburn” but you actually wanted to continue typing “how does ... a^2 relate to b^2 in Pythagoras’ theorem”. If explanations for automated decisions were provided as a standard, you would have been able to get an explanation of the inner workings of that search engine fairly easily. Due to a mixture of technical feasibility, communication challenges and strategic avoidance, such a standard does not exist yet. While a number of major providers and deployers of AI models have published takes on explainable AI (XAI) – most prominently [IBM](#), [Google](#) and Facebook – none of these efforts offer effective explanations for a lay audience. In some cases, lay people are simply not the target group, while in others, the explanations are insufficient. Moreover, collective interests are inadequately considered when it comes to how to explain automated decisions; the focus lies predominantly on individual or private interests.

This article will focus on how explanations for automated decisions need to differ depending on which audience is being addressed – in other words, on target-group-specific communication on automated decisions. In light of the neglected societal perspective, I will introduce the figure of the public advocate as a way of including collective interests in XAI frameworks.

TECHNICAL ELEMENTS OF AI SYSTEMS TO EXPLAIN

The technological complexity of AI systems makes it difficult to trace automated decisions. This is partly because the models have multiple layers, nonlinearities and untidy, large data sets. In response to this problem, there have been increasing efforts to develop so-called white-box algorithms or to use more simple model architectures that produce traceable decisions, such as decision trees.

But even if each element of an AI system were explainable, a complete explanation for an automated decision would consist of a fairly large number of elements. To give an idea of these elements, let me share an overview of possible elements (based on [Liao et al. \(2020\)](#)):

(1) The global model, which refers to the functionalities of the system that has been trained. This includes the specific training data that has been used and the architecture that has been employed (i.e. a convolutional neural network, linear regression, etc.). “Global” means that the functionality of the system is not case specific. (2) The local decision, which concerns a decision in a specific case. (3) The input data, which refers to the specific data that is used to make a local decision. (4) The output, which refers to the format and the utilisation of the output the system gives (5.) A counterfactual explanation, which shows how different the input would have to be in order to get a different output; such as (6) the performance of the system.

THE CHALLENGE OF TARGET-GROUP-SPECIFIC COMMUNICATION

If what you've read up to now has either bored or overwhelmed you, it could either mean that you are not the target group for this blog post or that I have missed the sweet spot between what you, as part of my target group, knew already and what you expect from this article. Target-group-specific communication is a struggle when explaining automated decisions as well.

To give you a schematic, but better, explanation, here are the elements listed above, applied to the search engine example from the beginning of this blog post:

- The global model in this case is the trained model that produces the autocomplete suggestions. The training data is most probably previous inputs by other users, what they were searching for and their entire search history.
- The input was what you typed in combination with your search history and other information the search engine provider has on you.
- The output is the autocomplete suggestion.
- The local decision is the suggestions you've been given, based on your input.
- A counterfactual could involve seeing what suggestions you would get when typing the exact same words but taking parts of your search history out of the equation or changing another parameter of the input data.

- The performance of the system would be based on how many people actually want to find out how it feels to die etc., as opposed to how Pythagoras' theorem works.

The system performance, for example, would probably not be of interest to the average lay person, but it may be relevant to the developer: people in different positions have different needs, expectations, and previous knowledge concerning explanations, and therefore, a different type of presentation would be needed for each target group.

WHO ASKED?

The standard target groups for explanations of automated decisions are the developer, the domain expert and the affected party.

The **developers** either build new AI models or further develop pre-existing AI models. This group basically needs to understand each element of the system, with a specific focus on the workings of the global model and data representation – to be able to improve and verify the system in an accountable manner. Such explanations have to be available for developers throughout the whole process of development, employment and maintenance of the system.

The **domain expert** is typically an employee of an organisation that uses AI systems. This could be a medical doctor assisted by an AI system when making a diagnosis or a content moderator on a social media platform who checks automatically flagged content. This person is assisted in their

decision-making with suggestions from an AI system, as a so-called human in the loop. Domain experts need to adapt to working with the system and need to develop an awareness of risks, of misleading or false predictions, and of the limitations. Therefore they do not just need explanations of local decisions (e.g. why did the system flag this content as being inappropriate); importantly, they need thorough training on how the global system works (e.g. regarding what data the system was trained on and whether the system looks for specific words or objects). Such training needs to take place in connection with a specific use context.

The **affected party** is, as the name suggests, the person (or other entity) that an automated decision has an effect on. Their needs range from knowing whether an AI system was involved in making a decision to understanding an automated decision in order to make informed decisions or to practise self-advocacy and challenge specific decisions or the use of an AI system altogether. Affected parties primarily need an explanation on the elements of the system that are connected to their case (the local decision). Counterfactual explanations can also be meaningful, as they would enable affected people to see what factors would need to change (in their input data) to produce a different result (the output).

A FOURTH TARGET GROUP: THE PUBLIC ADVOCATE

We propose considering a fourth target group: the public advocate.

The **public advocate** is a person or an organisation that articulates the concerns of the general public or of a group with special interests. The

general aim of all public advocate activities is to move closer to equality in our understanding of this target group. A public advocate might be an NGO/NPO dealing with societal questions connected to the use of AI systems generally – such as Access Now, Algorithmwatch or Tactical Tech – or an NGO/NPO with a focus on specific groups or domains such as medical representative organisations or organisations supporting people who are affected by discrimination.

Public advocates are concerned, on one hand, with lobbying and advocating for public or special interests – be it in deliberative processes in the media, in court, in policy-making or in collaboration with providers of AI systems. On the other hand, such organisations are well-qualified to educate others on AI systems, tailored to the needs of their respective community. These might be professional representatives of doctors, such as the German Ärztekammer (chambers of doctors), which provide radiologists (domain experts) with training and background information on the possibilities, risks and limits of technologies that could perform image recognition of a lesion in the brain.

To facilitate such support, these groups need access to general information on the AI system – on the global functioning of the model, input, and output. Further explanations of individual cases and the impact on individuals is crucial for this group, especially when their advocacy focuses on specific societal groups or use cases.

WHY IS A COLLECTIVE PERSPECTIVE IN EXPLAINABLE AI IMPORTANT?

The field of XAI is not free of power imbalances. Different actors' interests may interfere with one another. Against this backdrop, the need to have a public advocate becomes clearer: none of the traditional target groups are intrinsically concerned with collective interests and consequences. But a collective focus is important, especially with regard to seemingly low-impact decisions on what content is suggested to you on platforms or search engines. In isolation, these automated decisions may count as low impact, but they can become problematic as the number of users and/or decisions increases in scale. One such case arose when Facebook's recommendation tool contributed to the growth of extremist groups. While high-impact decisions for individuals – such as the often cited loan-lending case – are highlighted in XAI frameworks, “low-impact” decisions are much more in the shadows. Viewing them from a societal, collective perspective sheds some light on their importance. The content that is suitable for explanation from this perspective is different, and it can be formulated by considering the public advocate's target group.

Besides communicating and representing collective needs, public advocates can perform important tasks in the field of explainable AI. Training sessions on how specific AI systems work should be delivered by organisations that do not develop or employ such systems themselves and therefore do not have obvious conflicting private interests – which rules out commercial actors and governmental organisations. Public advocates can function as consultants to the developing teams if they are included early enough in

the development process and if there is a true interest in giving effective explanations.

Last but not least, public advocates can have more leverage than a singular affected person when lobbying for a collective. In comparison to the layperson, the organisations we have in mind have more technical expertise and a greater ability to understand how the system works – a capacity that increases their bargaining power further. Ideally, the work of public advocates should reduce the risk of providing ineffective explanations that are more a legal response than an actual attempt to explain – see Facebook’s take on explaining third party advertisements.

For all points mentioned above – automated decisions that become critical when viewed on a collective scale, the need to have a publicly minded entity to educate on AI systems and the benefits of joining forces with different affected parties – there needs to be a public advocate in XAI frameworks. This not only makes it possible to include the societal and collective dimension consistently when offering explanations to affected users but also helps make collective interests visible and explicit for the development of explainable AI in the first place.

FURTHER READS

Carvalho, D. V., Pereira, E. M., & Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. *Electronics*, 8(8), 832.

Liao, Q. V., Gruen, D., & Miller, S. (2020). Questioning the AI: informing design practices for explainable AI user experiences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-15).

Ribera, M., & Lapedriza, A. (2019). Can we do better explanations? A proposal of user-centred explainable AI. In *IUI Workshops* (Vol. 2327, p. 38).

Rohlfing, K. J., Cimiano, P., Scharlau, I., Matzner, T., Buhl, H. M., Buschmeier, H., ... & Wrede, B. (2020). Explanation as a social practice: Toward a conceptual framework for the social design of AI systems. *IEEE Transactions on Cognitive and Developmental Systems*, 13(3), 717-728.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence*, 267, 1-38.



PI

Public
Interest

AI

AI

PUBLIC INTEREST AI INTERFACE

Defining Public Interest AI is as difficult as it is crucial for our society. Academia and various societal stakeholders have entered a discourse to sharpen the edges of this explanation. Read here our proposed understanding.

[VISIT PUBLICINTEREST.AI](https://publicinterest.ai)



Ma

Project
Map

CONTRIBUTE

There is a lack of accumulated and detailed data on public interest AI projects, including their specific objectives, methods, and frameworks. This is why we created a survey.

Contribute by filling out the online survey and include your project in this interactive map and directory.

ADD YOUR PROJECT

In view 14/14 Total 17/17

PUBLIC INTEREST AI PROJECT MAP

This interactive global map aims to help foster research on Public Interest AI projects, demonstrate their self-understanding, and provide publicly accessible data about them to the broader public.

EXPLORE THE MAP



St

Stakeholder
Index

PUBLIC INTEREST AI STAKEHOLDER INDEX

The index gathers information about organisations and institutions involved in the discourse around Public Interest AI. Their involvement may take various forms – from giving funding to NGO work around the issue.

BROWSE THE INDEX



St
Stakeholder
Index



WOULD YOU LIKE TO GET AN OVERVIEW OF THE STAKEHOLDERS IN THE PIAI FIELD?

We have started to identify organisations and institutions that can play an important role in the development of the field or Public Interest AI (PIAI). Are you missing a stakeholder? Let us know and we'll be happy to add it!

[Research](#) [Think tanks/NGOs](#) [Politics](#) [AI Practice](#) [Funding](#)

Access Now

📍 New York, USA

[Research](#) [Think tanks/NGOs](#)

Ad hoc Committee on Artificial Intelligence (CAHA1)

📍 Brussels, Belgium

[Politics](#)

AI Campus (The Stifterverband)

📍 Essen, Germany

[Politics](#)

AI For Good (ITU Konferenz Initiative)

📍 Geneva, Switzerland

JOHANNA WALLENBORN

AI as a flying blue brain? How metaphors influence our visions of AI

Why is artificial intelligence so commonly depicted as a machine with a human brain? The narratives and metaphors of science fiction, media outlets and special interest groups about what is described by the term *artificial intelligence* (which is already a metaphor in its own right) have affected our view of AI systems in society. Metaphors are all around us and they are extremely powerful: they activate our memories, trigger emotions, influence our attitudes and shape our expectations about the future. This article shows why one misleading metaphor has become so prevalent.

[READ FULL ARTICLE](#)

ALEXANDRA BORCHARDT

Content moderation: What can stay, what must go?

Tweets by Donald Trump? Propaganda from the state broadcaster *Russia Today* spreading freely on Facebook, Instagram and YouTube? Must platforms take an anything goes attitude? What needs to happen so that social networks can continue to be central forums for debating and sharing information without providing a platform for propagandists? Automated deletion is a recent approach, but it is putting freedom of expression at risk. That's why we need a few rules. The research project Ethics of Digitalisation has worked out what these might be.

[READ FULL ARTICLE](#)

LAURA LIEBIG & ANNA JOBIN

AI made in X: the German AI policy landscape

Artificial intelligence (AI) is an international issue and so is its governance. You may have heard of the EU's AI Act, or the plethora of AI policy documents published by national governments and organisations from all over the world in recent years. But did you know that the German federal states care about AI too? Each of them has issued at least one document on AI policy. Because the subnational level of decision-making is often overlooked, we offer a federalist approach to AI in our research.

[READ FULL ARTICLE](#)

A LECTURE BY KANTA DIHAL

How the world sees intelligent machines

People have been imagining intelligent machines for millennia, in ways that vary greatly across cultures. Yet as artificial intelligence (AI) begins to fulfil its potential as a technology, having spread across the globe from its origins in 1950s' America, many of these non-Western perspectives have been marginalised. These stories, films and visions matter: they are interwoven with broader cultural attitudes and approaches to intelligent machines. In her lecture, Kanta Dihal introduces such visions from across the globe and elaborates on three main themes: the differences between “Eastern” and “Western” visions of AI, the ways AI is imagined in communist states and anti-colonialist narratives of AI. Why are these marginalised perspectives still of great relevance for today's societies and what can they teach us about ourselves?

[WATCH FULL LECTURE](#)

FREQUENTLY ASKED QUESTIONS

Artificial intelligence and society

What is artificial intelligence (AI)?

In general, artificial intelligence (AI) often refers to capabilities of technical systems that require intelligence similar to that found in humans. This can be, for example, the recognition of visual patterns or the analysis of language and texts. This proximity of technologies to human beings is also expressed in terms like “machine learning” or “neural networks”. Today, AI is often used to support or replace humans in certain activities. Therefore, it is also associated with an increasing automation of social processes. Behind the term AI, understandings around the world vary considerably from culture to culture.

FREQUENTLY ASKED QUESTIONS

What are the impacts of AI on society and vice versa?

In many areas of our society, technological automation helps to evaluate large amounts of data and thus to generate new knowledge. Examples include the analysis of diagnostic images in medicine or the pattern recognition of voice data. AI systems support humans in everyday life; they become more like us (social robots). Today, they can undertake some of the tasks that were previously only done by humans. These changes raise questions on many levels: about responsibility, trust, autonomy and ultimately, time and again, the question of what kind of society we want to live in.

How does artificial intelligence (AI) influence our lives (where do I already find AI in my daily life)?

We check the news on our mobile phones, and then briefly scroll through Facebook and Instagram, while Spotify is playing or, if it's evening, while the new Netflix series is on. Numerous media services use algorithm-based recommendation systems nowadays. But even beyond the screen at home, AI applications can be found in all fields of our lives: from medicine, finance, mobility, climate research to weapons systems.

SAMI NENNO & C. RICHTER

Sustainable AI – How environmentally friendly is AI really?

STEPHAN BOHN

Saving orangutan habitats – digital technology and rainforest protection in Indonesia

PANEL DISCUSSION

Wissen Macht Klima

FURTHER ARTICLES

A lion for sustainable AI: How to support a new standard for sustainability reporting?

Four key ideas for a sustainable platform economy in Europe

Digitalisation and sustainability

Su

AI

Artificial intelligence and society

Au

Autonomy, autonomous systems and robots

FOCUS

Digitalisation and sustainability

When it comes to sustainability, digitalisation opens up great opportunities and poses myriad challenges. It is important that societies, organisations and individuals begin to grasp the sustainable potential of digital technologies, because, currently, their negative effects often outweigh the positive ones due to enormous requirements for material and energy, but also due to discrimination against certain social groups. In our research at HIIG, we investigate how digital and sustainable transformation can be shaped together. How can technology be introduced and used in a responsible manner from a social, economic and ecological perspective?

[VISIT TOPIC OVERVIEW ONLINE](#)



SAMI NENNO & CHRISTOPHER RICHTER

Sustainable AI – How environmentally friendly is AI really?

Sustainable AI is becoming increasingly important. But how sustainable are AI models really? Big tech and smaller applications differ greatly in this respect. We have looked at how sustainable small-scale AI really is, what open questions remain and what recommendations can be made.

SUSTAINABLE AI AND AI FOR SUSTAINABILITY

For quite some time now, people have been writing about the potential of AI to aid in the fight against climate change. For example, Rolnick et al. (2023) conducted a survey on many (potential) use cases, such as enabling low-carbon electricity or managing forests. However, there are also concerns about the sustainability of AI. Often, scientists draw a distinction between sustainable AI and AI for sustainability. While the latter serves purposes like increasing the efficiency of renewable energy, the former is about making AI itself sustainable; because there is a point that often remains unsaid, namely that many AI models consume massive amounts of energy.

THE BIGGER THE BETTER

In a widely cited study, Emma Strubell and colleagues conducted experiments with natural language models. They found that computing some of these models emits as much CO₂ as five cars in their entire lifetime. In another famous [study](#), researchers at Open AI looked at the growth of AI models in recent years and found that they double their size every 3.4 months, thus contributing to a striking increase in energy consumption.

Based on these results, it is often claimed that the current trend towards bigger and bigger models is far from sustainable. A closer look at the numbers is startling: the data shows that the vast majority of the models have been built by big tech companies like Alphabet, Microsoft or Baidu. Indeed, it is no secret that AI research is not driven by “normal” universities anymore. The reasons for this are financial: training such models is extremely

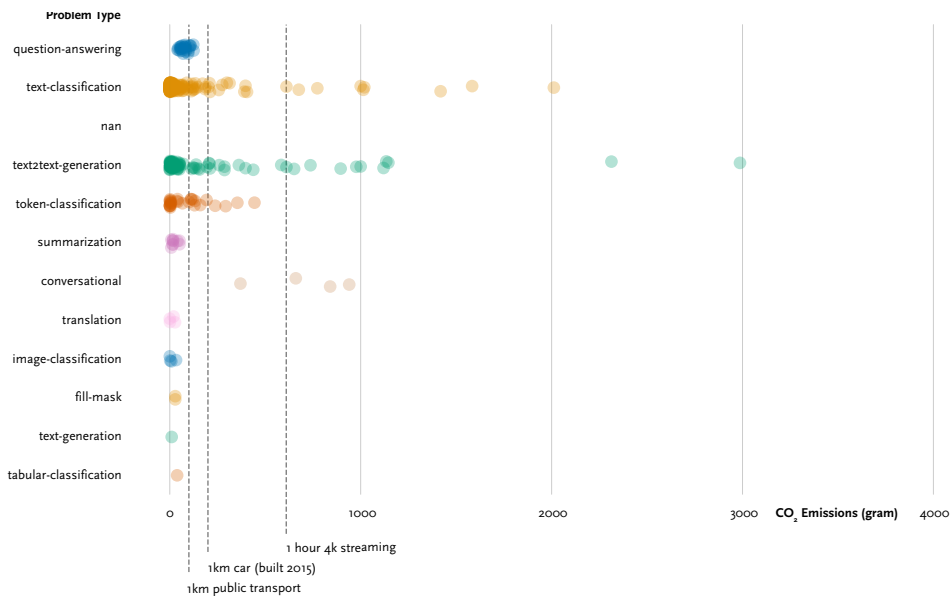
costly. For example, Strubell et al. (2019) reported that the costs for cloud computing can be around \$100,000.

SUSTAINABILITY ON HUGGINGFACE

Unfortunately, there are no reliable numbers on the energy consumption of smaller AI projects. The only exception is an experiment conducted by Marcus Voß from [Birds on Mars](#), who did this for the study on sustainability criteria for AI. We were able to reproduce and update the results (see visualisation on page 52: CO₂ Emissions of Models from Huggingface). The data source is the (self-reported) emissions by models found on [Huggingface](#), a hosting platform for AI models. These models can be downloaded and fine-tuned for the individual's own purposes, and since the models are freely available, they offer some insights into smaller AI projects.

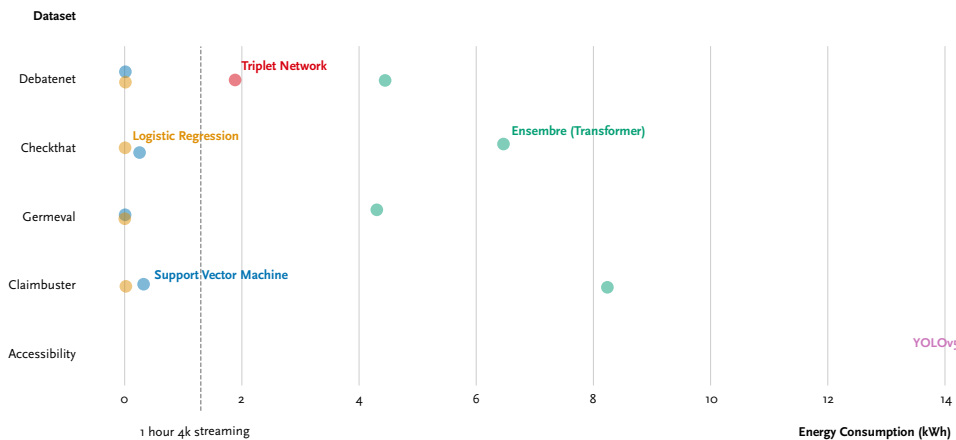
The experiment shows that the emissions for training a model are not always excessively high. Streaming in 4k, for example, has a greater impact than most of the models above. One might say that if you pass on watching the newest episode of *Bridgerton* today, you can train your Huggingface model with a clear conscience tomorrow.

However, the question is: who trained the models and where were they trained? Unfortunately, it is difficult to find out who exactly contributes to Huggingface. But this is crucial, since for our purposes, only smaller projects matter. Likewise, it is almost impossible to find out much about the geographical location of the training. And depending on the location, the emissions can vary significantly because renewable energy is much



CO₂ Emissions of Models from Huggingface

CO₂-emissions of Huggingface models for different tasks such as automatic translation or text summarisation. The numbers come from model cards. In recent times, these cards have not only been used for documenting accuracy metrics but also for documenting what the model emitted during training. The data on transportation can be found [here](#) and the data for streaming is [here](#).



Energy Consumption for Training Several Models

Electricity consumed to train several machine learning models on different datasets. For logistic regression, support vector machine and triplet network models, we used grid search for hyperparameter tuning. For YOLO, we used the already implemented hyperparameter evolution. We took the documented hyperparameters for the ensemble model. The language models were trained on four different datasets: Checkthat (approx. 47,000 sentences), Claimbuster (approx. 24,000 sentences), Germeval (approx. 4,200 sentences) und Debatenet (approx. 3,700 sentences). The computer vision model was trained on a dataset containing about 6,500 images.

more climate friendly than fossil fuel sources. This makes it difficult to compare the model emissions directly.

HOW SUSTAINABLE IS PUBLIC INTEREST AI?

For the Public Interest AI research project, we developed AI prototypes that are designed to serve the public interest. For this post, we measured the electricity consumed to train these models. Far from being representative, this is anecdotal; nevertheless, it should give some impression about what level of CO₂ emissions can be expected in a small- or medium-sized machine-learning project.

The first prototype is intended to map (in-)accessible places by object detection, which automatically recognises objects like stairs, steps, ramps and handrails (see [exemplary animation](#)). For this task, we annotated a [dataset](#) and chose YOLOv5 for the computer vision task. YOLOv5 is a widely used state-of-the-art deep neural network, and the technique for using it is called *transfer learning*, which is a pre-trained model that is fine-tuned to our accessibility dataset.

The second prototype is designed to support fact-checkers in their work in combatting disinformation. Before checking potential disinformation, the fact-checker has to find a claim to check. The purpose of this natural language model is to spot such claims automatically with a view to reducing the workload of human fact-checkers.

For this task, we tried different machine learning models. First, we used “old-school” models such as *logistic regression and support vector machines*, but we also used state-of-the-art models like an ensemble of transformers and a *triplet network*. These are instances of *transfer learning*, too.

The visualisation “Energy Consumption for Training Several Models” on page 52 shows the electricity consumption of the models. Two observations are central. First, the ensemble model has by far the highest electricity consumption. This is no surprise, since it is built from 60 individual transformers. However, second, it is striking that the electricity consumed to train these models is still only a little higher than the electricity consumed to stream one hour of television in 4k quality.

WHERE CAN MODELS BE TRAINED IN ENVIRONMENTALLY FRIENDLY CONDITIONS?

As mentioned above, depending on the geographical location, the same electricity consumption can cause different levels of CO₂ emissions. This is because of the energy mix in question, and due to the fact that renewable energy emits less than fossil fuels. An [interactive map](#) visualises the emissions of the ensemble model depending on the geographical location. Even though we trained our models in Germany, we can estimate the amount of CO₂ that would have been emitted if we had trained them in other countries.

The most striking observation is that the emissions may be several times higher depending on the geographical location, even though the electricity consumption stays the same. Furthermore, the [map](#) shows that, in some

countries, the question of whether the training happens during daylight hours makes a huge difference.

RECOMMENDATIONS AND OPEN QUESTIONS

What can we learn from this and what questions remain unanswered? Most obviously, it turns out that small AI projects come with low emissions. This does not mean that there is no room for improvement. Nevertheless, it shows that the shocking numbers presented by Strubell et al. cannot be generalised to all AI projects.

On the other hand, it becomes clear that the data situation is not good. Neither the numbers from Huggingface nor our own experiments are representative. We still need more documentation. This is as true for research as it is for real-world AI projects. To be fair, during the process of writing, we noticed that, each week, roughly 10 new models with proper documentation were added to Huggingface. It seems that sustainable AI is in fact becoming increasingly recognised. But documentation standards matter, too. As we saw earlier, emissions can vary significantly depending on the geographical location. Despite this fact, the location is only documented in a few cases. Details like this require further communication if we want to make sustainable AI a reality.

REFERENCES

- Ahmed, N., & Wahed, M. (2020). The De-democratization of AI: Deep Learning and the Compute Divide in Artificial Intelligence Research. *ArXiv:2010.15581* [Cs]. <https://arxiv.org/abs/2010.15581>
- Rohde, F., Wagner, J., Reinhard, P., Petschow, U., Meyer, A., Voß, M., & Mollen, A. (n.d.). Nachhaltigkeitskriterien für künstliche Intelligenz. *Schriftenreihe Des IÖW*, 220, 21.
- Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A. S., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C. P., Ng, A. Y., Hassabis, D., Platt, J. C., ... Bengio, Y. (2023). Tackling Climate Change with Machine Learning. *ACM Computing Surveys*, 55(2), 1–96. <https://doi.org/10.1145/3485128>
- Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 3645–3650. <https://doi.org/10.18653/v1/P19-1355>
- van Wynsberghe, A. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics*, 1(3), 213–218. <https://doi.org/10.1007/s43681-021-00043-6>



THERESA ZÜGER

A lion for sustainable AI: How to support a new standard for sustainability reporting?

The sustainability of AI is an issue that is growing in prominence – both in the research community and the political community investing in AI. However, existing research does not yet show a realistic picture of whether the different steps in an AI life cycle impact energy consumption and CO₂ emissions, and if so, by how much. What is missing is proper documentation and a standard that provides a basis for comparison and long-term research. One step towards a better standard of documentation in this and many other corporate sectors might be a new directive by the EU (CSRD). It requires an additional 2,000 big or capital-oriented companies to report on their corporate sustainability. But what kind of reporting animal is the CSRD going to be: a toothless paper tiger or a sharp and hungry lion?

[READ FULL ARTICLE](#)

LEA ERLLENWEIN

Four key ideas for a sustainable platform economy in Europe

As platform companies are changing the way business is done in many European markets, they face a number of unanswered legal and ethical questions. To confront the negative tendencies of platform businesses and ensure that they themselves foster a sustainable platform economy in Europe, HIIG hosted two roundtables in cooperation with the European Commission. In addition, a panel of experts shared their knowledge in a webinar, where they talked about possible solutions to the challenges that European platforms are currently facing. In this post, Lea Erlenwein draws upon opinions and statements voiced by participants at these events to highlight four key ideas on sustainability and fairness.

[READ FULL ARTICLE](#)



DIGITALER SALON

Wissen Macht Klima

Once a month, we publicly discuss the impact of digitalisation on society at Digitaler Salon. We invite special guests, engage in a dialogue with the audience and the online community and broadcast it on the local open channel Alex TV. As part of Open Access Week 2022, we discussed how equitable global access to knowledge really is in the case of climate justice.

[WATCH THE DISCUSSION \(GERMAN\)](#)



STEPHAN BOHN

Saving orangutan habitats – digital technology and rainforest protection in Indonesia

Digitalisation can do practically anything. Everyone is connected to everyone else, and anything can be measured by and transferred into data. But what effect does this have, especially outside of the digital world, for tangible issues such as climate change and nature conservation? To answer these questions, a HIIG research team is travelling across four continents to determine the potential of digital technologies for sustainability and development. The orangutan's desperate struggle with deforestation, fire and illegal logging in Indonesia's rainforests illustrates the challenge to find sustainable and applicable solutions in time.



Indonesia with its more than 17,000 islands has the third largest rainforests in the world, and it is quite beautiful and fascinating, with unique biodiversity and the last populations of orangutans. Indonesia's rainforests are home to thousands of plant and animal species, and millions of Indonesians depend directly on the forests for their livelihoods. Additionally, forests, and especially rainforests, are crucial in slowing down climate change as they can store carbon in their biomass. This is particularly important for Southeast Asia, which has the highest deforestation rate in the tropics ([Global Carbon Project](#)). So, the forest could be much larger and much more important for carbon sinks and local livelihoods if ever more forests were not being destroyed or irreparably altered by industrial developments such as mining, palm oil plantations and logging.

But what do digitalisation and digital technologies like drones, online platforms and AI have to do with it? Digitalisation provides a wealth of data on the rainforest as the first and most important element. These data help to monitor the growth and diversity, but also environmental destruction, and thus make both positive and negative developments visible. For example, today satellite data is analysed in real-time, and websites like [Global Forest Watch](#) use publicly available Geographic Information Systems (GIS) data and show which regions of the world rainforests are in trouble or being destroyed. No government in the world and no company can deny or conceal destruction, deforestation and slash-and-burn anymore, neither internationally nor vis-à-vis the population. Digitalisation, thus, creates transparency.

Likewise, the data collected will be used specifically from the field by wildlife and biodiversity conservation projects to protect and safeguard tropical rainforests. One example is the Borneo Nature Foundation (BNF), which uses GIS data to prevent fires in the rainforest at an early stage. “The potential of digital data and technology is huge because it can provide faster and more accurate data that can then help firefighters identify fire sources early”, says BNF’s Restu Aminullah. The BNF GIS coordinator is also collecting his own data on the ground using drones, which can also identify underground fires early through thermal imaging. Fires are a serious problem for the rainforest. Peatlands drained for palm oil cultivation, for example, catch fire easily and contribute to the spread of fires. In 2015 alone, an area the size of Saxony was destroyed by fire (Global Carbon Project). Such projects on the use of digital tools can set an example not only for Indonesia but also for Europe, which has been struggling with forest fires more and more in recent years.

Additionally, platforms such as the SMART – Spatial Monitoring and Reporting Tool are also being used to identify problems in the rainforest. Such tools provide a platform that conservation projects and local communities can use to collect, measure and evaluate data to improve the effectiveness of wildlife conservation efforts. It combines local knowledge of the population with GIS data, as local data about the health of the forest are collected with portable devices such as cell phones, which are increasingly common even in remote locations. Furthermore, social media posts are integrated to find out where crises, fires or deforestation are occurring.

Speaking of platforms and social media, the major platforms in Indonesia, such as Facebook and Instagram, play a critical role as news channels that should not be underestimated. Thus, they help to disseminate information that is otherwise difficult to access. The topic of rainforest protection plays almost no role in Indonesia's mainstream public discourse, for example, in the major television channels. However, bloggers and investigative journalists are working to change that. They are preparing green issues for a young and tech-enthusiastic public and disseminating them through social media. Websites and initiatives such as [kok-bisa](#) or [telusuRI](#) are working hard to raise awareness about the rainforest and create a new public that would be unthinkable without the internet and cell phones.

Of course, there are still many challenges, the use of data requires knowledge and intensive training, and more importantly not everywhere has internet or even electricity, especially when it comes to remote areas in Sumatra, Borneo or Papua. But even for this, there are solutions, for example, devices that collect data offline and then forward it regularly, so that the local communities can also be involved in the protection of nature, because they probably have the greatest interest in preserving the rainforest and protecting their habitat.

More about how rainforest conservation on the ground works with digital technologies can be found on the website of the project [Sustainability, Entrepreneurship and Global Digital Transformation](#) in cooperation with the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ).

FURTHER PROJECTS

- [BNF](#)
- [Hutan Harapan](#)

On global forest protection:

- [Global Forest Watch](#)
- [Global Forest Resources Assessment 2020](#)

GIS initiatives:

- [NASA Landsat data](#)
- [Sentinel data](#)



FREQUENTLY ASKED QUESTIONS

Digitalisation and Sustainability

How do we define (ecological, social and economic) sustainability?

The term sustainability is an (ideal) state and has been considered a guiding principle for economic, political and ecological action for many years. The sustainable development of our society, for example, refers to a situation in which our current needs can be met in such a way that future generations are not restricted. This is about acting in a future-oriented way, taking long-term effects into account and preserving the basis of human life. This applies to the sustainable use of the earth's natural resources (environment), peaceful coexistence with equal opportunities (people) and companies that take on social and ecological responsibility (economy).

FREQUENTLY ASKED QUESTIONS

How is sustainability related to digitalisation?

Digitalisation offers opportunities and risks for sustainable development. For example, the introduction, expansion and restriction of technologies usually has social and ecological consequences. So, whether they are sustainable or not depends on the respective framework conditions in which they are produced, used and ultimately disposed of. Furthermore, digital technologies can be used to demand and promote sustainability in our society: for instance, through more efficient use of resources as well as global communication and coordination. Therefore, it is essential to think of digitalisation and sustainability together.

What does digitalisation mean for the environment/nature?

The production and application of digital technologies requires a considerable amount of materials and energy. For example, the production of digital devices and their batteries consumes numerous rare metals. However, digital technologies also offer a wide range of opportunities, for example to optimise processes, to use resources more efficiently and to measure and monitor the health of ecosystems. Furthermore, digital technologies are an indispensable tool for the all-important communication on the climate crisis.

INTERVIEW

Artificial intelligence
and the human

FURTHER ARTICLES

Reclaiming digital autonomy
The age of machine autonomy?

LECTURES

On being autonomous in
the uncanny valley

Autonomy, autonomous systems and robots

Au

Su

Digitalisation and
sustainability

Wo

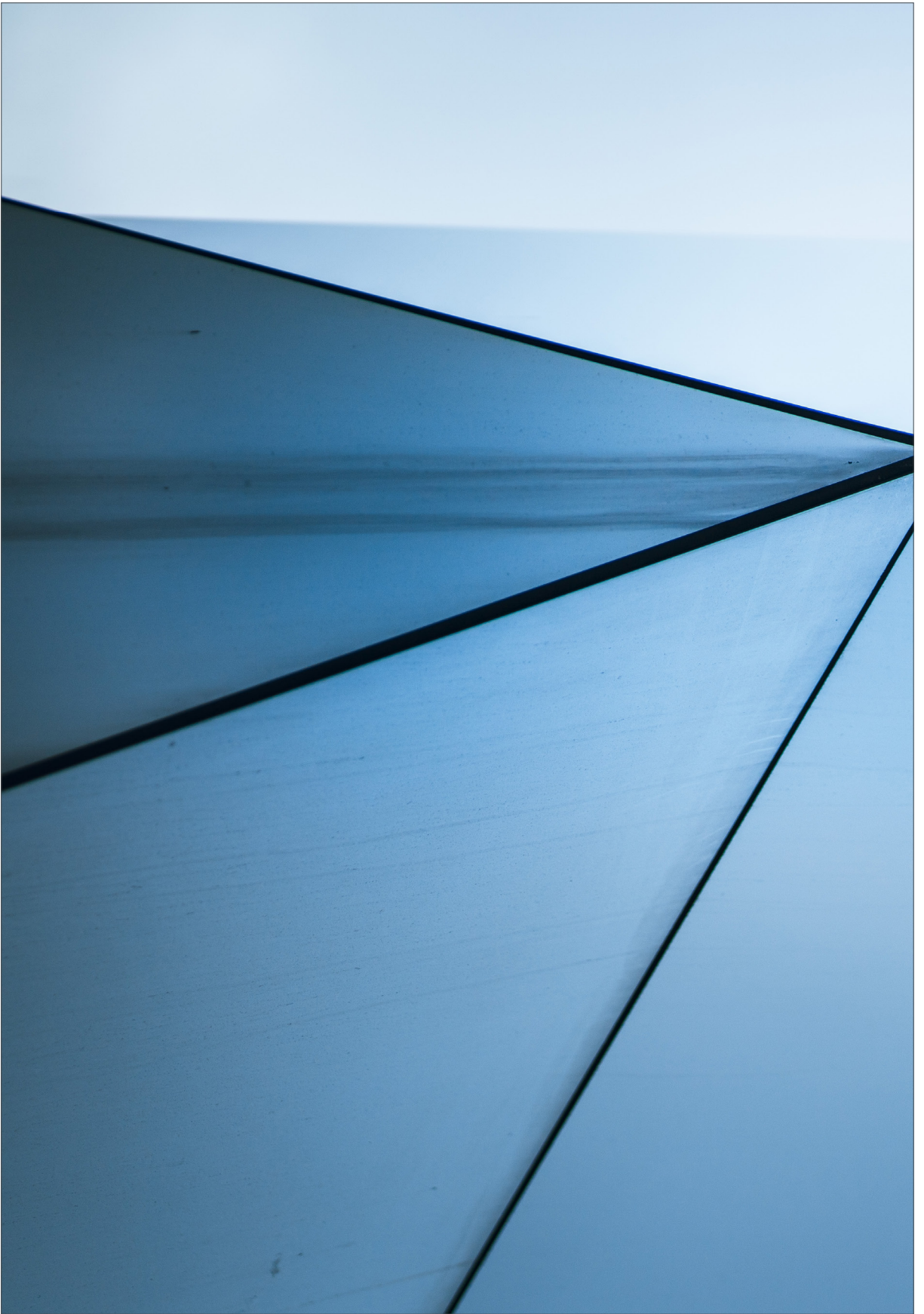
Digital organising and
the future of work

FOCUS

Autonomy, autonomous systems and robots

Self-driving cars, humanoid robots or weapons independent of human control – these so-called autonomous systems demonstrate the enormous influence of artificial intelligence (AI) on our society. They make decisions, move independently and interact with us almost like social beings. In our research, we investigate the complex and dynamic relationship between humans and intelligent machines. Autonomous systems challenge us to question the supposedly reliable notions of autonomy, intelligence, consciousness and control.

[VISIT TOPIC OVERVIEW ONLINE](#)



INTERVIEW WITH THOMAS CHRISTIAN BÄCHLE
AND YUKIE NAGAI

Artificial intelligence and the human

Imaginations about *the human* and technologies are far from universal – they are culturally specific. This is why a cross-cultural comparison is crucial for better understanding the relationship between AI and the human and how they are mutually constructed by uncovering the aspects regarded as natural, normal or given. Focusing on concepts, representations and narratives from different cultures, a conference on cross-cultural perspectives on science in fiction addressed two axes of comparison that help us make sense of the diverse realities of artificial intelligence and the ideas of what is human: science and fiction, East Asia and the West.

The conference Artificial Intelligence and the Human was held from 11 to 13 May 2022 in collaboration with Japanese-German Center Berlin (jdzb) and Waseda University. In preparation, jdzb conducted an interview with Thomas Christian Bächle, head of the research programme Evolving Digital Society at HIIG, and Yukie Nagai, project professor at the International Research Center for Neurointelligence at the University of Tōkyō, on the conceptual understanding of AI and the human.

Dr. Bächle, how do cultural representations of artificial intelligence (AI) and the actual conceptual and technical advancement of AI influence each other?

Thomas Christian Bächle: In fact, these two areas cannot be so clearly distinguished from each other. At the beginning of a technical development, there is always an idea, or at least a rough idea, of what function, what benefit or – more generally speaking – what meaning a technology has or should have. Especially in the case of developments that people like to call “revolutionary” – i.e. those that are supposed to be categorically new – references can quickly be made to cultural texts that have anticipated certain technologies. These concepts, developed and tried out in fiction, so to speak, are translated into scientific innovations and into research and development goals. They also take their inspiration from popular culture, such as comics, film, television. At the same time, however, the aspects discussed in politics, journalism, religion or ethics about possible future scenarios are also part of the imaginaries that exist around a technology.

How do these imaginary and cultural representations of AI in turn affect political and social discourses about AI – for example, in Japan and in Germany?

Thomas Christian Bächle: They have the important social function of making the hopes, opportunities or dangers associated with AI visible and understandable for as many people as possible. The crucial thing here is to open up a space for debate. Imagination therefore not only influences research and development but also guides political decisions, journalistic reporting and public opinion. Such visions are thus not mere entertainment but are always an essential and formative part of the reality of technologies. This is especially true for AI, because the technology associated with it is imagined as a form of intelligence that is similar to that possessed by humans but at the same time independent and autonomous of humans; it is viewed as an intelligence that is equal or even superior to ours. This naturally stimulates the imagination, in a broad spectrum from cooperation to competition, from AI as the solution to all our problems to AI as the feared instigator of the demise of humanity. In the country-specific view, imagined visions can of course not be understood independently of certain larger discourses and interests. The so-called Japanese Robot Culture, for example, not only corresponds to the self-image of many developers and companies. The Japanese government has deliberately used this buzzword as political branding, which, it is supposed, will be useful for it as a business location. In Europe, it is readily adopted: What can we learn from the Japanese? What developments in Japan should we be wary of? After giving lectures on robots in Japan, I am often asked whether we should delegate nursing work to the cold, emotionless machine, as is supposed to be the case

in Japan. This is a gesture of demarcation and fulfils a rhetorical function. Robots in Japan – they are then something special for both Japanese people and Europeans, and for both sides it serves its purpose.

In your research, you compare Japanese and European concepts of human-like robots. How is “humanity” in AI defined in the respective discourses, or are there apparent significant conceptual differences?

Thomas Christian Bächle: In comparative research, you have to be very careful not to tell stories about the exotic country in the Far East, especially when you are looking at these phenomena through European eyes. It is popular, for example, to always use Shintoism to describe the Japanese relationship to technical artefacts: for the Japanese, everything is animate, so it doesn't make much difference whether it is a robot or a human being. This view is a bit simplistic. Of course, there are very specific contexts in the history of ideas and different patterns in dealing with robots that can be recognised. This additionally becomes clear in interactions with robots and not just in imaginaries about them. A developer once told me that he has different experiences depending on whom he presents his robots to. While people from Japan mainly enjoy interacting with the machine, Europeans want to prove to the robot and its developer as quickly as possible that the machine is not really intelligent after all, that it has no consciousness, feels nothing, etc. This reflex certainly has to do with a European understanding of robots and of how humans are conceived, namely as rational, sentient, indivisible but also special individuals with autonomy and unique consciousness. Many fears and uncertainties formulated in Europe probably also stem from this

particular conception of humans. If you like, the Japanese attitude is more inclusive, less focused on the uniqueness of the human being. These many levels of meaning are a great challenge for research: What does AI mean in different cultures? What interpretations are triggered specifically, and for whom do which images of AI serve what purpose? It is enormously difficult to distinguish this without constructing cultural differences.

Dr. Nagai, you conduct research in the field of neurointelligence on cognition in robots - and in humans. What exactly is the subject of your research?

Nagai Yuki: My research aims at understanding the development principles of human cognitive function using a computational approach and thus at designing a support system for those with developmental impairments. Humans acquire cognitive abilities in the first few years of life, but it is unclear how the brain and body implement them. In contrast to AI, human intelligence is open ended, and humans acquire multiple cognitive abilities cooperatively and continuously, producing individual and group diversity. In experiments using a neural circuit model mimicking the human brain in a humanoid robot, I have been able to verify the neural network underlying this continuity and diversity of cognitive function. Such research, which until now has been largely empirical, can also be used to support people with developmental impairments.

How are cognition in humans and in robots related? What surprising findings were you able to make in your research?

Nagai Yuki: One result is the unified account of cognitive function development based on the theory of *predictive coding*. The brain is a *predictive machine* that combines bottom-up sensory signals with top-down prediction signals generated by the internal model acquired through experiences and seeks to minimise prediction error in these signals. I proposed that cognitive functions, such as self-other cognition, imitation, and prosocial behaviour, are acquired through the prediction-error minimization process and I have demonstrated this in robot experiments using neural circuit models. I found that an imbalance in combining sensory and predictive signals is potentially the cause of the developmental impairment found in autism spectrum disorder, etc. These results show that two aspects of cognitive development – continuity and diversity – can be explained uniformly based on predictive coding theory and have important implications for developmental (impairment) research. How is the relationship between humans and robots evolving as technology advances? By developing AI and robots that reproduce human cognitive functions, we can better understand human intelligence. The concept of neurodiversity was proposed in the late 1990s, and the idea of seeing developmental disorders as individuality that manifests as normal fluctuations in the neural architecture became more prevalent. It is difficult to evaluate what kind of individuality a person has, and even the cognitive characteristics of the self, especially in people with developmental impairment, are difficult to grasp. But, if robots can develop and learn like humans and reproduce and acquire human cognitive characteristics, human intelligence can be understood through the medium of robot intelligence.

We hope that the development of such artificial intelligence technologies will enable us to bring about a symbiotic society that makes use of the individuality of people and robots.

JAN SEBASTIAN GÖTTE

Reclaiming digital autonomy

In today's digital age, all of us depend on decent treatment from a number of large companies: we rely on them to guard our data, or at least to not abuse it too much without asking. Even if we wanted to, we cannot take things into our own hands and reclaim our digital autonomy, because too much depends on too many system operators. To help small-time users and systems administrators to fight this dependence, Jan Götte and Björn Scheuermann from HIIG have developed inertial hardware security modules, or IHSMs. IHSM-equipped low-budget servers provide high security even outside of well-protected data centres and enable us to regain our digital autonomy.

[READ FULL ARTICLE](#)

THOMAS CHRISTIAN BÄCHLE

The age of machine autonomy?

Can machines be autonomous – or is it a human prerogative? This categorical question dominates many discussions on our relationship to (the allegedly) intelligent machines. A human vs. machine rhetoric, however, doesn't get us very far. On the contrary, it even side-tracks the more important issue of how these autonomous systems change the ways we humans relate to the world at large and to each other. In a much broader sense, robots – as simulated humans in particular – are bound to shake up the fundamentals of what we regard as human autonomy.

[READ FULL ARTICLE](#)



KEYNOTE BY BEATE RÖSSLER

On being autonomous in the uncanny valley

From 23–25 Nov, HIIG hosted the Infrastructures of Autonomy conference together with Brandenburgisches Zentrum für Medienwissenschaften. The conference addressed the conditions, structures and relations that constitute both human and machine autonomy, given that autonomy has been a multifaceted term for centuries and has gained a renewed relevance in the context of technical innovation. The event was opened by philosopher Beate Rössler with a keynote on the topic of autotomy and its political infrastructures.

[WATCH FULL LECTURE](#)

FREQUENTLY ASKED QUESTIONS

Autonomy, autonomous systems and robots

What is an autonomous system?

In the context of AI, the term “autonomous” is mostly used to describe a computer-based system that performs complex tasks and to do so independently of human influence or control (to varying degrees). The individual steps in solving tasks are no longer prescribed or monitored by humans, but are chosen solely by the machine. For example, when a so-called autonomous vehicle is given a destination, the machine determines by itself the route, the speed or the driving manoeuvres it must execute according to the situation on the road.

What is the difference between autonomous and automatic systems?

The autonomous characteristics of a system are often understood as the next step towards automation, yet automatic systems originally were essentially about thinking ahead to all contingencies in the programming of a system. They can perform complex tasks, but they still follow the rules that humans have established. In other words, automatic systems are pre-programmed. Autonomous systems, on the other hand, are able to abstract patterns from unknown data (“machine learning”). They are programmed in such a way that they can also adapt step by step to what has not been determined in advance. This gives them the ability to adapt and act on their own.

FREQUENTLY ASKED QUESTIONS

What does the materiality of AI mean for the human influence on the autonomous system?

Autonomous systems are never fully self-driven (they are programmed by humans) or self-sustaining (they depend on energy sources and technical infrastructures). The term “autonomous system” is therefore somewhat misleading. The autonomous system has its own materiality (its hardware), which essentially determines what the system can do. Materiality also plays a major role in terms of the extent of human control: interfaces (screens, switches) determine how we interact with the system (speaking, gestures). These interfaces also offer us options for how to control the system – or can withhold such options from us. All these interactions determine to what extent humans are in the loop: Are humans part of the decision or not? When can humans intervene?

What cultural factors shape our understandings of AI?

AI is not a purely technical or abstract idea, rather the very notion is embedded in complex social and cultural contexts. The concept of autonomy alone – a concept that has been constantly renegotiated for centuries – shows how much our understanding of AI is shaped by European intellectual history. In addition, societal expectations and fears around AI are strongly linked to fictions and imaginaries (e.g. the representation of robots in popular culture). These cultural images have a strong influence on how we experience AI. Cultural comparisons, the kind we draw in our research, therefore sharpen our view of AI.

KUTOMA J. WAKUNUMA
AND TOM KWANYA

**Towards a socially just
gig economy in Kenya:
Stakeholder engagement
and regulatory processes**

FURTHER ARTICLES

**Shaping AI in the
interests of employees**

**Using AI for social good
in entrepreneurship:
promises and perils**

**The anywhere jobs are nowhere
near – how remote work is
moving towards the city**

DISCUSSION

**Workers' rights and
platforms' responsibility**

Digital organising and the future of work

Wo

Au

Autonomy, autonomous
systems and robots

In

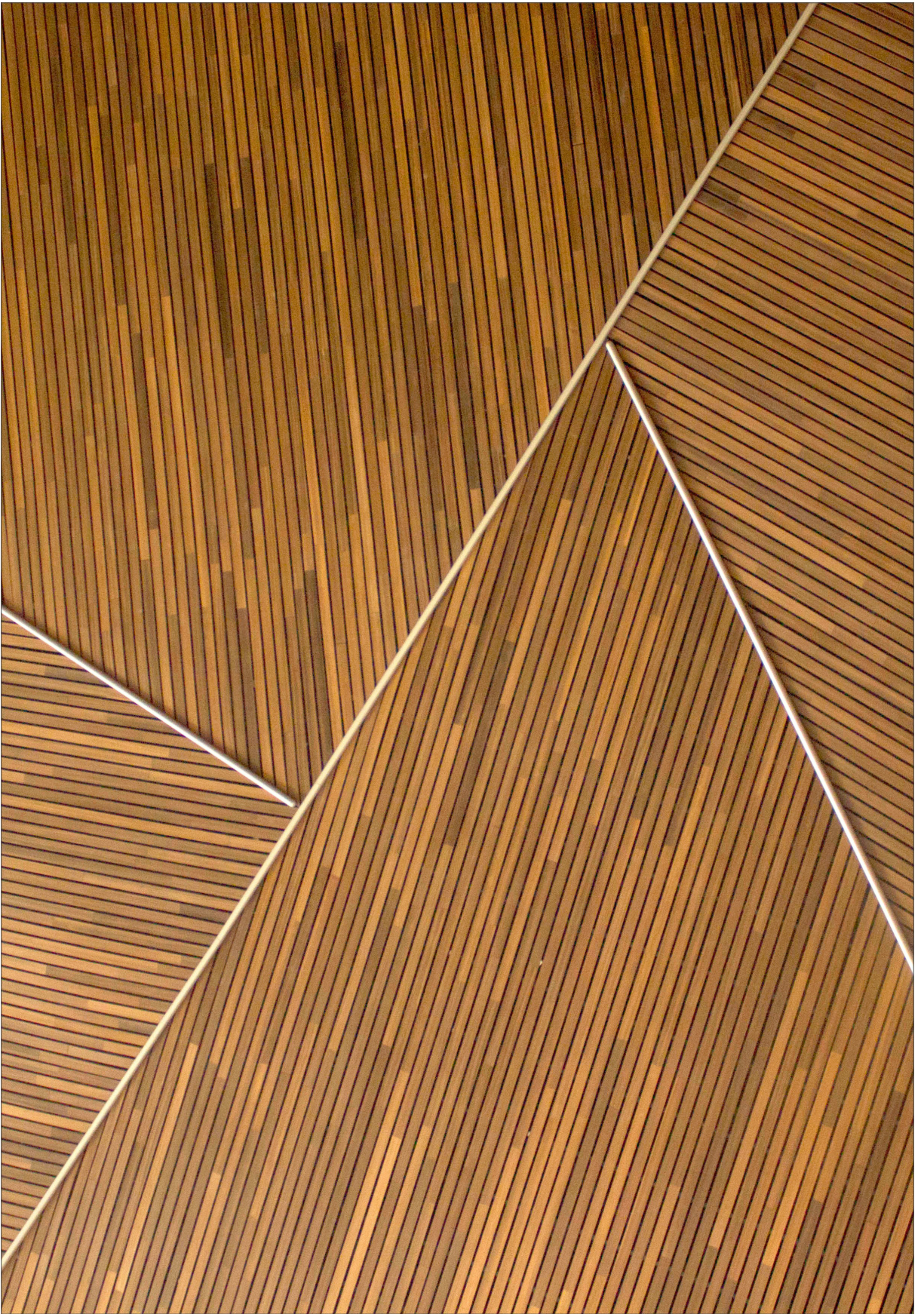
Digital infrastructures of
the knowledge society

FOCUS

Digital organising and the future of work

Digital organising encompasses a range of phenomena: from artificial intelligence (AI)-powered organisational decision-making to the global networking of workers in real time. Corporate boundaries are becoming more porous; work is becoming more fluid and less dependent on location and time. Therefore, the future of work is characterised, on the one hand, by spatial and temporal flexibility and, on the other, by the increasing interaction between people and technology. Through digitalisation, work is being replaced, relieved and restricted, or new work is being created.

[VISIT TOPIC OVERVIEW ONLINE](#)



KUTOMA J. WAKUNUMA AND TOM KWANYA

Towards a socially just gig economy in Kenya: Stakeholder engagement and regulatory processes

Digital platforms are fundamentally changing the world of work. At the click of a button, we can order food or services online to our doorstep in the so-called gig economy. While the platform economy opens immense opportunities for flexible, gainful and convenient entrepreneurship, the precarious livelihoods of workers and service providers often remain unaddressed. In particular, workers from economically developing countries are often subject to repetitive gig work, low levels of job security and high exploitation risks. Kutoma J. Wakunuma and Tom Kwanya have studied the livelihood of Kenyan gig workers, which allows us to better understand the perks and perils of the gig economy in the Global South.

About five years ago, getting a taxi in Kenya meant either standing on the roadside and waving one down or wandering around the town centre looking for taxi bays. Similarly, food delivery was limited to hotel residents who could request room service. Everyone else had to pop into a restaurant and order takeaway. Today, virtually all such basic services are available at the click of a button on digital platforms. Behind these platforms are thousands of people who deliver the services to clients conveniently and cost-effectively. Besides bringing services closest to the clients, these platforms also open immense opportunities for flexible, gainful and convenient entrepreneurship.

KENYA'S 1-MILLION-STRONG GIG WORKFORCE

A 2019 report by Mercy Corps put the number of online gig workers in Kenya at 36,500. However, there is evidence that the gig economy model in Kenya makes it difficult to count many online gig workers. For instance, most online gig work accounts have been set up by individuals who employ several others to deliver the services. Similarly, some online gig workers deliberately hide their identity for legal and tax reasons. Therefore, stakeholders place the realistic number of gig workers in Kenya at about 1 million¹. The stakeholders, including policymakers, civil society members, donors and platform owners also emphasised that the number of online gig workers in the country is growing exponentially. Indeed, the Federation

¹ The estimations were made during a multi-stakeholder dialogue event hosted by HIIG and GiZ in Nairobi, Kenya on 9th November 2022

of Kenya Employers (FKE) has projected that in the not-so-distant future, more workers in Kenya will perform multiple types of online and offline gig work with varying levels of formality and high levels of flexibility rather than rigid, formal employment, which is declining in popularity.

KENYA'S GIG ECONOMY – A PRECARIOUS AND EXPLOITATIVE FORM OF WORK?

Many online workers are attracted to the gig economy because it accords them an independence that they would not have in formal employment. For instance, they can choose who to work for, set their own working hours and work remotely. Despite these advantages, online gig workers face many challenges, including stiff competition for online jobs; a lack of a stable income, which makes them ineligible for financial services such as loans or insurance; a high cost of doing business, since gig workers often have to cover the costs of their own equipment, transport and other overheads that would otherwise have been borne by an employer; a lack of job security, since their platform accounts can be suspended or deactivated without notice in the event of customer complaints; and a lack of a safety net, such as health care benefits and pensions. The majority of these challenges can be addressed through regulation to ensure socially just gig work. Unfortunately, while Kenya has labour laws that aim to protect workers, these laws are tailored to the formal employment sector. The International Labour Organization (ILO) acknowledges that regulating informal labour presents a challenge for most countries because existing laws are often incomplete, too vague, out of date, or unclear regarding the definitions

of the players in the industry, thus creating loopholes for the exploitation of workers. The lack of an appropriate regulatory framework for informal online workers exposes them to high compliance costs due to multiple licensing demands. Similarly, they are vulnerable to harassment, unclear terms of recruitment, punitive dismissals and unfair work environments.

SHEDDING LIGHT ON THE LIVELIHOODS OF KENYA'S GIG WORKERS

Empirical evidence from a recent study of 314 online gig workers by the Alexander von Humboldt Institute for Internet and Society, in cooperation with the Digital Transformation Center Kenya of the Gesellschaft für internationale Zusammenarbeit (GIZ), revealed weaknesses in the regulatory mechanisms for the online gig economy in Kenya. It emerged that most (89.8%) of the gig workers in Kenya regard the gig work as their primary source of income while a minority (10.2%) use it to generate supplemental income. However, the majority (75.8%) of the gig workers earned low wages, ranging between 10,000 and 30,000 Kenyan shillings monthly (80–240 EUR). Furthermore, most of the gig workers (38.5%) indicated that the income was inconsistent and varied from week to week. Although the majority (81.5%) of the respondents said that they had a contract with the platforms they worked on, 18.5% of the respondents had no contract. Similarly, most of the respondents (93.0%) were aware of the terms and conditions of their engagement but some (7.0%) of them did not know their terms of engagement. Additionally, 13.0% of the respondents were not aware of any policies on the platforms they used. Importantly, most of the respondents (60.8%) were not aware of any regulations used to manage

gig platforms in Kenya. In terms of regulations, most of the respondents (56.0%) indicated that the Government of Kenya is best placed to regulate gig platforms with the support of the other stakeholders, including donors and investors, civil society organisations, and the platform owners or employers. The findings of the study also revealed that most of the gig workers (50.3%) felt that the stakeholders did not support the gig economy. There was generally a low awareness of labour unions and welfare societies among online gig workers.

HOW TO ENSURE JUSTICE AND FAIRNESS IN KENYA'S GIG ECONOMY

These findings show that the Kenyan gig economy is controlled by platform owners who operate using a business model that is steeped in making profit. This model works to benefit the employer the most. For example, Uber drivers are monitored using GPS technology, but the drivers themselves cannot monitor their employer. Also, the platform dictates the terms of transactions for the worker. The gig platforms also allow customer ratings of employees rather than direct supervision by the employer. Furthermore, most of the gig workers' employers have the capacity to set wages, which is influenced by factors like the number of gig workers on the platform, the rate of unemployment and work preferences. There is a need to regulate the platforms as a means of guaranteeing justice and worker welfare. Given that platform owners would naturally be interested in having a say in regulation, the other stakeholders should take the lead in fostering a regulatory regime that balances the needs of the online gig workers with the business interests of the platforms. The Government of Kenya should recognise gig work

as employment and gig workers as employees whose rights need to be protected legally. Similarly, the Government of Kenya, in collaboration with the other stakeholders, should review, revise or update the legal and policy frameworks governing terms of employment, workers' rights and welfare to include the interests of the burgeoning number of online gig workers in the country. Gig workers in Kenya should be sensitised to their rights as employees and should be encouraged to join unions that can lobby and advocate for their rights. Civil society organisations and unions should spearhead initiatives in this regard.

STEPS TOWARDS A SOCIALLY JUST GIG ECONOMY IN KENYA

The Government of Kenya should also identify and implement strategies that diversify and promote gig work as an alternative income generation mechanism for both employed or unemployed citizens. Civil society organisations and unions alike should provide affordable or free legal advice or representation to gig workers whose rights are violated but who have no means of seeking legal redress. The judiciary of Kenya should create a section for gig workers under its Labour and Employment Division. This would strengthen the capacity of the judiciary to handle disputes emerging from gig work competently and expeditiously. Programmes for capacity building should expand their horizons to include mentorship, hand-holding and acceleration mechanisms to create and strengthen the practical skills needed by the gig sector.

ABOUT THE STUDY

The findings presented in this article are based on a study conducted by the authors as part of their research in the Sustainability, Entrepreneurship and Global Digital Transformation (SET) project in Kenya in 2022. It is going to be published in spring 2023.

The SET project is organised by HIIG in cooperation with the Federal Ministry for Economic Cooperation and Development (BMZ) and the Deutsche Gesellschaft für internationale Zusammenarbeit (GIZ). In Kenya and multiple other countries, HIIG supports the BMZ's Digital Transformation Centers (DTCs) as a scientific partner and carries out exchange and research formats in partner countries.

GEORG VON RICHTHOFEN,
SONJA KÖHNE & HENDRIK SEND

Shaping AI in the interests of employees

In Germany, a variety of AI applications have been used in the context of knowledge work to date: from transcribing and indexing texts to identifying hate speech in comments, up to organising accounts of customers. AI offers opportunities and risks for employees. What can managers and works councils do to enable potential positive effects and avoid negative ones? Based on a handbook about AI in the context of knowledge work, this article outlines fields of action and approaches for an employee-oriented design.

[READ FULL ARTICLE](#)

NICOLE SIEBOLD, GEORG VON
RICHTHOFEN & ALI ASLAN GÜMÜSAY

Using AI for social good in entrepreneurship: promises and perils

Artificial intelligence (AI) is seen as a new breeding ground for entrepreneurial opportunities. While research agrees that AI has great potential for economic growth, one question remains: How can AI be used for social good? A study interviewed European entrepreneurs and found that the application of AI for social entrepreneurship comes with both specific promises and perils. This blog post briefly summarises the key findings of the study and provides important recommendations for start-ups.

[READ FULL ARTICLE](#)

FABIAN STEPHANY

The anywhere jobs are nowhere near – how remote work is moving towards the city

Remote work is here to stay. But while many of us could work from anywhere, the urbanisation of digital work continues. Cities are becoming the mega-hubs for online-mediated work. But what will this change bring to our cities and the rural regions that are left behind?

[READ FULL ARTICLE](#)

PUBLIC PANEL DISCUSSION

Workers' rights and platforms' responsibility

Online platforms play an increasingly important role in today's digital economy. One example of the growing relevance of the platform economy is online labour markets. These two-sided platform marketplaces allow buyers and sellers of online mediated work to connect around the globe and transact tasks in various occupations, such as graphic design, translation work or software development. Platform labour markets deliver value to customers and provide economic opportunities for suppliers. While these markets can be a means for economic empowerment and growth, they also generate substantial economic benefits for the platforms themselves and often introduce novel and precarious forms of work and economic interaction. Hence, the emergence of these digital marketplaces calls for novel and sustainable approaches to regulation.

HIIG organised a multi-stakeholder dialogue and a public panel discussion with renowned experts on the challenges and opportunities towards a sustainable regulation of Kenya's online labour markets.

[WATCH FULL DISCUSSION](#)

FREQUENTLY ASKED QUESTIONS

Digital organising and the future of work

What are the effects, advantages and disadvantages of digitalisation on the world of work?

Digital technologies can be used for increasingly complex activities in the world of work. For example, AI applications have already been used for years in the service sector and increasingly also in knowledge work. This offers opportunities and risks for employees. For example, digital innovations can relieve workers of strenuous, dangerous and monotonous tasks. They can also perform tasks in certain areas that exceed human capabilities. However, the use of digital technologies can also lead to a restriction of employees' scope of action or even to a loss of relevance of their work in the context of automation. In addition, the increasing mediation of work via platforms goes hand in hand with the emergence of new employment relationships. These offer opportunities, such as greater flexibility for workers, but also involve risks, such as a lack of security.

FREQUENTLY ASKED QUESTIONS

What do we mean by the future of work?

The future of work is a collective term for a field of research that deals with current and future changes in the world of work. We assume that some of the formative developments for the future of work are already emerging today in certain areas and can be empirically investigated accordingly. Examples include the use of AI in knowledge work, the mediation of work through platforms or the management of employees by algorithms.

What is digital organising?

A variety of digital tools and technologies, such as AI, blockchain and big data, have driven the development of new forms of digital organising. Digital organising encompasses different areas: for example, the use of data for AI-powered automated decision-making, but also the global real-time coordination of diverse stakeholders to organise and combat social challenges, such as in the context of global climate protests or in open-source networks.

INTERVIEW WITH
KATERYNA SUPRUN

Digital transformation ahead of schedule – How Ukrainian universities are surviving and transforming in times of war

FURTHER ARTICLES

What the COVID-19 experience has taught us about the limits of evidence-informed policy making

DEFINING CONCEPTS

Data justice

Artificial emotional intelligence beyond East and West

TOOLBOX

Future(s) of the digital world

Digital infrastructures of the knowledge society

In

Wo

Digital organising and the future of work

FOCUS

Digital infrastructures of the knowledge society

Digital infrastructures refer to the use of digital innovations to create, share and disseminate knowledge. In our everyday lives, we use technical tools and organisational systems such as digital platforms to communicate, collaborate and circulate information. Thus, digital technologies play an increasingly crucial role in areas such as research, education, business or administration. In our research, we investigate how these digital infrastructures can be used for our society to share and make accessible the knowledge we generate.

[VISIT TOPIC OVERVIEW ONLINE](#)

INTERVIEW WITH KATERYNA SUPRUN

Digital transformation ahead of schedule – How Ukrainian universities are surviving and transforming in times of war

Several universities in Ukraine have been either completely destroyed or damaged. Most recently, the central building of the Taras Shevchenko National University, located in the centre of the Ukrainian capital Kyiv, was damaged by a nearby explosion. At the same time, the interview itself illustrated the dramatic circumstances of this process: the online talk was disrupted by air raid siren alerts. Kateryna Suprun, the acting head of the Digital Transformation of Education and Science Expert Group at the Ministry of Education and Science of Ukraine, discussed possible ways to support the Ukrainian academic system.

At the Open Space event of the Stifterverband für die Deutsche Wissenschaft on 28 October 2022, Christoph König from the Stifterverband and Nataliia Sokolovska from HIIG conducted an interview with Kateryna Suprun of the Ukrainian Ministry of Education. Providing insights on how universities in Ukraine are dealing with the current situation, Suprun also elaborated how the EU and Germany could help rebuild the research infrastructure. This interview was posted both on *Elephant in the Lab* and *Hochschulforum Digitalisierung's* blog.

Christoph König: Katerwyna, could you give us some insights into the current situation of universities in Ukraine, especially with regards to occupied and embattled territories in the East?

Kateryna Suprun: Most importantly, right after the war started, all academic institutions were advised to go on several weeks of vacation in order to gain a bit more clarity on the situation and how it will change and how to deal with it. The vast majority of them continued teaching in a blended mode, mostly through online teaching. The details really depend on the region and institution. If we take the Eastern regions, then, of course, it's out of the question, they simply are not able to teach face-to-face. But in the Western part of the country, those institutions that have the resources to provide an air-raid shelter and additional security measures started to work in usual face-to-face mode from April or May. We also have to understand that the situation is changing quite rapidly. And yes, the biggest challenge we have to deal with during this academic year is safety concerns, but with switching to the online mode, a number of other problems popped up.

Christoph König: What kind of challenges appeared when switching to the online mode?

Kateryna Suprun: After two years of the COVID-19 pandemic, of course, we understand that online teaching has become more common in our universities than before. But some problems still remain. Unlike school education, in higher education institutions we do not decide on the content of the curriculum in a centralised manner. Still, we have to ensure that the standards of higher education are met and that the learning outcomes are obtained during the course of study. When we're talking about schools during the pandemic, the "All Ukrainian Online School" was developed – a platform with several thousand video lessons on different school subjects covering all grades from 5th to 11th. We simply cannot do something similar for universities, because the curriculum and the programmes differ a lot. And even in times of war, we want the universities to enjoy this level of autonomy, so we do not aim to bring them all to teaching the same stuff.

What we did is this: We made some deals with international teaching platforms like Coursera, Udemy and EdX; they provided a number of free-of-charge courses that are typically not free of charge to all our students. At Coursera, for example, we have 20,000 unique users of the platform who have completed at least one course. We are trying to engage more and more students to spread the news to actually make sure that these opportunities are available; there is a dedicated team that works on the deployment of all these online courses. The second thing we are doing is to make sure that there are enough technical tools available for teachers, university staff and students so that they can teach and learn not only with

the help of prerecorded, delivered and developed resources, but also if they want to provide synchronous online teaching on their own. We've reached a number of agreements with Zoom, Google and Microsoft, who provided software solutions and who are running a number of webinars to make sure that everyone at universities knows how to use these instruments. We have not had much success yet in equipping higher education institutions with laptops, computers and web cameras. We do have huge support in this sense from partners and donors for schools, but with higher education, it's a bit different, because it's not compulsory.

Finally, a fundamental problem is the lack of understanding of the general situation. How much infrastructure was destroyed? What's the cost of this destruction? How many students and teachers have gone to different regions of Ukraine and how many of them have gone abroad? On this matter, we've recently conducted a survey together with UNICEF and will be looking into the details more thoroughly soon.

Nataliia Sokolovska: Could you give us some general insights on the status quo of the research infrastructure and what's happening to maintain it at the moment? Maybe some laboratories are being evacuated?

Kateryna Suprun: At the beginning of the war, we launched the interactive map "Education under attack" where you can see how many schools, kindergartens, universities and other educational institutions have been either damaged or destroyed.



Destroyed Kharkiv National University documented on the interactive map "Education under attack"

In most cases, the university gets evacuated and remains formally a separate entity, but all staff and students are relocated and temporarily based at the premises of other universities. This was the case, for example, with Mariupol State University, which is now relocated to Kyiv.

Anyhow, deciding what to do always depends on the situation, so it's a permanent "work in progress" situation and occurs on a case-by-case basis, so no generalisations are really possible here. We definitely need to change our funding mechanisms, because in the last few years, we applied the formula "the more students you have, the more funding from the state you receive". For many reasons, this is not how it should be, and two years ago we diversified this formula a bit and made the state funding dependent on the quality of the "services" these universities are providing. We introduced performance-based funding that is based not only on the number of students but also on research outputs, international rankings and regional differences etc. Right now, we cannot stick to this formula either, because it would be unfair to compare the performance of universities. Irrespective of the funding logic, we have to endure deficits, as the first state priority are weapons and military spending.

Christoph König: You mentioned that two years of the COVID-19 pandemic already prepared the education system a bit for working digitally. What changed after the Russians invaded and how critical was the situation?

Kateryna Suprun: It escalated dramatically. Speaking about universities during COVID-19: It was only March, April and maybe early May 2020 that

the majority of universities had to go completely online. Later it was seasonal and there were times when teaching returned to face-to-face mode. Of course, in times of war this cannot happen. So, how well were we prepared? Not really, but the fact that the situation was not homogenous at all was also a challenge.

In Ukraine, it's quite difficult to say what the average level of digital readiness of universities is, especially because it goes back to how financially well off they are and what resources they already have in place. Also, it's important to understand whether or not they have had any previous experience with digital learning. Around a dozen universities were quite well placed in this sense. For example, the Sumy State University had digital laboratories with virtual reality technologies and a studio to record its own video content, and they had all these tools even before COVID hit. Consequently, when the pandemic started, it was not such a huge challenge for them, because they had the tools, they had the infrastructure, but still in terms of teaching and learning techniques it was challenging. Overall, we can say that our universities managed to switch to online learning, but with different successes and different student experiences. When it comes to what kind of support we could provide, we did some training sessions for the teachers on how to use digital tools and we helped to meet their needs concerning the digital infrastructure with different investment projects. Given the scientific autonomy of universities, all we could do is to make sure that the teachers have the skills to teach online and that the infrastructure is available.

Once the war started, of course, when all universities had to switch to online mode, it was smoother, because even those that had not had a lot of

success knew how to do it anyhow. We still have to figure out what to do with those programmes for which it's obligatory to have some face-to-face interaction; I mean, for example, tutoring in medicine first and foremost. As long as the war is ongoing, it's not possible to provide any activities in laboratories. Of course, there will be learning gaps, and we have to find a way to address them, especially for students who started their studies during the pandemic and only experienced a very brief time in class. Our main challenge is to ensure that it does not prevent the students from being successful on the job market. But so far, we are still considering how to even assess this learning gap, because as I said, the curriculum is very versatile.

Christoph König: How about the very basics, for example, connections to the internet; is it still functional everywhere?

Kateryna Suprun: In terms of the internet, it's actually quite promising, because it looks like the internet connectivity in Ukraine overall, in times of war and in times of crisis, is the least of our concerns. The Ministry of Digital Transformation is working on all aspects in the context of internet connectivity, also for education, and this does not seem to be a problem.

Christoph König: You mentioned that there was some training offered by the ministry to enhance digital skills, because you need different didactic approaches to teach online. Have you gotten feedback from universities and teachers on what their experiences were?

Kateryna Suprun: Everything started with the pandemic. The training provided by Microsoft and Google was quite a success; we received positive

feedback and the number of participants was high. Our partners are doing an amazing job in this. When the war started and we intensified our cooperation with online-learning platforms, such as Udemy, Coursera and EdX, they provided “how to” training, where potential participants could learn the basics about the platforms and how they operate, giving them some hints and showing some common tricks. We also had workshops with experts on teaching and methods discussing how to integrate these non-formal education courses into the formal education process. The timing of the collaboration with the non-formal education providers was really good because of a new decree passed by the ministry. Starting from March, officially 25% of the programmes delivered by universities can be covered by a non-formal provider. So, to put it simply, in a four-year bachelor’s degree, one year can come from Coursera. For the IT sector, we even increased this number to 35%. Still, the responsibility for recognising these credits and for recognising these learning outcomes lies with individual universities.

Many universities had already had recognition processes in place since 2017 but were hesitant because they feared doing something wrong. So we issued particular recommendations to encourage the recognition of non-formal courses but also to provide some kind of quality assessment. It is not easy, but we really try to encourage universities to let students go and pursue a Coursera certificate from a very good university, let’s say in Germany, the UK or any other country that will prevent them from having a learning gap and will also prevent them from struggling to connect to an online lesson from a bomb shelter where there’s no internet. We are working

with institutions, trying to persuade them and help them understand how to make the most of the current situation.

Nataliia Sokolovska: Aren't universities and university teachers afraid that their lectures will be replaced by such digital learning platforms?

Kateryna Suprun: Of course, if you have other sources to provide learning materials than the professors, the latter are put under stress. Logically, they ask themselves: What do I do if I'm replaced by an online platform? There are two aspects to this question. The first one is that we're trying to communicate to universities that you can actually collaborate with those platforms. So, the universities might not use platforms to substitute for 100% of their courses, but they may integrate the online course into their lectures. What I'm saying is that there is room to combine and modify; we don't want to simply substitute our teachers with Coursera. And this is what we're working on at the ministry currently – we are trying to pass on innovative methods of teaching from those who are more experienced and can share insights on how to combine different modes of teaching.

The second aspect is more challenging, and here we don't actually know how to deal with it for now. It comes down to the issue that professors are remunerated according to the number of students they are teaching. And yes, when performance-based funding, as I described earlier, is introduced, this should no longer be a problem. But right now, the linkage to the number of students is in place, hence these dependencies.

Nataliia Sokolovska: Do you intend to stimulate the production of MOOCs provided by Ukrainian universities?

Kateryna Suprun: During the war, an interesting grassroots initiative came into being, namely the Open Ukrainian University. Several universities that enjoy quite a high standing in Ukraine have come up with the idea to establish a platform, and to produce and publish high level content that can be used by those universities who are not in the position to produce online courses themselves. The plan is to have a low-key platform accessible to all other universities; these lectures could be used by other universities Ukraine-wide and could be officially recognised. Most importantly, the content will be produced mostly in Ukrainian. They do have the ambition to start making content in English, but we understand that this excludes many Ukrainian students. Right now the Open Ukrainian University is in the process of applying for funds.

The second aspect, or, more precisely, ambition, is that we'd like to make sure that such online courses are high quality and could potentially be used by universities abroad and on the job market. What is important to understand is that this process will take some time. When we developed the All-Ukrainian Online School, it took us quite some months to get it up and running, to find the partners and to understand how to record lessons and provide a quality check. The other initiative worth mentioning is the Ukrainian Global University, which is a platform for matching universities from abroad and students and staff from Ukraine. Basically, students and staff can apply and submit what they want to research or study. At the same

time, foreign universities that have the capacity to onboard students and employ staff from Ukraine can also publish their requests.

Nataliia Sokolovska: Let's say you received some help from the EU and/or Germany tomorrow. Where would this help go?

Kateryna Suprun: One area that is quite worrisome right now is the institutional capacity of our network. We are receiving a lot of help right now and it's often targeted at individual researchers, students and university staff. What we are very worried about in this context is the potential brain drain, or as we call it, the "undesired academic mobility". Let me elaborate on that: of course, it's not about making people come back when missiles are flying, just like some minutes ago when one hit the city centre of Vinnitsya and killed twenty people. We are perfectly aware of all the security and the economic challenges. But in the future, people who want to come back need to have a place to come to. Something we'll be looking into is to receive some structural funding support from the EU for universities to rebuild their institutional networks. In Ukraine, we have a lot of good universities and often they are underestimated and not treated as equals in the international community. In this sense, we'd like to internationalise our universities, make them more visible in other countries and make sure they know how to promote themselves.

When it comes to the structural higher education reforms, we've come a long way. For example, from 2005 we've been part of the Bologna process, we have introduced the ECTS system, a National Qualifications Framework and have a quality assessment system that is being improved all the time. But when it

comes to individual institutions, they're not visible enough. When they have success, they don't always know how to communicate this properly. Right now, with all the existing challenges, like the damage to the infrastructure and the potential brain drain, universities might be weakened. And they will need not only to catch up with what they had to do before war, but move several steps further. I think getting Ukrainian universities involved in European initiatives is crucial right now. In Erasmus+, for example, we are not yet a programme country, which is why we are not eligible to fully take part in the European University Alliances. This would be a great opportunity to reduce our universities' isolation. We are trying to approach and motivate the European Commission and other European partners to have some dedicated measures in place that would make it possible for our universities to sustainably be a part of the wider European network.

One particular measure that we're very thankful for has been provided by the European University Association, which waived tuition fees for all our universities till the end of this year. Because of this, our universities are now able to attend peer events and conferences, they get access to high quality materials and they can learn how to further develop and so to speak "join the club". Long story short, it is important, in my view, not only to help individuals, but also to help the institutions and show them that they are part of a bigger network that will not disappear in one or two years. This way they can rebuild and come out even stronger and be regarded as equal partners rather than being reduced to their need for basic materials or basic infrastructure. Of course, we have these basic needs as well, and they are our biggest challenge. For example, the Karazin Kharkiv National



University has been destroyed almost completely by the Russians. So of course, we have to rebuild our universities, but it is my personal wish and desire that our universities are seen, are heard and are treated equally, with an understanding that we have something to learn but also something to showcase.

PAUL ATKINSON

What the COVID-19 experience has taught us about the limits of evidence-informed policy making

The COVID-19 pandemic experiences of many different countries have shown how an exceptional need for research evidence can lead to a rapid new deployment of scientists in advisory roles to aid policy makers in making evidence-informed decisions. This can help us to understand how science advisors experience, and try to manage, the challenges of insufficient, evolving, and conflicting evidence as they work to inform public health decision-making. Of course a fast-moving crisis like COVID-19 poses some very specific challenges to evidence-informed policy making, but there are some general lessons to learn here as well. In this short analysis, Paul Atkinson highlights the uncertainties associated with the field of evidence-informed policy making, especially in crisis situations such as COVID-19.

[READ FULL ARTICLE](#)

DEFINING CONCEPTS OF THE DIGITAL SOCIETY

Data justice

Data justice has emerged as a key framework for engaging with the intersection of datafication and society in a way that privileges an explicit concern with social justice. Engaging with justice concerns in the analysis of information and communication systems is not in itself new, but the concept of data justice has been used to denote a shift in our understanding of what is at stake with datafication beyond digital rights. In this essay, we trace the lineage and outline some of the different traditions and approaches through which the concept is currently finding expression. We argue that, in doing so, we are confronted with tensions that denote a politics of data justice, both in terms of what is at stake with datafication and what might be suitable responses.

[READ FULL DEFINITION](#)

DEFINING CONCEPTS OF THE DIGITAL SOCIETY

Artificial emotional intelligence beyond East and West

Artificial emotional intelligence refers to technologies that perform, recognise, or record affective states. More than merely being a technological function, however, it is also a social process whereby cultural assumptions about what emotions are and how they are made are translated into composites of code, software and mechanical platforms that operationalise certain models of emotion over others. This essay illustrates how aspects of cultural difference are both incorporated and elided in projects that equip machines with emotional intelligence. It does so by comparing the field of affective computing, which emerged in the North Atlantic in the 1990s, with *kansei* (affective) engineering, which developed in Japan in the 1980s. It then leverages this comparison to argue for more diverse applications of the concept of culture in both the development and critique of systems with artificial emotional intelligence.

[READ FULL DEFINITION](#)

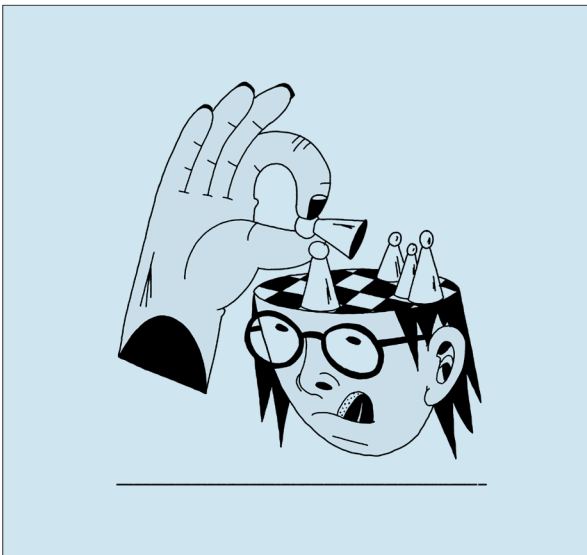
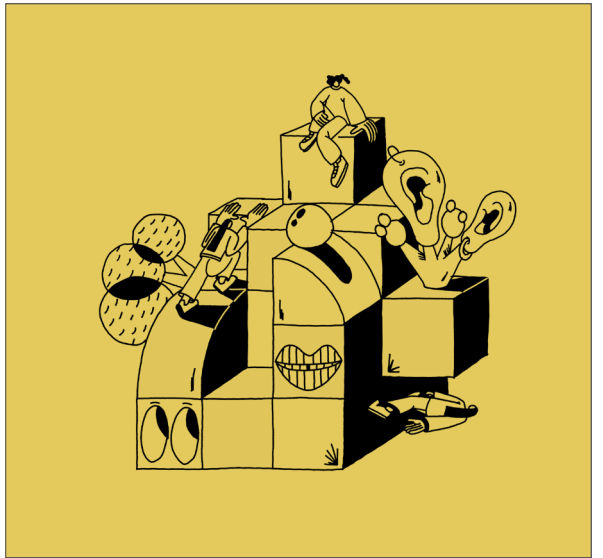
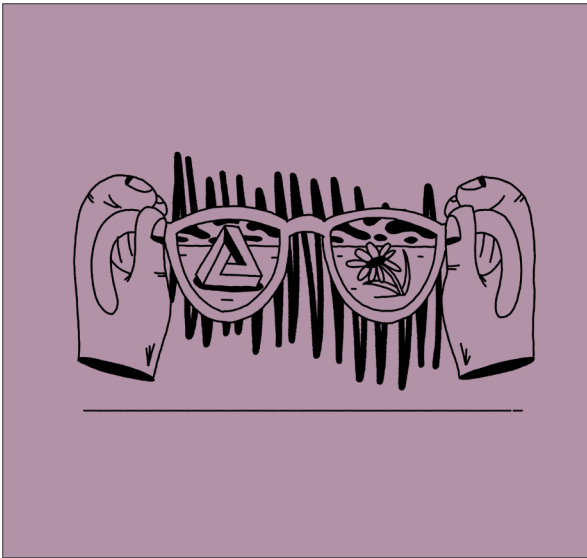


TOOLBOX

Future(s) of the digital world

Our toolbox *Making sense of the future* aims to promote transdisciplinary, multi-perspectival and creative (future) thinking. The six exercises encourage digital visionaries to engage with the paradigm of uncertainty that constitutes the future. The tools encourage a creative engagement with possible futures of the digital society. This OER can be used in knowledge-based engagement with possible futures in schools, universities and organisations. It can also be used privately for education, inspiration or as an intellectual exercise.

[EXPLORE TOOLBOX](#)



FREQUENTLY ASKED QUESTIONS

Digital infrastructures of the knowledge society

What are digital knowledge infrastructures?

Digital knowledge infrastructure includes tools for data management, collaboration and communication, as well as platforms for sharing research results and digital content. The use of these digital technologies can facilitate research by making it easier for researchers to access data, to collaborate with each other and to share their results. These open and transparent research practices also enable new forms of science. Thus, digital knowledge structures have the potential to increase the speed, efficiency and impact of academic research.

FREQUENTLY ASKED QUESTIONS

What do we mean by a (digitally-networked) knowledge society?

In a knowledge society, knowledge is the most important resource for social, economic and media coexistence. Here, digital tools and systems offer better access to information and facilitate entirely new forms of collaboration. However, the digital knowledge society also harbours risks: for example, in questions of data security or in the case of unevenly distributed access to digital resources worldwide. In this context, research also plays a crucial role by producing new knowledge and training future generations of scientists. Therefore, academic institutions need to redefine their role in the knowledge society and adapt their practices to the digital age.

Why do we need to measure and evaluate the impact of scientific knowledge in society?

In a digital knowledge society, the methods for measuring the quality as well as impact of research must change, for digitalisation has the potential to transform all fields of research. It underlines the importance of taking greater account of the use of knowledge in practice. How do science and society interact with each other? How do stakeholders from politics, media, business and civil society process scientific findings? Only with these questions in mind can we meet the diverse challenges of today.



ALEXANDER VON HUMBOLDT
INSTITUTE FOR INTERNET
AND SOCIETY

encore

Volume 2022/2023

Published February 2023

PUBLISHER

Alexander von Humboldt Institute
for Internet and Society
Französische Str. 9
10117 Berlin

EDITORIAL TEAM

Stefanie Barth, Stephan Bohn, Frederik Efferenn,
Larissa Wunderlich, Yuanwen Zhong

DESIGN

Larissa Wunderlich

Digital ISBN 978-3-9821760-8-6