

# Research, Innovation, and Education Roadmap for Digital Humanism

Erich Prem, Lynda Hardman, Hannes Werthner, Paul Timmers (eds.), The Digital Humanism Initiative.

*With contributions from Carlo Ghezzi, Peter Knees, Julia Neidhardt, Walter Palmetshofer, Viola Schiaffonati, Stefan Woltran, and all participants of the Fourth Workshop on Digital Humanism, Vienna, March 2022.*

*Cite as* E. Prem, L. Hardman, H. Werthner, P. Timmers (eds.). Research, innovation, and education roadmap for digital humanism. The Digital Humanism Initiative. Vienna, 2022.

<https://dighum.ec.tuwien.ac.at/>

## Background

Digital technologies are changing **human behaviour**, significantly changing our **society**, and our **environment**. Digital humanism observes and describes these changes and aims at shaping and influencing the development of digital technologies and policies towards the values of **human rights, democracy, participation, inclusion, and diversity**. It is a broad concept combining technical and social innovation and ranging from research to politics.

This roadmap is based on talks and discussions held at the 4th Workshop on Digital Humanism, TU Wien and online in March 2022. Following up on the Vienna Manifesto on Digital Humanism, the purpose of this document is to provide strategic guidelines to realise the vision of digital humanism and integrate its ideals into society via three important pillars: **Research and Development (R&D), Innovation, and Education**. It

The roadmap is a snapshot and of course has gaps. It will have to be refined and improved step by step, through discussions and contributions from the international community. It is presented starting from the problems and challenges to the topics and research necessary to move towards digital humanism. As such, it should be a guide to help define and communicate the project of digital humanism and strengthen the collaboration of researchers, especially across discipline boundaries.

The primary purpose of this document is not to create a new lobby group or to establish a new 'hot topic' for funding research. Instead, the roadmap should help include digital humanism topics in all kinds of research. In addition, the document should help shape and prioritise the development of policies that can further the realisation of digital humanism in our society.

## Challenges for digital humanism

A more positive perspective on digital humanism is to focus on constructive contributions of digital technologies, e.g. by empowering people and support shared values in society:

### *Digital empowerment and basic rights*

As an example, to digitally empower people, increase access to knowledge, participation and inclusion in society, support diversity, and guarantee fundamental rights. Other forms of empowerment are:

- Put the human in control
- Strengthen the social contract, lend a personal voice
- Mitigate risks of categorization and classification
- Provide personal data privacy
- Detect malicious behaviour
- Develop and deploy trustworthy systems
- Develop AI to work in partnership with humans

### *Support societal values such as:*

- Morality
- Human dignity; the human as worthy of respect; the *human condition* in its contrast to machines
- Value-based online content curation
- Sustainability and environmental protection
- Democratisation and transparency in recommendation and personalization
- Varied stakeholder views in the design of digital technologies
- An inclusive society

### **Digital humanism - a constructive endeavour**

Digital humanism starts from a positive and constructive perspective of digital technologies. We *can* design ICT in ways that are much better than what exists today. This calls for computer science and other fields of engineering to adopt principles of digital humanism and improve systems beyond how they are designed today. We believe that the way technologies are designed can have great power. At the same time, we are aware that it is not enough to design them in accordance to positive values to always guarantee positive outcomes. This requires a truly interdisciplinary effort, including different approaches and perspectives. Importantly, and beyond informatics, this constructive perspective also extends to fields such as law, political science, or philosophy that include constructive activities. These disciplines need to come together to meet the often complex questions and many interacting facets of problems in digital humanism. In digital humanism we are challenged to co-construct technology and society.

As an example, consider the case of privacy whose nature is still debated in philosophy. There is a need for constructive proposals about the concept of privacy as much as we need improved concepts for regulation, education and other social constructs on privacy aspects that actually work to empower users. Computer science, in turn, needs to further improve on existing tools and techniques for privacy-preserving technologies to make them computationally efficient, easily deployable, and widespread and support users in defining and realising their own privacy preferences.

This is not to say that the more analytical fields of scholarly research are useless. Quite to the contrary, there are important insights, analyses, and principles that need to be contributed from social science, science and technology studies, political science etc. But it will be important for the success of digital humanism to enter into a constructive phase of designing IT systems to

meet and support values aligned with the principles of digital humanism including through the contributions of more analytical disciplines instead of merely lamenting their shortcomings.

Finally, digital humanism should not shy away from making a constructive contribution to address the hard problems that decision-makers and policymakers are struggling with. These include war and peace at our borders or inside our own country, safeguarding perhaps even rescuing (digital) humanism in the face of authoritarianism, injustice caused by climate change with all its consequences from starvation to migration, and dealing with dependencies on other countries when sovereignty is at stake. Addressing these requires interdisciplinary constructive digital humanism.

### **Digital humanism: research, development, innovation and more**

Digital humanism requires efforts on various levels of technology readiness. It demands basic research into the nature of computation and its impact on humans as much as it requires the application-oriented development of tools and easy-to-use services. Given the many complex interactions of IT systems with humans and society, it will be necessary to approach some questions in an experimental fashion with proper attention given to ethical research and evaluation.

Digital humanism can only succeed with the help of knowledgeable experts who are open to the exchange of ideas between research fields. Education of future scientists and engineers is as important as is the education of technologically open-minded scholars in the humanities.

**Problem areas: the potential negative impact of digital technologies on our lives**

Digital humanism addresses a broad range of concerns about digital technologies and how they impact on the individual and society. Important problem areas include:

- AI and automated decision making
- Work and automation, including the gig economy
- Privacy, security, and surveillance
- Platform monopolies, market concentration

- Online media, fake news, and their impact on freedom and political discourse
- Digital sovereignty, geopolitics, and the role of states
- Environment, and sustainability

This is not a complete list but helps to clarify that digital humanism aims to address the ethical and societal challenges of digital technologies as a whole and is not solely focused on single specific topics such as explainable AI, surveillance, or digital sovereignty. The following table provides a first overview of the problem areas and research topics that need to be addressed for furthering digital humanism.

*Table 1 Overview of digital humanism problem domains and research topics  
(see text for further description)*

Problem domains	AI and automated decision making	Labour and work automation	Privacy and security	Platform monopolies	Online media, political discourse and NLP	Digital sovereignty, and sustainability <sup>1</sup>
Research topics						
explainability	x					
transparency	x		x		x	
data privacy			x			
personalisation, including in recommender systems	x				x	
fairness	x	x		x	x	
accountability of systems and providers	x	x			x	
human control	x	x	x			
AI/human coop. & distribution of power	x	x			x	
norms and ethics	x	x	x		x	x
security			x			x
new regulatory approaches					x	x
DigHum business models						x
resilient systems and algorithms						x
low carbon footprint, sustainability						x
work design and labour policy		x				
content moderation, freedom of speech					x	
productivity paradoxon		x		x		
dependencies measurement (power rel.)		x		x		x
architecture resilience				x		x
distributed vs. centralised approaches				x		x
software development			x	x		
interoperability			x	x		
open systems, open data, etc...			x	x		

<sup>1</sup> Sovereignty and sustainability are separate problem domains put in one column to improve readability.

## ICT research issues

### Artificial Intelligence and automated decision making

AI has become a tool that supports decision and creates systems that are capable of acting to some extent autonomously. These systems are impacting the daily lives of consumers, workers, employees – simply anybody interacting with and through electronic devices or ICT-based services. This has created a large number of challenges and many open questions and debates. We have barely scratched the surface of these debates, let alone found technical answers. From the point of digital humanism, the following research topics are particularly important today:

- Developing improved ways to make AI systems more understandable or explainable AI.<sup>2</sup> This includes work on logic/machine learning combinations and tools for analysing trained AI systems and work on the context-dependent nature of explanation.
- Novel concepts and approaches to productively distribute power between humans and AI or autonomous systems. This could lead to the development of frameworks that reinforce people's autonomy when dealing with AI systems.
- Novel approaches for dealing with emotions, supporting the human emotional dimension and the role of emotions in moral decision making.
- Improved strategies, frameworks, and concepts for supporting human-machine partnerships and how to build such hybrid intelligence systems.
- A continuous effort to map the landscape and ideologies of AI, including efforts towards generalised AI, best-practices and bad practice, i.e. AI-stupidity.
- More interdisciplinary research on accountability of systems and their

providers including concepts, frameworks and tools is required.

### Labour and work

The role of AI for labour and work requires further research attention. This includes the topic of work design, policies, and union achievements for the digital world as well as measurements regarding dependencies and power distribution. It will be important to also address societal-level phenomena of work emerging from the individual perspective. In addition, on both the digital and economic / innovation perspective, the productivity paradox needs more attention.

Generally, research should be shifted from human-replacing AI to AI supporting humans in partnership, i.e. in societally beneficial ways. The importance of human control and the distribution of power between humans and AI or AI/human systems will remain key challenges as will questions of norms, ethics, and politics. Given their broad range of applications, many AI topics overlap with research challenges in other areas such as recommender personalisation and fairness (see below).

### Privacy and security

Research into privacy and surveillance needs to be better matched with fairness considerations and power relationships emerging from digital technologies. Topics such as architectural resilience, distributed versus centralised architectures, interoperability, open systems, and ethical software development require more attention in research and the development of guidelines, established practice, and standards that support practical work. Additional research is necessary to improve our understanding and management of the individual and societal aspects of privacy and surveillance including privacy integrity in various situations and contexts, e.g. IoT or smart objects.

---

<sup>2</sup> See also <https://www.informatics-europe.org/component/phocadownload/category/10->

[reports.html?download=74:automated-decision-making-report](https://www.informatics-europe.org/component/phocadownload/category/10-reports.html?download=74:automated-decision-making-report)

Maintaining the privacy of users and keeping their information secure is a central value in digital humanism. Individual privacy is not only instrumental to other values, such as human autonomy, but also to the maintenance of collective values such as democracy. Significant progress has since been made in the legal realm, especially since the advent of Europe's General Data Protection Regulation and similar regulation in other regions of the world. However, there are still significant challenges regarding the legal situation and even more so regarding the practice of how to protect personal data. Unfortunately, there is ample evidence of continued privacy breaches even to the extent of straightforward illegal behaviour of companies. Other practices may not be strictly illegal ("dark patterns") but use verbose descriptions to conceal how personal data may be used. More research is required, to develop new approaches that make privacy options understandable, adaptable, and certifiable. This involves both computer science and legal research. It may also mean to go beyond mere informed consent, especially where data is used for inferences.

Enforcing privacy requires action both at the legal and engineering front, for example improved privacy concepts and technologies as well as tools and techniques for investigating data transfers and documenting suspected breaches.

- Further research will be needed in fields such as differential privacy and related techniques (e.g. homomorphic encryption) to make these technologies computationally efficient and even more secure. In addition, development activities are required to create easy-to-use tools and services that make these technologies more widely known and available.
- There is a need to inform broader audiences about privacy-preserving technologies including public authorities and policy makers. Frequently debates centre on an all-or-nothing conception of privacy that does not adequately represent the current state-of-the-art.
- Today, only few tools and dedicated technologies are available for the identification of data leaks and for tracking

data destination, e.g. to be used as smartphone monitors. There needs to be more research on improved technologies for the control of international data transfer, especially of personal data. This should include tools for the analysis of browser behaviours.

- In close cooperation with user groups and legal bodies, research, development, and certification activities are needed to develop and introduce privacy metrics, certification schemes, and methods for the verification of system behaviour regarding privacy, data transfers etc.
- A larger research effort will be required to develop novel privacy-enabling architectures, e.g. for smartphones.

With the massive increase in internet-connected objects, the amount of personal information transmitted over networks connecting these objects is growing steadily as are the concerns regarding security and safety. The large number of internet-connected devices and an increasing reliance on their functions in the private, business, and public domain the geopolitical dimension regarding national sovereignty. Research is particularly needed in the following topics:

- Tools and techniques for developing a more secure internet of things, e.g. regarding surveillance, firmware updates, passwords and other forms of authentication, secure transfers over unsecure connections etc.
- Permission management for IoT devices that guarantees not only privacy, security, and safety, but also facilitates ease-of-use and user-centric approaches.
- Development of tools and techniques in accordance with values, such as the minimization of collected personal data, deletion of personal data etc.

### **Social platforms, monopolies and personalization**

The challenge of monopolies and social media platforms includes questions of dependencies and power relationships and how to overcome or change those challenges. The role of architectures, software development, open architectures or open data, resilient systems for counteracting tendencies of market concentration has only been insufficiently addressed to date.

We have come to live in an age of evaluation; from restaurants to doctors, from things we buy to the music we listen to, recommender systems are at the core of online businesses. They not only help users choose products and services but have become key components of business models. A tightly woven network of recommendations, feedback, scoring, personalised offerings etc. is driving highly individualised interactions to the extent that it is unclear who benefits most from recommendations and what are the objectives used to optimise online recommendations. Personalised systems need a radical increase in transparency, democratisation, and control. This can be achieved through:

- Reporting the optimization objectives of systems
- Empowering users to make their own choices
- Higher transparency on data used and how it is processed
- Increase the accountability of systems and system providers

Personalization, preferences, and context-specific user preferences and values should be under the control of the user, ideally managed and applied outside of the providing service. This includes values embedded in the systems. To achieve this, the unbundling of the interaction and personalization from system optimization / service provision, and management of content and data will be essential.

Research therefore needs to include the following:

- Improved transparency tools for users and other stakeholders regarding the objectives of recommendation and scoring systems.
- Novel ways to increase the governability of recommendations thereby putting the user in control, e.g. on secure personal devices outside of platforms.
- Techniques and tools for the democratisation of recommendations including ways to go beyond individual-level aspects, e.g. including societal aspects.
- Open and public debates on the way in which recommendation and other tools can be designed.
- Methods for the unbundling of personalization and technologies to ensure that users can switch providers of personalization services easily. This will also require interfaces or trusted third parties for personalization labels and certification schemes.
- More generally, tools, techniques, and regulatory efforts for the improved separation of personalization and content or service offerings.

To ensure interoperability with a range of services and to avoid overly strong dependence on a small set of services, users must own their data and profiles. There need to be standard interfaces and services to make user preferences etc. easily transferable and interoperable between different systems. This is equally important from an economic perspective to avoid technology lock-in.

### **Online communication, political discourse, and natural language processing**

Social networks and online communication are one of the most clearly visible problem domains of digital humanism. Phenomena such as echo chambers, limitations on freedom of speech, and how to support productive online discourse are among the most important and widely visible challenges that digital humanism needs to address.

For example, AI-supported content moderation will remain a central research topic for digital humanism. Other topics such as transparency, fairness, accountability, as well as the distribution of power between humans and systems have already been mentioned but are of central importance in online social networks as well.

Although all types of media are relevant, language plays a particularly important role. Language- and speech-based interfaces have become widely available and massively used, for example as interfaces with electronic assistants, controlling cars, and for text-based interaction with virtual agents. This surge in applications is mostly due to new statistical tools and the harvesting of large online data sets and it is available in a range of common languages. On the downside, many tools are not available in good quality in rarer languages or for dialects. Despite its fancy name, there is little understanding in “language understanding” as most techniques rely on statistical pattern matching. And AI-based decision making has come to play a massive role in the shaping, control, and the exclusion of online speech.

Research to ensure human-centred and value-based natural language systems should include the following topics:

- Tools and techniques for online content moderation that implement principles of digital humanism. Novel concepts for ranking, deletion, and calming discourse in line with principles of free speech and other human rights. Tools for re-instantiating content, for gender- and diversity-sensitive moderation, and for actively supporting productive and constructive discourse.
- This will require research and development for the creation of NLP data sets that are relevant, interesting, balanced, and shareable with broad audiences - especially for the case of less frequent languages.
- Novel techniques are needed for an improved understanding and following concept shifts, such as those happening over extended periods of time in historical data, e.g. regarding sexism.

- Researchers and research funders need to give increased attention to minority languages and dialects. Also, much more effort should be invested in improved NLP interfaces for persons with linguistic handicaps as well as for proper interactions with younger audiences etc.

## **Sovereignty**

Digital sovereignty at the national level, geopolitics and the role of states requires consideration of norms, ethics, and politics of digital systems. Questions of regulation will be as important as the issues of sustainable business models and resilience of systems in a world that is increasingly dependent on IT support in all areas of life. The topic of centralised versus distributed architectures and the resulting power relations are also important. In terms of environmental aspects and sustainability, more research in resource efficiency and low-footprint solutions are required. In the military realm, fast decision making, emotions, and transparency will also be important research topics from the viewpoint of digital humanism.

‘Digital’ is not only becoming pervasive but also ubiquitous in the sense of manifesting itself in an enormous number of instances. Companies and authorities are dealing with millions of connected devices and connected humans. The sheer complexity of the digital world is at the same time overwhelming for protection against cyber-incidents or reigning in exploding electricity consumption, an enormous opportunity to extract value from the generated data, and a purposeful tool for some to exercise control. Digital complexity already leads, as an unintended consequence, for many to a feeling of being powerless, perhaps even loss of personal sovereignty, and thereby risks eroding humanism.

## **Sustainability**

Computer science has not sufficiently considered the ecological footprint of computing, especially of complex processing such as training of AI models.

However, energy-efficiency and energy-aware computing are important fields in informatics, e.g. in internet-of-things and powerless systems. Digital humanism requires more attention to the environmental aspects of computing including energy efficiency and related aspects such as Green electronics, sustainable resources and energy-smart systems.

### Education and Teaching Digital Humanism

Progress in digital humanism requires the contribution of creative minds that are not only competent in developing novel concepts, frameworks, and architectures for IT systems. They need to be aware of the challenges, intricacies, and complex societal interactions of those systems. Digital Humanism cannot just rely on established scholars but requires novel contribution from young minds including entrepreneurs. To achieve this, education and training in the field of digital humanism need to be developed, rolled out, and intensified in the following ways:

1. Change the narrative: the story of what a discipline “is” is to include “Digital Humanism” issues and notions given the heavy societal impact of digital technologies. Quality cannot be achieved if a view on the field is outdated and lacks societal relevance.
2. Influence goals of curricula from there (also formally): becoming a “responsible” scientist or professional as a recognized goal of education; to be included also in review frameworks and accreditation`
3. Embed multidisciplinary perspectives in teaching and learning in a natural and fitting way as they derive from key digital society themes and challenges. This means both including humanities in training of engineers and information technology in the humanities. Many universities provide environments that can foster the multidisciplinary required for effective Digital Humanism research and initiatives. Documenting and disseminating good practices will be key.
4. Pursue multiple pathways: grow Digital Humanism issues education bottom up as evolutionary growth from existing situations

and contexts in different disciplines and curricula. There is not a single educational model; an approach may be to start from the existing curricula, then ask the associated difficult questions, and then ask for consideration how they can and should be more satisfactorily addressed.

5. Foster collaborative / (multidisciplinary) *team-based* approaches to teaching and learning
6. Create low threshold sharing mechanisms for interdisciplinary research and teaching. Academics from different disciplines do not know each other and speak different languages (concepts, terminology, methodology, theories, scientific goals and philosophy); only by continued exchange over a long period will these issues slowly disappear.
7. Create events with value for young researchers. Young researchers such as PhDs and postdocs are generally more open to the changing disciplinary landscape as they have less vested career/tenure track monodisciplinary interests and are less socialised into such frames compared with “established” academics; also, there is not much available internationally for young researchers in a cross-disciplinary way; and COVID has made this isolation of young researchers even worse.
8. Form “coalitions of the willing”. A somewhat unfortunate term given its history, but it does serve to illustrate a key point: there will be academics really interested in this, but there will also be some that don’t (for a variety of reasons). Just let them be. The basic principle here is: We do not have to convince everyone, just start with the 10-20% that is *really* interested, motivated and capable.
9. Education in digital humanism should not be a hobby or treated as such. A typical “repressive tolerance” reaction (from for example local ivory-tower deans) is that this is all potentially interesting, but please do this just as an additional on-top effort for teachers and as an extracurricular activity for students. This is not the way to go, and short of the needed official recognition of the importance of these matters as a central and not peripheral concern. Researchers and teachers should be awarded for their efforts in dealing with Digital

Humanism type issues. This would include mechanisms to evaluate academic careers so that interdisciplinarity work and publications are rewarded and positively evaluated. For students, all this should be credit bearing. For universities, all this should be part of a proper financial model or business model that is financially sustainable.

## Research issues in related fields

Although the focus of this roadmap is on ICT aspects, research challenges in other fields arise naturally when taking a digital humanist perspective.

### Philosophy

Many topics of digital humanism require not just a technical solution or a quick legal fix. In fact, a range of problems are still poorly understood (e.g. the role of AI in politics) or conceptually unclear (e.g. the notion of privacy). Philosophy can contribute as a critique of scientific and technical practices as much as a discipline to analyse and design concepts. Research to support digital humanism will include the following:

- Philosophy and up-to-date theory of information and information systems; clarification and debate about core concepts of digital humanism.
- A continued and differentiated discussion of human-machine differences including with a view of AI, robotics, systems science etc.
- An improved understanding of the nature of explanation (e.g. for XAI)
- Work on the relation of politics, democracy, and technology (including AI)
- Contributions to the role and nature of “good” discourse in society and in democracy
- Critique of research contributions to information technology; prioritisation and de-prioritization of ICT research, such as “general AI”

- Frameworks and approaches to make space for the subjective and the human in “hard-core” engineering disciplines including ICT
- A debate on liberty, freedom, but also on paternalism ICT systems, their deployment, and use
- Continued development of theories of meaning, including areas of natural language processing, symbol (AI) grounding, representation, and intentionality
- A constructive analysis of the foundations of engineering and technical sciences and their methods with a view to the important role they play in modern society.

Such work should lay the foundation of a framework or components for a theory of digital humanism.

### Ethics

Digital humanism raises a large number of ethical issues that are not limited to AI-based trolley problems or automated weapons. Arguably, a whole reconceptualising of ethics for the digital age may be required that develops frameworks and establishes common principles for practitioners as well as for laymen. More research is particularly necessary in the following fields:

- Improved identification and awareness of ethics-washing, tech-solutionism, ivory towers etc.
- Ethics of synthetic data
- Prudence, practical virtue, and practical wisdom in the use of AI and other IT systems – sometimes called phronesis
- Practices and policies for dual-use digital humanism research. Civil research often finds military application and conversely, requirements and funding from defence can accelerate civil research. Digital humanism should develop technological and deontological checks and balances for dual use.
- Existing or new ethical frameworks to support the conceptual clarification essential to fill in the gaps in the existing policy vacuums.

- Ethical frameworks embracing an incremental approach to some ethical issues, recognising that it's not always possible to anticipate problems at the design phase, but that some of these problems emerge when digital technologies are put in their context of use and interact with humans.

## **Law**

Regulation of the private and public sphere will be essential to strengthen digital humanism in practice. A practical focus today is the area of privacy. While many countries have put in place privacy policies, there is a need for simplification and, of course, for improved enforcement. A potential way forward could include verified privacy labelling schemes.

## **Putting Digital Humanism in practice**

### **Standardisation, innovation, and outreach**

Apart from the more academic, university-type research needs, there is a huge opportunity already now to put principles of digital humanism in practice. This includes, for example, labelling schemes, frameworks, and standards for more human-centric systems that respect and further societal values. It also includes improved verification of system behaviours, certification or simply good practices. Also, while many of the digital humanism research challenges require much more work, there are already important results, tools, and techniques that are not broadly known and insufficiently utilised, e.g. the use of synthetic data for protecting privacy or more democratic ways to provide access to data for all.

*Innovators* have started to “just do things” such as training underprivileged groups, deploying ethical principles, disseminating good practices etc. Such bottom-up activities are important for preparing the grounds and collecting experiences. We need more of such grassroots activities, collect data and good practices, and broadly disseminate results and experiences. It is particularly important to include businesses – both small and large – in the practice of

digital humanism. Given the importance that companies play in the digital world, changing the game will require onboarding businesses. To succeed, new and improved business models will also be required that do not rely on harvesting personal data and on other detrimental practices of today's ICT businesses.

For digital humanism to succeed it needs to significantly increase outreach and awareness about the challenges of issues. The European General Data Protection Regulation (GDPR) can be regarded as a successful piece of legislation that has impacted globally and changed the perception about basic rights and values in relation to information technology. Similarly, raising awareness and changing perceptions are also needed for other areas of central importance to digital humanism. This could include, for example, the demystification of AI, the role of ICT systems in geopolitics, or the current social silences regarding ICT that require much more attention in the future.

### **Research methodology and interdisciplinarity**

Progress in digital humanism clearly requires contribution from many different disciplines. In the past, contributions to critical thinking about the impact of digital technologies have come from many different perspectives including but not limited to philosophy, political science, sociology, and science and technology studies. It is central to digital humanism that such collaboration is reinforced with a synthetic and constructive aim to design improved socio-technical systems. Computer science has a long tradition of interacting with domain experts in other fields and such interactions need to be reinforced. This is particularly the case in AI and machine learning where sometimes the narrative goes in the opposite direction, i.e. that no domain knowledge would be required for developing systems. To the contrary, more interactions and systematic frameworks for improved dialogues between various domains and data-driven approaches are needed.

It is fruitless in this context to complain about the challenges or even impossibility of integrating over distant disciplines. The history, concepts, theory,

and practice of computer science and sociology, cryptography and philosophy, machine learning and political science are indeed fundamentally different. In digital humanism, contributions from various disciplines can productively work together given the shared objectives of a human- and society-centric view. An easy and straightforward way to start collaboration could come from a “digital anthropology”, for example: Such investigation will start from observing how people use IT systems and developing an improved understanding about their beliefs, intentions, and objectives. This could create important foundations for various disciplines and a basis for developing improved socio-technical systems rather than merely optimised computer systems.

Research in digital humanism needs to be inclusive, inclusive of diverse groups of people and inclusive of societal aspects. This will require research in inclusion/exclusion of groups of users, citizens, special groups and constructive approaches towards more inclusive solutions – either technical or socio-technical ones.

Finally, digital humanism also needs to address current shortcomings of *research methodology* and *research practices*. It is important to understand and shed more light on the fact that ICT research today is to a large extent privately funded. Companies – both large and small – are driving research in new AI systems, online services, and in microelectronics, computing systems etc. In some areas such as natural language technologies, big companies control significant portions of the available infrastructure for research, e.g. data. There is a need to balance this disproportion, for example through open data collection, by targeting the needs of minorities and disabled people, or simply by improving access to private research infrastructure. European efforts to promote open science can be taken as a good example, both at the level of the EC and member states.

Finally, digital humanism needs to step-up its influence in research and technology policy and in political debates around digital systems in general. This includes the continuous information about the impact of digital technologies on society and also the

opportunities of digital systems for fostering democracy, inclusion, education, etc. Research and science policy needs to be informed about the many open research challenges in digital humanism and about its huge potential for society. There is a clear need to improve the collaboration between various disciplines also in research funding models and in peer evaluation. Furthermore, researchers in digital humanism need to advertise both the long-term basic research aspects as well as the opportunities for short-term, more applied developments. Today there are only a few places where digital humanism meets with policy makers and more such fora are required.

### **Acknowledgements**

Work on this roadmap was supported by WWTF, the Vienna Science and Technology Fund <https://wwtf.at/>

We would like to thank all participants of the Fourth Workshop on Digital Humanism, Vienna, March 2022 and the presenters: H. Werthner, G. Metakides, C. Ghezzi, J. Neidhardt, H. Nowotny, G. Kappel, V. Schiaffonati, E. Prem, R. Burke, P. Inverardi, E. A. Lee, P. Knees, M. Vardi, M. Lindorfer, B. Krenn, S. Woltran, R. Pichler, M. Stampfer, S. Winter, J. Gordijn, E. Nardelli, H. Akkermans, A. Ebert, K. Karachalios, A. Kaspersen, D. Schabus, C. Neppel, A. Stanger, J. Guszcza, M. Hauswirth, S. Wyatt, C. Reinprecht, I. Nalis.