HAL open science

Toward a Research Agenda on Digital Media and Humanity Well-Being

Chavalarias David, Antonio A. Casilli, Alexandre Delanoë, Melanie Dulong de Rosnay, Beatrice De Gelder, Divina Frau-Meigs, Bertrand Jouve, Diminescu Dana, Nahla Ben Amor, Anna Boros, et al.

To cite this version:

Chavalarias David, Antonio A. Casilli, Alexandre Delanoë, Melanie Dulong de Rosnay, Beatrice De Gelder, et al.. Toward a Research Agenda on Digital Media and Humanity Well-Being. CNRS. 2023. hal-04091733v1

HAL Id: hal-04091733

https://hal.science/hal-04091733v1

Submitted on 9 May 2023 (v1), last revised 14 May 2023 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

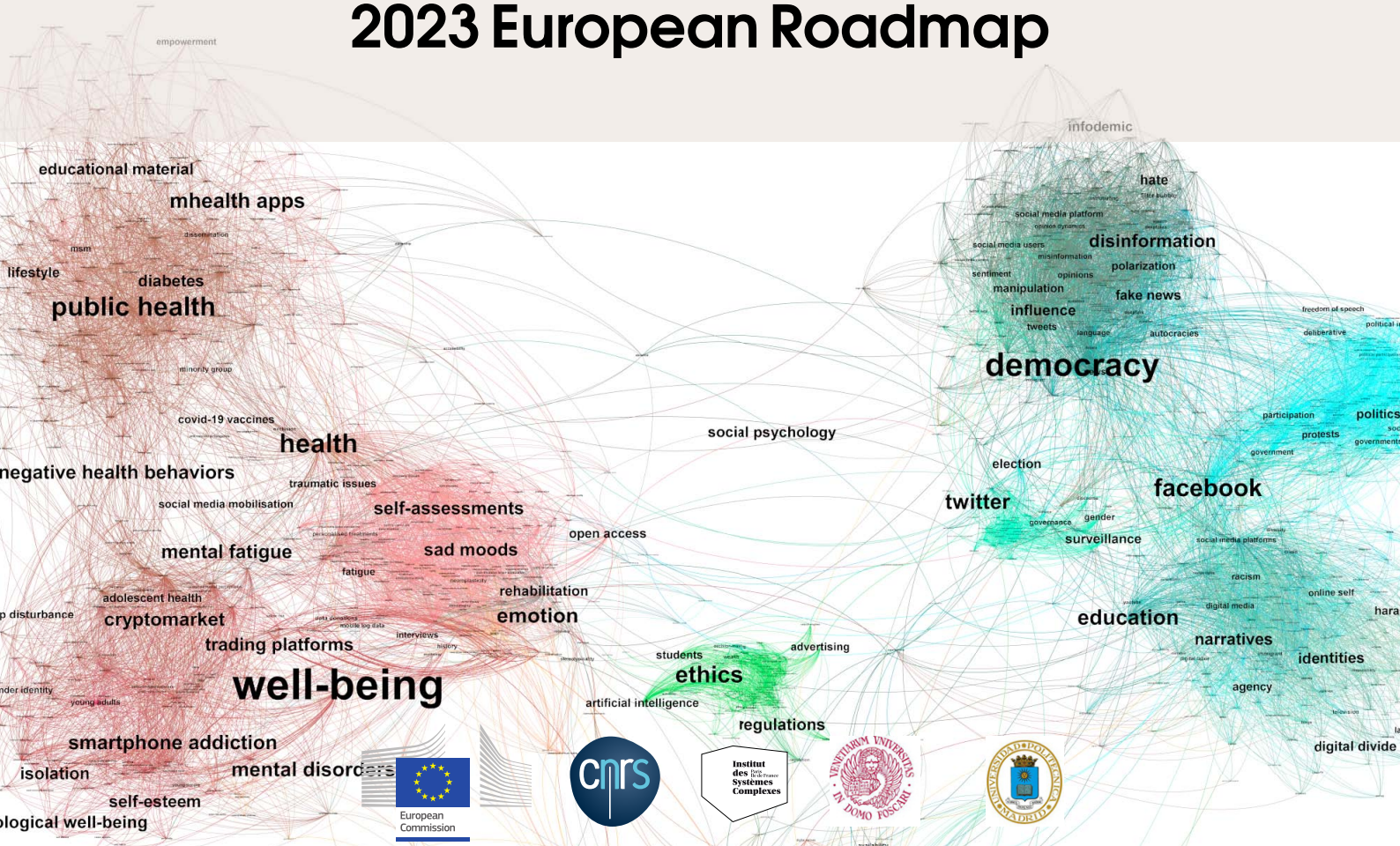
Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

Toward a Research Agenda on Digital Media and Humanity Well-Being

DIGEING Project

Fourty contributors from fourteen EU countries

2023 European Roadmap



Report coordinated by David Chavalarias and produced by forty contributors (see Ch.A for the list) as part of the project DIGEING – Digital Media and Human well-being. The DIGEING project was funded by the EUROPEAN COMMISSION Directorate-General for Communications Networks, Content and Technology, project VIGIE 2020-661 - LC- 01607979.

The information and views set out in this report are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this Study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein.

To cite this roadmap: *Toward a Research Agenda on Digital Media and Humanity Well-Being* (2023), Chavalarias D. Ed.
To cite a chapter: Chapter's authors (2023), Chapter's title, *in* Chavalarias D. Ed., *Toward a Research Agenda on Digital Media and Humanity Well-Being*.

Images Credits: David Chavalarias or David Chavalarias@Midjourney, except banner Ch. 4.

First printing, April 2023. Distributed under CC BY-NC-ND 4.0 licence

Foreword

In the 2020*s, an American citizen will spend an average of 6h35 a day on social media, compared to 3h35 for television. As for social networks, which were non-existent less than 20 years ago, about 40% of US citizens use them at least once a week as source of news*[1] *and they now have an estimated 60-70% penetration rate worldwide.*

This means that in less than a generation, digital media have radically transformed the way we inform and socialize, and that this transformation is still ongoing as older generations are gradually replaced by digital natives. From a scientific point of view, this transformation generates many phenomena to be studied, and even "unknown unknowns" whose effects will be revealed only with time.

This roadmap covers the issues, impacts and future challenges of digital media as they relate to human well-being in the broadest sense, from mental health to the health of democracies.

Its objective is to initiate a new interdisciplinary research community in this field, to define a research agenda, to formulate recommendations for future digital media policy and design, and to inspire future EU calls for projects to develop innovative and transdisciplinary research on these societal challenges.

The roadmap is the result of the EU-funded project DIGEING conducted by an international consortium with the help of an interdisciplinary advisory group of renowned experts. Its writing was based on an hybrid methodology developped at CNRS and

[1] Nic Newman, Craig T. Robertson, Richard Fletcher, Kirsten Eddy, Rasmus Kleis Nielsen, 2023. Digital News Report 2022. Reuters Institute.

*powered by **GarganText**[2], where the advisory group acted both as catalyst and guide for a larger collaborative mapping of the state-of-the-art and identification of challenges of that emerging field. More than forty researchers from fourteen European countries have contributed to the writing of this roadmap.*

To recruit these scholars, the advisory group conducted an iterative mapping of the scientific literature to identify expertise not initially represented in the group and thus solicit additional expertise in the relevant fields (see section I.C). A call for contributions was also issued, allowing additional contributors to join the collaborative work. An international conference concluded this roadmap exercise in Paris, the videos of which are available online.

This roadmap is complemented by online interactive maps of the literature and challenges identified in this roadmap. They can be used by researchers to situate themselves in this evolving scientific landscape and by research funding agencies to launch new calls for projects.

David Chavalarias
Roadmap coordinator
CNRS/ISC-PIF & EHESS/CAMS

[2] ***GarganText*** is an intelligent ICT tools developed at the CNRS Complex Systems Institute of Paris Île-de-France for the collaborative writing and mapping of large set of unstructured documents. It is a free software distributed under the aGPLV3/CECILL licence gargantext

Contents

I About this roadmap

A Roadmap contributors 8
A.1 Editors 8
A.2 Members of the roadmap's advisory board 8
A.3 Chapters contributors 9
A.4 Contributors to the Rome high-level symposium 10
A.5 Science mapping team at ISC-PIF 11
A.6 Partner institutions 11
A.7 Fundings 11

B Roadmap rationale 12

C Methodology 16
C.1 Where does "digital media" meet "well-being"? 16
C.2 Scientific landscape 20
C.3 Trends 23
C.3.1 Example: Evolution of VR and AR research 25
C.3.2 Example: Impact of digital media on children and adolescents 26
C.4 Collaborative writing of the roadmap and its challenges 27

II Roadmap

1 The interwoven on-line/off-line world 29

2 Digital embedding, mental health, cognition & identity 36

3 Challenges to personal freedom and agency 42

4 Digital literacy, social and individual well-being 47

5 Privacy based on pseudonymity vs. anonymity 56

6 Immersive social media and the metaverse 62

7 Interdisciplinary design for an inclusive and trustworthy digital public sphere 70

8 Impact of the digitalization of society, digital labor and well-being 76

9 Digital commons 84

10 Digital media in a sustainable world: towards inclusivity, equality and autonomy 90

11 Indentification of potential risks related to the Future of digital media 95

12 Conclusions and recommendations 101

Bibliography 106

Index 113

III Supplementary information

Appendices 117

.1 Queries used for the state-of-the-art 117

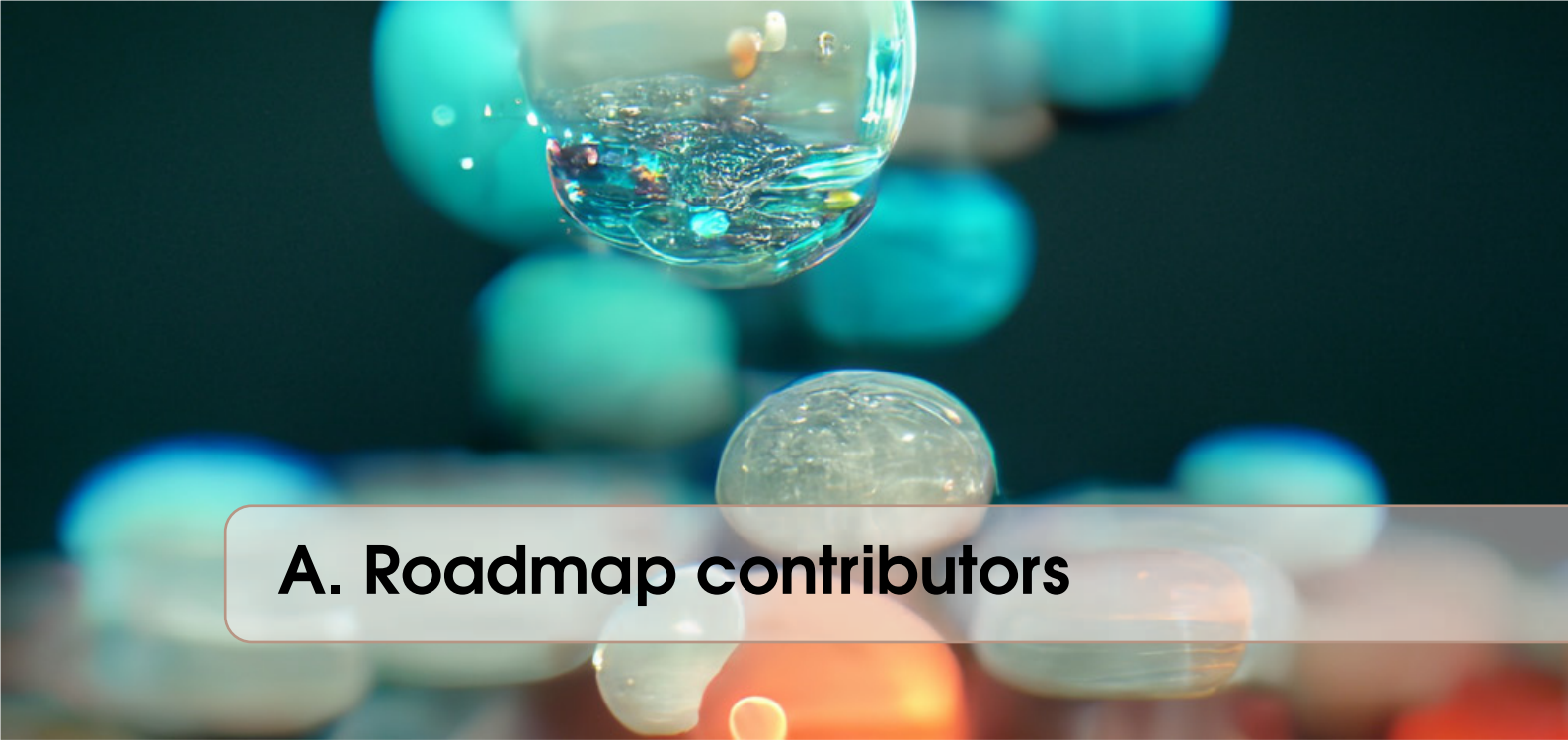
.2 Semantic perimeter of the research on *digital media and humanity well-being* 117

About this roadmap

A Roadmap contributors 8
A.1 Editors 8
A.2 Members of the roadmap's advisory board 8
A.3 Chapters contributors 9
A.4 Contributors to the Rome high-level symposium 10
A.5 Science mapping team at ISC-PIF 11
A.6 Partner institutions 11
A.7 Fundings 11

B Roadmap rationale 12

C Methodology 16
C.1 Where does "digital media" meet "well-being"? 16
C.2 Scientific landscape 20
C.3 Trends 23
C.4 Collaborative writing of the roadmap and its challenges 27



A. Roadmap contributors

This roadmap is the result of an 18-month consultation of the scientific community, during which two workshops (Universidad Politécnica de Madrid & Ca' Foscari University of Venice) and an international conference (CNRS, Complex Systems Institute of Paris IdF) were held. More than forty scientists contributed to its drafting.

A.1 Editors

Coordinator and main editor: **David Chavalarias**, CNRS research Director, Complex Systems Institute of Paris Île-de-France and Centre d'Analyse et de Mathématique Sociales (EHESS). Coordinator of the DIGEING project.
Associate editor: Mark Buchanan, physicist and science writer.

A.2 Members of the roadmap's advisory board

- **Beatrice de Gelder**, Maastricht University. *Expertise*: Cognitive neuroscience and neuropsychology, Augmented and Virtual reality
- **Guido Caldarelli**, Professor in Theoretical Physics at Ca' Foscari University of Venice. *Expertise:* Statistical Physics and complex networks
- **Melanie Dulong de Rosnay**, CNRS, head of the Center for Internet and Society. *Expertise:* law, open data, digital commons, technology
- **Fosca Giannotti**, Director of research of computer science at the Information Science and Technology Institute "A. Faedo" of the National Research Council, Pisa, Italy. *Expertise:* Privacy-Preserving Data Mining, AI Decision Making
- **Vittorio Loreto**, Sony CSL & Sapienza University of Rome. *Expertise:* complexity science and its interdisciplinary applications
- **Andrzej Nowak**, Department of Psychology, University of Warsaw, *Expertise:* social psychology and group dynamics,
- **David Pearce**, Professor at Universidad Politécnica de Madrid, Departamento de Inteligencia Artificial. *Expertise:* Artificial intelligence.
- **Richard Rogers**, University Profes-

sor and holds the Chair in New Media & Digital Culture at the University of Amsterdam. *Expertise:* Digital Methods
- **Camille Roth**, CNRS/CMB, Berlin, *Expertise:* socio-semantic networks, computational social sciences, sociology
- **Mel Slater**, University of Barcelona, ICREA Research Professor. *Expertise:* Augmented and Virtual reality

A.3 Chapters contributors

By alphabetical order of names:

- **Nahla Ben Amor**, Institut Supérieur de Gestion Tunis, Tunisia
- **Anna Boros**, University of Warsaw, Poland
- **Michela Brunori**, University of Urbino, Italy
- **Maria Jose Brites**, Lusófona University, Portugal
- **Guido Caldarelli**, DSMN, Ca'Foscari University of Venice, Italy
- **Antonio A. Casilli**, Institut Polytechnique de Paris, France
- **David Chavalarias**, Complex Systems Institute of Paris Île-de-France, CNRS, France
- **Caroline Datchary**, Laboratoire Interdisciplinaire Solidarités, Sociétés, Territoires, Université Toulouse Jean Jaurès, France
- **Alexandre Delanoë**, Complex Systems Institute of Paris Île-de-France, CNRS, France
- **Dana Diminescu**, Institut Polytechnique de Paris, France
- **Melanie Dulong de Rosnay**, Centre Internet et Société, CNRS, France
- **Luisa Fassi**, MRC Cognition and Brain Sciences Unit, University of Cambridge, UK
- **Divina Frau-Meigs**, Sorbonne Nouvelle University, France
- **Beatrice de Gelder**, Université de Maastricht, Nederland
- **Laura Hernandez**, Laboratoire de physique théorique et modélisation, Université de Cergy-Pontoise, France
- **Pawel Horodecki**, Gdańsk University of Technology, Poland
- **Sirkku Kotilainen**, Tampere University, Finland
- **Bertrand Jouve**, Laboratoire Interdisciplinaire Solidarités Sociétés Territoires (LISST), CNRS, France
- **Jean Lassègue**, LIAS CNRS, France
- **Emmanuel Lazega**, Sciences Po CSO, France
- **Quentin Lobbé**, Complex Systems Institute of Paris Île-de-France, CNRS, France
- **Paul Lukowicz**, German Research Center for AI (DFKI) and University of Kaiserslautern, Germany
- **Julian McDougall**, Bournemouth University, UK
- **Darian Meacham**, Maastricht University, Nederland
- **Andrzej Nowak**, University of Warsaw, Poland
- **Elisa Omodei**, Central European University, Austria
- **Amy Orben**, MRC Cognition and Brain Sciences Unit, University of Cambridge, UK
- **Geoffroy Patriarche**, Saint-Louis University, Belgium
- **David Pearce**, Universidad Politécnica de Madrid, Spain
- **Maria Gabriella Pediconi**, University of Urbino, Italy
- **Víctor Rodríguez-Doncel**, Universidad Politécnica de Madrid, Spain
- **Savino Romani**, University of

Urbino, Italy
- **Camille Roth**, Centre Marc Bloch, Berlin (CNRS/MEAE), France
- **Jerôme Sackur**, Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS, France
- **Valérie Schafer**, Université du Luxembourg, Luxembourg
- **Fabrizio Silvestri**, University of Rome, Italy
- **Mel Slater**, Institute of Neurosciences of the University of Barcelona, University of Barcelona, Spain
- **Aureli Soria-Frisch**, Neuroscience BU, Starlab Barcelona SL, Spain
- **Erika Stael von Holstein**, Re-Imagine Europa, Belgium
- **Luca Tummolini**, Institute of Cognitive Sciences and Technologies, CNR, Italy
- **Mirko Zichichi**, Universidad Politécnica de Madrid, Spain

A.4 Contributors to the Rome high-level symposium

Participants to the Rome high-level symposium *Are social media changing our brains? - Towards an interdisciplinary view on Social Media*.

- **Megan Boler**, University of Toronto
- **Guido Caldarelli**, IMT Lucca
- **Anne-Linda Camerini**, USI, Switzerland
- **Stéphane Chaudron** , JRC, Italy
- **David Chavalarias**, CNRS Paris
- **Beatrice de Gelder**, University of Maastricht
- **Derrick de Kerckhove**, University of Toronto
- **Beatriz Fagundo**, Centro de Estimulación Cerebral Barcelona
- **Marina Geymonat**, Telecom Italia, Milano
- **Kai J. Jonas**, University of Maastricht
- **Vittorio Loreto**, Sapienza Università di Roma
- **Geert Lovink**, Institute of Network Cultures, Amsterdam
- **Markus Mobius**, Microsoft
- **Andrzej Nowak**, Warsaw, Poland
- **Petra K. Novak**, Jožef Stefan Institute, Ljubljana
- **David Pearce**, Universidad Politécnica de Madrid
- **Oreste Pollicino**, Universita' Bocconi, Milano
- **Radu Prodan**, University of Klagenfurt
- **Camille Roth**, Centre Marc Bloch, Berlin
- **Mikko Salo**, Faktabari, Finland
- **Ulf-Dietrich**, Reips University of Constance
- **Mel Slater**, Universitat de Barcelona
- **Laura Smillie**, European Commission, Brussels
- **Bernard Stiegler**, Centre Pompidou, Paris
- **Mavi Sanchez Vives**, ICREA, Barcelona

A.5 Science mapping team at ISC-PIF

David Chavalarias (CNRS, ISC-PIF), **Alexandre Delanoë** (CNRS, ISC-PIF) & **Quentin Lobbé** (CNRS, ISC-PIF).

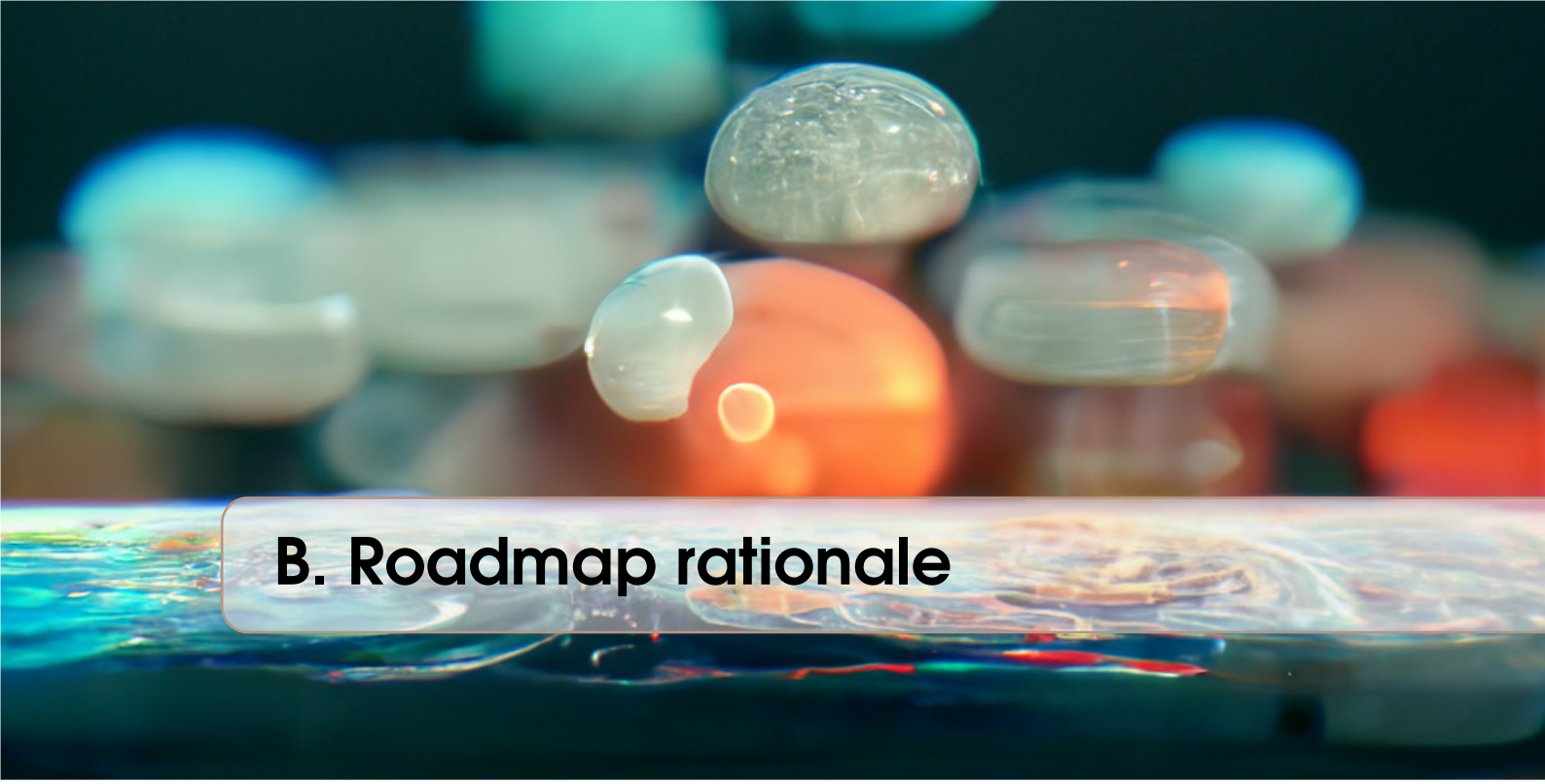
A.6 Partner institutions

Members of the DIGEING project:

- Centre National de la Recherche Scientifique, Complex Systems Institute of Paris Île-de-France and Centre d'Analyse et de Mathématiques Sociales, David Chavalarias (coordinator), Anne-Boucher (project officer, CNRS/ISC-PIF)
- Ca' Foscari University of Venice, Professor Guido Caldarelli.
- Universidad Politécnica de Madrid, Departamento de Inteligencia Artificial, Professor David Pearce.

A.7 Fundings

This roadmap and the DIGEING project were funded by the EUROPEAN COMMISSION Directorate-General for Communications Networks, Content and Technology, project VIGIE 2020-661 - LC- 01607979. EU officers : Peter Friess and Ralph Dum.



B. Roadmap rationale

Given the prominence in our lives today of digital online media and technologies, there is a clear need to better understand their impact on individuals as both users and as citizens. Deeper scientific understanding will be important in developing a next-generation of conscious users of digital technologies and digital or screen-mediated forms of communication. We must go beyond the default notion of "Designed in California, made in China, criticised in Europe"! Europe must bring its strong base of scientific knowledge and humanistic values to bear on discussions about the future of online media, and social media in particular.

The human brain was shaped by its evolutionary history and its properties reflect that history. The brain is in many respects "wired" for the social world, for several reasons. First and foremost, human infants experience a long dependence on others during ontogenesis. The primate brain is also wired to prioritize social information. Face, voice, and body expressions need to be decoded quickly and fluently as they provide crucial information required for future actions. Last, social interaction is intrinsically rewarding. Whether cause, consequence, or a phenomenon, human and non-human primates prefer to spend time in the company of others.

Because of the centrality of social interactions to human well-being, digitalization of communication, dominance of visual media and virtualization now create major challenges. Social media are effectively replacing the proverbial "village square," traditionally a place for meeting friends or foes, for catching up on news and gossip or counting ones' allies, for announcing public events, and for measuring consensus among the like-minded and debating others. Such social-meeting spaces existed in a defined time and space, and were rooted in the past, present, and future of the community, offering context – and filtering mechanisms – for the interpretation of statements or actions taken.

Much research notes dramatic changes to human information flows brought about by social media platforms such as Twitter, Facebook, YouTube, Instagram or TikTok which create communities with indifference to time, place, and history. These platforms are available anywhere, at any time, and for anybody. Here, understanding is by its nature relative, tied to the conditions that give meaning to words, and easily results in misunderstanding. Unsurprisingly, within such virtual platforms, semantics is up for grabs. The shared meanings and social norms humans have relied on for centuries are increasingly fragmented, with the rise of political polarization and intense in-group vs out-group animosity among the consequences.

It will be a challenge to build a fundamental understanding of how digital media are changing

individuals at both the personal and social levels, and to find pathways to building a more socially beneficial media. To this end, this roadmap will explore how an interdisciplinary collaboration of researchers across a very wide range of fields might help to address the opportunities and challenges presented by digital online media and, specifically, social media and by the ubiquitous presence of digital media and screens in our environments. Relevant areas of research include neuroscience and cognitive sciences, psychology and social psychology, sociology, behavioural economics and law, ethics and philosophy, computer science and computational social science, artificial intelligence, network sciences and complex systems science.

This roadmap aims at providing an important starting point for the collaboration of researchers from many disciplines and all over Europe and beyond to further a science-based approach to digital media. This approach should help to inform the decisions of governments and stakeholders in digital and online media.

The following points briefly summarise the main lines of thought and identify fields of interdisciplinary scientific collaboration outlined by the participants to the Rome symposium and further developed and elaborated by the contributors to this roadmap.

Research on Digital media and well-being should involve at least three levels:

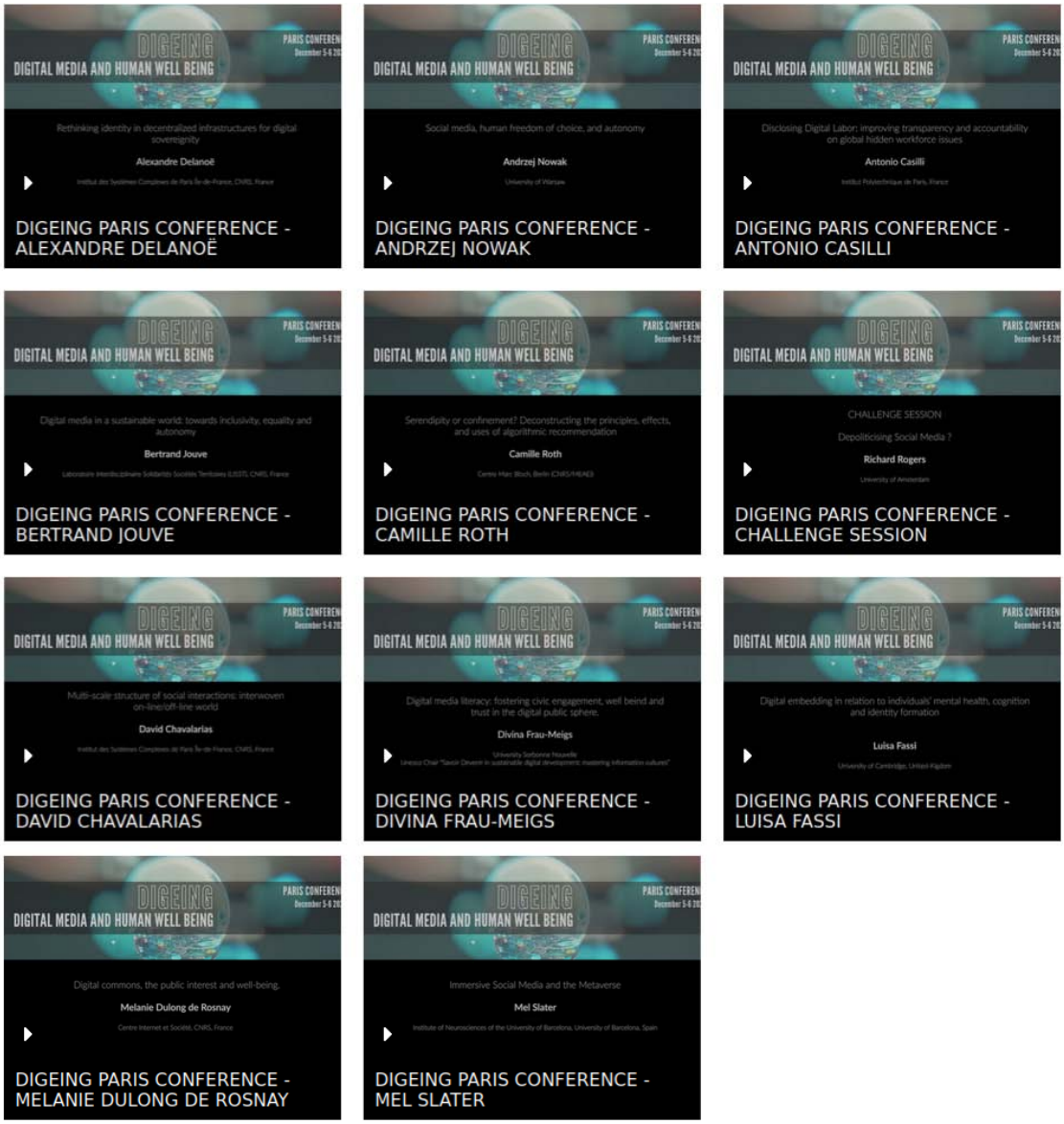
- **The individual ('microscopic') level:** our behaviour as individuals; how online media affect us as individual users and what determines our social media use.
- **The collective ('macroscopic') level:** the way individuals interact and form a collective and, ultimately, a society, and the way social media change how we interact online as compared to our offline interactions and communications.
- **The intermediary ('mesoscopic') level:** how the individual and collective levels influence each other; how we collectively create new structures and how these structures, in turn, affect the individual level; what is new in the way we create such structures in the age of online and social media.

In practice, these levels are not clearly separated, but share a deep interdependence. Keeping this interlinked, multilevel structure in mind, the chapters of this roadmap address distinct, though sometimes overlapping, domains characterised by specific interdisciplinary challenges:

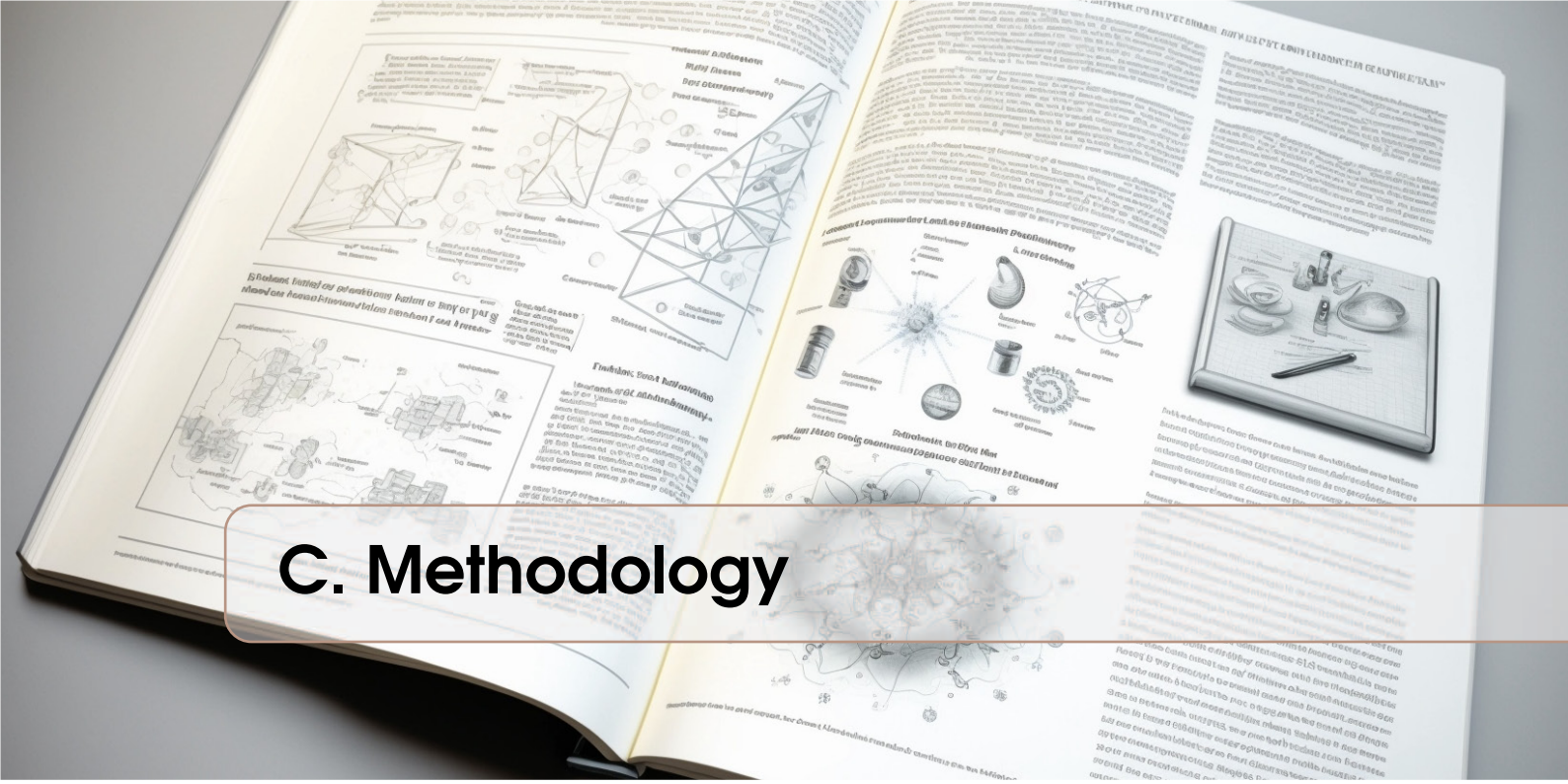
- **1. Multi-scale structure of social interactions: interwoven on-line/off-line world** Embedding individuals into the digital environment changes how we interact with one another, with the potential to affect how our societies create and agree on shared norms and knowledge. Can we forge a deep understanding of what the consequences of such digital embedding may be our cognitive and social dynamics in our interwoven on-line/off-line worlds?
- **2. Digital embedding in relation to individuals' mental health, cognition and identity formation.** These new digital environment have different effects and different categogies of population. Growing concern over the impact of social media use on adolescents' mental health, cognitive development and identity formation has stimulated scientific research to identify the potential risks from the use of digital technologies. This research remains largely inconclusive, however, motivating an urgent need for further investigation in this area.
- **3. Challenges to personal freedom and agency.** Human autonomy, or the ability to make independent decisions and turn them into actions, is among the central elements of being human. The rise of digital media poses a number of direct challenges to human autonomy and freedom. We require research to learn how such crucial values can be preserved and supported.

- **4. Digital literacy, social and individual well-being.** Our human activities are increasingly embedded in digital environments which influence all aspects of modern life, including education, politics, work and leisure. Despite this trend, a focus on Digital Media and Information Literacy (DMIL) is absent in curricular activities and in the formal education sector. We urgently need to create innovative pedagogical and research postures to create a literate citizenry ready to face global uncertainties and crises.
- **5. Privacy based on Pseudonymity vs Anonymity, rethink Identity towards decentralized infrastructures for sovereignit.** Privacy can be simply defined as the ability to control who has access to personal information and how it is used. In this era of digital media, there is a growing recognition of the need to protect privacy in various contexts, such as online communication, social media, and e-commerce. Research in this area can help to identify and address potential threats and to develop effective solutions for protecting personal information.
- **6. Immersive Social Media and the Metaverse.** Virtual reality (VR) has moved from being an esoteric and expensive tool of university labs and industry into a low-cost consumer product. As VR enters the mass market, however, there is little or no ethical supervision despite a lack of empirical data concerning the effects of VR on individuals or groups. While companies envision a world changing "metaverse" unleashing creativity, we also urgently need a deeper understanding of the risks involved and how such technologies may change both individuals and societies.
- **7. Interdisciplinary design for an inclusive and trustworthy digital public sphere.** The public sphere increasingly exists online in digital media spaces designed for profit rather than to support public values. We need new regulatory tools to help steward best practices for the European digital public sphere and to ensure socially beneficial outcomes.
- **8. Impact of the digitalization of society, digital labor and well being.** The rhetoric surrounding the displacement of jobs by automation fails to acknowledge the role played by human labour in AI systems. Yet an expanding body of literature documents the adverse impacts of digital labour on well-being, especially mental health. The Digital Labour Disclosure challenge provides a unique opportunity to evaluate the social costs of this type of work. By formalising digital labour's role and mapping its global operations and supply chains, this framework can help establish a coalition of institutions and companies to improve corporate accountability and transparency regarding digital labour workforce issues.
- **9. Digital commons.** A socially beneficial system of rights and policies is required to support a sustainable digital commons ecosystem, ensuring fair access rights to data, knowledge and information, and also enhancing collaboration. Currently, different legal regimes regulate access to data and more broadly rights pertaining to the reuse of personal data, privately-owned or hosted data. As the possibilities for new legal structures are evolving rapidly, it is an urgent challenge to understand how, in practice, we can fully realise the potential of the digital commons.
- **10. Digital media in a sustainable world: towards inclusivity, equality and autonomy.** Digital technology is rapidly transforming all of our lives, yet the way our digital media are being developed does not reflect our ambition to pursue a learning, inclusive, responsible and autonomous society. To ensure that digital technology can contribute to global well-being, it is urgent to radically change our usage habits as well as the tools developed.
- **11. Indentification of Potential Risks related to the Future of Digital Media.** New

digital technologies have already caused numerous and significant problems linked to human well-being, yet potential risks and future problems remain largely unexplored. We urgently need to create research pathways to foresee such risks, for they will broadly impact media users and the society at large.



Some of these challenges have been presented during the Paris conference on Digital Media an Human Well-Being. The corresponding videos are online at https://iscpif.fr/digeing/digeing-paris-conference-replay.



C. Methodology

C.1 Where does "digital media" meet "well-being"?

Digital media and *Human Well-Being* are to board areas of research. Consequently, the first challenge of this roadmap was to identify the diversity of research communities that had adressed their intersection.

For this purpose, we mobilized ***GarganText***[1] [95], a collaborative tool for dynamical literature mining, in order to help experts of the field to map the state-of-the art related to *digital media and humanity well-being*. ***GarganText*** also features a new-generation collaborative environment, invented by Alexandre Delanoë. It allows collaborative writing within a group of people of any size to be organized seamlessly and their productions to be included in the mapping in real-time. With a knowledge mapping engine at the heart of the roadmap exercise, the strategy was therefore to first identify the current trends and disciplinary diversity of the field by analyzing the global literature (60*k* documents in total). In a second step, we built on this to describe the challenges of this domain of research and produced their map in ***GarganText***. Last we could made recommendations for future digital media policy and design (*cf.* Fig. C.1).

Simple quantitative analysis

Let's start with a standard bibliometrics analysis from *Web of Science* (WoS) to get a first flavor of the momentum and diversity of the research on digital media and human well-being. What is for example, the characteristics of the research related to the research area defined by the following query in the WoS ? :

> **Query** $Q1$ **[47,8k docs up to 2023-03]:**
> ("social media" OR "virtual reality" OR "augmented reality" OR "Web 2.0" OR "social networks" OR "facebook" OR twitter OR Internet OR "digital media" OR "smartphones") AND ("Well-Being" OR wellness OR happiness OR wellbeing OR democracy OR opinion OR politics OR election)

[1] ***GarganText*** is developed at the CNRS's Complex Systems Institute of Paris Île-de-France by Alexandre Delanoë and David Chavalarias. It is a free software distributed under the aGPLV3/CECILL licence.

Figure C.1: Steps to produce this roadmap.

First, it is a fast growing research area (Fig. C.2-left) with an average of 19,7% growth per year in number of publications between 2004 (creation of Facebook) and 2022 (to be compared to a general growth of the overall publications volume of about 7.9% in the WoS in 2019) ; and a peak at +42% in 2016 (the year of the advent of the so called « post-truth era »).

Research in this field is very recent, with less than 300 articles per year in 2004 and the previous years, while the current volume of annual publications is twenty times higher.

Moreover compared to other research communities such as *graphene* or *deep learning*, research on *digital media and humanity well-being* can still be considered as an emerging field (*cf.* table C.1 and section C.3).

Topic	$Q1$	**Graphene**	**Deep learning**
#pub over 2020-2022	18,700	116,000	127,000

Table C.1: **Comparison of the relative weight of the research output in the Web of Science.** Even if $Q1$ is not supposed to cover all the research related to digital media and well-being at this stage, and even taking into account the bias of the WoS where hard sciences and technology are over-represented, the gap is still significant and research related to $Q1$ can be considered as emerging (see also sect.C.3).

However, as seen on Fig C.2-right, the community is growing rapidly, both in terms of its total size and the size of its core of scholars (scholars who have published more than two papers over a three-year period). As this growth accelerates, *it is expected that this research community will continue to grow very significantly over the next few years.*

Second, focusing only on the literature that mentions research contributions covered by the query $Q1$ (Fig. C.3), the relevant papers are strikingly dispersed across a large number of academic disciplines. Altough computer sciences represents about 25% of the whole, there is no clear dominance by any of them. Moreover, "hard sciences" and social sciences and humanities are both represented. This points to a strong potential for the development of interdisciplinary research in this field. To leverage this potential, a research agenda on *Digital Media and Human Well-Being* should therefore promote a truly interdisciplinary research and develop research

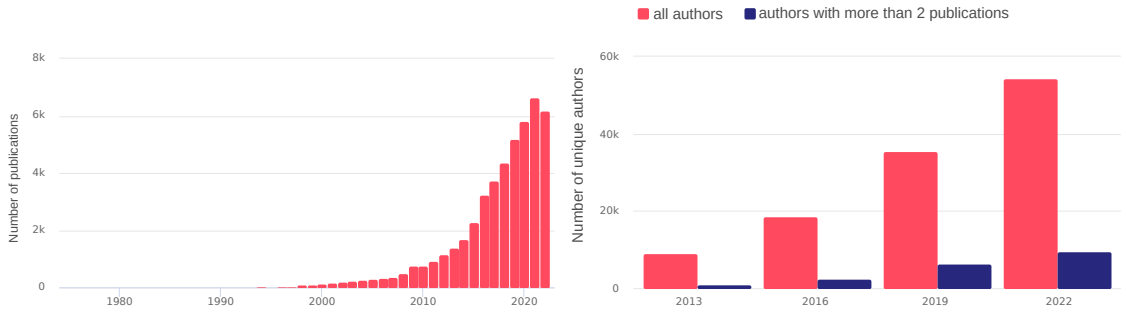


Figure C.2: *Left:* **Evolution of the number of publications** retrieved by $Q1$ in the Web of Science ; *Right:* **Number of unique authors as identified by the WoS for $Q1$.**

paths at the interfaces between these disciplines.

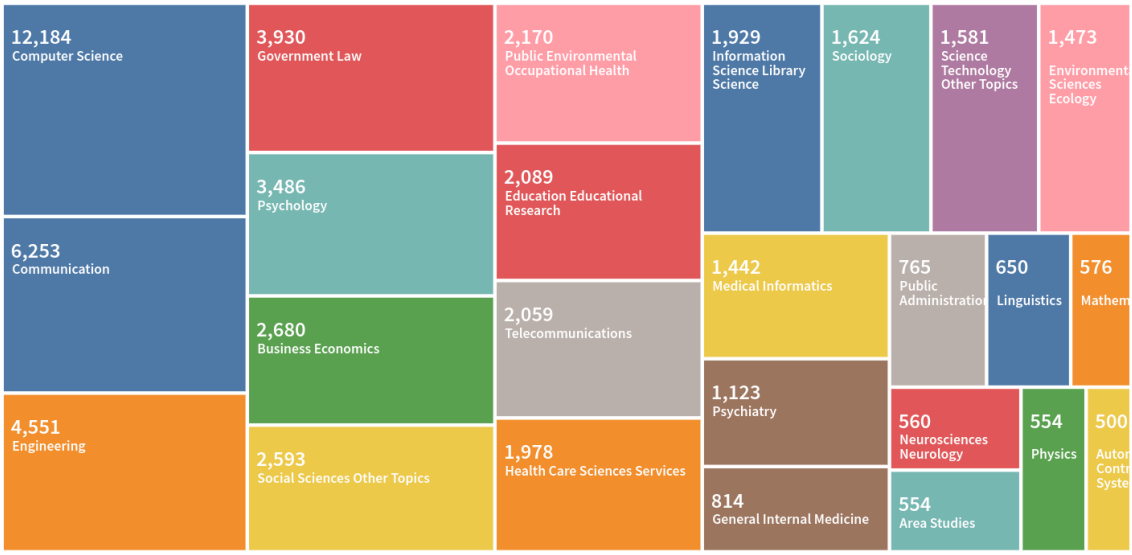


Figure C.3: Disciplines of papers related to the query $Q1$ in the Web of Science database (47k records) as computed by the WoS. This corpora is only a part of the literature related to this roadmap. *Warning:* areas in this WoS tree map are not proportional to the number of papers.

Finally, while several EU countries appear in the top 25 of the most publishing countries on this research area (*cf.* Fig. C.4), there is considerable heterogeneity across EU countries in their involment in this type of research (*cf.* Fig. C.5). For exemple, Finland has a relative research focus on this research area of $+62.7\%$ compared to other countries in this set while France has a relative research focus of -50.6%. For comparison, the research focus of U.S.A in this domain is exactly in the average ($+0\%$ of relative research focus).

The first challenge for a *research agenda on digital media and humanity well-being* is to transform this multidisciplinary research field into an interdisciplinary and transdisciplinary field, by *promoting interdisciplinary collaborations and supporting researchers with mixed scientific cultures*. It is therefore necessary to create bridges and spaces for dialogue between these disciplines in order to foster the emergence of an interdisciplinary research field and community, and to be able to identify researchers working at the interface between these disciplines.

In any case, it is necessary to start by identifying, as precisely as possible, the various research currents involved in this field and then reflect on the interdisciplinary challenges that they could

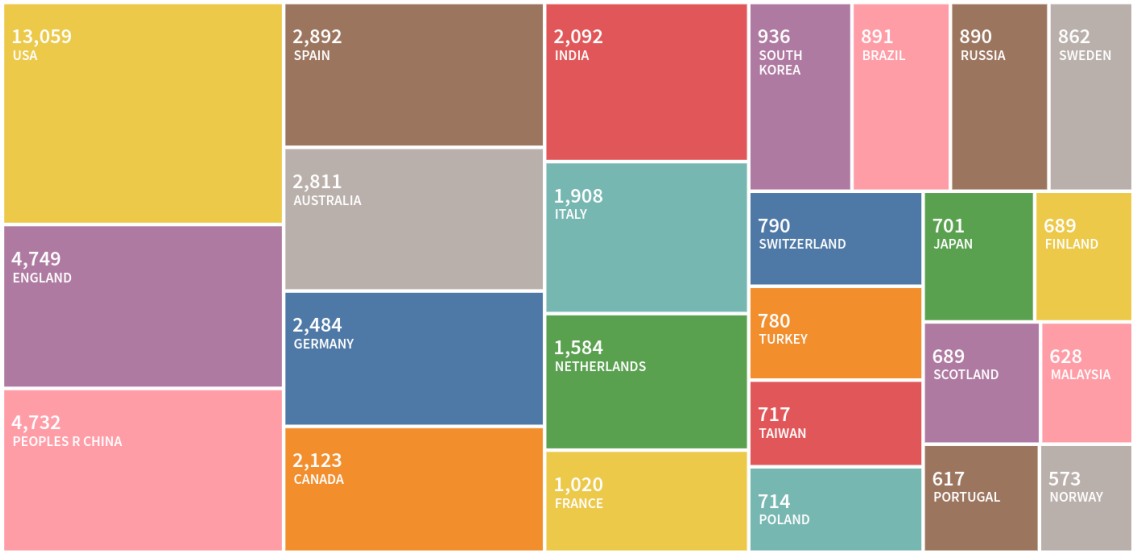


Figure C.4: Number of publications per countries for query $Q1$. *Warning:* areas in this WoS tree map are not proportional to the number of papers.

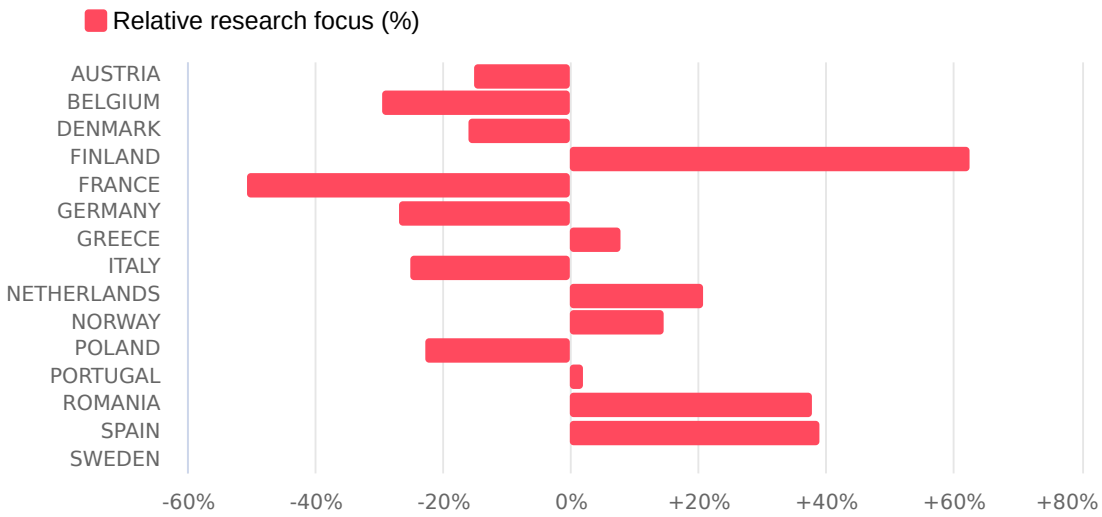


Figure C.5: Relative weight of $Q1$ publications in the production of the top publishing EU countries compared to the average weight among those countries (the relative research focus is defined as the number of $Q1$ publications per country, normalized by the number of WoS publications of these counties in 2022 (any topics) and the average of the normalized number of $Q1$ publications over all these countries).

address.

Iterative collaborative mapping of a scientific field

In order to get a bird eye view the research on *digital media and humanity well-being*, we have mobilized text-mining technologies to map this scientific area.

Science mapping is a method used in scientometrics and bibliometrics to visually represent the research landscape of a specific field or topic. It combines advanced text-mining, information

visualization techniques and statistical analysis to identify patterns, relationships, and trends in scientific literature.

The collaborative platform ***GarganText*** allows a group to leverage the expertise of its members through AI and text-mining methods. It allows for interactive annotation of a large set of documents to define its specific vocabulary while keeping the human in the loop at all times.

Once this vocabulary has been identified, ***GarganText*** uses co-word analysis and network analysis (cf. [10, 15]) to identify topics and their relations.

The state-of-the-art mapping for this roadmap was done in a recursive way all along the 18 months of the DIGEING project, as described on Fig. C.6, in order to progressively incorporate the visions and expertises of new contributors to the roadmap. 58,000 scientific papers have been analysed at the end of the process[2] and 1,650 core expressions have been identified to describe this interdisciplinary field of research (the full list is provided in SI.2).

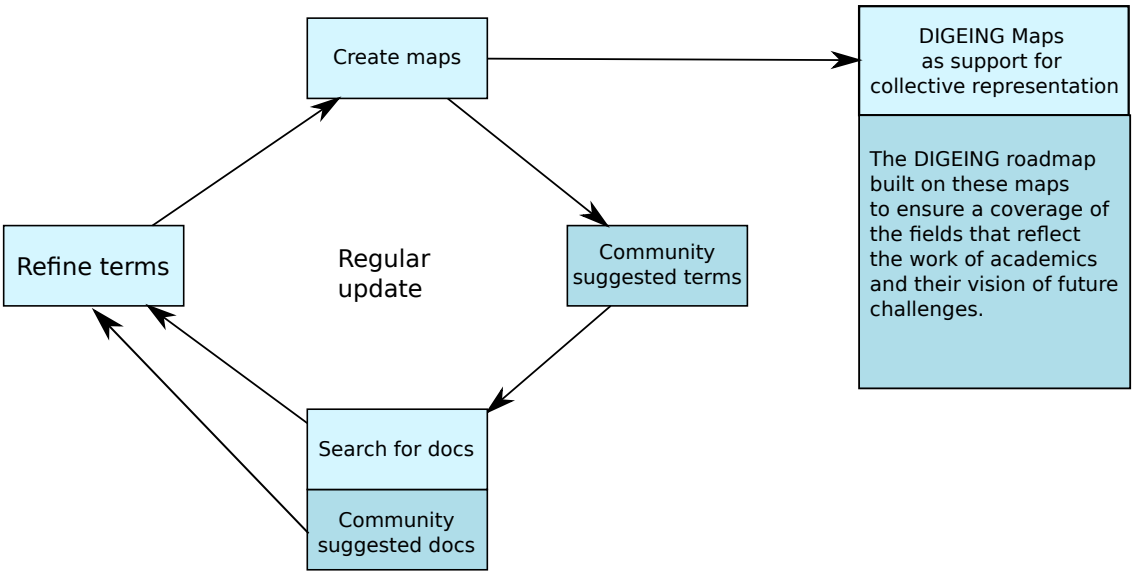


Figure C.6: Worlflow for the iterative science mapping with ***GarganText***.

Although the whole literature related to *digital media and humanity well-being* is definitely not fully contained in these 58*k* papers, and even if the WoS database is somehow biased toward hard sciences [40], the scientific landscape obtained by the screening of these documents offered a solid ground to depict the different research areas concerned by *digital media and humanity well-being* and coordonate the prospective activity of roadmap contributors.

Two types of interactive visualizations have been generated from this science mapping, that help us to identify important topics and trends in the *digital media and humanity well-being* literature : static maps and phylomemies (dynamics structures, see below and [11, 12, 33]).

As described in the following, these two types of visualization extend the description of the multidisciplinary aspect of a *research agenda on digital media and humanity well-being* (cf. Figure C.3) and point to emerging trends.

C.2 Scientific landscape

Scientific literature is the digital record of the contributions of a myriad of scientists who interact in a decentralized manner in advancing science. As such, it is a complex system that admits several

[2]The detailed queries used to retrieve those paper is available in SI .1.

levels of observation: the document and its concepts, the research topics and their epistemic communities, the branch of science in which these topics are embedded constitute a scientific domain, the general articulation of domains or branches of knowledge [12].

The mapping of the 58*k* papers related to *digital media and humanity well-being* with *GarganText* highlights these different levels of observation as well as the great diversity of the scientific questions addressed, which entangle individual concerns with collective issues.

The map of Fig. C.7 –also **available online** in an interactive version–, shows that *digital media and humanity well-being* research is articulated around three large domains: **health**, impacts of digital media on **society at large**, and **technology** issues and opportunities. All these fields interact via research conducted at their interface and a fourth domain bridge them all: **ethics and regulation** (see Fig. **??** & Ch. 1, 5, 6, 7, 9, 11)

The three main domains are themselves composed of sub-domains as followed (the roadmap chapters that extend the reflection on these subjects are indicated in brackets):

- **Health**
  - The opportunities of digital media in the field of public health, especially for personalized health monitoring via mobile applications (mhealth) or for education and awareness,
  - The mental disorders induced by the intensive use of mobile devices, in particular in adolescent and young adults (Ch. 2),
  - The opportunities but also the risks of digital media for the improvement (or deterioration) of cognitive functions, especially in case of neuroplasticity (Ch. 2).
- **Technology**
  - The opportunities but also the risks of *augmented reality* and *virtual reality* (Ch. 6),
  - *Privacy issues*, *personnal data* management and the new opportunities brought by *decentralized protocols* such as blockchains (Ch. 5),
  - The issue of *recommendation* and profiling, including the question of *trust* in online services and the new opportunities for matching offer and demand (ex. on job markets) (Ch. 7),
  - The developement of digital media for *e-services* in a wide range of economic domains, from tourism to online courses.
- **Society at large**
  - The issue of *digital media literacy* and their impact on *identities* and the evolution of *cultural norms* (Ch. 4),
  - New form of *citizenship* and *political participation*, and in particular the impact of social media on political mobilisation, partisanship, protests and radicalization (Ch. 1),
  - The role of digital media in the *disinformation* or *manipulation of opinion*, and their impact on the dynamics of opinion and democratic processes, as well as their use by totalitarian regimes to monitor and control public opinion (Ch. 1, 6).
- **Ethics and regulation:** The ethics of AI, the recommandation, social data reuse and VR technologies ; regulation of the BigTech sector and its algorithms (explainability, transparency, etc.).

The map of Fig. C.7 was used in the early months of the DIGEING project to ensure that most important subdomains were covered by at least one advisory board member or external contributor. The goal was to begin the foresight phase by ensuring that all of the expertise needed to consider future challenges in an interdisciplinary manner would be engaged.

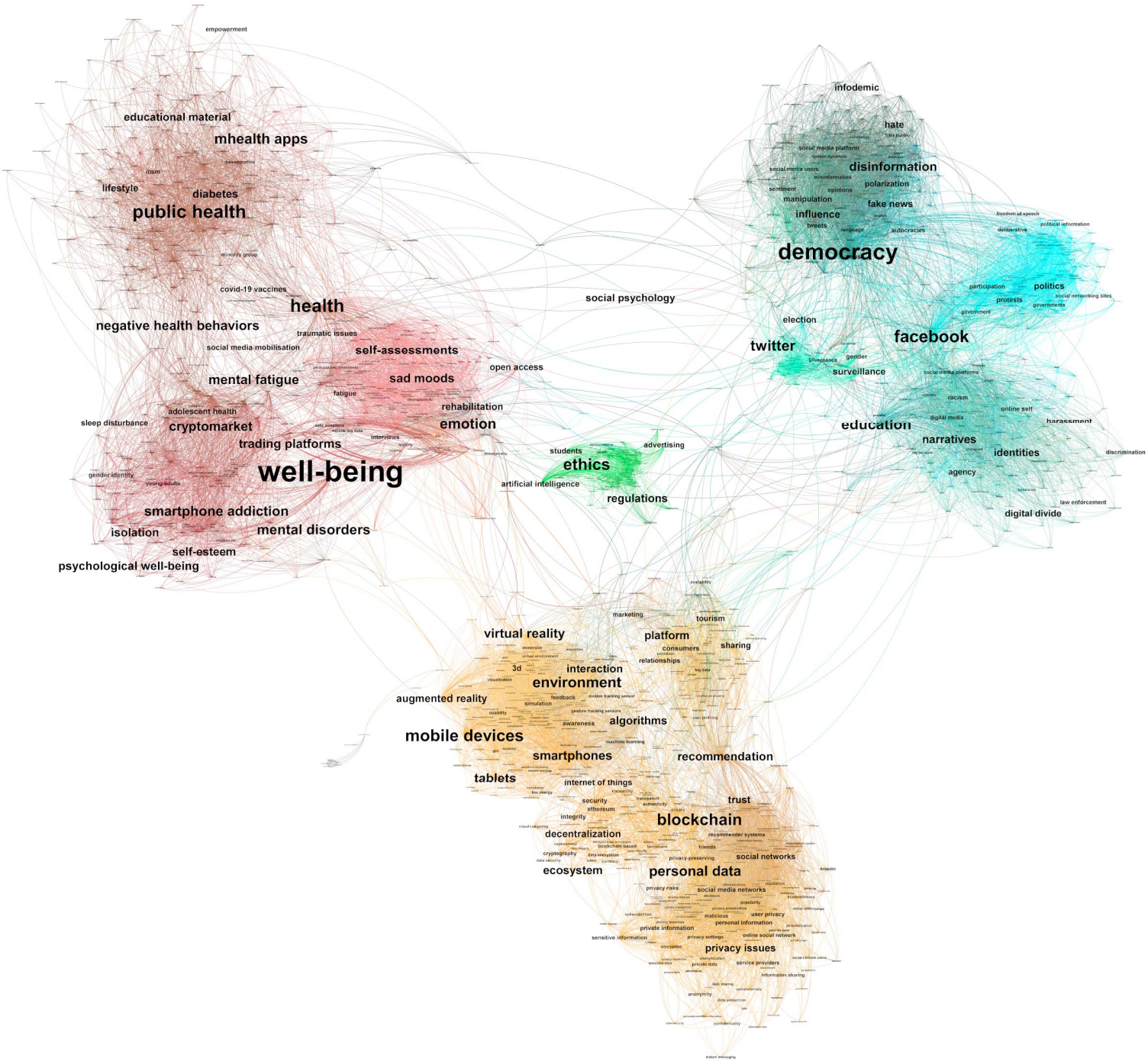


Figure C.7: Mapping of 58k publications related to *digital media and humanity well-being* over the period 1991-2022. Links between terms represent their relation measured with the *confidence proximity measure*. Most term sizes are proportional to their importance in the network of terms (Page Rank), but some have been manually increased to allow for better correspondence with the roadmap chapters. Map generated with ***Gargan Text*** [95] and spatialized in Gephi [89] with the Force Atlas algorithm [29]. The *confidence proximity* between two terms i and j is the max of the estimation of the two probabilities of having one term given the presence of the other in the same document.

C.3 Trends

The map of Fig. C.7 is usefull to uncover the diversity of issues around *digital media and humanity well-being* but does not provide any trends as it is computed over the whole period 1991-2022. To go beyond this static snapshot, we generated the *phylomemy* of the field with *GarganText*.

A *phylomemy* is roughly speaking the reconstruction of the dynamics of a field of knowledge from a series of science maps [71, 12, 33]. It allows to visualize the evolution of knowledge branches in a given corpus of any type (academic articles, patents, tweets, etc.) provided that its documents are time-stamped.

Applied to science, phylomemy reconstruction gives an overview of the dynamics of knowledge at the scale of an entire research field. It allows us to identify key periods in the development of the field, to detect emerging topics and to follow the evolution of key concepts.

In the example Fig. C.8, two research topics related to online harassement have been identified from the analysis of the relations between our 1,650 core expressions in the documents published on the periods 2013-2018 and 2014-2019. They are represented by keywords in circles. The phylomemy reconstruction has identified theses topics as related, which is visually represented by a vertical link between the two circles. We can see from the keywords that the focus seems to have evolved from the general question of harassement related to identity to the harassement and victimization in adolescent populations. The terms describing those field are in the same area and in the same color on C.7. When they are connected through different periods, they form branches of science as depicted on Fig. C.9.

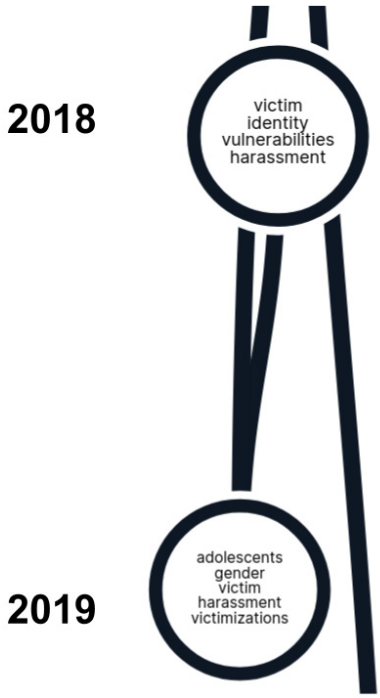


Figure C.8: Example of two fields of research at two consecutive periods that present conceptual similarities and have been associated by the phylomemy reconstruction algorithm.

It must be emphasized that a phylomemy is always drawn for a given level of observation (cf. [12]), a bit like geographical maps which are more or less detailed depending on the zoom level. This implies that we see trends at a certain resolution and that the fact that a branch emerges or disappears indicates that the research area has gained (resp. lost) enough intensity to appear (resp. disappear) from the overall view but not that it is totally inexistant when not featured in the phylomemy.

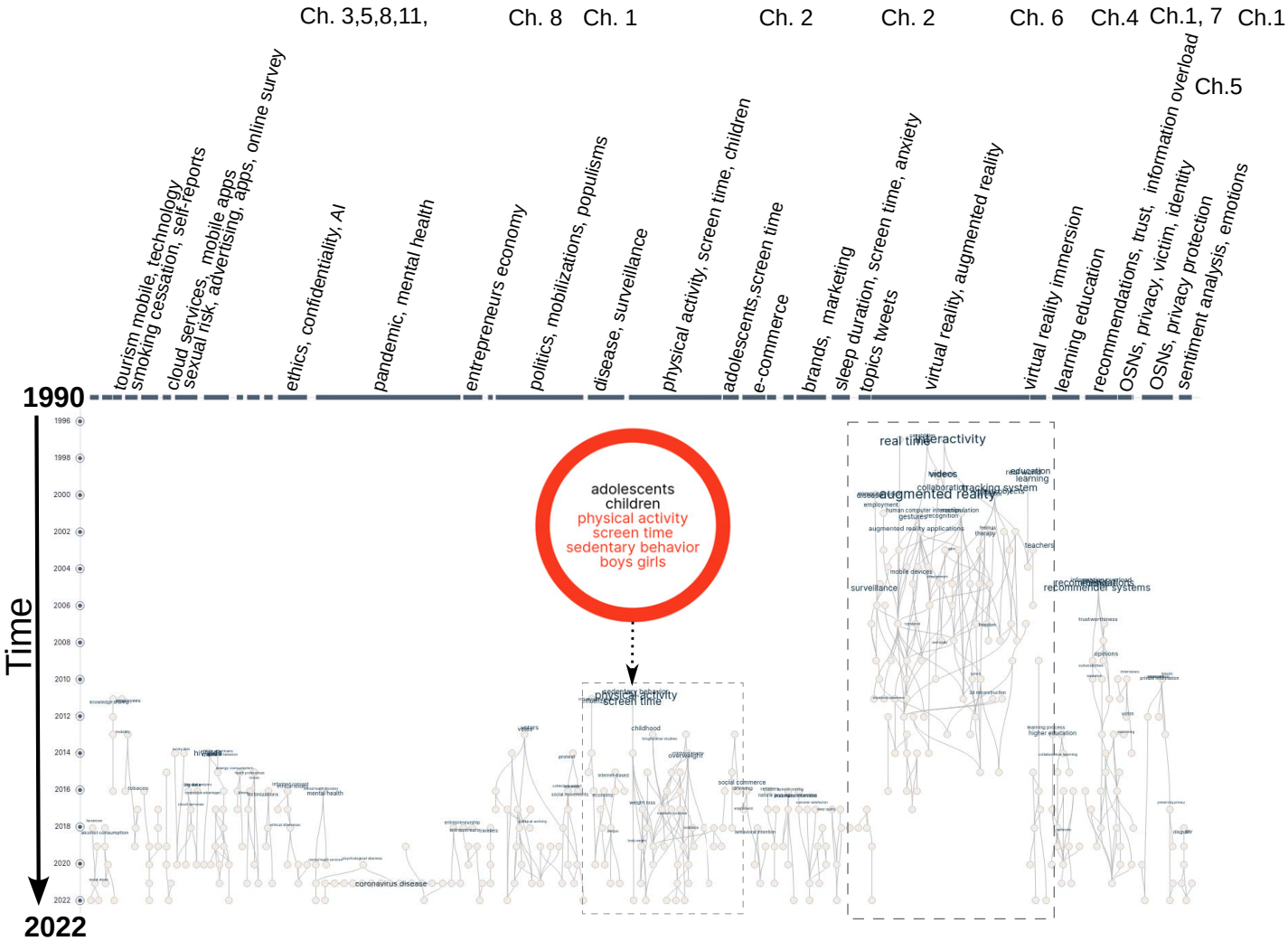


Figure C.9: Phylomemy of the research on *digital media and humanity well-being*. Time flows from top to bottom. Each bubble is a research area automatically identified by *GarganText* over a given period and defined by a few keywords in a circle. An example is given with the enlarged circle, which corresponds to research on the impact of screen time on children and adolescents. Circles are organized into branches of science that depict the evolution of a given research field. The names of the most important branches are given on the top as well as the roadmap chapters that intersect with branches topics. This phylomemy reveals that the research on digital media, which was mainly focused on VR/AR until the end of the 2000s, then diversified considerably with the arrival of social networks and mobile terminals to touch on a very wide range of issues, from mental health in children and adolescents to political participation and misinformation, including applications to education, tourism or marketing. Note the relatively recent emphazis on ethical issues, particularly in relation to AI, which corresponds to the central green cluster of the map Fig. C.7. The two branches highlighted by dotted line are detailed in sections C.3.1 & C.3.2. An interactive version of this visualization is available online.

The detailed analysis of this phylomemy helped us to give context to the current trends in *digital media and humanity well-being* and orient the chapters of the roadmap. Let's comment the branches highlighted with dotted lines in Fig. C.9 to give examples of how such visualization can guide the interpretation.

C.3.1 Example: Evolution of VR and AR research

Virtual and augmented reality technologies are emerging as the first areas at the intersection of digital media and human well-being, particularly in the field of surgery, therapy, education and game, with an intense development from 1990 (the start of our corpus) to 2017 (Fig. C.10). It had several offshoots before this branch gradually fade out between 2012 and 2017 in favor of a new branch, still focused on VR, but with an emphasis on head-mounted displays. We can think that this corresponds to the arrival on the market of new cheap devices such as Occulus Rift, bought by meta in 2014, which have considerably boosted the applications of this type of technology. Research challenges related to this technology in relation with individual and collective well-being are adressed in Chapter 6.

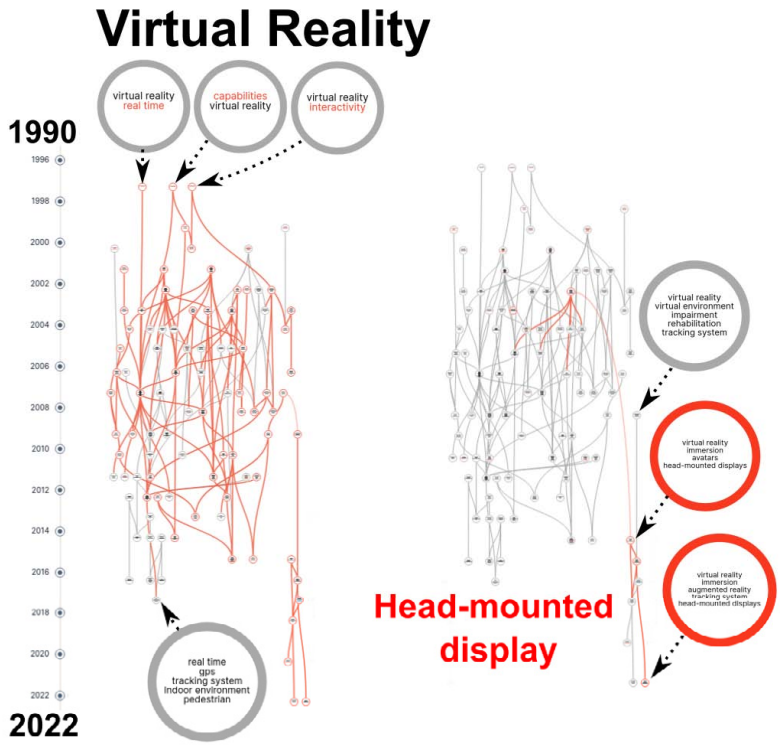


Figure C.10: Detail of two branches mentionning "*Virtual reality*". On the left, fields mentionning "virtual reality" are highlighted in red. On the right, those mentionning "*head-mounted display*" are highlighted in red. We can see that VR is present in both branches (contrary to AR that is mostly in the biggest one) and that the second branch is characterized by the presence of *head-mounted display* technology as well as an emphasis on immersion (compared to AR that mixes virtual and analogic worlds).

Main terms by order of occurrences in the oldest VR branch : *virtual reality, augmented reality, real time, interactivity, tracking system, videos, virtual objects, virtual environment, real world, visualization, gestures, patients, robot, augmented reality, applications, manipulation, user interface, hardware, recognition, computer vision, surgery, immersion, therapy,...*

Main terms by order of occurrences in the newer VR branch : *virtual reality, immersion, virtual environment, avatars, interactivity, tracking system, head-mounted displays, augmented reality, 3d interaction, impairment, 3d user interfaces, rehabilitation, human subject*.

C.3.2 Example: Impact of digital media on children and adolescents

Several phylomemetic branches address the issue of the impact of digital media on children and adolescents. One of these, shown in Figure C.11, specifically addresses the issue of excessive screen time, which impacts physical activity patterns and may be a cause of overweight and obesity. The issue of sleep also became a significant concern since 2017.

More recently, parents, educators, policymakers, and the general public have become concerned about the impact of social media use on adolescents' mental health, cognitive development, and identity formation, which is emerging in the 2016-2021 period. However, despite these efforts, the picture that emerges from the current literature remains inconclusive and additional research is needed in this area to provide a more thorough understanding.

The cognitive developpement issue has been identified as critical for the research on *digital media and humanity well-being* and has been addressed in the chapter 2.

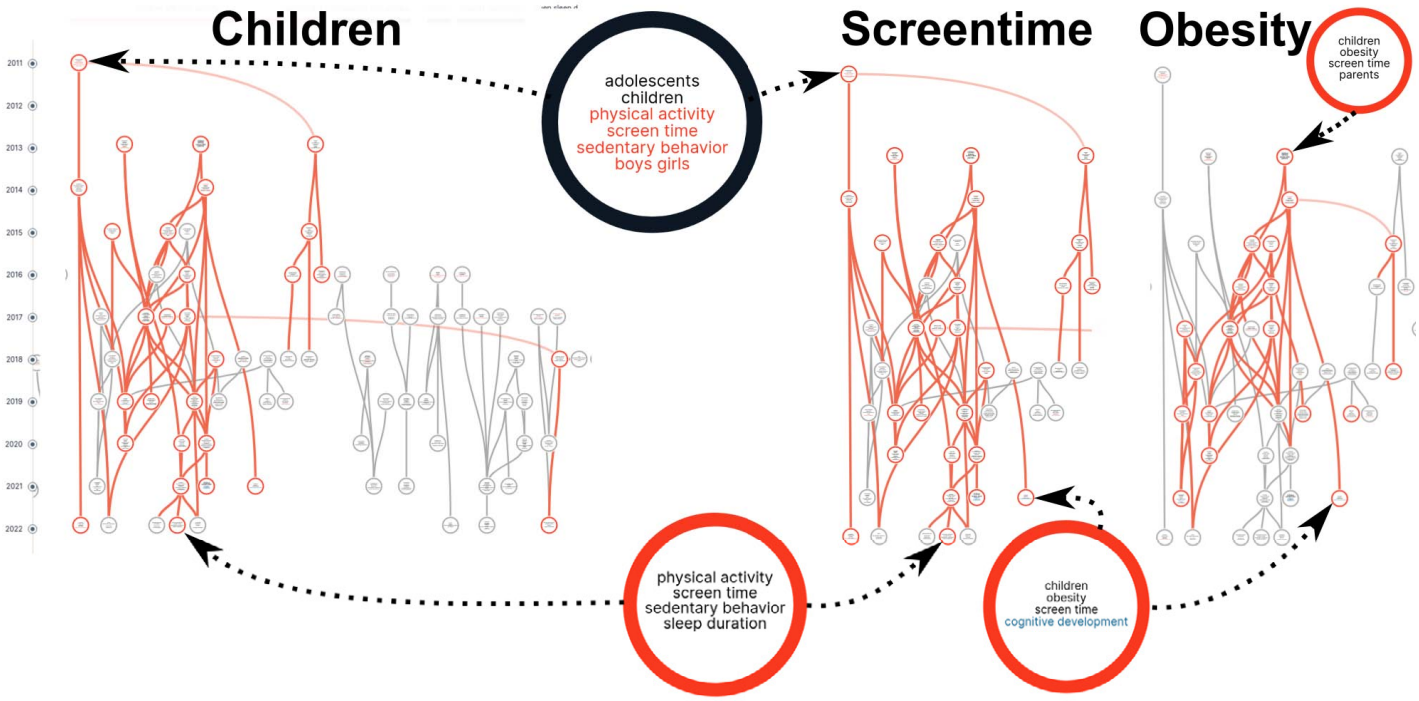


Figure C.11: Detail of some branches dealing with *"children"* and *adolescents T*. On the left, fields mentionning "children" are highlighted in red. On the right, the same branch is diplayed, with fields mentionning *"sceentime"* or with fields mentionning *"obesity"* highlighted in red.

Terms by order of occurrences in topics: *physical activity, screen time, children, obesity, parents, overweight, diet, bmi, sedentary behavior, childhood, lifestyle, young children, adolescents, weight loss, health, childhood obesity, longitudinal studies, boys, girls, preschoolers, toddlers, sleep, diabetes, schools, mobile technology, health behaviors, body weight, hypertension, blood pressure, sleep duration, metabolic syndrome, education, gender, belief, self-reports, health promotion, smartphone apps, screen-based, mediator, caregivers, quality of life, foods, infants, chronic diseases, cognitive development, self-tracking, excessive screen time*.

C.4 Collaborative writing of the roadmap and its challenges

The state of the art map and the phylomemy were both used to structure the collective writing of the roadmap. Two writing workshops defined the chapters and groups of contributors who then continued to collaborate on writing the chapters under the responsibility of a chapter editor.

Scientists with complementary skills to those of the advisory board were invited to these workshops in order to cover the different topics of the state-of-the-art as well as possible and to allow a real interdisciplinary reflection on the promising avenues of research at the intersection of *digital media and humanity well-being*. These topics were used as a starting point to identify gaps in the current mainstream research and future perspectives to be addressed by roadmap chapters. The chapters were written in *GarganText* under the responsibility of one or several editors.

Each chapter includes few challenges and there is a certain circulation of problems between the chapters. In total, more than 150 challenges have been identified and mapped by *GarganText* (see Fig. C.12). The topics covered by the roadmap chapters are summarizes in section B.

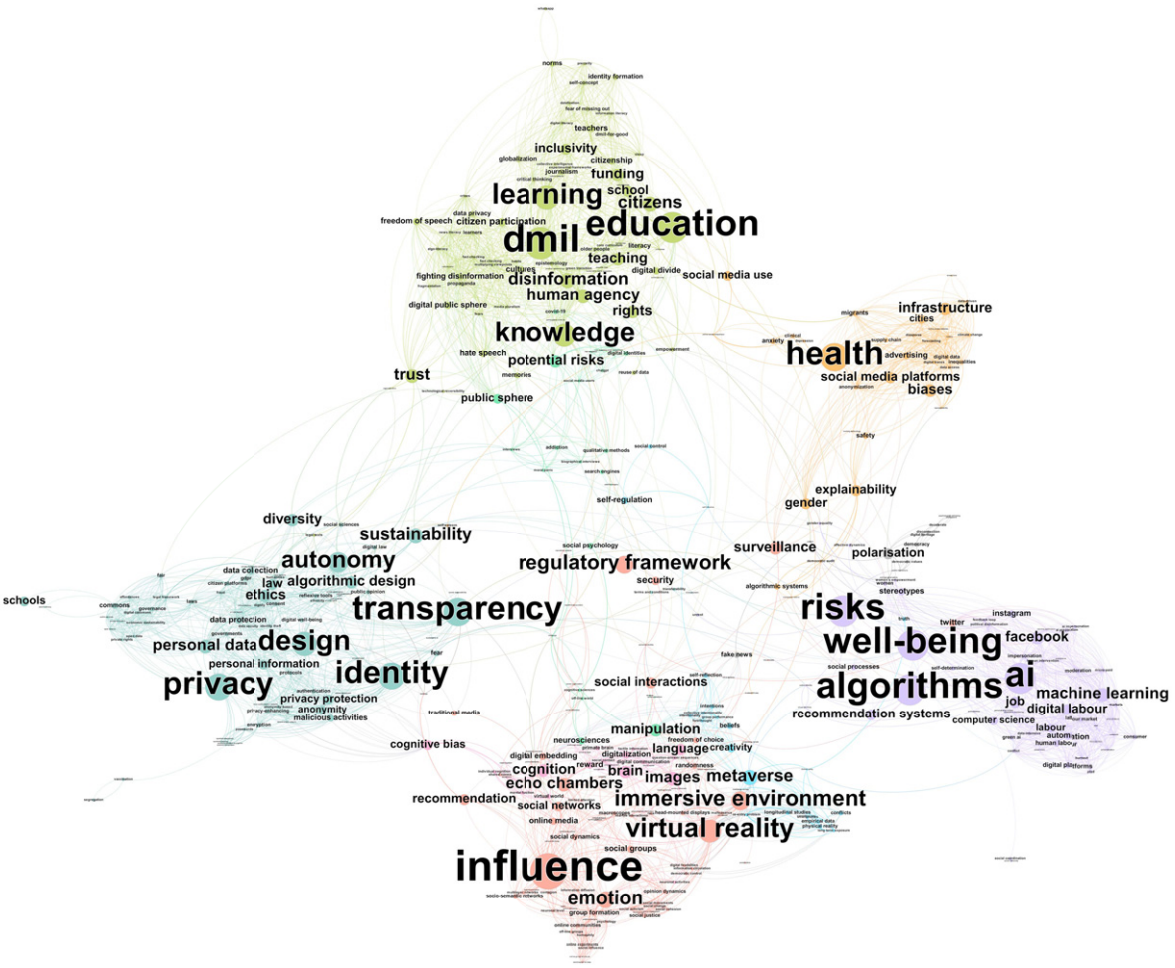


Figure C.12: Map of the challenges proposed by roadmap contributors and external contributors. An interactive version of this map is available online.

Roadmap

1 The interwoven on-line/off-line world 29

2 Digital embedding, mental health, cognition & identity 36

3 Challenges to personal freedom and agency 42

4 Digital literacy, social and individual well-being 47

5 Privacy based on pseudonymity vs. anonymity 56

6 Immersive social media and the metaverse 62

7 Interdisciplinary design for an inclusive and trustworthy digital public sphere 70

8 Impact of the digitalization of society, digital labor and well-being 76

9 Digital commons 84

10 Digital media in a sustainable world: towards inclusivity, equality and autonomy 90

11 Indentification of potential risks related to the Future of digital media 95

12 Conclusions and recommendations 101

Bibliography 106

Index 113

1. The interwoven on-line/off-line world

The multi-scale structure of social interactions

Editors: David Chavalarias
Authors (alphabetical order): David, Chavalarias, Beatrice de Gelder, Laura Hernandez, Emmanuel Lazega, Víctor Rodríguez-Doncel, Camille Roth, Aureli Soria-Frisch, Luca Tummolini
Keywords: social processes, group formation, socio-semantic networks, digital embedding, recommendation systems.

Rationale

Digital media often act to modulate or amplify existing cognitive biases, thereby influencing cognitive processes (see Ch 2). The embedding of individuals in the digital environment changes how we interact with one another, altering how communities form, grow and dissolve, thereby shifting the nature of political dynamics [70]. This has the potential to undermine basic processes affecting how our societies create and agree on shared norms and knowledge. A key question is: How does the embedding of individuals in these new digital environments change the way we interact with others, and what are the consequences for cognitive and social dynamics in interwoven on-line/off-line worlds?

Topic 1: Group dynamics and empirical observations

Psychology describes a number of biases and regularities in how people form social ties. Sociology tells us that homophily and selection are powerful factors in group formation in all societies, which results in strong regularities across cultures. But on-line groups or communities form under very different principles of selection and homophily, influenced by different cues influencing human decisions on social interactions. Over the past decade, many activities that involved personal contact have migrated to online environments. This ranges from simple information exchange between friends to dating sites, medical consultations, professional interactions and

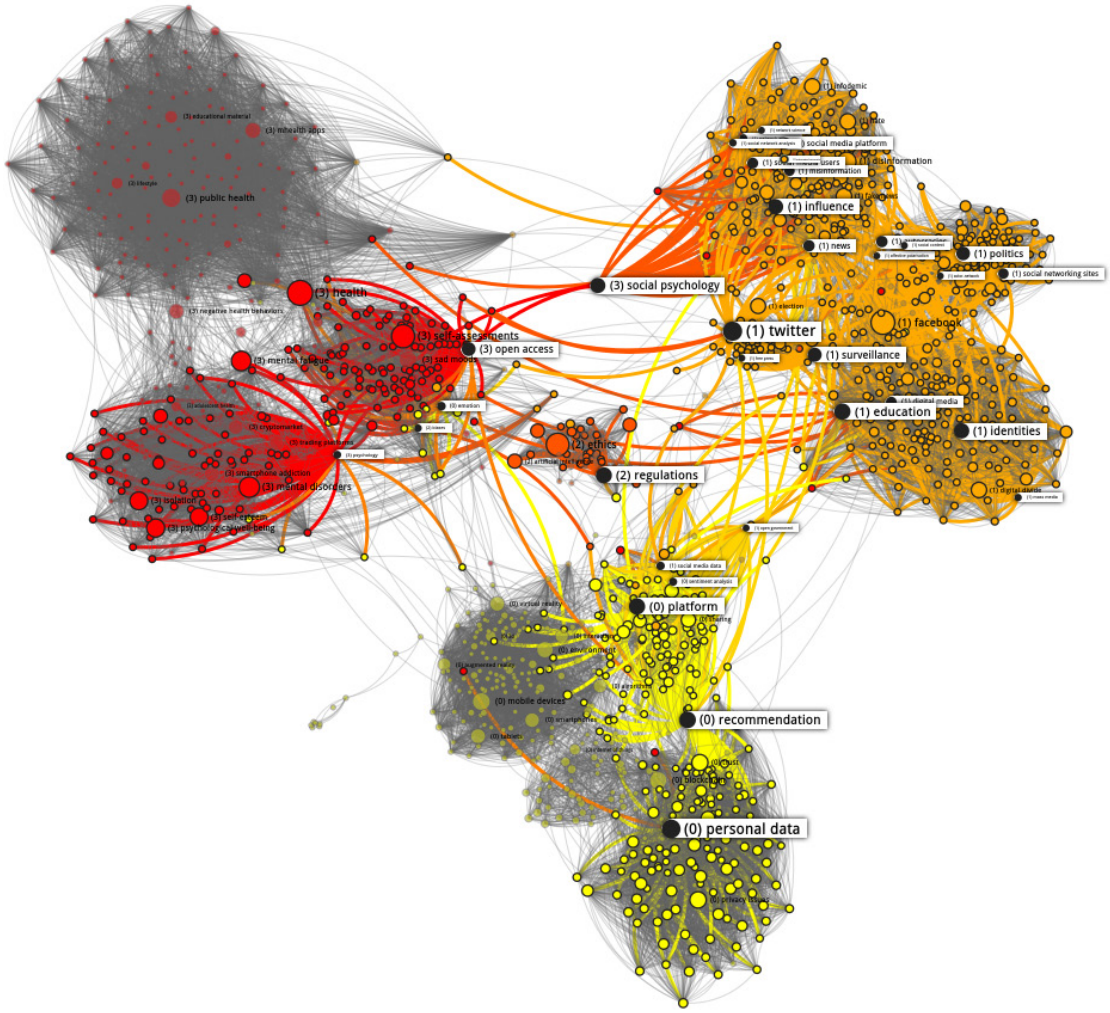


Figure 1.1: State-of-the-art domains concerned by this chapter.

even psychotherapy sessions. At the current rate, it may only take a few more years for physical social contact to become a minority fraction of our social interactions.

There are several consequences to this digital narrowing of the channels of social communication both on individual and the collectives they form. What are for example the impacts of these differences on online group structures? Are there qualitative differences, such as greater polarisation, or greater homogeneity? Do these online differences then spill over into real-life? Do they impact individual well-being?

Addressing the issues described above will require detailed scientific study of the differences between the structure of online- and off-line groups, both in formation and in their ongoing dynamics and the place of individuals in these new collective dynamics. Studies will need to exploit observation in online experiments or reconstructions and sociological surveys. Among other tasks, this will involve studying the dynamics of collective action.

Keywords: homophily, social influence, online communities, collective action.

Key challenges and questions

▪ Challenge 1.1 — Identifying how digital media can enable different mechanisms to support the solution of collective action problems and social dilemmas by shap-

ing social organisation. We need to explore a variety of mechanisms and understand how they drive or influence the development of particular collective outcomes. Example mechanisms are temporal embedding (the creation of repetitive encounters), spatial embededdness (which induces assortment), the sharing of common knowledge (which facilitates coordination) and reputation systems and observability (which encourage formation of shared norms).

▪ **Challenge 1.2 — Learn to measure and understand how social ties form, and the roles of emotion and contagion from a neurological and psychological point of view.** We need to assess the difference between on-line and off-line selection, and undertake sociological studies of online groups. We must also understand the alignment/overlap of online and off-line groups, and the role of media environments (recommendation systems, TV, etc.) in shaping group characteristics.

▪ **Challenge 1.3 — Learn to tailor large scale on-line experiments aiming to explore the influence of different types of platform designs and structures (different rewards or ways of exposing information to the audience, etc.) on overall online social outcomes.** Online social networks and other digital media platforms constitute a unique opportunity to observe social systems "in vivo" in a controlled way to measure the influence of different types of social interactions and affordance on the social fabric. There is a challenge to make this kind of large scale controlled studies easily accessible to researchers.

▪ **Challenge 1.4 — Understand the importance of non-verbal channels.** What passes through the channels of virtual communication is visual and auditory information. What is missing are non-verbal cues from tactile communication, dynamic changes in individual whole-body expressions and the way participants in an interaction position themselves, reflected in the varying personal distance between the participants. Nonverbal behavior contributes substantially to the success of social interactions. These are all aspects that contribute to optimizing interpersonal contact and they get lost in virtual translation of social interactions. We need to understand how this lack of non-verbal context makes misunderstandings arise more easily.

▪ **Challenge 1.5 — Understand how digitization influences the formation of social capital.** Social contact is not only about exchanging information Often the need for information is a pretext for initiating an interaction (a typical example is asking about the weather). Social contacts often seek to gather social capital and to assess, test and confirm social rank and positions. In virtual media, however, the informational content of exchanges gets stripped of this social dimension which is carried mostly by nonverbal cues. As a consequence, the desire for the social reward, naturally expected from person-to-person interactions, is left unsatisfied because it is not balanced at the cognitive level by successful information exchange. We need to understand how this dissociation between the cognitive dimension (what is said) and the affective dimension (the why or the social contact desired) creates a vacuum where participants (symbolically) cry out to be heard at whatever price.

Topic 2: Models and protocols for the study of multi-layer networks for the social sciences

Online systems and social media are multi-layer networks – overlapping and intersecting networks of elements linked through a variety of different kinds of connections and pathways. Studying the structure and dynamics of such networks is a major topic in contemporary science, but much remains to be understood about the link between micro level rules and dynamics and the macro level outcomes.

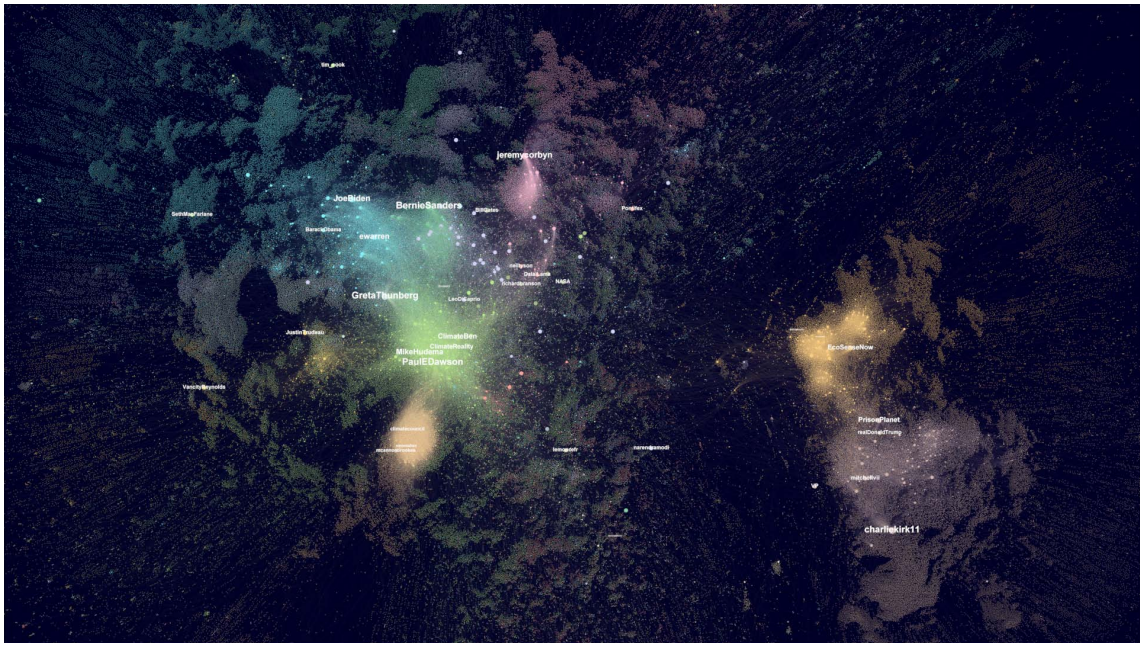
Keywords: multi-layer networks, evolving networks, socio-semantic networks, online social networks modelling

Key challenges and questions

▪ **Challenge 1.6 — Learn how to build and explore useful first-principles models of dynamics in artificial societies.** Such models must be informed by concepts from the cognitive sciences, social psychology and neurosciences and aim to explore the main interactions which lead to the qualitative behaviour observed in the real world. If possible, these should be refined and improved by detailed comparisons with measurements in the off-line world (as in Topic 1).

▪ **Challenge 1.7 — Systematically explore models of the temporal evolution of multilayer networks.** The aim should be to understand how the diffusion of opinion across layers is influenced by the actions of actors at individual nodes. Understand the influence on diffusion of features such as different levels of connectivity (density of connections) within each layer, and in general understand how platform structure or architecture influences network behaviour (see also Ch. 3 and Ch. 7).

▪ **Challenge 1.8 — Understand the main structural influences on information diffusion online such as recommendation algorithms, ease of platforms accessibility, etc.** This will require modelling online social networks and diffusion processes therein. Also, we need to understand how individual factors such as rational thinking *vs.* emotional reactions may contribute to such diffusion, or influence different kinds of content or communities in distinct ways. Understand the interplay between traditional media and online media.



Macroscope example. The structure of worldwide climate change debates on Twitter as depicted by the Climatoscope. *Left:* the pro-climate accounts, *right :* the climate deniers. *Source :* [13].

Topic 3: Macroscopes for the observation of the digital world

Efforts to understand the basic functioning of online social media and their impact on social reality require reflexive tools to effectively observe and measure social dynamics at different scales. Research must be conducted on the impact of the public use of such macroscopes on collective dynamics, decision making and collective action (see also Ch. 7).

Keywords: tangible socio-semantic networks, macroscopes, data streams, opinion dynamics, emotion contagion, mass surveillance.

Key challenges and questions

▪ **Challenge 1.9 — How to make collective dynamics tangible?** We should develop data-based macroscopes able to extract information about the dynamics of diffusion in on-line platform and make them available for use by non-specialists. Currently, this would mainly involve textual, image and video analysis, but may be more general in the future.

▪ **Challenge 1.10 — Characterisation of formation and diffusion of opinions.** How do online interactions among individuals, as well as the influence of off-line media, determine the formation and diffusion of opinions? How do opinions expressed in social platforms act back to influence the off-line treatment of information? How does social media shape human behaviour at multiple layers and scales from small-scale social groups to social movements or entire populations?

▪ **Challenge 1.11 — Characterisation of collective social reflexivity.** What are the effects of making a macroscope available to the public? How is it likely to change the fixed points of collective dynamics? How can we learn to anticipate such effects and which ideas or concepts about social coordination might be most helpful in doing so?

▪ **Challenge 1.12 — Macroscopes and surveillance society.** Where is the frontier between public macroscopes and mass surveillance tools? Is it possible to build understanding

of social dynamics at a fine level of detail while also preserving privacy?

▪ **Challenge 1.13 — Impact of virtuality on collective action.** Does the placement of individuals within social networks impact on the collectives they form? In particular, do emotional responses differ when interaction is mediated through technology? What are the differences between social interactions in real and virtual worlds? How can we objectively characterize these different types of interaction? How do important differences influence individuals' capabilities in building social structures or making group decisions? At which point of technological evolution (if any) might the difference between real and virtual disappear?

Topic 4: Understand the impact of digital media on the enforcement of social and moral norms both online and offline

Digital culture may have important effects in both the short- and long-term on fundamental aspects of social reality, in particular, as technology fundamentally changes the human social environment [70]. We need to systematically study the effect of virtual reality environments on social engagement.

Keywords: social cohesion, social activism, social change, freedom of speech, social justice, emotions.

Key challenges and questions

▪ **Challenge 1.14 — Social cohesion and digital media culture.** We need to explore the impacts of digital environments on social activism and social movements which enact social change. What are the long-term effects of digital media culture on social cohesion? How does social networks and virtual reality moderate the effects of norm enforcement or view on freedom of speech and social justice? From a methodological perspective, we also need large-scale experimental techniques to establish causality or interdependences or to identify the factors affecting social norms.

▪ **Challenge 1.15 — Collective intelligence and group decision making in on-line and off-line settings** We need to understand collective intelligence, group decision making and group behaviour. This will require developing techniques to compare phenomena in real and virtual domains, and to identify how groups of human beings differ in each of these environments at the neuronal level. We should develop methods to monitor multiple brains of individuals while interacting in real and virtual environments. This will help identify the most important attention mechanisms linked to emotional responses which may play a role in supporting or undermining mental health. To achieve these aims, we need synchronization technologies to allow effective study of multiple individuals in distributed networks.

▪ **Challenge 1.16 — Emotion and group formation** Group formation is strongly influenced by emotional responses mediated through channels such as attention or belief. What is the role of emotional response in online group formation and are there differences with group formation in a real environment? Are there distinct neuronal markers when online groups form?

Topic 5: Finding technical tools to understand and control the agenda-setting problem in social networks

Broadcasters have a strong influence over the topics of greatest interest to the public. Currently, both traditional media and social network platforms carry influence, but their combined influence appears to be narrowing interests and discussions into a small set of topics through positive

feedback processes. It is important to examine feasible tools which might help redistribute the public's limited attention so that it is not only concentrated on the issues promoted by the largest media, while also avoiding the creation of echo chambers. That is, can we find ways to favour greater connectivity while preventing concentration on dominant or mainstream broadcasting control?

Keywords: agenda setting, causality in online social media, public attention, echo chambers, algorithm transparency, accountability.

Key challenges and questions

▪ **Challenge 1.17 — Economic model for a healthy digital society.** How do we define a healthy society in terms of information circulation? How might this be achieved in economic terms? This outcome naturally runs against the dominant business models of media platforms based on advertising and marketing of user data, both traditional and online. There is a challenge to reconcile the business model of digital companies with the short and medium term interests of their users in a way that is not at the expense of the long term interests of users and society.

▪ **Challenge 1.18 — Algorithm transparency and accountability.** At a time when the modification of few lines of code can affect access to information for billions of people, recommendation algorithms of large digital platforms must be the focus of attention. What kind of regulation mechanisms should be introduced to encourage redistribution of attention when required? How can this be implemented without sliding into some kind of censorship? (see also Ch. 7).

Topic 6: Information processing in the brain, cognition or affect

Digitalization of interactions favors information exchange at the risk of ignoring the context of social embedding which helps give it meaning. It is often taken for granted that in the course of processing information the 'objective', informational content is easily and unambiguously separated from the affective, social significance. However, our current understanding of the primate brain does not support this view. The primate brain does not process the cognitive and the affective aspects of situations separately, nor does such a separation make evolutionary sense [73, 46]. Cortical and subcortical processes impact the brains' information activity starting with early visual processing stages and progressing to frontoparietal areas [56]. This means that normal cognitive assessment of the truth takes place in interaction with brain structures involved in affect and reward [63]. Consequently, how information is gathered, processed and used is always a function of its specific meaning for the participant.

Keywords: information exchange, social embedding, affect, reward, emotions. Key challenges and questions.

Key challenges and questions

▪ **Challenge 1.19 — Understanding how informational and affective content are inextricably interwoven.** Individuals process information based on a continuous assessment of its relevance, salience, and novelty. These are essential dimensions of information, and they are assessed in continuous interaction with others because they are defined in relation to the social context. Consequently, group membership influences expectations about what is novel and what is most salient. These are also the dimensions of information that bring into play brain areas known for their role in social processes such as prefrontal cortex, and temporo-parietal cortices, anterior and posterior cingulate cortex as well subcortical areas.

2. Digital embedding, mental health, cognition & identity

Editors: Luisa Fassi, Amy Orben and Jerôme Sackur
Corresponding author: Luisa Fassi
Authors: David Chavalarias, Michela Brunori, Caroline Datchary, Beatrice de Gelder, Luisa, Fassi, Maria Gabriella Pediconi, Laura Hernandez, Quentin Lobbé, Víctor Rodríguez-Doncel, Savino Romani, Aureli Soria-Frisch, Luca Tummolin.

Rationale

In recent years, parents, educators, policymakers and the wider public have raised concerns about the impact of social media use on adolescents' mental health, cognitive development and identity formation. In response to these rising concerns, the scientific community has devoted extensive time and resources to investigating the potential risks arising from the use of digital technologies. Yet, despite such efforts, the picture that emerges from the present literature remains inconclusive. Specifically, while converging evidence has established that the association between time spent on social media and well-being is small and not alarming [2, 45], more divergent results come from studies using indicators of social media use other than time spent or indicators of mental health other than aggregate measures of wellbeing [23, 43, 62].

For example, research has found [62] both weak and strong, positive or negative links between common indicators of social media use and diverse outcomes including symptoms of anxiety or depression. Because of these contradictory findings, the field has been unable to provide informative answers to the public or to guide evidence-based policies to help safeguard adolescents. Consequently, researchers have sought to identify some of the reasons which might underlie the existing inconsistencies in research findings. We urgently need further research in this area to provide a deeper understanding.

Topic 1: Differential impact and mechanisms linking digital media use and mental health

Corresponding authors: Luisa Fassi (challenge 1-3), Maria Gabriella Pediconi (challenge 4)

We draw particular attention to four gaps in the literature which need to be addressed. We

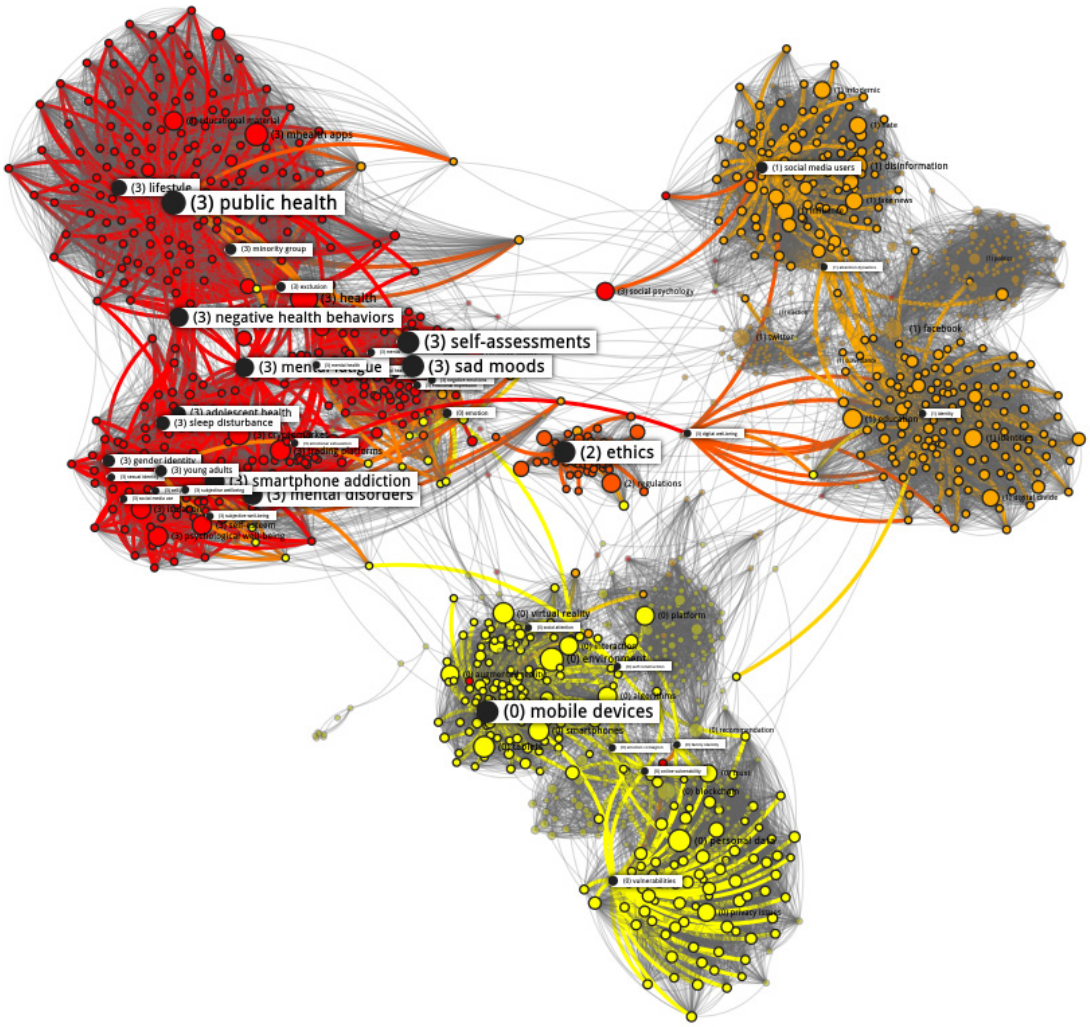


Figure 2.1: State-of-the-art domains concerned by this chapter.

currently lack adequate research on: 1) clinical and vulnerable groups, 2) mechanisms, 3) identity formation and adolescent development and 4) approaches able to integrate evidence from both qualitative and quantitative studies.

Key challenges and questions

▪ Challenge 2.1 — Quantifying the effects of social media use in clinical populations. Most studies in the social media literature have been limited to analysing data from school- or population-based samples. They have, therefore, consistently overlooked those adolescents on the clinical side of the mental health spectrum. Existing studies need to be complemented by research on people with mental health diagnoses, as the association between social media and mental health might differ from links found in the general population. Hence, it is crucial for researchers to 1) acknowledge variability in mental health states among the general population, 2) split individuals into groups based on such variability (e.g., by differentiating that experience mental health conditions at the clinical vs non-clinical level), and 3) consider the differential effect of social media use separately for these groups. For example, suppose we do not distinguish individuals who suffer from different degrees and forms of mental distress and study them as part

of the same sample. In that case, it is no surprise that mixed and contradictory findings might emerge, with some people experiencing positive effects and others negative ones. Addressing this challenge requires large-scale data collection of observational and longitudinal cohort studies, including clinical groups (e.g., groups diagnosed with depression, anxiety, eating disorders, attention deficit hyperactivity disorders etc.)

▪ **Challenge 2.2 — Exploring the differential mechanisms of vulnerability to social media use.** A second key question is: what are the mechanisms that underlie the relationship between social media use and mental health? We cannot assume that the mechanisms linking social media use and adolescents' mental health issues are homogenous across different clinical and non-clinical populations. Rather, we need to understand how social media is differentially related to a wide variety of mental health outcomes across groups with various degrees of vulnerability. For example, a person with an eating disorder might use and feel impacted by social media differently than a person with no mental distress, with social anxiety or attention deficit hyperactivity disorder. To address this challenge, experimental and large-scale observational studies need to test and compare the role of different mechanisms in explaining the link between social media use and mental health across groups. Examples of potential mechanisms might be lack of sleep, time displaced from offline activities (e.g., physical exercise) to online ones, or lack of self-esteem resulting from online social comparison.

▪ **Challenge 2.3 — Digital technologies in relation to self-concept and identity development.** Adolescence is a crucial transition phase from childhood to adulthood that involves identity formation and self-concept development. Social media might play a particularly salient role during this transition, as they offer continuous opportunities to curate one's image and receive feedback from others. Therefore, a third key question is: How does the use of social media impact the development of self-concept in adolescents? How does having multiple vs a single social media profile impact identity formation? Does the availability of quantifiable feedback on social media platforms impact the development of the concept of self? How does such development relate to mental health?

▪ **Challenge 2.4 — Affect and emotions in relation to social media use.** When examining the relationship between adolescents' mental health and social media, the scientific literature primarily focuses on the most tangible aspects of users' practices, such as time spent engaging with certain online content. This leaves little space to explore the affective dimension which inevitably accompanies digital experiences. For instance, exposure to content such as images of food might result in opposing emotions for different users, being neutral or positive for some while triggering or negative for others. An integrated analysis that accounts for the emotions and affective dynamics experienced by adults and teens while using social media would allow a deeper understanding of how technology use relates to mental health. Hence, a fourth key challenge is understanding the role and affective function that digital devices play in teens' daily lives. How does the use of social media platforms relate to individuals' emotional states? What reasons lead teens to prefer online communication and relationships rather than face-to-face ones? What are the factors that lead users to invest affectively in a specific social media platform such as Instagram, Facebook, Tik Tok, etc? Are there peculiar emotions and/or psychological dynamics that lead teens to prefer passive use of social media, such as viewing the posts of others, rather than active uses such as sharing stories, images and thoughts? Lastly, future research should explore the meanings teens give to each social media platform and the origins of problematic behaviours, such as fear of missing out (FoMO) and nomophobia, which are increasingly common among adolescent users.

▪ **Challenge 2.5 — The topic of "digital well-being" considers the impact of digital technologies on human beings and their ability to lead good and meaningful lives.** Discussions in this area involve issues in several domains including healthcare, education, governance, and media and entertainment. There are many open areas needing further research. Can human–computer interactions be tailored or personalised to improve well-being? How can digital media be designed to support autonomy and self-determination? (see also Ch. 3) What is the impact of digital media on cognitive development?

Topic 2: Methodological challenges

Addressing the challenges outlined above will also require concerted research on some methodological issues, that involve both the collection and analysis of social media data.

Key challenges and questions

▪ **Challenge 2.6 — Quantitative research methods.** Quantitative analysis is used to assess the relationship between social media use and mental health in a way that generalises from the examined sample to the target population. Studies can apply quantitative methods to experimental or observational data. Experimental data involves directly manipulating the independent variable of interest – for example, changing the amount of time people spend online or manipulating exposure to specific content on social media.

This can be done as part of an experiment, for example, by simulating a social media platform and studying the effect of manipulating one of its features in relation to participants' mental health. However, due to ethical concerns and logistic constraints, most social media literature analyses observational rather than experimental data. In observational studies, the independent variable is not controlled by the researcher. This type of research design makes it hard to establish causality, particularly when the data is cross-sectional (collected at a single time point). In order to examine the bidirectional effects of social media use on mental health and vice versa, more research is collecting experimental and longitudinal data (collected from the same person at multiple time points) is needed.

▪ **Challenge 2.7 — Qualitative research methods.** Qualitative analysis is often used alongside quantitative analysis to formulate hypotheses or add nuances to the results. Hence, integrating qualitative and quantitative approaches would allow a detailed analysis of adolescents' subjective experiences as social media users. In this regard, we encourage the use of both declarative and non-declarative qualitative methods. Declarative methods include self-reports, focus groups and interviews. Participants are asked to verbally narrate their experiences and thoughts when using these methods. In contrast, non-declarative methods include word associations, pictures or drawings.

These methods allow researchers to go beyond the conscious/declarative territory and explore the affective aspects of digital contexts. Notably, nonverbal stimuli or projective techniques are particularly useful as they bypass adolescents' typical resistance to explicit communication. Overall, research that uses these qualitative methods would allow exploring and subsequently quantifying the most intimate aspects of subjects' experiences as digital users (e.g., tapping into implicit motivations, feelings, fears and desires) in a non-intrusive way.

▪ **Challenge 2.8 — Data access.** Directly accessing social media data from companies is particularly challenging for researchers, as data is often not made accessible. Further, even for users, downloading personal data is often burdensome. These barriers restrict the possibility of researchers to access and analyse data either from companies or directly from users. As a result,

most research relies on self-reports (subjective data), whereby individuals are asked to report the amount and type of social media use they engage in. This type of data might be unreliable, and should be complemented with log-based data (objective data). For this reason, it is crucial that companies are prompted to 1) increase accessibility to the data arising from the use of their products, and 2) become more transparent about the impact of their products, which prevents users from making informed decisions and researchers from identifying problems and solutions.

Topic 3: Impact of digital media on cognitive development

Corresponding author: Jerome Sackur

> The primate brain that neuroscientists now study was shaped by its evolutionary history. In ways we don't understand yet, its properties reflect that environmental and social history and the species-unique ways of dealing with it. We can so far only speculate to what extent new technologies create environmental and social conditions that represent more or less radical departures from the previous ones. Digitalization of communication, dominance of visual media and virtualization now create major challenges. A recent systematic review [64] examining the link between technology and executive functions among healthy young adults highlighted the lack of robust research in this area. Overall, preliminary results show that smartphone use is negatively associated with inhibition and decision-making. Further, working memory performance improves by increased time engaging in video games and by refraining from smartphone use prior to bedtime.

Key challenges and questions

▪ **Challenge 2.9 — Understanding the core characteristics of individual cognition, how to model them and how they are modulated by digital environments.** Is cognition different in real vs virtual environments? Digital media modulate or amplify the consequences of existing cognitive functions, such as attention, memory, intelligence and cognitive biases. Understanding how will require devoted psychology/neurophysiology studies [2]wilmer2017smartphones. In virtual worlds, feedbacks between perceptions and actions are limited, which will make cognitive processes different in the virtual world. What are the consequences for mental function?

▪ **Challenge 2.10 — A pervasive feature of digital media is an intense battle for attention.** Humans have limits to their attention and face trade-offs in attention economy. Are individuals capable of using individual reward systems to cope with the scarcity of attention? What are the links between the use of digital technologies and memory, emotional regulation and impulse control? Is the link between social media use and mental health mediated by changes in cognitive processes? How do we measure changes in cognition (e.g., memory of attention) as a results of technology use?

▪ **Challenge 2.11 — What are the cognitive effects of the shift to images over words? Digitalization is bringing about the dominance of image over text.** Traditionally, information was conveyed with the written word, especially in communications of serious matters. Unlike images, language and verbal communication unfold over time. Interpretation is built over time, even during short question-answer sequences and consensus builds up during the verbal exchange. With digital communication we now witness an extreme condensation of the verbal messages with associated changes in verbal style including tendencies towards slogans. On

the other hand, daily communication is increasingly based on visual images in ways we could not have imagined one or two decades ago. Consequently, the gap between verbal communication and more specifically the written word (short text on social media) and communication with images is getting bigger and bigger. What are the cognitive effects of this change?

3. Challenges to personal freedom and agency

Editors & corresponding author: Andrzej Nowak
Authors (by alphabetical order): Pawel Horodecki, Paul Lukowicz, Andrzej Nowak, Erika Stael von Holstein.

The ability to make independent decisions and turn them into actions is among the central elements of being human. Human freedom, autonomy, and agency are recognized in social psychology as both fundamental values and empirically demonstrated as conditions required for well-being and forming stable, satisfying relations (e.g., [50]). "Freedom from" refers to the absence of external pressures. "Freedom to" refers to the capacity of individuals to achieve their objectives [77]. *Human agency* refers to an individual's capacity for reflective and purposeful action [4]. *Autonomy* means self-governance, self-determination, and intrinsic motivation [16].

Human autonomy and agency have become important criteria for responsible AI (e.g., [24, 47]. Autonomy has been recognized as a fundamental human value to respect and protect by the Organisation for Economic Co-operation and Development (OECD). Human autonomy has been named as the top priority for responsible AI by the European Commission's White Paper on Artificial Intelligence.

Randomness in the digital environment of an individual is an important factor affecting human agency and autonomy. There likely is an optimal level of randomness for human agency and autonomy. Too much randomness makes the environment chaotic, the integration of information difficult, and limits the potential for informed choices and effective actions. As fortuitous events could amplify the potential for the human agency [4], too little randomness may limit human agency.

The information collected about the users and used for personalization may limit the user's autonomy and agency, allowing for more precise methods of influence that are custom designed for the specific user and that minimize the randomness in the digital environment of a person. These opportunities and challenges for human agency are likely to grow with the rapid emergence of AI and virtual reality, especially the Metaverse.

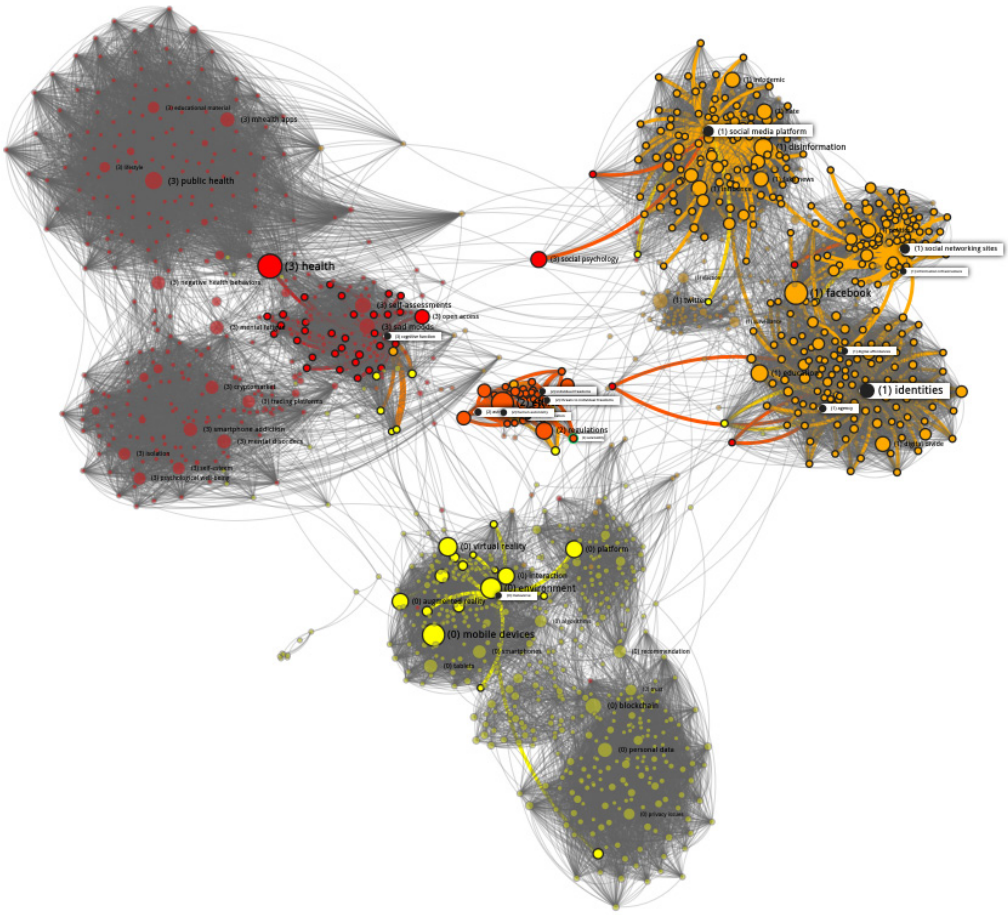


Figure 3.1: State-of-the-art domains concerned by this chapter.

Topic 1. Human agency and autonomy vs. heteronomy

Human Agency

The four aspects of human agency [4], which refers to an individual's capacity for reflective and purposeful action, are:

- Intentionality refers to the fact that human action is based on intentions that include action plans and strategies for realizing them.
- Collective intentionality is a guide for effective group performance.
- Forethought – the ability to foresee the consequences of action and to base the current action on the anticipated consequences. It is the basis of purposeful action. According to Bandura [4] it provides direction, coherence, and meaning to one's life.
- Self-reactiveness provides the basis for self-regulation. Humans are able to self-regulate to plan actions, motivate and execute them.
- Self-reflectiveness is the potential to reflect on oneself, one's thoughts and actions and to use this reflection to modify oneself and correct the action,

People who are characterized by a higher degree of agency, in contrast to people with a lower degree of agency, have a higher level of competencies, self-regulatory skills, and beliefs in their efficacy. They can perceive more options, which expands their choices, and they are more successful in achieving their goals [4]. Agency is a value, and at the collective level, people

sanction unauthorized forms of social control.

Human Autonomy

The concept of human autonomy refers to self-regulation, choice, self-determination, and will [16, 49, 50]. The autonomy-oriented regulation is driven by reflecting on possibilities and choices. Heteronomy, the opposite of autonomy, is defined as controlled regulation or regulation that occurs without self-endorsement [50]. The control-oriented regulation is driven by overt or salient rewards and punishments, usually in the form of social contingencies.

Numerous studies (*cf.* [50]), have revealed that autonomy, in contrast to external control: enhances performance and creativity and improves the quality of relationships. It facilitates attachment and intimacy as well as enhances well-being. In contrast, increased external control decreases well-being. External rewards lead to reporting less autonomy, happiness, and quality of relationships, as well as reduced internalization. As Kaiser (2002) notes, "when rewards are contingently offered, people can easily lose sight of important values, needs, and social concerns".

The availability of meaningful choices and informational feedback facilitate intrinsic motivation. In contrast, limiting choices and external control decrease agency and autonomy. Empirical studies have demonstrated that coercion, rewards and punishments, controlling conditions, social controls, minimization of choice, surveillance, and frequent evaluations can significantly reduce agency and autonomy. Human behavior may become subject to external control, often without awareness of the person being controlled.

The influence of digital media on agency and autonomy

Digital media are a double-edged sword for human agency and autonomy. On the one hand, they strongly enhance the human potential for autonomy and agency due to the unprecedented expansion of possibilities resulting in in almost unlimited access to information (within the constraints of human physical or biological limitations), contacts on social media, totally new tools for realizing one's projects (*e.g.*, crowdfunding) and enhancing bottom-up social processes.

The richness of the informational environment paves the way for informed decisions. The ease of finding new contacts and opportunities broadens the scope of choices which facilitates agency and autonomy. Digital media offer many tools for making choices, planning actions, financing them (e.g., crowdfunding), and carrying them out. The ease of getting diverse feedback on one's thoughts and actions may enhance self-reflexivity and agency. Sharing information with others, group discussions, and possibilities to coordinate actions with others pave the way for group actions and collective intentionality. There might be a potential for the development of AI-based tools that help guard human autonomy and enhance human agency. In general, digital media, in many ways, enrich the informational environment, offer new options for action, and nurture bottom-up processes, thus enhancing human agency and autonomy.

On the other hand, digital media represent a threat to human autonomy and agency by disturbing the quality and diversity of information (*e.g.*, fake news), manipulation, and strong influence on human choices.

Digital media and their AI-based algorithms can result in the loss of agency and autonomy as well as a decline in cognitive capacities [42]. Such effects arise in a variety of ways, including the reduction of the diversity of information (*e.g.*, filtering, echo chambers), the presence of overt and covered manipulation (*e.g.*, micro-targeting and misinformation), shaping human behaviors by reinforcement, limits on the scope of choices, loss of cognitive competencies resulting from delegating the decisions to AI (*e.g.*, recommender systems) and the reduction of randomness in the informational environment. Information technology and AI have become powerful tools of manipulation; search engines and social media algorithms can deprive individuals of information

needed for informed decisions and exposure to diverse views. This can result in increased external control and a reduction of agency and autonomy.

Information collected about users and exploited for personalization may limit users' autonomy and agency by allowing for more precise and more effective custom-designed methods of influence. Personalization limits access to information that diverges from already existing personal knowledge or that could question the default judgments and decisions. (see also Ch. 1 and Ch. 7)

Key challenges and questions

▪ **Challenge 3.1 — How to increase freedom of choice, human agency, and cognitive abilities thanks to digital media and IA.** How can we learn to capitalize on the potential produced by digital media and AI for increased freedom of choice, human agency, and cognitive abilities, and yet also avoid the dangers to the same values resulting from these technologies? Answering these questions will require advances in the basic understanding of how digital media influence human agency and autonomy. What are key digital media features which can support or encourage human autonomy and agency, and how do such features interact with one another? Can we develop a set of recommendations and requirements aimed at protecting human agency and autonomy, which might be a basis for developing the comprehensive European approach to digital media?

▪ **Challenge 3.2 — Can we construct AI-based tools which will guard human autonomy and facilitate human agency?** We need to understand the likely effects of increasing the organization of information and decreasing randomness on human agency. In so doing, we should also develop practical tools to help protect individuals from manipulation and automated reinforcement systems. Can we find tools to help individuals detect hidden influences and shield themselves from them? Are there tools for improving collective human agency, and how might these be provided? In practice, how can we best preserve the diversity of the human ecosystem of information? In particular, can we identify a general optimal relation between digital and analog information carriers in human life? Can we determine whether, when, and how it depends on a particular information character?

Topic 2. Randomness as a factor in human agency

Increasing the organization of information increases human understanding of the environment in which an individual acts, makes decisions easier, and makes actions more efficient. However, some degree of randomness in an environment is an important condition for human autonomy and the self-organization of mental and social processes. The elimination of randomness in information presents a challenge for human agency.

In quantum physics, certain events happen in a fully random way, with no dependence on any prior event in the universe (cf. [6]). Such randomness reflects more than a lack of knowledge on the part of an observer. Such deeply random events imply, at the quantum level, that events happen that are not caused by fully deterministic causal rules.

A few researchers (*e.g.* [3, 18, 30]) suggest that similar quantum-like randomness may also characterize some brain processes, with the human mind being another source of deep randomness in addition to quantum processes, although the mechanisms would likely be different. Randomness might be the precondition for creativity and a potential source of causation. If so, this might also imply that, in general, factors that reduce the role of randomness might limit human agency. On a phenomenological level, these hypotheses harmonize with an analysis of scientific creativity as a constrained stochastic process [52].

The role of fortuity in the human agency was discussed by Bandura [4]. He pointed out that fortuitous, seemingly insignificant events can alter the course of life. Examples include decisive events such as meeting a marriage partner or finding a job. Such events can have disproportionally large effects and make prediction difficult by altering simple, predictable linear life trajectories. Bandura noted that fortuitous events could amplify the potential for the human agency because individuals have the capacity to decide if they want to capitalize on a chance event or instead let it go unnoticed. Likely, there is some optimal level or a small range of randomness that is maximally beneficial to human agency and autonomy, although their specific values may also depend on other individual factors, the role of which should be studied in the future. Too much randomness in the environment may overwhelm an individual with disorganized information. In contrast, too little randomness may restrict agency, with choices being strongly dictated by the environment. Humans may differ in their capacity for generating randomness, and human capacity also depends on the situational context [7].

Information in digital media that is too highly organized and personalized may reduce human agency. Some limits on the potential for autonomy may come from features such as algorithmic filtering, echo chambers, systems fitting the content to users' interests, recommendation systems, and search engine algorithms, all aimed at maximizing the users' engagement. The use of reinforcement algorithms, for example, which learns how to shape users' behavior, may directly limit human autonomy. Virtual reality, where the whole environment is designed and controlled, may present an especially important threat to human agency and autonomy. There are already search algorithms (*e.g.* [58, 66]) that involve randomness. Such algorithms can locate information that would not be found using a purely deterministic search, thereby broadening the scope of search results. Adding random results to search algorithms offers an option for the users to transgress their information niche. Adding some random contacts may help users to go beyond their echo chambers. Encountering information that does not fit the structure of someone's beliefs may lead to questioning of already held beliefs and opinions and may result in more autonomous decisions. Further studying the properties of the informational environment modified in the above ways might also help in finding other important factors related to agency and autonomy.

Key challenges and questions

▪ **Challenge 3.3 — What is the optimal value for randomness in digital environment?** As digital media become more organized and the information individuals get becomes increasingly personalized, randomness in the digital environment may diminish below the optimum value. Can we understand where the crucial threshold is? What are the most evident signs of a loss of human autonomy? Is there any sign of such loss in either the human brain directly or in human decisions when performing particular tasks? Which digital technologies pose the greatest threat to the erosion of autonomy? Which technologies and algorithms enhance autonomy and agency?

▪ **Challenge 3.4 — If the introduction of some randomness in digital media may aid human autonomy, can we identify the best mechanisms for doing this?** Are there ways to monitor digital media to maintain awareness of changes in the environment that may affect human autonomy? How, on a meta-level, can one monitor research on human freedom, agency, and autonomy to ensure that its methodology is sufficiently independent of the specific logic of digital-media-type data processing, communication, and goal formulations? In addition to randomness, are there other important signatures of erosion/increase of agency and autonomy?

4. Digital literacy, social and individual well-being

Editor and corresponding author: Divina Frau-Meigs
Authors (alphabetical order): Maria Jose Brites, Lusófona University (Portugal); Divina Frau-Meigs, Sorbonne Nouvelle University (France); Sirkku Kotilainen, Tampere University (Finland); Julian McDougall, Bournemouth University (UK); Geoffroy Patriarche, Saint-Louis University (Belgium)

Keywords: digital competences, digital uses, digital literacy, media and information literacy, critical thinking, agency, engagement, well-being, learning, teaching, communities of practice, lifelong learning, disinformation, curriculum, schools

Rationale

Our human activities are increasingly embedded in digital environments which influence all aspects of modern life, including education, politics, work and leisure. A growing research literature emphasises the need for greater awareness of the impacts technology is having on an individuals' mental health, cognitive capacities, socialization and identity formation. Vulnerable users and especially young people merit extra attention in learning to live safely in a digital world, managing screen time as well as their representations of self and others, and also facing other attendant risks to well-being.

Such issues demand more attention as digital technologies and media frame our relation to space and time and our access to culture, education and jobs. In particular, the Covid-19 crisis put into sharp relief the dependency and lack of preparedness of teachers, parents and other caregivers when it comes to conducting education with such media and data-driven technologies. It is urgent to build a deeper understanding of the relationship between formal and informal ways of learning and new models of education and schooling in the 21st century.

Coordinated jointly by DG-EAC and DG-Connect, a 2022 taskforce on "fighting disinformation and promoting digital literacy" pointed to the need to integrate information, data and algorithms as they drive social media and other upcoming immersive media. Analysis finds that a focus on Digital Media and Information Literacy (DMIL) is absent in curricular activities and in the formal education sector. We lack innovative pedagogical and research postures to create a literate citizenry ready to face global uncertainties and crises.

However, there is likely no simple technological solution, given that disinformation is a principal feature of unhealthy digital media ecosystems. Here DMIL is not neutral. It can be seen as contributing to disruptive digital media practices, or, alternatively, as helping to harness new media uses as key infrastructure for work, leisure and citizenship. Some may see DMIL as an ambitious support for citizens' agency; others may see it as a recurring threat to the conformity to social norms. In the debate on education and technology, DMIL tends to promote critical thinking and creativity, which should lead to more equity and inclusion. However, DMIL varies greatly across countries and within countries, as it is dependent on the local digital infrastructure and political structures. It also depends on the global geopolitics of the Internet, which can create unhealthy situations such as disinformation by interfering foreign agents.

Such disparities may lead to new forms of social mobility as privileged connected citizens can embrace the new educational and civic opportunities afforded by DMIL. They can also lead to new risks as citizens become audiences targeted for commercial, persuasive, manipulative, predatory, abusive and exploitative digital interactions. DMIL can thus be used successfully by users with sufficient access to technology and networks but it can paradoxically be used by actors with malign intent, against users' self-interests.

The main challenge for DMIL is learning to exploit the positives while steering around the negatives. This is especially true as new predictive uses of ubiquitous individual and social data emerge, and cross-border flows of information influence our social actions. This situation calls for a new restructuring of the notion of DMIL to make it responsive and to foster beneficial uses of DMIL. Among the general public, the Covid-19 pandemic has greatly boosted awareness of our need to develop informational resilience.

Topic 1. Defining the nature of digital media and information literacy (DMIL) and its opportunities

Keywords: learning, literacy, co-construction, transmission, participation, knowledge acquisition, epistemology, agency, engagement, change. Key challenges and questions

Driven by new technologies and practices, many new literacies have appeared since the turn of the millennium: news literacy, visual literacy, information literacy, search literacy, data literacy, algo-literacy, among others. This "literacy turn" reflects the needs of society, as all sectors of society - from work to politics and culture - are pervaded with new sets of skills, both soft and hard. As schools are required to go digital and increase the digital competences of their students as future citizens and workers, Digital media and information literacy (DMIL) is increasingly solicited to deal with new skills, attitudes and values. These are generally different from the instrumental realm of Science, Technology, Engineering and Mathematics (STEM) and seek to encompassing digital citizenship for future citizens.

The literacy turn is not neutral. It stands in tension between the necessity to "adapt" to digital culture, and the opportunity to actively shape the future digital culture of our societies at large. In the various competing paradigms for learning, DMIL tends to move away from top-down knowledge acquisition in school institutions to embrace bottom-up development of

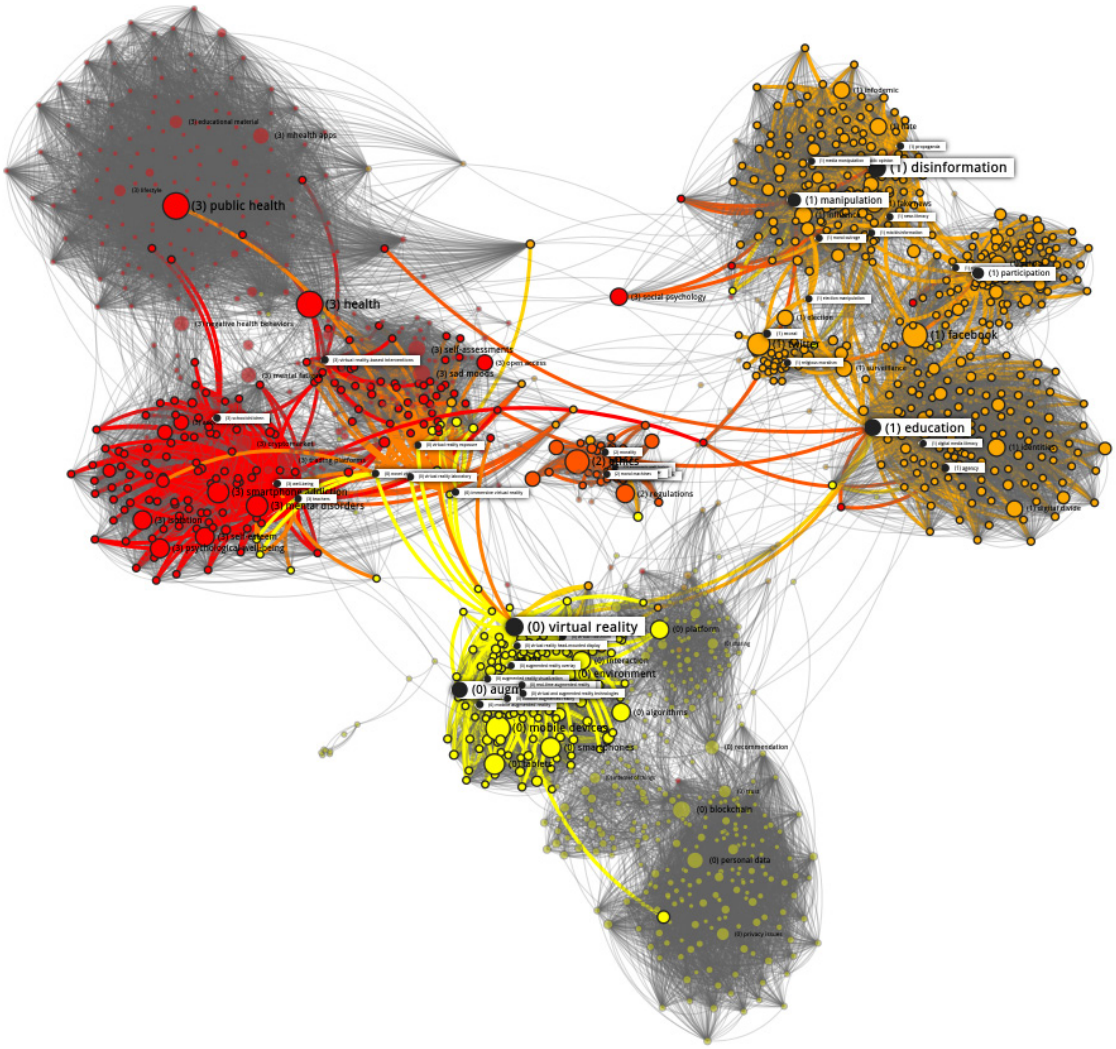


Figure 4.1: State-of-the-art domains concerned by this chapter.

competences via shared tools for mediation. It also emphasizes participation through horizontal and vertical circulation of knowledge within communities. This opens up a huge potential for transformative change, but also opportunities for resistance. Particularly urgent is the need for DMIL to enable the Green transition and to foster Green competences in relation to digital media, especially around the sustainability of information commons and the materiality of the digital footprint.

To generate positive change, DMIL needs to move from mere access to emphasis on understanding and competency. It also needs to address capability through civic engagement or employability and consequences including rights, responsibilities and social justice. The four elements of this theory of change for DMIL are inter-related and combine to form a holistic ability to engage fully with digital media in citizens' connected daily lives. This means engaging with media safely and healthily, critically and actively, with positive social consequences.

Key challenges and questions

▪ **Challenge 4.1 — What are the contours of DMIL?** What are its core concepts and epistemology? How does it differ from STEM? How does it fit in with training in other areas

such as computer science and ICT? As there are many different and competing approaches to DMIL, how is the "debate" over DMIL to be framed? By whom and for what purposes? For which ideological/political orientation? How do the different DMIL discourses compete with each other and try to gain ascendency? To what effect?

▪ **Challenge 4.2 — What is the relationship with other sciences and research on education, youth studies, communication and information sciences, digital humanities and data studies?** How can we produce and fund research on DMIL without diluting it in traditional subject matters?

▪ **Challenge 4.3 — How can we focus on the positive uses of DMIL for improving the health of digital media ecosystems?** What is the link with participatory cultures and sciences? How to deal with creativity and collaboration? What new conditions for civic agency, engagement and empowerment may arise? How can DMIL support the sustainability needed to bring about the Green transition?

▪ **Challenge 4.4 — How can we ensure protection from risks and promote safety while encouraging play and simulation and immersion?** Globalization and trans-border media networks have loosened bonds of nationality and brought citizens to exert multi-level loyalties – to their selves, their local communities, their countries and regions. In this context, how can we reframe citizenship and decision-making at large with the construction of an array of subject-positions and expression of opinions?

Topic 2. Identifying key competences of DMIL and their impact on learning and teaching

Keywords: competences, life-long learning, critical thinking, problem-solving, information, communication, creativity, safety, public service value, commercial internet service providers

The many literacy discourses can create confusion or even dilute the strength of DMIL. But such profusion may also help create sufficient momentum to introduce DMIL as part of the core curriculum. The educational goal remains the same: fighting illiteracy and information illiteracy, as well as the cultural and digital divides which come from a lack of DMIL basics.

The drive for DMIL Education focuses on competences, understood as a combination of knowledge, skills, attitudes and values supporting citizens in their daily lives. There are several frameworks and models in existence, including the European DIGCOMP 2.2. This aims to be comprehensive and has five defined dimensions: information and data, communication, creativity, safety and problem-solving. In parallel, and with an emphasis on lifelong learning, the Council of the European Union adopted a recommendation on eight key competences for lifelong learning. Among the eight, those relevant to DMIL are literacy competence, digital competence, personal, social and learning-to-learn competence, citizenship competence and cultural awareness and expression competence. These are the most relevant proposals intending to prepare children, young people, adults and older people to become better critical thinkers, effective communicators and active and empowered members of their societies. The challenge now is to connect these key competences to the four elements of DMIL for positive change. Access, awareness and capability are integrated already into the DIGICOMP and other frameworks. Consequences is the category in need of extension.

Outside of schools, DMIL has benefited from a proliferation of "sensible" practices and initiatives across Europe for at least two decades. But these have reached little visibility, failed to scale up or struggled for funding, and often lacked any relation to established competence

frameworks. Beyond the challenges that include lifelong learning for DMIL, they tend to focus on youth and less on adults and older people, despite their information illiteracy. These initiatives, and the diverse context of key competences of DMIL they foster, have created debates and controversies between different DMIL projects, discourses and actors. Additionally, on the Internet, multiple DMIL "providers" are emerging, with commercial and proprietary solutions and opaque agendas. Such providers often position themselves outside of educational institutions – as being "independent" or "close to the people" – with potential risks to educational independence and public service value.

Key challenges and questions

▪ **Challenge 4.5 — How are DIGCOMP applied and if not, why is there resistance?** Which developments and trends have taken place? What has worked well and what has not? Which priority areas can be identified for moving forward? How can we translate them and operationalize them in different European linguistic and teaching contexts?

▪ **Challenge 4.6 — How are competences articulated with the need for a basic knowledge of technology?** How do ICTs affect education and learning? How can we make DMIL competences responsive and agile so as to equip learners with the mindset they will need for lifelong digital literacy updates? For instance, what are the mechanisms behind algorithms that affect access to information? What key privacy issues need to be mastered for safe online e-presence and digital identity?

▪ **Challenge 4.7 — How can we bridge formal and non-formal learning environments and perspectives?** How can we use the best of both worlds? How can we bridge the gap that learning and teaching communities face in relation to media education, educators' knowledge and students' engagement with media and democracy? How is lifelong learning – through its eight key competences – being settled in learning communities throughout life courses that are increasingly diverse and fragmented? And how can the key competences be integrated with and extended to a theory of change for beneficial DMIL employment and ecosystem health improvement?

▪ **Challenge 4.8 — What are the DMIL projects of commercial "providers"?** How do they present themselves? Do they have an audience? How do such initiatives differ from, and compete with more institutionalized forms of DMIL as fostered by schools and media education organizations?

▪ **Challenge 4.9 — In DMIL initiatives and programs, how can we reach the most underrepresented and underserved communities such as low income people, migrants, rural area dwellers?**

Topic 3. Fighting disinformation and other information disorders

Keywords: disinformation, risks, fact checking, news literacy, algo-literacy, agency, polarization, trust, radicalization, hate speech

Our media-saturated and rapidly evolving information cultures are increasingly prone to risks and crises. Schools need to foster resilience, so as to enable learners to resist anxieties and fears caused by excessive algorithmic surveillance and systemic doubt and distrust about democratic institutions. Other emerging threats come from social polarization, fragmentation and radicalization, as well as climate change and pandemics such as Covid-19. These all call for

new approaches to nurturing innovation in pedagogy to mitigate information disorders, among which disinformation, hate speech and undue surveillance.

Nations and international entities such as UNESCO, CoE and the EU have identified DMIL as a priority, typically supported by fact-checking expertise. The democratic vision underlying DMIL, resting on freedom of expression, media pluralism, data privacy and safety, makes it a means for fostering civic engagement, well-being and trust in the digital public sphere while mitigating against disinformation. The fight against social and cognitive biases and the promotion of critical thinking can help counteract the potential risks related to the future of digital media arising from propaganda, external interference and mass surveillance.

Key challenges and questions

▪ **Challenge 4.10 — Disinformation involves not only factuality and cognitive deception, but emotions, imagination, prejudices and stereotypes, moral panics, collective memories and social life.** Disinformation also has technological roots (e.g. bots, algorithms) and can be used in multiple contexts (e.g. clickbait, hate speech, propaganda) and have many consequences for well-being. It operates at the individual (e.g. mental and physical heath disorders) and societal (e.g. violent extremism) levels . How should this complexity be addressed in DMIL? And how should the risks of disinformation be balanced with the opportunities for valuable information? How can DMIL extend to positive action to reduce pollution in the digital media ecosystem by contributing trustworthy information and positive representations?

▪ **Challenge 4.11 — What are the best "sensible" practices for fighting disinformation?** On top of fact-checking and source evaluation, other practices such as reporting suspicious content, sharing fact-checks and correcting disinformation should also be promoted, as well as the awareness of algorithm-based logics of spreading disinformation. How should such practices be fostered as part of DMIL initiatives? How can they be transferred? To what age-relevant publics (juniors, older people...) should such efforts be directed, and through which channels or in which settings?

▪ **Challenge 4.12 — Propaganda and disinformation are increasingly (mis)used in public discourses for argumentative and manipulative purposes.** Citizens should be able to critically decipher such strategic discourses. How does DMIL approach this issue? How can DMIL improve understanding of scientific enquiry vs opinion and of deliberative processes? How can we foster a critical approach to information without creating a breeding ground for conspiracy thinking?

▪ **Challenge 4.13 — How can we promote alliances between DMIL practitioners and other disinformation fighters such as journalists, fact-checkers or data scientists?** How can this also be done outside of schools? What are the possibilities in DMIL "third spaces"?

Topic 4. Capacity building and Teacher training

Keywords: assessment, training, in-service, pre-service, resources, metrics, funding, capacity-building, resources, institutions, transfer

Teacher training is currently a weak point of DMIL. With the rise of disinformation, DMIL practitioners are pressed as front-line workers and expected to build informational resilience to risks and citizenship preparedness. This puts constraints on DMIL practices, which typically lack training, resources and materials, sufficient time and institutional support. Teachers often lack

the ability to transfer their private digital skills and practices into their professional environment.

Assessing the impact of DMIL interventions is another challenge for education systems. DMIL task-based pedagogies cannot be assessed just in terms of how students perform on assignments. Efficacy may be reflected in longer-term changes in students' attitudes and digital-skills awareness. Teachers and administrators also lack baseline measures against which to judge improvement.

A lack of reliable, independent proof of DMIL effectiveness has an impact on funding, as funds tend to be evidence-based for attribution. However, capacity building should not stop at teacher training. DMIL is a life-long process and all citizens should be upskilled. This is particularly true of the senior population that is less educated in DMIL, and seems to contribute most to the amplification of disinformation. DMIL training can move from the traditional ways of understanding mass media to taking a critical stance on platform-based digital media. People need to understand basic facts about online tracking, recommendation systems and content filtering, for example, as well as the business models underpinning these practices. Such basic knowledge is essential for understanding the sharing and production of disinformation.

Key challenges and questions

▪ **Challenge 4.14 — How are curricula designed to develop DMIL competences in teacher training?** Which examples of curricula constitute sensible practices? How can we ensure equal access for all and scale up? How to introduce DMIL competences in existing programs? How can we create a continuous, coherent and progressive full curriculum for DMIL in education (from K1 to K12)? What about the insertion of DMIL in the core curriculum as a subject in itself?

▪ **Challenge 4.15 — What are the conditions for effective teacher training, pre-service and in-service at universities on DMIL?** What are the existing programs in universities and other higher education institutions and how do they relate to secondary education? To lifelong learning?

▪ **Challenge 4.16 — Are there valid national and supra-national assessment benchmarks for DMIL?** What are the appropriate metrics and measures? How can they be improved for more effective measurability?

▪ **Challenge 4.17 — What are the new ways of preparing teachers to adjust to future educational opportunities via digital media?** Are MOOCs and online training an option for mitigating lack of teacher preparedness? If so, on which conditions (micro-credits, open badges…) and to what career development goals? What kinds of up-dated curricula contents and learning material exist for teachers to understand the logics of algorithm-based communication?

▪ **Challenge 4.18 — What are the conditions for upskilling segments of the population that are far from education, such as seniors? Can e-learning (MOOCs, distance teaching...) provide effective solutions and reach out to adult populations in a meaningful way?** How can we scale up such interventions?

Topic 5. Transforming the whole school approach: coordinating partnerships and monitoring funding

Keywords: collective intelligence, whole school approach, commercial media, for-profit platforms, immersion, data-fication, metaverse, funding, transparency

Increasingly, holistic, long-term approaches are considered necessary to achieve real impact.

The challenges on disinformation, for example, must involve multiple stakeholders including journalists, librarians, data scientists. This implies promoting cooperation between schools and opening up schools to local and national partners such as ministries, municipalities or civil society organizations. We also need new conceptual and experimental frameworks for piloting new forms of networked civic participation and collective coordination. Finally, we need means to ensure school integrity and independence in relation to private and civic partners and "providers". External partners need to be vetted for their ability to conduct DMIL interventions along strict ethical, inclusive and non-promotional lines.

The role of policy-makers and their relation with educational technology is under-studied. Technology companies and online media providers determine their own sense of responsibility for the content they provide schools and receive little feedback on their actions. Meanwhile, they keep results for themselves, especially when it comes to disinformation and their role in its amplification. We need a better understanding of the public and industrial infrastructures that may support DMIL. This is also about "future-proofing" DMIL, as educational technology moves into the metaverse and technology such as augmented reality multiplies immersive applications. As ubiquitous AI pervades the digital educational spheres, more attention needs to be paid to the way data are used in schools.

Funding is key to the successful future implementation of DMIL. Given the possible consequences for democracy, we need guarantees that DMIL is not out-sourced large internet for-profit platform "providers" and operators. Otherwise, digital education will be lost from the remit of public interest authorities. Keeping clear of partisan and commercial interests and ensuring full funding transparency and diversity of DMIL may require careful monitoring tools from independent actors and researchers. This could also ensure that recipients are shielded from any undue pressure from educational technology enterprises and their increasing need of data for automation of education and scalability of its applications.

Key challenges and questions

▪ **Challenge 4.19 — How to best engage the entire educational institution community in tackling DMIL challenges and making both teachers and students more digitally literate?** To what extent can we involve external partners and actors such as journalists, fact-checkers, AI experts, NGOs and media outlets in education while preserving its public service independence? What are the current best transparency standard-setting solutions? How do they effectively foster donor and recipient transparency?

▪ **Challenge 4.20 — What are the risks of opportunities of connected schools and data?** How does data monitoring impact DMIL? With what consequences for student well-being? How to effectively enforce data privacy for protection of minors while supporting data for good in education and research?

▪ **Challenge 4.21 — What are the tools for effective monitoring of funding in digital education?** How to ensure partnerships while maintaining transparency and independence? What solutions to make the Edtech accountable without undue chilling effects?

▪ **Challenge 4.22 — What is the future of the metaverse and how will it likely impact DMIL? How does virtual / augmented reality affect learning and teaching?** How are and should teachers and learners be involved in the design of such connected tools? What opportunities are there for new kinds of DMIL third spaces?

Community

- ALL DIGITAL (Belgium)
- CEMP: Centre for Excellence in Media Practice (UK)
- CERI (OECD)
- CrAL – Creative Audiovisual Lab (Croatia)
- EDMO (EU)
- ENGAGE (Belgium)
- GREMS (Belgium)
- MEDIA AND INFORMATION LITERACY ALLIANCE (UK)
- MEDIA & LEARNING (Belgium)
- MEDIALAB (France)
- MeLCi Lab – Media Literacy and Civic Cultures Lab (Portugal)
- MILID (UNESCO)
- ORBICOM (UNESCO)
- Salzburg Academy on Media and Global Change (Austria)
- Savoir*Devenir (MILID and ORBICOM Unesco chair, France)

List of web resources (not exhaustive)

- ARCOM (France)
- Better Internet for Kids
- CANOPE (France)
- EDUSCOL (France)
- PIX (France)
- SELFIE (EU)
- Safer Internet Centres
- Safer Internet Day

List of public policy documents (not exhaustive)

- A Global Framework of Reference on Digital Literacy Skills for Indicator
- Code of Practice on Disinformation
- Digital Education Action Plan (2021-2027)
- European Council conclusions of May 2016
- Guidelines for industry on online child protection
- The European Council Recommendation of May 2018
- The European Digital Competence Framework for Citizens
- The European Framework for the Digital Competence of Educators (DigCompEdu)
- Recommendation of the Committee of Ministers to member States on developing and promoting digital citizenship education

List of projects/sensible practices (not exhaustive)

- Digital Citizenship Academy (2019-2022)
- Educational games
- Digital Citizenship Education
- Digital citizenship education - Trainers' Pack (2020),
- The Digital citizenship Education Handbook
- Improving Media Literacy campaigns on disinformation (ERGA Report)
- Mapping of media literacy practices and actions in EU-28 (Council of Europe)
- MIL CLICKS Social Media Initiative
- SMART-EU
- YouVerify!

2

5. Privacy based on pseudonymity vs. anonymity

Rethink Identity towards decentralized infrastructures for sovereignity

Editor: anoe
Authors: alexandre.del[anoe] and b9fb10e5edab70751bea54a781960087 ano[nym]e

keywords: privacy rights, GDPR

Rationale

Privacy can be simply defined as the ability to control who has access to personal information and how it is used. In this era of digital media, there is a growing recognition of the need to protect privacy in various contexts, such as online communication, social media, and e-commerce. Research in this area can help to identify and address potential threats and to develop effective solutions for protecting personal information.

Unfortunately, most digital infrastructure as currently implemented does not adequately protect privacy. Moreover, risks linked to this issue could grow more severe with the arrival of future technologies. A key question is: How can we better prepare our digital systems to protect privacy?

One promising avenue of research for privacy protection is to design new infrastructures based on anonymity and using privacy-enhancing technologies, along with implementing a variety of strong privacy policies and practices.

Current technological infrastructure failures and lack of regulation As currently implemented, digital infrastructure based on identity does not adequately protect privacy. This approach relies too much on passwords and other authentication methods which can be easily be defeated or compromised. Additionally, there is a lack of global regulation to establish uniform standards for data protection, leaving users vulnerable to misuse and abuse of their data

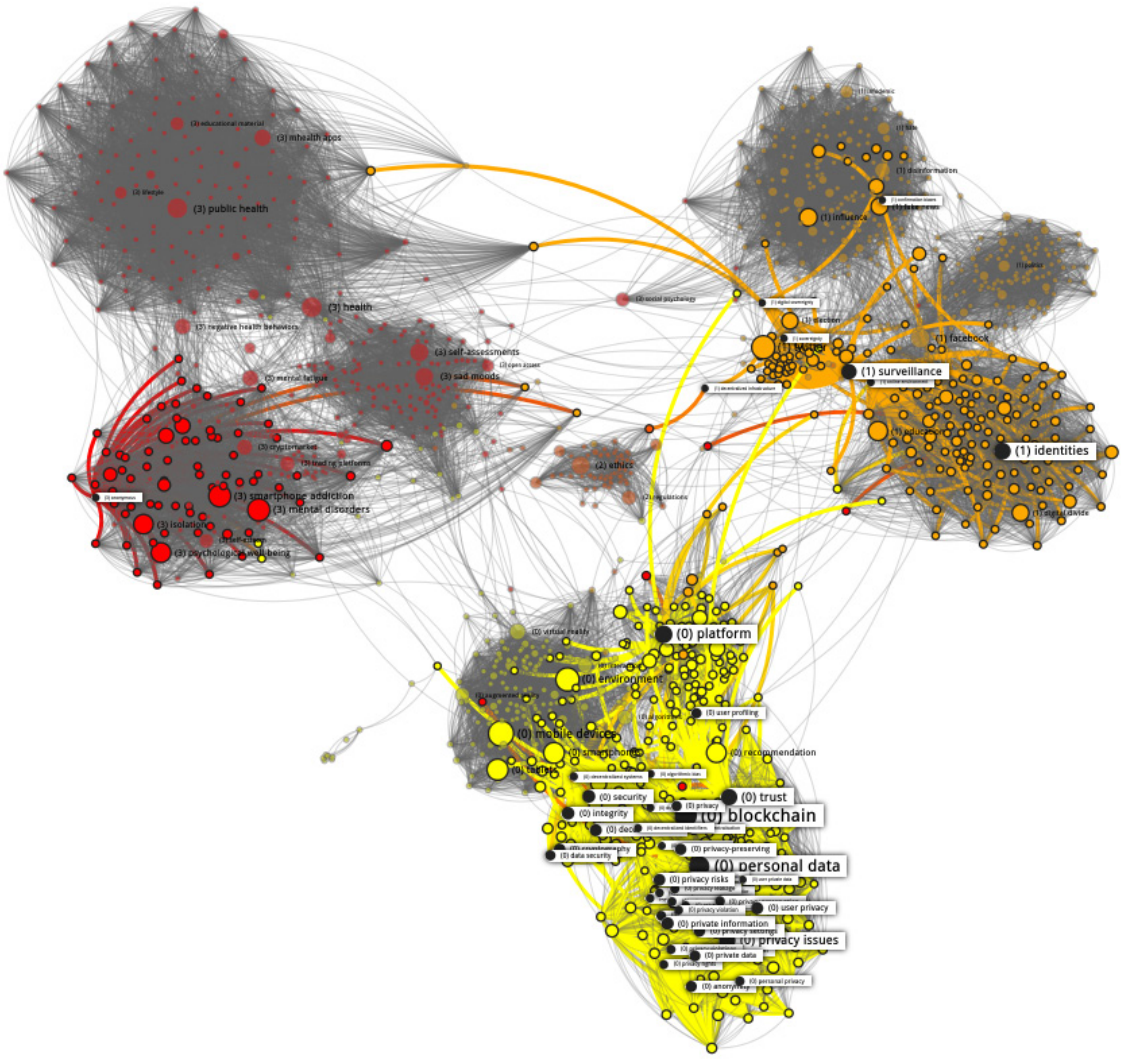


Figure 5.1: State-of-the-art domains concerned by this chapter.

by malicious actors. Many actors fail to provide users with sufficient transparency about how their data is being used.

Moreover, next technological revolution associated with quantum computing has the potential to break many of the encryption algorithms currently being used to secure online communications and protect data privacy.

The Legal and Policy challenge: acknowledging privacy as a fundamental right in the context of digital media. Article 12 of the Universal Declaration of Human Rights mentions the fundamental right of privacy. In practice, how this right takes shape in the legal environment of digital media is not so clear.

The legal context for privacy in Europe is largely governed by the General Data Protection Regulation (GDPR). The GDPR is a comprehensive data protection law that applies to all individuals and organizations in the European Union (EU) and the European Economic Area (EEA). It sets out the rights of individuals to control their personal data and the obligations of organizations to protect it. It also establishes a framework for international data transfers. The GDPR also requires organizations to report data breaches to the relevant supervisory.

Despite this general legal framework, many issues remain confusing. The legal implications of data collection depend on the type of data being collected and the laws of the jurisdiction in which the data is collected. Generally, data collection must comply with applicable laws and regulations, including those related to privacy, data security, and data protection. Also, data collectors must ensure that they have the consent from individuals whose data is being collected, and that the data is being collected and used in a lawful manner. Failure to comply with applicable laws and regulations can result in sanctions.

The problem with GDPR and similar legislation is that such laws are written too broadly. Hence they try to cover too many unrelated issues, and do not address many unforeseen consequences. We need further research to understand how privacy rights, as fundamental human rights, might be better respected and protected. Also, how can we ensure that people have the right to decide who has access to their data and how it is used?

Topic 1. Understanding why design matters

keywords: identity, anonymity, design.

Protecting privacy is a matter of design in digital infrastructure. How digital systems are designed and built can either enable or hinder the protection of personal information. When privacy is considered at the design stage of a digital system, it is much easier to build in privacy-enhancing features and safeguards that can protect personal information from being accessed or misused.

On the other hand, if privacy is not considered at the design stage, it can be very difficult or even impossible to add privacy protections to a digital system after it has been built. This can leave individuals and organizations vulnerable to privacy breaches and the unauthorized use or disclosure of their personal information.

There are two main design to protect privacy: privacy based on identity, or privacy based on anonymity. Privacy based on identity is generally considered more expensive than privacy based on anonymity because it requires more resources to maintain.

On the other hand, privacy based on anonymity does not require the same level of identification and authentication, and therefore may be simpler and less costly to implement. In an anonymous system, the identity of the user is not tied to any specific data or activity, so there is less need for the system to have robust safeguards in place to protect personal information. Both approaches to privacy have their own trade-offs.

Key challenges and questions

▪ **Challenge 5.1 — Understand the possibilities of infrastructure for the digital world.** Can we understand how to build infrastructure that effectively protects privacy and promotes trust in the digital world? Real understanding must confront a host of technical factors including data minimization, data security, data transparency or data control at the design stage of digital systems.

Topic 2. Fund research to support the axiological neutrality

keywords: cognitive biases, anonymity, identity, privacy, infrastructure.

Axiological neutrality refers to the idea of being neutral or unbiased with regard to values or value systems. This is important in efforts to solve many problems where prevailing biases or preconceptions make some approaches seem “obvious” and others much less so, despite the lack

of any real basis for such a perspective. Axiological neutrality allows individuals and institutions to approach issues objectively and without preconceived notions or biases, which can help ensure that decisions are based on evidence and facts rather than personal beliefs or preferences. According to the metrics of papers dealing with Privacy based on Identity or Anonymity in the Digeing corpus (68k docs, see section C.2), axiological neutrality is strongly needed in this field due to the influence of powerful cognitive biases entrenched by years of earlier practice.

Cognitive biases in favor of identity. A number of prevailing biases act to favour identity as an approach to security over anonymity. Addressing such biases will be crucial to making real progress in solving the privacy problem. Key biases include the following:

1. **Anchoring bias:** This bias occurs when people rely too heavily on their initial impressions or estimates, leading them to overlook better solutions. For example, if someone initially believes that the best way to protect privacy is to use privacy based on identity technologies such as authentication, they may not consider other options such as de-identifying data or using anonymization techniques.
2. **Confirmation bias:** This bias occurs when people seek out information that confirms their existing beliefs and ignore evidence that contradicts them. For example, someone who believes that an approach based on identity is the best way to protect privacy may selectively seek out malicious and illegal activities which confirm their belief that other approaches are inferior. In particular, they may focus on problems linked to anonymity and ignore information about the potential drawbacks of any infrastructure based on authentication and identity.
3. **Groupthink:** This bias occurs when people in a group make decisions based on the desire to maintain group harmony and cohesion, rather than considering all options objectively. For example, a group discussing privacy protection may seek to preserve group harmony by maintaining the revenus of copyrights of authorship over finding the best solution to protect Privacy based on anonymity.
4. **Framing effects:** This bias can occur when the way information is presented or framed influences people's decisions. For example, if options for privacy protection are presented in a way that emphasizes the potential risks of using anonymity, people may tend to ignore that option, even if it might be the most effective. For example, people often usually fear anonymity because it can be used to hide identity and intentions, which can lead to malicious activities such as cyberbullying, identity theft, and other forms of online harassment. Additionally, anonymity can be used to spread false information or to manipulate public opinion.

Key challenges and questions

▪ Challenge 5.2 — Can we learn to reduce the impact of cognitive biases when considering possible solutions to privacy?

Topic 3. Support privacy based on anonymity

When people feel that their privacy is being respected, they are more likely to trust the company or organization handling their data. This increased trust could lead to more people using the company's products or services, thereby boosting revenue. This additional revenue can then be used to invest in even stronger privacy protections, which can further increase trust and attract even more users. Over time, the organisation benefits from a virtuous circle of privacy protection, trust and success.

A key insight is that privacy based on anonymity can help protect privacy based on identity. For example, anonymity based design could help protect personal information from being shared or accessed without permission. It could also help protect people from being targeted or discriminated against based on their identity.

More generally, we need research to understand how we might build a system with anonymity by design. Two concepts that will likely be of major importance are decentralized networks and distributed ledgers. Decentralized networks are designed to provide anonymity by design. They are built on a peer-to-peer network, which means that there is no central server or authority that can track user activity. A distributed ledger is a type of database that is shared and synchronized across multiple computers. It is designed to be secure and anonymous, as it does not store any personal information about its users. keywords: decentralized network, distributed ledger, trust, discrimination

Key challenges and questions

▪ **Challenge 5.3 — Is the option to build infrastructure based on anonymity by design using technologies such as anonymous communication networks really effective?** How can we use such networks to help protect users' privacy by routing their communications through multiple nodes, making it difficult for anyone to track their online activity?

▪ **Challenge 5.4 — Anonymity-based infrastructure may use privacy-enhancing technologies (PETs) such as zero-knowledge proofs and homomorphic encryption.** Can these provide a solution to protect privacy? We need to understand how these technologies might help to protect privacy by allowing users to prove the correctness of a statement without revealing any additional information.

▪ **Challenge 5.5 — How can organizations implement strong privacy policies and practices, such as encrypting personal data and limiting access to only those who need it?** Can we devise training and resources to help employees understand the importance of protecting privacy and how this might be done effectively?

Topic 4: Improve Privacy based on Identity

keywords: decentralized network, distributed ledger, trust, discrimination.

Research on anonymity can help to identify best practices for protecting personal information in situations where it is not possible or desirable to rely on identity-based approaches. For example, researchers might study how anonymized data can be used for research or other legitimate purposes without compromising the privacy of individuals.

Research on anonymity can also help to inform the development of technologies and approaches that can be used to protect personal information in identity-based systems. For example, researchers might study how anonymization techniques can be applied to personal data to protect privacy while still allowing for the use of that data for specific purposes.

In another example, research on anonymity can help to identify the trade-offs and limitations of anonymity-based approaches to privacy, which can inform the design and implementation of identity-based systems. For instance, researchers might study the limitations of anonymity in terms of accountability, or the potential for anonymity to be abused for malicious purposes (which could of course avoided in a well designed system).

Key challenges and questions

▪ **Challenge 5.6 — We must understand how research on anonymity might provide valuable insights and lessons that can be applied to the design and implementation of privacy-based systems.** These might also be based partically on identity, helping to ensure that such systems are effective and responsible in protecting personal information.

6. Immersive social media and the metaverse

Corresponding author: Mel Slater
Authors: Mel Slater & Beatrice de Gelder
Keywords: Metaverse, Virtual Reality.

Rationale

From the 1990s onward, research on Virtual Reality (VR) has mostly been limited to university laboratories and industrial settings, concentrated on areas of benefit to society. This has changed in the past decade, however, as developments in VR hardware – in particular low-cost stereo head-tracked head-mounted displays and associated tracking technology – have transformed VR from being an esoteric and expensive tool of university labs and industry into a low-cost consumer product. Global companies have transformed the situation so that devices today can deliver high-quality experiences for the home user at costs equivalent to a smartphone.

Whereas earlier university and industry research was under tight ethical control through university ethical boards and company health and safety guidelines, now, as VR enters the mass market, there is little or no ethical supervision. The companies and organisations developing and distributing VR applications are also deciding whether or not these are suitable for mass consumption. This despite a lack of empirical data concerning the effects of VR on individuals or groups.

Global companies have also embraced the concept of a 'metaverse'. This far-reaching idea asserts that today's web will ultimately be replaced by virtual platforms wherein large numbers of people carry out their everyday work and leisure activities in an immersive shared world, interacting with digital representations or "avatars" of one another. Individuals in such a metaverse will work, shop, attend medical appointments, obtain news, visit relatives, attend concerts, and so on in an environment which also encompasses all the functions of today's social media.

While this could unleash a new era of creativity by freeing people from some constraints of physical reality, it also contains potential dangers. For example, the historian Yuval Harari [80] has argued that the metaverse could realise the 'ideal' totalitarian society, wherein everything that everyone does is fully monitored at all times, with machine learning analysing masses of data

as a means for manipulation and control. Such total surveillance and control was not possible even in George Orwell's classic 1984. Here we outline a few topics of interest arising from this potential alternate reality.

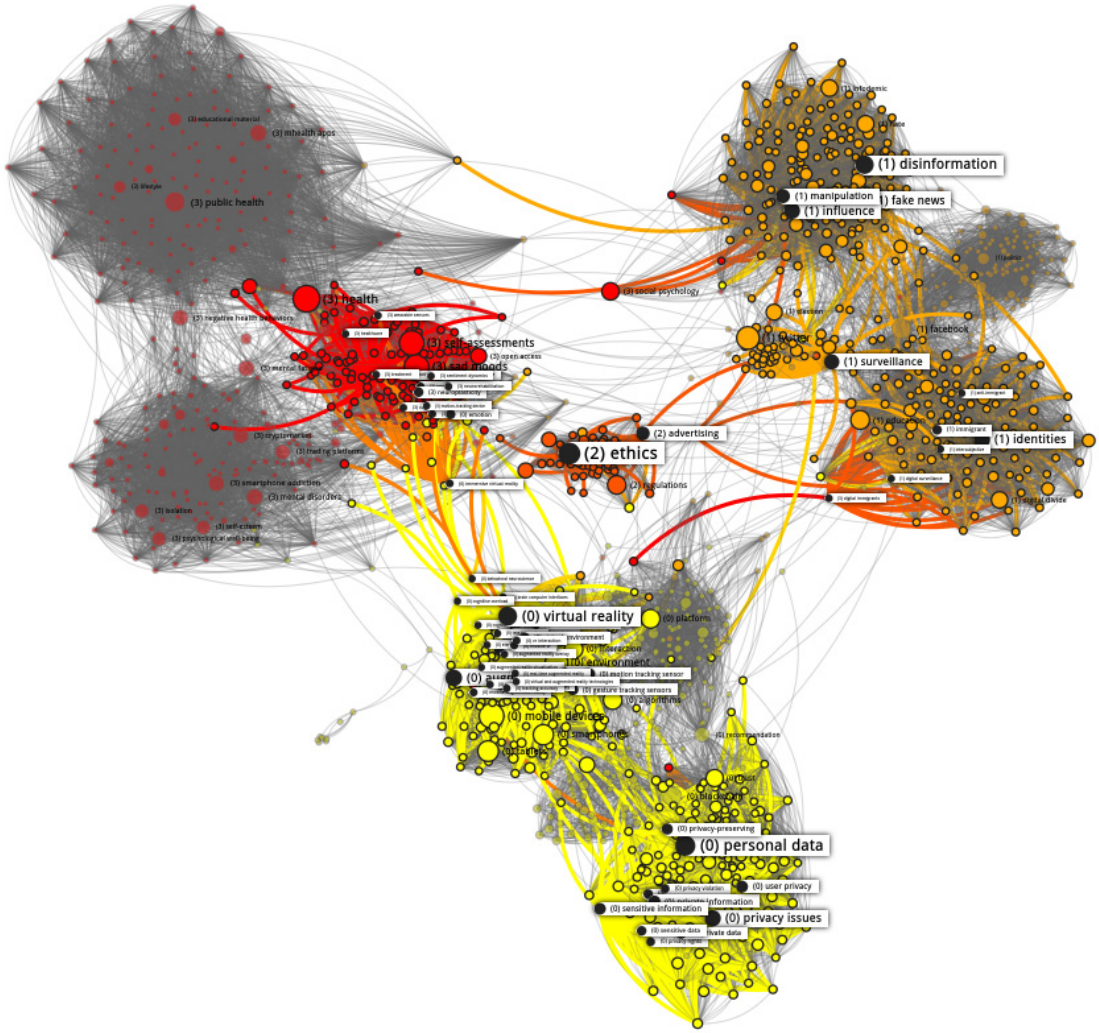


Figure 6.1: State-of-the-art domains concerned by this chapter.

Topic 1. Total surveillance

Keywords: surveillance, tracking, prediction, advertising, control

Modern consumer oriented and relatively low cost head-mounted displays and associated devices can track a user's head-gaze direction, eye movements and gaze direction, as well as movements of facial muscles, mouth, lips, hand and fingers. This all happens within the head-mounted display and handheld controllers, or without any controllers, and without the need for an additional computer to drive the system. In principle, the whole body can be tracked from the head to foot if a computer is also used together with external tracking devices. Additional peripheral devices can monitor heart functioning (ECG), arousal (skin conductance), respiration, blood pressure and temperature.

In the relatively near future it is likely that all of this functionality and more will be encapsulated within the head-mounted display without the need for additional peripheral devices or a separate computer. Research also shows that the identity of a person can be inferred even with limited tracking data [39, 53]. And, because gaze direction, body movements and physiological activity are all in response to immersive digital content (i.e., the participant is in an environment and responding to situations events and initiating interactions) there is a direct connection between what the participant perceives and how they respond.

This emerging technology may offer many opportunities for human benefit, but it also clearly raises risks for pervasive human surveillance. Imagine data from millions of participants capturing the association between content and their responses at multiple levels, and machine learning applied to this data to build predictive models. These could operate at the mass level where social group and societal activity might be predicted, or at the individual level where responses to content can be predicted. This could be highly useful for company business models based on advertising.

Tools for prediction can also be used for control, however, as participants can be manipulated into attitudes and behaviours to which they might not normally subscribe. While this can be put to positive use (e.g., health campaigns) it can obviously be used for political purposes, for fomenting hate, gaining support for particular individuals or political programs. It has been argued that illegal use of Facebook data, for example, influenced the Brexit outcome [32].

The data on Facebook is limited to clicks and likes. Imagine how much more powerful this will be when the data available consists of the activities and physiological responses of people in response to immersive stimuli. A major challenge for future research is to foresee some of the major risks in this area, and understand how these technologies can be developed in a way which ensures broad social and individual benefit.

Key challenges and questions

▪ **Challenge 6.1 — How can technology enable the multiple positive advantages of mass social immersive systems such as the metaverse while avoiding the negative consequences for surveillance, manipulation and control?** The real-time registration of tracking data is essential for the operation of immersive systems, and in some applications (medical or mental health) real-time physiological data may also be useful. But there is no specific need for the transmission of such data to external servers. The data is used as it is encountered – for example, a head-turn is used immediately to update the visual displays – but need not be stored outside of the display system itself. If it is stored outside of the display system, for example, on central company servers, there is no need for the analysis of such data. Even if anonymised data might be useful for constructing future improved systems, there is no need for it to be used for advertising or political manipulation. We require a world-wide regulatory framework able to balance the true needs of the technology for its own internal development with societal and individual privacy and protection.

▪ **Challenge 6.2 — Ensuring personal identity security.** If a large proportion of human activity is to move to immersive online systems such as the metaverse, how can each individual participant be 100% sure that the individuals they are interacting with really are those individuals and not imposters? How can individuals be sure that their own identities are not 'hacked' for nefarious purposes? Of course this already happens on social media with, for example, some Twitter users having been banned after impersonators stole their identities and posted hate speech. How much more powerful will this threat become when not just text, images and video, but the 'whole person' including their appearance, movements, voice and behaviour can be hacked and used by with others in the metaverse? Developing methods to protect individual personal identity is an extremely high level requirement for the safe functioning of metaverse-like systems.

▪ **Challenge 6.3 — Avoid the breakdown of shared reality.** In principle, an immersive environment might display the same content to every individual participant. Those individuals will then perceive objects, situations and events from their unique embodied perspectives, and what they perceive will depend on their own particular dispositions, just as in physical reality. However, it need not be this. The content displayed to each individual might be made different without anyone knowing that this is the case. Two people interacting might be displayed as an assault to one person, but as a friendly conversation to another. Through such effects, shared reality can begin to break down without anyone knowing why. If the metaverse becomes the norm for conducting mass human interactions, ranging from work through medical care and leisure, then maintaining a shared social reality is an absolute requirement to avoid societal conflict and breakdown. Is it possible to regulate content to the extent that there is some minimal level of truth invariant across participants?

Topic 2. Blurring the distinction between reality and immersive digital events

Keywords: distinction between real and digital content, bias, confusion, fictional reality.

Participants experiencing some scenario in a virtual immersive environment may later have the feeling that the events in which they were involved actually happened to them. While most scenarios may be quite innocuous, others could have profound implications. For example, a negative interaction with members of a particular racial group in virtual reality could bias the person against that race in reality. Such virtual mis-representations may lead to individuals retaining their experience as if it were first hand [5]. This could be used to purposely manipulate people towards hatred of a particular person or social group, or it might take place by accident. However, addressing this potential problem may be problematic, since deliberate interventions to make clear that "this is not really happening" would defeat the purpose of an immersive application. The same would be true if false events were identified by the use of different colours or types of shading.

It could be argued that the same applies to movies, which also portray a fictional reality. How do people maintain the distance between movie fiction and reality? A significant difference, however, is that movies are about other people and are experienced as if observed from the outside. In contrast, immersive systems involve the participant in the action, and important events occur in peripersonal space. Even if the participant observes an event from a distance, it is possible that the event could also happen in their immediate vicinity just through a change of location. This is impossible while watching a movie, since events cannot cross the screen barrier. It is in the nature of immersive systems that situations and events are perceived as real [69, 54], so addressing this potential issue is problematic.

Key challenges and questions

▪ **Challenge 6.4 — To obtain empirical data on boundary confusion.** While many speculate that immersive systems can lead to the breakdown of the boundary between reality and virtual reality, there is no data on this. Therefore experimental studies on this issue need to be carried out, with the resulting data contributing to enhanced knowledge of this issue and how to overcome it.

▪ **Challenge 6.5 — Regulatory control is not possible.** As argued above, it is in the very nature of virtual reality that the human perceptual system tends to take them as really occurring, even though cognitively people know that nothing real is happening. It is likely, therefore, that negative interactions with a virtual character portraying a real actual person could have negative consequences for real-life interactions with that person. It is not possible to regulate against this in a general way. However, it may be possible to protect the digital representations of people and organisations so that portrayals leading to a negative interaction between one participant and the virtual representation of another would not be possible. Yet this still does not avoid negative interactions between participants and other people such as members of a particular race or class. In the context of an immersive story, a virtual criminal must be represented somehow as a particular gender, race, nationality, class, appearance. Such portrayals may influence people in real life in their relationships with actual members of that group. A difficulty is that controlling this means controlling everything that is possible in virtual reality. Perhaps the only way out of this problem is 'good practice' principles where deliberate malfeasance is avoided.

▪ **Challenge 6.6 — The use of warnings.** For movies we have a classification system which

pre-warns audiences about possible harmful content, with guidance about suitable age ranges for viewing. This is possible with VR, but it is not clear if it is would be useful. Since in VR people are directly involved, not passive observers watching through a screen, the impact can be far more powerful than watching a movie; hence, people may 'forget' or ignore warnings. For example, it is surprising how often it happens that when people encounter negative content in VR they forget that they can just take off the head-mounted display. Another way to signify negative content which may blur over to reality is to have conventions for portrayal that signify 'this is not real'. Again, this may be unfeasible, since it would undermine the very purpose of immersive systems.

Topic 3. The re-entry problem and the effects of long term exposure

Keywords: re-entry problem, adaptation to reality, dangerous activities.

The problem of transitioning out of VR involves the challenge of readapting to reality [5, 55]. This may include low level perceptual aspects, such as the participant having become adapted to the visual aspects of the head-mounted display, then needing to readapting to normal vision. Or it might involve higher level aspects such as not caring about the physical environment or other people because they may be considered as only virtual [36]. However, Lanier[81] noted in his discussion of the early days of VR that on 're-entry' to the real world, people marvelled at the sensory richness of reality compared to VR: the multitude of colours, the sharp resolution, the wide field of view, the sounds and sights of physical reality in contrast.

In other words 're-entry' improved their appreciation of physical reality. Whether the re-entry problem actually occurs or not is an empirical question. However, we can imagine that it might lead to some dangers. Suppose, for example, that someone spends hours a week in a virtual environment where they drive a car as part of a game that involves crashing cars into one another: might they carry out the same actions in physical reality by mistake? Or suppose in virtual reality they become used to flying, jumping out of a window and floating to the ground. Might they mimic such behaviour in physical reality? This is especially important for the metaverse, where if this immersive world comes to be as expected, people will be spending many hours a day in virtual reality. The re-entry problem is really a sub-aspect of the general problem of long term exposure to immersive systems.

Key challenges and questions

▪ **Challenge 6.7 — To obtain empirical data on the re-entry problem.** It is unknown whether the re-entry problem actually exists. This will require long-term experimental studies because the problem is precisely concerned with long-term exposure. Therefore we need longitudinal studies of this issue. This ideally would require a large cohort of individuals to be followed over several years.

▪ **Challenge 6.8 — Regulatory limits.** In some professions there are limits to the amount of time in which people are allowed to be continuously engaged in the required activity. For example, there are strict controls on how many hours a pilot may fly continuously, or over various time periods such one month. Enforcing such rules in an immersive system would be impossible and undesirable. However, as with some smartphones, participants may be provided at least with data about the numbers of hours of exposure. There may be standardised test programs that could be used to measure the degree of re-adaptation to reality, and the problems of the effects of long term exposure.

▪ **Challenge 6.9 — The problem of addiction.** It has been argued that exposure to VR, especially long-term use through gaming, might lead to addiction and withdrawal from society. There has been a phenomenon in Japan and other countries referred to as 'Hikikomori' [41] where continuous exposure to digital media has led to withdrawal from society. Could this occur with prolonged and continuous use of immersive systems? This again is an empirical question which can only be answered by longitudinal studies.

Topic 4. Understanding how much VR can help restore full multi-sensory communication, and the problems likely to arise

Keywords: multi-sensory communication, non-verbal modes, neuroscience, natural environments

New technologies will allow VR and AR to become a media of communication, replacing newspapers, magazines, and traditional TV news shows. In principle, this could promise a restoration of some of the conditions of real life communication, bridging the gap between social and informational dimensions. In addition, VR can now be enriched with information about the participants in the course of the ongoing interaction that is not normally available in daily life. Such information includes body temperature, cardiac function (ECG), arousal (skin conductance), respiration and blood pressure. But it is still unclear what role they could play when provided online in real time as part of a virtual interaction. Supplementing virtual characters with real physiology and adding that to the interactions in VR presumably increases the realism of the virtual experience.

But this virtual experience will also place new cognitive demands on people, with unpredictable consequences. VR is rooted in an ensemble of engineering achievements that create and sustain somebody's digital presence, the plausibility of the specific environment and the digital embodiment of the participant in it. The implementation of each of these conditions reflects our current understanding of determinants of human existence in the natural world. This scientific understanding is still very limited; however, the neuroscience literature is rich in examples of dissociations with their associated pathologies between the different strands that make up the texture of daily life in natural environments. This is an important issue for the enrichment of virtual social communication. Combining images with words, adding to that tactile information, possibly also smell and taste should be based on an understanding of how the sense merge which we do not have.

Key challenges and questions

▪ **Challenge 6.10 — How will virtual environments disrupt natural human modes of perception?** Recreating natural existence in the virtual realm will inevitably bring the fragile harmony of multiple modes of perception (visual and auditory) to the surface and exacerbate imbalances and dissociations between the strands. For example, in the presence of visual and auditory information, the human brain has subtle and complex means of either combining both or giving precedence to one over the other as a function of context and behavioral needs. There is no one rule that fits all for how information from different sensory changes is combined. Are there ways to anticipate VR environments which cause problems? Can these be avoided? Can the brain be supported with tools for aiding the successful combination of different information pathways?

▪ **Challenge 6.11 — How will individuals use or misuse novel physiological information?** The interpretation of various physiological signals is not obvious for the layperson as

none of these measures on its own has a direct meaning. Indiscriminate enrichment of digital communication with such additional sensory channels may have a negative effect on communication. How in practice will people use or misuse such information? How will this change interpersonal perceptions or the dynamics of social interactions?

7. Interdisciplinary design for an inclusive and trustworthy digital public sphere

Editor and corresponding author: Camille Roth.
Authors (alphabetical order): Anna Boros, Luisa Fassi, Quentin Lobbé, Darian Meacham, David Pearce, Camille Roth, Mirko Zichichi, Guido Caldarelli, Jerôme Sackur, Erika Stael von Holstein.

Keywords: inclusivity, trust, public sphere, transparency, digital law, algorithmic explainability, design transparency, digital law compliance, platform and algorithmic values, inclusive design, explainable artificial intelligence.

Rationale

The public sphere encompasses a set of social practices through which individuals collectively identify and discuss societal problems, and thereby influence political action. This sphere now exists largely online in digital media spaces and platforms which support deliberation, opinion and identity formation, collective action and content creation. These platforms are not designed to support and further public values, however, but rather profit.

Moreover terms and conditions are largely set to establish conditions of private governance. As a result, existing regulatory tools have thus far shown themselves to be insufficient or incomplete, revealing a need for a more bottom-up approach to the design of the European digital public sphere. Understanding and stewarding best practices for the public sphere is extremely important to addressing a number of intertwined high-level challenges. These include climate change and the green-energy transition, as well as management of affective political polarisation in some EU countries.

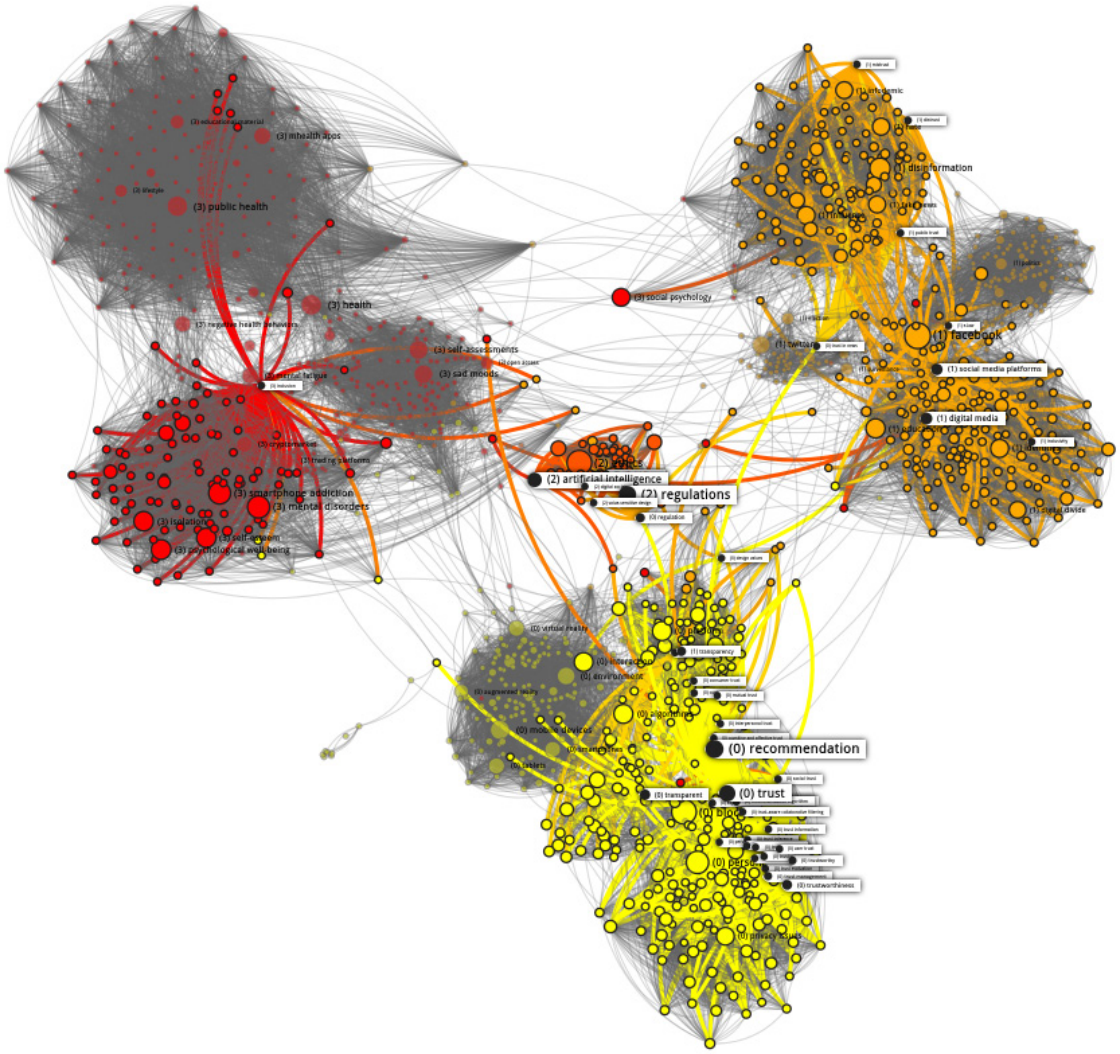


Figure 7.1: State-of-the-art domains concerned by this chapter.

Topic 1. Trust in the algorithmic implementation of law

Keywords: digital law, legal compliance of algorithms, algorithmic regulation.

We currently lack adequate transparency and trust in the digital public sphere, especially in the domain of digital law and justice. Digital law, encoded in software, is often not easy to interpret or understand. Many legal experts are not trained in computer science and software development, even though digital law is now a prominent feature of legal systems in many areas. This situation raises a number of key issues and challenges that future research needs to address.

Key challenges include finding ways to 1) combine old and new forms of legislation, 2) formalise legislation in software readily intelligible to legal experts and the general public, 3) verify the correctness of legal software, and 4) ensure that digital procedures and interactions that are mediated by software are transparent and fair to all parties.

Addressing these issues in practice will require ensuring broad collective participation in software development, but doing so will also mean exploring further questions. What kind of collective involvement or collaborative processes will be needed? Should these be internal or, more broadly, external to a given organization? How does the level of software development –

the scope and size of the project – influence the possible achievement of useful collective activity?

Further important issues arise in connection with key European principles and specific information technologies.

Key challenges and questions

▪ **Challenge 7.1 — How can the scope of software be appropriately defined in relevant legal texts** (e.g., the "AI Act" [22]), especially as it often relies on development through artificial intelligence and machine learning? How can legal protection be provided, where and when required, for new types of employment and/or regulation? Can the use of existing legal tools (e.g., GDPR, DSA) be improved or adapted for the effective regulation of AI?

▪ **Challenge 7.2 — Is there potential for these tools to have an impact well beyond the boundaries of the EU** (i.e. through the Brussels effect [67])? More broadly, what are the guidelines for international dialogue related to 'digital law'? How can dialogue effectively allow and encourage different approaches? How much coordination and coherence must be assured to achieve smooth functioning on a global scale?

▪ **Challenge 7.3 — At a more technical level, we need to develop methods able to bridge the gap between software code and legal texts and legislation** , the latter being formulated in natural languages. There is potential for declarative, rule-based, formal languages to formalise regulations and legislation in a machine readable form, while staying close to ordinary language, and thereby remaining intelligible to non-experts. It remains unclear, however, how we might harness the methodological advantages of these languages over traditional formulations of legal texts so as to ensure the consistency of a given set of laws or regulations and their correct application.

Topic 2. Explainability of algorithmic systems and rules

Keywords: algorithmic intelligibility, democratic audit of platforms, appraisal of algorithms, transparent metrics

Recent advances in AI have led to the emergence of a wide range of systems that are able to perceive and learn, make decisions and act on their own, on some tasks even more capably than human beings. However, the trustworthiness and efficacy of these systems is limited by a lack of transparency — people are generally unable to understand the basis on which these algorithms' outputs and performance rests. This problem of "explainability" has been recognized as a key barrier to the further development and exploitation of AI technology in society. Discussions in the scientific literature (*e.g.* [37]) suggest that both intelligibility and fidelity are required to achieve explainability: intelligibility captures how comprehensible an explanation is to humans, while fidelity refers to how accurately an explanation describes the behavior of the model.

Explainability is notoriously problematic for devices relying on machine learning techniques, as developers often do not fully reverse-engineer the core principles driving a given algorithm. This issue extends to symbolic rule-based systems as well. Beyond the difficulty of assessing the outcomes of complex interactions between the rules of a given algorithm, behaviour in practice also depends heavily on data that is time-, individual- and context-dependent. As a result, even with full access to source code, ensuring the intelligibility or appraising the fidelity of an algorithmic device is likely to require a tremendous research effort, whose conclusions may only have short-term validity as a result of the constantly fluctuating nature of regularly-updated algorithms.

Building upon some of the challenges associated with the previous topic, the computations that underlie the inference to a certain conclusion or a decision may be too complex for non-experts to understand. This may remain the case even if logical, rule-based languages enable code to be written that is closer to ordinary language and so more easily intelligible. In this way it is not enough that a logical system reasons correctly according to its own criteria of adequacy. For a logical system to be acceptable to a rational agent, we also require methods to make explicit and readily intelligible the essential factors leading to a given conclusion or decision.

Key challenges and questions

▪ **Challenge 7.4 — How can we achieve an effective level of algorithmic readability by non-technical people?** Which technologies might be able to support citizens in conducting legal and democratic audits of an algorithm?

▪ **Challenge 7.5 — How can we learn to regulate platform algorithms, and understand their likely impacts, without reverse-engineering them?** Is this possible without engaging in a thorough case-by-case study of the likely complex interplay between code, data and behaviors? Would it be enough to identify and disclose the macro-level criteria guiding the development of algorithms, rather than the micro-level of the software implementation? More broadly, what level of transparency is needed to achieve adequate explainability?

▪ **Challenge 7.6 — Can we build reflexive social media wherein users could access intelligible feedback on impact metrics computed at the platform level, and use these to improve their own experience?** How can users be informed about the main consequences of their consent to the use of such metrics, or to system-level processing of their information? How can dynamic consent approaches be used to provide greater user autonomy and transparency? (see also Ch. 1)

Topic 3. Transparency in the algorithmic design and values

Keywords: ethical design, digital well-being, metrics transparency, algorithmic freedom, value-sensitive design.

Digital public platforms emerge from a scaffolding of specific technical design choices determining everything from the back end of data storage and ontologies to the front end of a platform's interface, and the myriad algorithms operating at and between each of these levels. Many important decisions on the corresponding software architecture are often made by small teams of engineers or designers. Little is known, however, about how these actors take into account regulations when making these choices, how the pursuit of certain goals (including upholding normative values or reaching commercial targets) influences their decisions, and how well-equipped they are to understand the impacts of those decisions. These issues underline a need for more design choice intelligibility for platform developers, and greater design choice transparency for users.

In more detail, we need more attention to design choice transparency and intelligibility because designs impose consequences in three major areas: 1) the design of data bases and collection has important consequences for data protection and GDPR compliance, 2) the design of information processing pipelines can influence the effect of recommendation and ranking algorithms on the balance and diversity of consumed content, and 3) the design of the interaction grammar imposed on users (such as the possibility of liking, following, broadcasting content; the placement of information in such and such a way, and of signals here rather than there) may trigger or favor certain types of behaviors, cognitive biases or nudges.

Key challenges and questions

▪ **Challenge 7.7 — How can we promote inclusive design practices aiming to foster well-being while also recognizing the importance of attention and self-esteem disorders?** Similarly, how can we promote inclusiveness, while still acknowledging the different impacts of age, ethnicity, personality, or further minority-friendly designs while avoiding the embedding of offline discriminations into digital design choices? Members of the public are generally affected by various types of digital divides or risks related to the protection of their personal data. Given this context, how can we ensure that different usage types and vulnerabilities are fairly considered? How can we help users to situate themselves within social media through the development of tools for increased self-awareness, self-mapping and self-visualization? What kind of digital work should be encouraged and how should it be designed so that positive effects outweigh any eventual disadvantages?

▪ **Challenge 7.8 — The emergence of a new generation of alternative and more inclusive digital public sphere platforms will require the support of an ecosystem of legal, economic and technical services.** How can we encourage the creation and growth of such platforms, especially on a small scale and while remaining responsive to local views? How can we create financial support to test new models for civic participation within digital platforms? How can we develop experimental frameworks to test the relationship between feature combinations and human well-being?

▪ **Challenge 7.9 — How can we support value-sensitive design, perhaps as an alternative model to solutionism?** How can we seek to stimulate the collective and collaborative definition of certain kinds of pro-social values at a large scale within a given platform and possibly across platforms? What kind of broader scale participatory methods might be used, or which soft or hard legal tools can be utilised? How can we foster interdisciplinary design practices and

the constitution of what may be called "socio-technical teams," bringing together perspectives from law, computer science, behavioral sciences and the social sciences and humanities?

▪ **Challenge 7.10 — Which set of values would be minimally desirable or consensual for the guidance of design?** How can we ensure their persistence and coexistence, even if some may be partially contradictory (users may be able to choose, for example, between principles supporting more or less politically-balanced information diets) [20]? How can we encourage platforms to disclose the criteria and metrics governing the evaluation of the performance of their platforms' design and technical choices?

8. Impact of the digitalization of society, digital labor and well-being

Editor and corresponding author: Antonio A. Casilli

Keywords: digital labour, digital platforms, artificial intelligence, data work

Rationale

Algorithm-driven systems and data-intensive information technologies often overlook the importance of labour and the challenges associated with it. In particular, the rhetoric surrounding the displacement of jobs by automation fails to acknowledge the role played by human labour in AI systems. Nevertheless, an expanding body of literature examines the adverse impacts of digital labour on well-being, especially mental health. Some forms of digital labour are linked to work-life permeation, mental burnout and PTSD, while others expose workers to financial strain and precarious life conditions, which increase stress levels.

Digital media should be considered systemically at the societal level, not just in terms of their effects on society. How people use and perceive them can be influenced by how they are designed and manufactured. Online advertising, subscription models, and data collection practices can affect how search results and content recommendations are personalized and marketed.

The design and production choices that determine the overall effect of social media go beyond the algorithms used to personalize search results. The invisible work of data producers and

annotators on crowdwork platforms, as well as users on social media, must also be recognized. By tagging, filtering, annotating, and creating user-generated content, users enrich social platforms' data. Similar tasks are performed by crowdworkers, for generally insufficient compensation. Additionally, it is crucial to consider production aspects of the environmental impact of digital media. Global energy consumption and CO_2 emissions are largely accounted for by the information technology industry. Both qualified and unqualified labor are involved in the manufacture of these processes.

The Digital Labour Disclosure challenge provides a unique opportunity to evaluate the social costs of this type of work. By formalising digital labour's role and mapping its global operations and supply chains, this framework builds on initiatives in green AI and workforce disclosure along multinational supply chains. The establishment of a coalition of institutions and companies is necessary to improve corporate accountability and transparency regarding digital labour workforce issues.

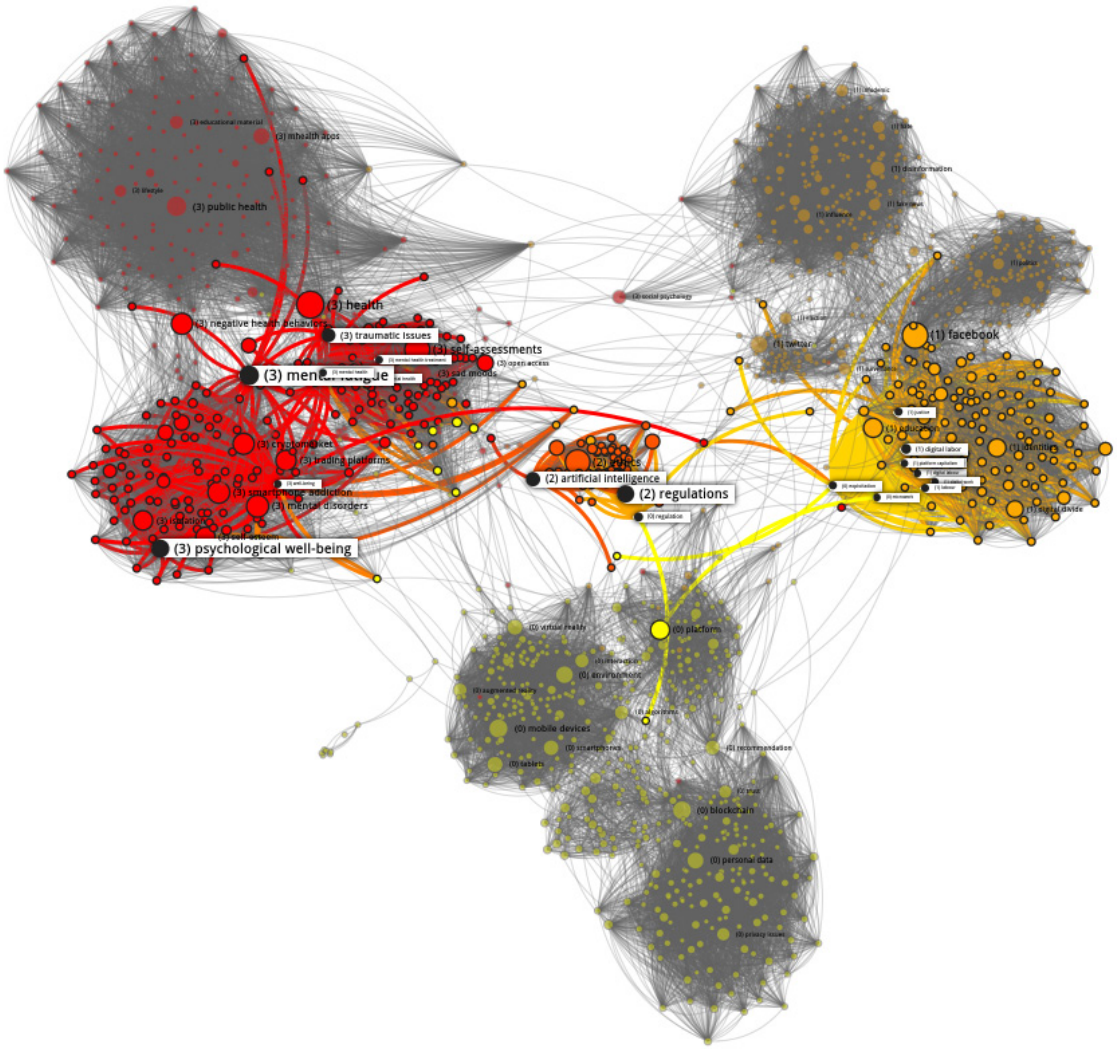


Figure 8.1: State-of-the-art domains concerned by this chapter.

Artificial intelligence and the hidden role of human labour

The push for artificial intelligence involves an attempt to automate large portions of human societies. By utilising data, processes can minimise or eliminate the need for human intervention. In the last few decades, a great deal of research has considered the link between smart technologies and the displacement of jobs, with some studies foreseeing a loss of almost 50 percent of jobs to mobile robotics and machine learning [26]. Yet robotics and machine learning affect human employment in different ways. For physically demanding tasks, robotics seems to be a potential replacement for human labour. In contrast, machine learning solutions seek to integrate themselves substantially into the systems of human labour [91].

In this chapter, we focus on the continued reliance of AI technologies on human labour. New concepts, such as "heteromation" [74], "fauxtomation" [57], and "ghost labour" [79], have emerged over the past few years to highlight the human cost of supposedly labour-saving technologies. These notions capture how the success of an AI system requires a significant amount of human input during development. Apart from computer scientists, data scientists, and system engineers (CSL, or computer science labour, in the following paragraphs), other workers contribute "unskilled digital labour" [78]. Despite low wages, atomisation, and invisible status, this is essential labour and DL providers involved in AI systems are often undervalued. Current research obscures or overshadows the challenges facing this workforce in working conditions, remuneration, equality, and access to the labour market.

The combination of artificial intelligence and digital platforms creates digital labour, oriented around the use of data both during the production process and as an output. In addition, this process is broken down into small tasks that are delegated to a large number of individuals. DL typically refers to unrecognised, scattered work. On today's labour market, there are three main sets of DL activities:

- On-demand location-based platforms work, used to provide services such as food delivery, logistics, and transportation;
- Micro-work, usually paid as piecework, these short tasks are used to produce and annotate data and last from one minute to a few hours;
- Socially networked labour, i.e. any activity, paid or unpaid, that supports content-based digital platforms.

The final element that must be stressed is that DL sits on a continuum between underpaid (e. g. Uber drivers, couriers, etc.), micro-paid (e. g. micro-workers, moderators, clickfarmers) and unpaid labour (e. g., content and data producers on social media and on messaging apps).

AI platforms and mental health

AI-driven platforms are especially harmful to mental health since they increase the risk of burnout by infiltrating our daily lives with work. Similar effects also apply to virtually any type of home-based or remote work. The use of algorithmic management stresses workers and users, who suffer from information asymmetries, unfair incentives, and manipulation if an algorithm decides who they have to deliver to, meet for a date with, or chat with [92]. In other instances, financial strain adversely affects the physical and mental health of individuals. In order to earn a higher wage, for example, couriers, freelancers, and digital pieceworkers compete against one another and against algorithms [51].

Although platform workers are economically dependent, they also do not receive the benefits

and social protection typically associated with dependent employment. The precarious and contingent nature of such work can have particularly negative impacts, especially for work performed at home or in isolation. This category includes DL tasks vital to AI production such as data annotation and AI output evaluation, which expose workers to risks such as social isolation and conflict. Severe forms of PTSD have been observed [86] when the DL is related to content moderation on social media.

This challenge on digital labour seeks to counteract the adverse consequences for well-being of AI solutions that permeate today's technological landscape. There is a high human cost associated with the production and deployment of these solutions. Any evaluation of AI systems should include a social impact assessment which focuses on the labour that is required to produce and market them.

Resonance with emerging trends towards labour disclosure and social cost assessment

The Workforce Disclosure Initiative inaugurated by the British NGO ShareAction informs the need for DLD [65]. A growing number of companies are participating in the Workforce Disclosure Initiative in order to collect data on human capital. The disclosure of social costs allows companies to compare themselves and establish comprehensive benchmarks about factors such as gender composition, working conditions, compliance with regulations and worker voice. Regular reporting that includes information on how companies perform on human capital aims to help policymakers to formulate policies, companies to keep track of their supply chains, and investors to be more responsible.

The DLD challenge can build on initiatives such as this and pursue corporate transparency around AI-related workforce issues. These principles would be applicable to virtually all economic sectors using AI. Regular surveys, inquiries, and reports on DL workforce should be conducted by both public and private organisations. The data produced can then be incorporated into performance analyses, and companies can gain insight into how to address pressing workforce issues. As with the Workforce Disclosure Initiative, the disclosure of digital labour can be particularly useful when it is targeted at multinational technology companies.

Another approach informing DLD is Green AI, an emerging trend in computer science and machine learning. As opposed to "Red" AI, which relies on polluting and wasteful technologies to train algorithms, Green AI advocates an approach to machine learning that considers not only economic or computing benchmarks, but also how they minimise environmental impact. By encouraging environmental cost disclosure, proponents of this trend aim to create environmentally sustainable AI that does not use excessively energy-intensive methods. The simultaneous disclosure of environmental and social costs is essential to sustainable AI. Researchers will have to improve on existing disclosure tools like corporate social responsibility reports and independent studies to take on the challenge of identifying the human costs of AI technologies.

A specific purpose of this prospective task is to examine the benefits and challenges of implementing digital labor disclosure initiatives that involve creating tools for protecting digital rights.

Topic 1. Formalising the Role of Digital Labour

Formalising the role of digital labour in machine learning solutions is the first step toward addressing the challenge of digital labour disclosure. Training, testing, and verifying solutions

such as deep learning algorithms and large language models requires huge amounts of data. In addition, producing AI also requires numerous runs of the same program to adjust for different parameters and correct errors. The overall cost of a particular AI solution depends not only on the cost of creating a dataset, but also on the number of trials necessary to determine the optimal hyperparameter settings. Green AI advocates formalise this using the equation:

$$Cost(r) \propto e \cdot d \cdot h$$

The cost of training an AI solution to achieve a certain result, **r**, grows linearly with three variables: the cost of executing the model on a single (**e**)xample; the size of the training (**d**)ataset; and the number of training runs needed to reach the optimal control of (**h**)yperparameters [59].

Key challenges and questions

▪ **Challenge 8.1 — Can we update this equation to consider the social costs of producing inputs for machine learning, including specialised computer science labour and digital labour?** The variables e, d, and h can in principle be described as functions of these other two variables. The cost of computer science labour has a direct impact on the cost of executing a single example and the cost of running trials to reach optimal hyperparameters ($e = e(CSL), h = h(CSL)$), while the cost of data annotation influences the size of the datasets ($d = d(DL)$). It is therefore possible to rewrite the equation as follows:

$$Cost(r) \propto e(CSL) \cdot d(DL) \cdot h(CSL)$$

In this approach, digital labour and computer science labour intervene at different points in the production process, affecting input costs differently. By rewriting the initial equation, we see that the resulting cost grows linearly with the cost of inputs that are a function of computer science labour multiplied by the size of the datasets which in turn is a function of digital labour:

$$Cost(r) \propto k(CSL) \cdot d(DL)$$

Rather than minimising labour costs, formalising these technologies should seek to recognise both the highly paid and lowly regarded contributions to their production.

Topic 2. Differentiating the Types of Digital Labour

We should move away from the linear view of AI systems as being in two distinct phases: production and deployment of AI. Rather, according to [60], it is rather useful to analyse three phases of AI: preparation, verification and impersonation.

Data generation and annotation are both necessary components of AI preparation, as users and DL workers create videos, pictures and texts together. The data is uploaded to paid platforms as well as to social media platforms, which may be "free" services such as *TikTok* or *Whatsapp*. Worker and user input must also be enhanced, triaged, sorted and filtered for the collected information to be of the highest quality. This is accomplished, for instance, through the use of tags, such as those used on Twitter or Instagram, or through Facebook's reaction feature. In some cases, these activities occur organically as a result of the platform's use, so they are not compensated. Other cases require the concerted effort of individuals who are rewarded, though minimally.

Machine learning pipelines continue to require human intervention even after solutions have been produced and released. The process continues with AI verification, which involves ensuring that the AI solution is functioning according to its specifications and standards. A product is also verified when it is marketed and used by millions or billions of users (another use of unpaid digital labour). Furthermore, for AI solutions such as image filters and LLMs, the evaluations and feedback provided by beta users are critical. Other tasks include debugging games and platforms, testing interfaces, flagging messages for moderation, evaluating automatic translations, solving reCaptchas. Such tasks are so ubiquitous and frictionless that they do not register as work, yet are crucial to the monitoring of systems and the detection of systemic failures at an early stage. Micro-paid or unpaid validation ensures that there are no inaccurate results and that the outputs are consistent with the norms, rules, and cultural values that apply to a given community of users by minimising inaccurate outcomes and checking the outputs for consistency.

When a model appears to be malfunctioning during the verification phase, what is sometimes called "real-time supervision" or "AI impersonation" takes place. A popular phrase used to describe AI impersonation is "artificial artificial intelligence", which is based on a quip made by Amazon's Jeff Bezos in 2006. While some claim to have fully automated solutions, many have only partially implemented AI systems, or haven't implemented them at all. Impersonation by artificial intelligence is a complex phenomenon that must be understood on a systemic level. Several workers perform real-time tasks to correct outputs and assist users. While they operate from the back office of a platform, part of their job involves maintaining that they are part of the AI systems without revealing that the tasks are being performed live by humans. There is a tendency to consider this to be a form of imitation of artificial intelligence, but AI impersonation has been around for decades, following its formalisation and patenting by IBM in 1984 under the name "Wizard of Oz computing".

Key challenges and questions

▪ **Challenge 8.2 — AI preparation does not always naturally compensate individuals involved in preparing content.** In some cases, these activities occur organically as a result of the platform's use, so they are not compensated. Other cases require the concerted effort of individuals who are rewarded, though minimally. Are there ways to adjust the preparation process so that digital labour is more appropriately rewarded?

▪ **Challenge 8.3 — How can nearly invisible processes involved in AI verification be properly rewarded?** Many tasks in verification are so ubiquitous and frictionless that they do not register as work, yet are crucial to the monitoring of systems and the detection of systemic failures at an early stage. We require new ideas on how such digital labour can be recognised and rewarded.

▪ **Challenge 8.4 — We should establish greater awareness of AI impersonation by real humans.**

Topic 3. Mapping Digital Labour Supply Chains

The world map of AI producers and users shows most AI developers, marketers and users located in the global North. There are two notable exceptions: China and India, which are customarily classified as emerging countries. If, however, we take into account where digital workers (those who create, generate, annotate, and triage data) reside, the geographical composition of the workforce changes significantly.

For this reason, we need a more complete mapping of digital labour supply chains and production networks in countries in both the northern and southern hemispheres. Research on digital labour identifies several AI data flows. A first data flow originates in South East Asian countries such as the Philippines, Vietnam, and Nepal and travels through India to Europe and the United Kingdom. Afterward, it reaches North America and the United States, which has a large internal market as well. Typically, workers recruited in the Philippines or Nepal are paid significantly less than those recruited in the United States. This explains why many companies view international digital workforces as a significant business opportunity [28, 86].

As for the French- and Spanish-speaking countries, the situation is vastly different. Historical differences and distinctive post-colonial patterns explain why data labor markets cater to different clients [68, 0]. Startups and tech companies in France look for data trainers in countries such as Madagascar, Côte d'Ivoire, Senegal, Morocco, and Tunisia. This is because these countries have a large pool of potential data workers. However, not all African countries were former French colonies; Kenya, Nigeria, Ghana, and Uganda were once British colonies and still fall under their economic influence in AI production [1]. As countries such as Egypt cater increasingly to Gulf countries and China, new dependences are emerging.

In Latin America, the political landscape is shaped by the United States' influence. There is a minority of data produced by workers in Argentina, Colombia, Chile, Bolivia, and Venezuela that is sent to Europe, particularly Spain. But North America is the principal destination for data exports [38, 60]. In a political setting that can be described as the "Monroe Doctrine for Machine Learning", digital labour flows to US-based AI companies.

Finally, China has a large domestic market of primarily rural workers living in Chinese inner cities. Digital labor is sold by these workers to AI companies located in wealthy coastal cities such as Shanghai and Shenzhen. Chinese-speaking diaspora communities in many other Asian countries also work in the huge tech hubs on the China Sea coast.

Completing and improving our understanding of these AI and data flows is key to understanding the AI supply chain and digital labour.

Key challenges and questions

▪ **Challenge 8.5 — By distinguishing the three phases of AI production, and systematically charting them across the surveyed countries, we can gain insight into the current situation.** One could observe, for example, whether AI preparation takes place both in high-income countries where data are generated and in low-income countries where data are annotated. Additionally, data verification can occur at all stages of the supply chain: workers verify the AI's functionality in countries along the supply chain, while end-users test final products in high-income countries. The final test can be whether AI impersonation takes place primarily at the end of the supply chain (i.e., in countries with a low level of income). We need to test these hypotheses against supplementary evidence in order to confirm them.

Topic 4. Launching a Multi-Stakeholder Digital Labour Disclosure Initiative

The launch of a multi-stakeholder initiative would create a coalition of organisations and companies which are impacted by this social transformation and need to tackle specific issues associated with it. Such a coalition would aim to improve corporate transparency and accountability in the workplace by focusing on the digital labour they buy and sell.

In the first phase of the initiative, companies and investors would be assisted in gathering the necessary data and sharing it with stakeholders, along with tools for analysis.

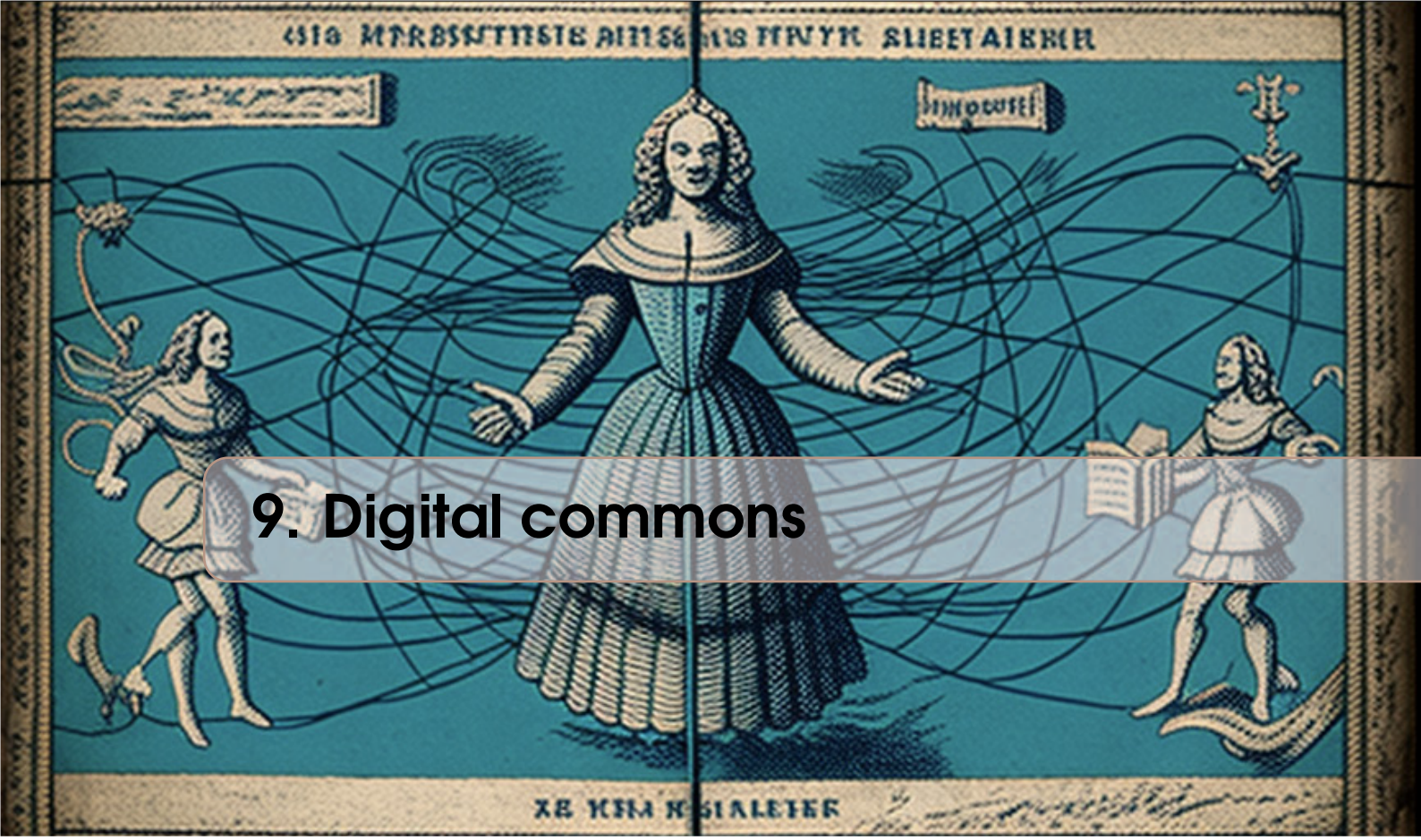
Key challenges and questions

▪ Challenge 8.6 — To federate and centralise as much data as possible about digital labour disclosure, the project should conduct a global digital labour survey. The purpose of this effort is to persuade companies to volunteer their workforce data annually and to contribute computing power to analyse the data. Using this information, companies and platforms can avoid externalities and liabilities by reflecting the social cost of their products. Data can thus be integrated into their investment and productivity analyses. The rights of workers involved in the production of digital technologies can be better protected by establishing policies to systematically monitor the amount of work. In order to accomplish this goal, government initiatives need to be supplemented by the actions of civic associations and local communities creating data trusts and digital commons.

▪ Challenge 8.7 — An additional objective of this initiative is to inform public policy. Public and private sectors can use the data disclosed to gain practical insight into digital workforce issues. Developing such initiatives would require the collaboration of a diverse community of stakeholders, ranging from research institutions to networks of scientists and activists. Also, they would rely on global NGOs, international research centres, and European unions which advocate for platform workers' rights.

Community

- Oxford Internet Institute (OII)
- University of Cambridge Centre for Research in the Arts, Humanities and Social Sciences (CRASSH)
- Global Network of Internet & Society Research Centers (NoC)
- Humboldt Institute for Internet and Society (HIIG)
- Weizenbaum Institut Berlin
- International Network on Digital Labor (INDL)
- COST Platform Work Inclusion Living Lab (P-WILL)
- Red Tierra Común
- Fairwork initiative
- Reshaping Work
- Partnership on AI
- Transnational Forum on Alternatives to Uberisation
- IG Metall
- Confederazione Generale Italiana del Lavoro (CGIL)
- Unión General de Trabajadoras y Trabajadores (UGT)
- Force Ouvrière (FO)
- Ugict CGT Union générale des ingénieurs, cadres et techniciens
- European Trade Union Institute for Research (ETUI)
- UNI Global Union
- IndustriALL
- International Transport Workers' Federation (ITF Global)
- ILO Research
- Institut de recherches économiques et sociales (IRES)



9. Digital commons

Editor and corresponding author: Melanie Dulong de Rosnay.
Authors (alphabetical order): Melanie, Dulong de Rosnay, Valérie, Schafer

Keywords: access to public information, access to scientific data, public participation, public platforms, digital rights

Rationale

A sustainable digital commons ecosystem requires an effective and socially beneficial system of rights and policies to ensure fair access rights to data, knowledge and information and to enhance collaboration. Access to data is crucial for science, education, public services and citizen participation. It also furthers private innovation, as many platforms rely on full access to users' data, without necessarily returning value to society beyond access to a service.

Different legal regimes, sometimes laws, sometimes terms and conditions, regulate access to data and more broadly rights pertaining to the reuse of personal data, privately-owned or hosted data. Between the two extremes of public domain without control for the users, scientists or the producers, or full exclusivity, there are other options for producing and sharing data, information and knowledge under commons-based conditions.

In particular, a commons governance model has been applied to digital works and data in various fields. These include free software, community media, collaborative encyclopedia, open data and open science. Such governance generally requires a favourable legal framework supportive of the needs of the general public and the technical possibilities of digital media processing, while preserving a balance between public and private rights. It is important to identify and recommend policies to guarantee better access to data, public or private, and to reconcile this with the privacy of data debate.

However, the full potential of digital commons has yet to be realised. One key issue, identified by previous research, has been a lack of legal and economic sustainability of projects exploring

alternatives for digital commons. In many cases, data stopped being produced after the end of the project's funding period or the departure of key volunteers, or it was stored but never reused. As a result, the full impact of the investment, sometimes with public funding, could not be assessed with specific metrics.

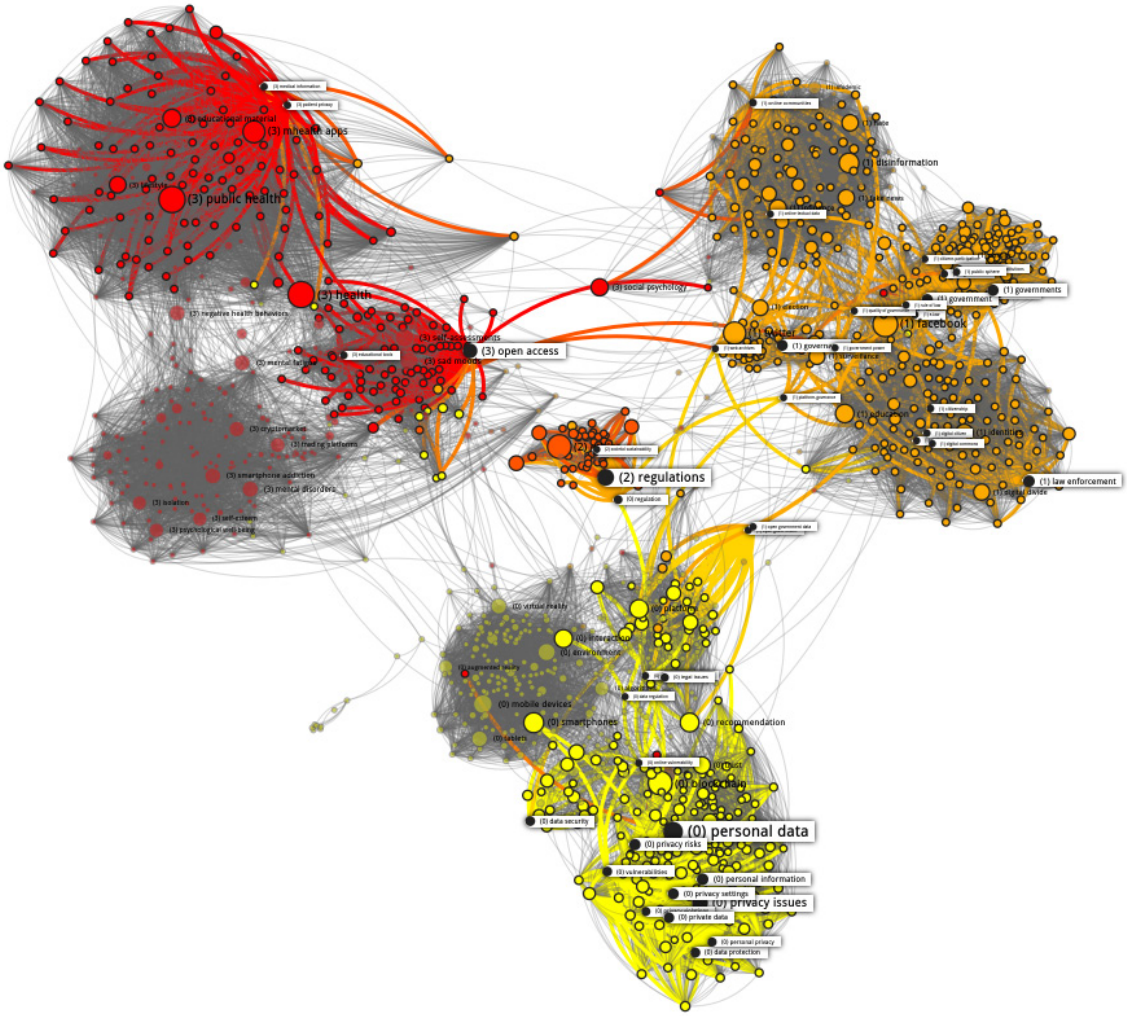


Figure 9.1: State-of-the-art domains concerned by this chapter.

Topic 1. Access to public services, public sector information, public participation, digital rights

Keywords: connected cities, public broadcasting, community media, media archiving, vulnerable citizen

Well-being and informed citizenship relies on individuals having access to knowledge as well as all public services and public sector information, including through digital media and education online. Human rights and freedom of expression rely on citizens having access to and being able to participate in the production of information which can be broadcast to the public, or co-produced as community media. Digital citizenship relies on participation platforms and applications, including to access public services.

Funding for digital public services and digital rights must not be neglected, especially for infrastructure maintenance, updating or digital archiving. This is true even if such efforts do not fit with the latest technological innovation or media hype, attention and subsidies, as such factors do not necessarily correlate with real social impact. This is especially important as our public media, public services and administrative trail are rapidly becoming digital-native heritage only (and not just digital versions of a pre-existing paper archive).

Key challenges and questions

▪ **Challenge 9.1 — Citizen participation and access to public information.** Digital applications and platforms already exist for the preservation of general medical information and confidential health data, public information and personal data related to individuals' travels and web searches, and information on cultural heritage. These platforms are crucial for individual and collective well-being, and must be frequently updated, adapted, curated and nourished with more content and metadata. Future EC calls should require that all funded developments, applications and platforms must be released under free software licences. Similarly, data should be managed under open data schemes if possible with all the relevant porting documentation to allow smaller hospitals or cities to also offer apps in their local jurisdiction without having to redevelop the wheel.

▪ **Challenge 9.2 — Digital archiving of community and public media.** Digital native web archives of social networks constitute cultural research resources. As such, archiving practices and methods require attention to issues of governance, access, openness and inclusiveness, which make them interesting case studies for understanding digital heritage patrimonial issues. Archived content can be used to analyse recent crises. Examples include collections linked to the "Me Too" movement, data collected in Europe during the COVID 19 crisis or studies on the expression of online memories or the development of digital uses. This challenge applies more broadly to the preservation of our collective memory by libraries, universities and public institutions in the service of historical and patrimonial objectives. These resources can become a cultural commons if a user community helps to define governance rules, for instance, by selecting the topics and categories which should be preserved. These are crucial challenges for cultural diversity.

▪ **Challenge 9.3 — Access to information on rights and public information with a focus on vulnerable populations.** Even after public sector information platforms have been established, they must be continually adapted to the specific needs of the most vulnerable populations through activities such as translation and development of specific contents. Rights, policy information and applications should be made available in all formats and languages, inclusive of all genders and ages, including children and teenagers. Specific adaptations are required for the elderly, perhaps with mediators for the digitally non-connected, easily accessible services for persons with disabilities, as well as non-native speakers, minorities and migrants. The latter groups have specific concerns related to issues such as emergencies, banking, transportation, visa processing, healthcare and childrens' education.

Topic 2. Supporting the development and the sustainability of commons-based citizen platforms

Keywords: legal framework, public and private funding, governance rules

Commons-based platforms can be defined as alternatives to public or private initiatives, but can also form from partnerships of public and private institutions. They are produced

collaboratively and managed according to rules defined in common by a community with defined shared access rights. Many such initiatives have been developed in the field of digital media and digital services, including the well-known encyclopedia Wikipedia, and many original community expanded projects related to data, media, news or travel. Commons-based platforms developed and maintained by citizens also exist for maps (OpenStreetMaps), citizen science and internet services, as well as for digitally-mediated services such as tourism, car-pooling, food delivery, community gardening, city collective metrics for birds or pollution, etc.

After recent governance changes at Twitter, many have called for the broad adoption of Mastodon, a decentralised commons-based social media alternative platform. While this option has been assessed as easier for certain communities (digitally-skilled academics and IT workers), it has been evaluated as less suitable for others such as the disabled, chronically sick, or members of the global South. Further motivation for turning to commons-based platforms as opposed to those of surveillance capitalism may come from world crises. The global pandemic and the current energy and environmental crisis raise the idea that more distributed platforms or communities could be more broadly able to address bigger challenges of the digital world.

Key challenges and questions

▪ **Challenge 9.4 — Legal sustainability.** The legal environment should at least not impede the development of commons-based alternatives to public or private digital media services, and ideally should support their legal needs with an appropriate legal framework. Regarding copyright, the public domain and usage rights should be protected from private interests, and creators empowered. In the field of telecommunications, which includes community networks as alternatives to big internet service providers, the legal framework and fees to access public infrastructure should also accommodate the emergence of small players with privacy-friendly data retention policies.

▪ **Challenge 9.5 — Economic sustainability.** Access to a diversity of funding models, including public-commons-private partnerships, can ensure that a platform will survive after the initial design phase. This is true for European projects, as well as for community-led efforts such as an agreement linking municipal public procurement with a private service provider, or for a mix of such sources of support. Governments should develop lifelong training schemes where contributions to a digital commons can be recognized and performed as part of one's job. Public investment to support innovation could be designed in a way to include small commons-based initiatives, and not only larger, competitive, commercially-sustainable projects. Grants could be repurposed to allow a longer duration of funding without having to develop more functionalities, and this way will support the continuation of a project as it is without overburdening human resources with the need to apply for further funding.

▪ **Challenge 9.6 — Social and environmental sustainability.** Commons are governed by a community, which decides on production and governance rules based on shared values. Access to social and environmental information should also be promoted, as should respect for workers' rights or the right to repair and recycle digital content and devices. The measurement of environmental impacts can also guide consumers choosing to use commons-based platforms for their services rather than the usual dominant private provider. More platforms are addressing social and environmental concerns, and also need institutional legal and economic support to survive and provide sustainable alternative digital services to informed citizen-consumers.

Topic 3. Access and reuse of data for scientific purposes

Keywords: text and data mining regulation, data protection, open access, impact assessment

Methods for text and data mining, scraping or machine-learning are becoming increasingly indispensable in many scientific disciplines. Unfortunately, such research is often constrained by legal barriers to the access and reuse data by scientists. There are exceptions, such as a recent text and data mining exception in the last European copyright directive revision which improved the legal framework for research purposes. But legal changes to allow such work in the scientific context, with appropriate safeguards, need to go further.

As stated by Communia, an association devoted to furthering public domain production and sharing of knowledge, "Public dissemination of research activities, including for purposes of verification of research results and for engaging in collaborative research, are essential to the scientific method. Europe needs a mandatory exception to copyright and other exclusive rights that facilitates all non-commercial research activities, including sharing of protected materials between researchers."

Key challenges and questions

▪ **Challenge 9.7 — Identify the scope of the scientific needs for open data commons.** There is great variation in the level of detail required by different research tasks, and the privacy risks such research may entail. While some research will not require access to an entire database including personal information, other research may preserve links to identifiable persons even after all data processing, mining, scrapping or machine-learning automated tasks have been performed. Hence, depending on the goals and methods used, reuse of data may be entirely safe, or could be potentially infringing some limited set of rights. We need guidelines and practices to help identify the kinds of risks which might be most relevant to different research protocols so safeguards can be adjusted for each particular case.

▪ **Challenge 9.8 — Consider data protection regulation.** Some usage of personal data can clearly be damaging for the privacy and security of vulnerable persons. But the usage of some data linked to individuals can also have a huge positive potential and present only limited risks, if appropriate protection measures were in place. Impact studies of risks, sector by sector, could distinguish those areas where data protection rights cannot be accommodated even for research purposes, and contrast them with other areas where a balance can be reached between private rights and the common and public interest.

▪ **Challenge 9.9 — Develop best practices, policies and platform recommendations.** Policies to govern data reuse can be designed and applied by governments, researchers and the industry to facilitate scientific access. This will require legal reform by governments, for instance, to establish exceptions and limitations to exclusive rights and circumvention of terms of use for research purposes. For academics, educational material and best practices guides should be produced to explain to each research communities how data can be reused without legal friction and what kind of research requires permission. These are time-consuming and uncertain legal processes beyond the capabilities of many research teams. To persuade actors to support the reuse of data for scientific purposes, impact assessment studies should demonstrate how academic reuse of data for research could benefit the economy and society at large, even if that data was originally produced by users for free.

Community

- Global Network of Centers for Internet and Society with 100+ academic centers and some

civil society centers leading R&D on digital commons
- Open Future Foundation policy think-tank
- Communia policy recommendations
- Open Knowledge Foundation and its network in 40+ countries
- ODECO – Towards a sustainable Open Data ECOsystem

10. Digital media in a sustainable world: towards inclusivity, equality and autonomy

Editor and corresponding author: Bertrand, Jouve.
Authors (alphabetical order): Nahla, Ben Amor; Bertrand, Jouve; Elisa Omodei.

Keywords: sustainability, autonomy, smartness, inclusive networked societies, equality.

Rationale

Digital technology is transforming all of our lives, as well as the nature of our societies, and has an impact on climate change in a multifaceted manner [19]. Addressing its role in environmental sustainability is a necessity that cannot be done without considering all the complexity of our globalized society. It is now recognized that the trajectory taken by our societies, in particular concerning technological innovation, must be modified but the nature of the transformations is still subject to debate. However, a society that is most likely to lead its own beneficial transformation will be a learning, inclusive, responsible and autonomous one. This may emerge not through opposition to innovation, but by encouraging it to preserve autonomy and to maintain reversibility as a feasible alternative. "Technological solutionism", an attitude which asserts that digital technologies will solve all problems worldwide, is sometimes applied to the world of digital media. By addressing the effects of the problems instead of their causes, however, technological solutionism embraces a philosophy of "no turning back" regardless of the danger and is a dangerous hypothesis for the future of mankind. In this context, we must address many issues concerning the relation between "smartness" and "sustainability".

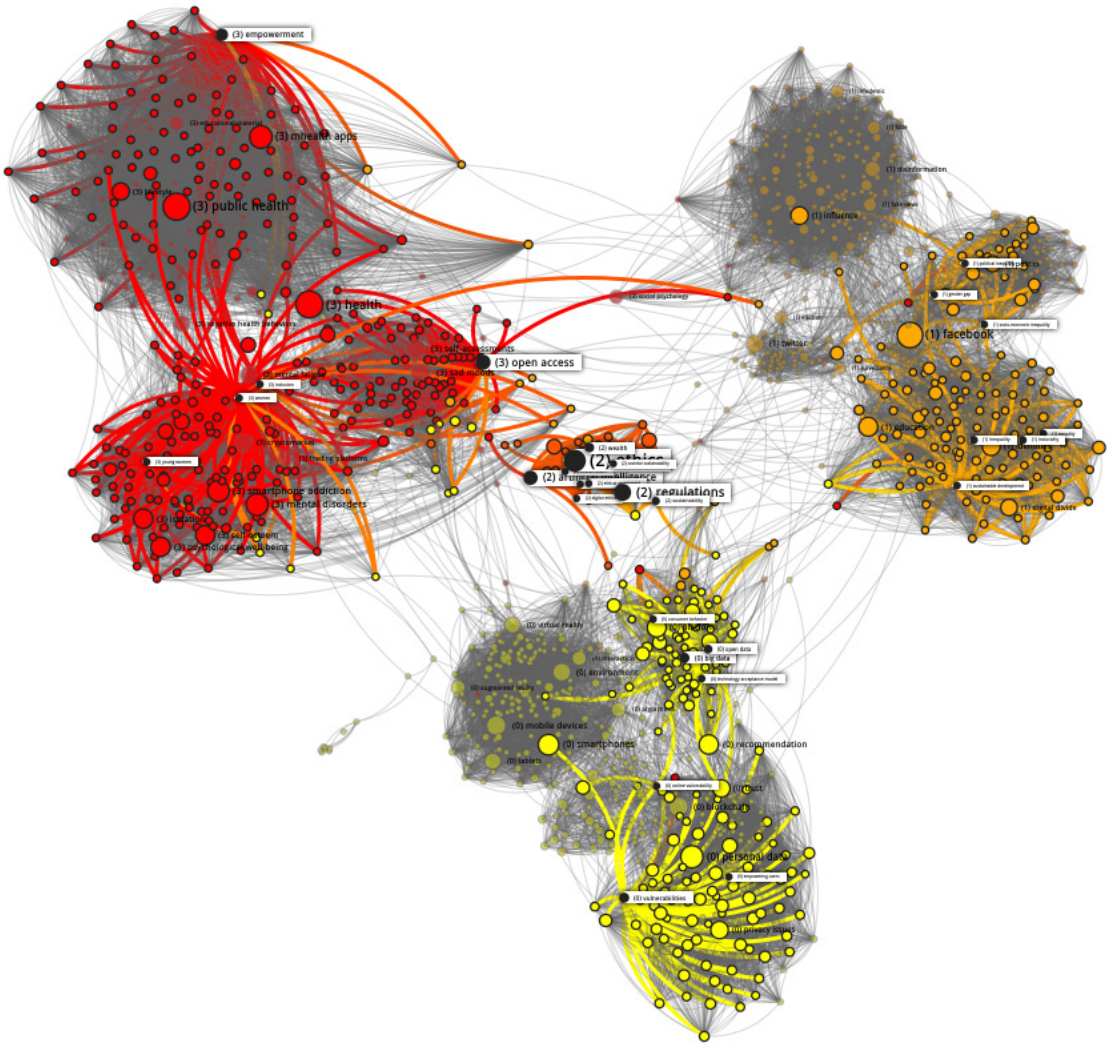


Figure 10.1: State-of-the-art domains concerned by this chapter.

Topic 1. Increasing inclusivity in the digital world

Keywords: inclusive networked societies, vulnerable people, technological acceleration

The information society we are building, founded on a world of global networks, is altering the fundamental categories we use to understand our world [8]. Within its complexity, few can comprehend the ultimate consequences of their participation in networks where synchronised actions interact across many scales. In the process, many people are effectively excluded and marginalised.

Key challenges and questions

▪ **Challenge 10.1 — Establishing the conditions required for citizen participation in a sustainable construction of our digital media.** These networked worlds make incredible volumes of information available to essentially all people. Yet the overwhelming volume of information also threatens to undermine the diversity of our knowledge. Ideally, available globalised information should serve to multiply viewpoints and benefit our "living together" by including ever more people and cultures [82, 83]. In this way, digital media could open the way to greater transparency for a more durable society. However, in filtering the information

each person receives, algorithms operate like a magnifying glass and may lead each to construct his/her own reality. It is crucial to understand if such a society can be sustainable. Is such a society capable of sharing and developing a collective project?

▪ **Challenge 10.2 — Learning how to control technological acceleration and establish the necessary conditions to guarantee technical reversibility at the individual and societal levels.** "Speeding up" is associated with modernity and anyone opposing it may be accused of "opposing progress" [96, 48]. Yet it is legitimate to consider the possible benefits of slowing down for the well-being of an individual living in an ever-accelerating society. Slowing down could be advantageous for our capacity to create knowledge and to prevent it from being delegated to machines. Once technologies have become established in a society, criticism of their development is often portrayed as resistance to progress or rejection of innovation [84]. It becomes difficult to discuss options for reversing or slowing negative trends. We need to understand how individuals can preserve their autonomy as well as technological reversibility as an open alternative.

Topic 2. Minimize tension between digital media and the environmental emergency

Keywords: environmental emergency, digital frugality, rebound effects.

The current digital environment is locked into economic models which naturally encourage further innovation, whether or not it truly leads to improved individual or social outcomes. But our digital world environmental footprint is critically increasing. Finding ways to develop our digital world differently, and with a smaller environmental burden, will require exploring novel economic models and changing our habits.

Key challenges and questions

▪ **Challenge 10.3 — Explore and test scenarios for transforming our digital habits and assessing their individual and social impacts.** The growing scale of digital media consumption places an increasing burden on the environment. Yet environmental awareness is also growing, both individually and collectively, creating an urgent tension. In this context, it is necessary to explore and test the most efficient scenarios which might allow some resolution of this opposition. Because habits are difficult to change, this will require an active search for pathways to reducing environmental impacts which can be deployed in the short term. Digital producers have expertise in engineering tools which change our habits, but what is now needed is changing our habits towards more efficient tools to reduce usage. Political regulation will also be required to bring about such changes. Can this be achieved despite the natural tension with the economic objectives of big international digital firms?

▪ **Challenge 10.4 — Devise models to predict rebound effects triggered by large-scale technological innovations.** Rebound effects are now widely recognized in digital technology: new services with higher speed and easier access encourage digital consumption and therefore reduce the possible energy efficiency gains. This issue is a key point that must be part of the process of development and promotion of technological innovations. We require efficient models which allow exploration of rebound effects at various scales with possible bifurcation and cascade effects.

Topic 3. Digital data for sustainable development

Keywords: big data, sustainable development, validity of computational models

Traditional data sources such as census and large-scale surveys are a fundamental source of information to monitor the 17 goals defined by the UN 2030 Agenda for Sustainable Development, and to design prompt interventions and long-term plans. However, these data collection methods are costly, both in monetary and human resources terms. During the last few years, researchers have begun exploring how large-scale digital data can be used as an efficient and low cost complementary source of information. These novel approaches may enhance both forecasting and the exploration of scenarios, but they should be used with caution, as they have some important limitations [9, 44].

Key challenges and questions

▪ **Challenge 10.5 — Data access and representing the most vulnerable.** Digital data is not being actively designed and collected to answer a specific research question, but is rather generated for other purposes, such as billing or online social exchanges. As a result, scenarios and models constructed from digital data often under-represent parts of the world, especially disadvantaged geopolitical and socioeconomic areas involving the most vulnerable people. The increasing use of digital data for building our future societies requires us to evaluate and overcome these biases. This requires more access to the technological black-box, but also to improve our modelling methods. We need to understand if there are ways to infer the activity patterns of those segments of the population which are underrepresented or not present at all in the available data. How can we combine different data sources to allow for a more comprehensive understanding of the phenomena under study? To what extent is it possible to use digital traces when no ground truth information or other data exist?

▪ **Challenge 10.6 — Spatio-temporal transferability of models.** Computational models are trained and tested on data for a limited geographical area and covering a defined time period. To what extent are they replicable and transferable to other geographical regions and time periods? Complexity science teaches that, in a fully connected world, we should expect nonlinear dynamics and emergent phenomena which are difficult to predict. Therefore complexity scientists should focus their research efforts on less explored vulnerable settings to further investigate these limitations in order to develop models that are both scalable and adaptable to different contexts.

Topic 4. Gender, emerging country and digital media

Keywords: Gender stereotypes, Equitable AI, Egalitarian digital media.

Media pluralism and diversity take on special importance as the border between the "digital" world and the "physical" world continues to blur. In particular, women remain under-represented and often discriminated against in professions of the digital technology sector, including Artificial Intelligence (AI). In a booming digital world, a key challenge is to combat the continued dominance of social and cultural stereotypes which have prevailed in the "physical" world. Even more, digital media offers an opportunity to spread awareness of important issues related to the struggle for gender equality. The success of the #MeToo movement is an outstanding example. These issues may have a particular resonance in emerging countries.

Key challenges and questions

▪ **Challenge 10.7 — Digital technology, a lever of egalitarian digital media.** While Information and Communications Technology (ICT) has the potential to promote gender equality and womens' empowerment, it suffers from a gender divide for two reasons. First, human users naturally reproduce in the digital world the gender relations characterising the familiar physical world. Second, AI algorithms themselves encode gender as well as ethnic and cultural stereotypes, as AI is what humans design it to be. AI systematically reflects the biases of the (mostly men) who build it, reinforcing stereotypes based on those biases. Reducing such bias will require answering a host of questions, such as: Is gender inequality in the "digital" world a simple transposition of that of the "physical" world? How can we break this gender-bias continuum which propagates from the physical to the digital world? What are the main causes of algorithmic bias? In combating gender stereotypes, can algorithms contribute to creating a more egalitarian digital media and society? How can we develop AI so as to ensure a more equitable future?

▪ **Challenge 10.8 — Digital media: a tool for women's empowerment in emerging countries.** Digital technologies are already acting as tools to help bridge the gender divide both in developed and emerging countries. ICT offers significant potential for women's economic empowerment through more education, including online courses or software-based literacy programs, or by building employment opportunities through telework, e-business or similar channels. As ICT professionals, women still face strong gender bias and stereotypes regarding their entrance into the labour market. Surprisingly, however, many emerging countries have a higher proportion of women in digital professions than do similar developed nations. What can we learn from this phenomenon? Might it offers lessons or opportunities to help encourage economic and social transformation all over the world? We should seek to understand the likely economic and social consequences of such emancipation of women in the digital world in emerging countries. Which obstacles might slow down this economic empowerment of women? In general, what can Europe do to encourage and empower women in ICT?

11. Indentification of potential risks related to the Future of digital media

Editor and corresponding author: Víctor Rodríguez Doncel
Authors (alphabetical order): David Chavalarias, Caroline Datchary, Dana Diminescu, Jean Lassègue, Quentin Lobbé, Víctor Rodríguez-Doncel, Fabrizio Silvestri

Keywords: trends in digital media; risk management; risk assessment; risk mitigation; long-term threats

Rationale

Previous chapters have considered the impact of current digital media on human well-being and outlined an associated research agenda. This chapter completes the picture with an attempt to identify effects of media that, as of today, have not yet been fully materialized, but are likely to appear. Trying to foresee these future effects using risk-analysis tools is of the utmost importance, for they will broadly impact media users and the society at large.

The methodologies to carry out this analysis should be object of research and they will be explored in **Topic 1** in this chapter. Although transformations are happening at an unprecedented pace, anticipation of likely future trajectories is possible at least for the short term future. The most notable of these risks relate to how groups form online and their associated social dynamics (**Topic 2**), or how social processes affect the public perception of risks (**Topic 3**). Finally, the most difficult risks to be evaluated and anticipated are related to global, long-term and systemic problems, overviewed in **Topic 4**.

We can only speculate about the future development of new media technologies. However, extrapolation of current trends suggests the near-term future will see more active social media users (reported to be more than 4.5 billion as of 2022), spending more time online (currently more than six hours per day) and pervading more professional activities beyond leisure [97]. Image and video based social neworks such as TikTok and Instagram, currently the fastest growing networks [93], may continue to gain momentum. Voice search and the interaction through

smart assistants may also continue growth. However, other changes are to be expected beyond the extrapolation of current trends. Major transformations could take place due to regulatory changes (e.g. the European Digital Identity Wallet), technological developments (new disruptive technologies such as ChatGPT may emerge) or even have a black-swan nature (e.g. remote education and working brought about by COVID-19).

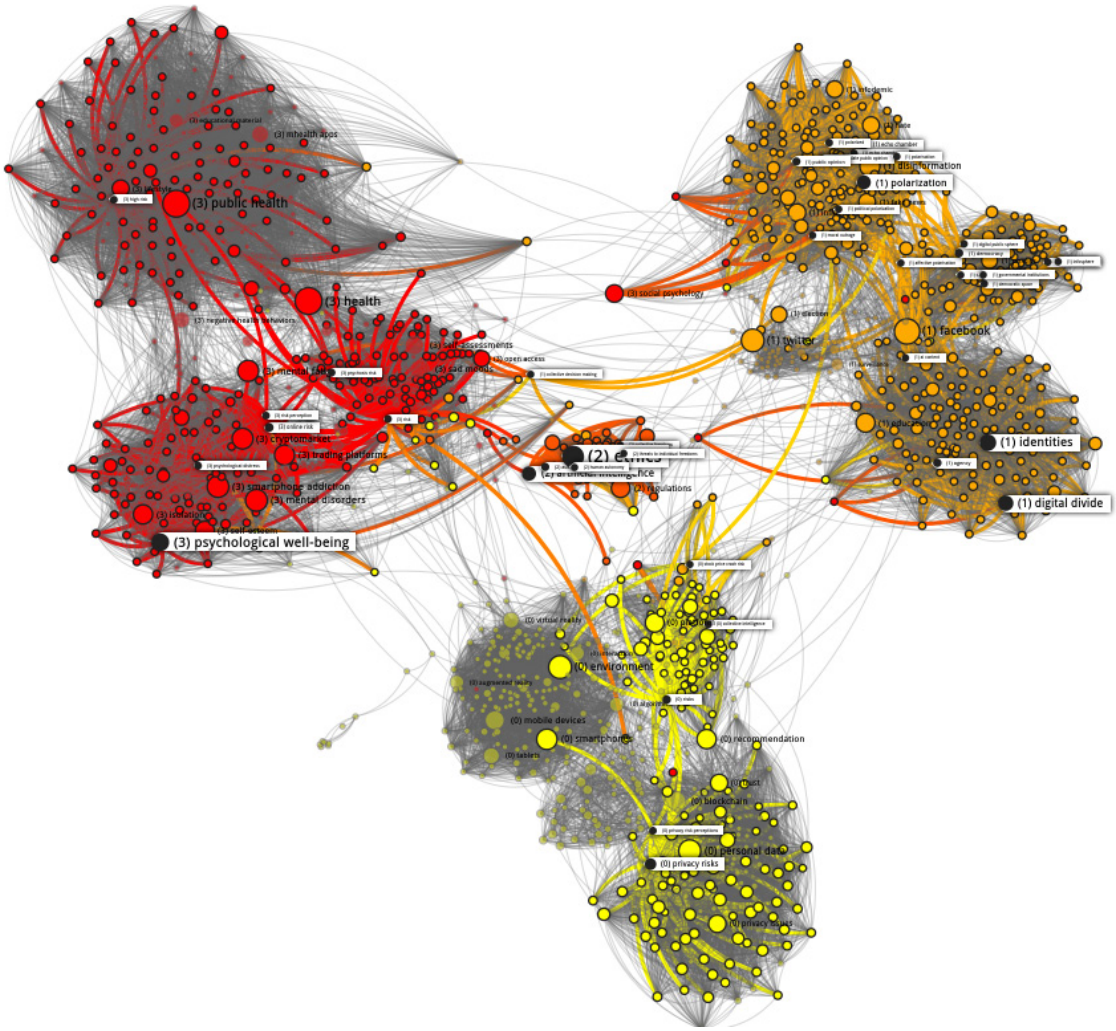


Figure 11.1: State-of-the-art domains concerned by this chapter.

Topic 1. Methodologies for risk assessment in future digital media

Keywords: foresight analysis; ethical foresight analysis; risk assessment; risk management; incident databases

There is a generally acknowledged need to assess the desirability of new and emerging technologies early in their development. It is better to anticipate developments when they are still malleable, even if at this stage the future is still uncertain. Foresight Analysis is a methodology used since the 1950s for predicting the outcome of potential policy decisions, emerging technologies and artefacts, as well as economic and societal trends. Methodologies for doing Ethical Foresight Analysis have also been proposed, with the goal of identifying and predicting the most

salient ethical issues likely to arise from new technological artefacts, services, and applications [25, 34, 87]. These techniques can also be used to anticipate potential risks in future digital media, but they need a place on their own in the research agenda.

Key challenges and questions

▪ **Challenge 11.1 — Foresight analysis methodologies.** Specific methods and techniques must be used to identify individuals and groups affected by future digital media developments, to anticipate and predict the risks affecting these individuals and groups and also to plan mitigation measures. Research questions include: How should these methodologies be adapted to the specific case of future media? How can we integrate both online and offline methodologies? How can we best anticipate unlikely technological, social or political developments? Are standard risk management procedures (such as those in ISO 31000) also applicable in this domain?

▪ **Challenge 11.2 — Datafication of digital media-related incidents.** In the last few years incidents related to AI systems have been systematically collected in databases. One such collection, the Artificial Intelligence Incident Database (AIID) [94], classifies incidents collected from the news with a taxonomy of seventeen harms or near-harms categorised by five severity levels. Equivalently, databases of harms related to digital media might be collected. These databases may contain not only news but also other resources such as biographical interviews, data related to self-confrontation psychology experiments, field knowledge and other kind of ethnographical research data. The availability of this connected network of datasets may lead to the creation of knowledge graphs and social digital twins [21], allowing statistical analysis of individuals, groups, media and their relations. Research questions are: Can such deeply qualified datasets be used to reliably infer potential risks? Is simulation of different scenarios (what-if) possible? Other research questions around datasets of incidents related to future digital media include: How can we categorise risks related to digital media? Which databases can be systematically collected for the rigorous study of risks? Which features should be reflected in the media-related incident databases?

Topic 2. Risks associated with group formation and dynamics

Keywords: echo chambers; filter bubble; polarization; human autonomy; collective decision making; digital divide; digital literacy.

Digital media change the way we make social ties and influence one another, profoundly altering the processes of social groups formation and the circulation of information within and between these groups. Backstrom et al. [88] identified the structural determinants in social networks that most influenced these processes in the early 2000s, but the significance of the media and the formed groups has changed since then, and a broader analysis is now necessary.

New areas of study such as the sociology of the internet or the digital anthropology still have to explore the transition from a group-based to a network-based society that is decoupling community and geographic proximity, and thus requiring new understanding [17]. Exclusion from groups how has more severe consequences, and new digital divides require more attention. Meanwhile, being too strongly influenced by a group leads to opinion polarisation, a likely effect of group-formation dynamics that locks media users into closed communities. The global reach of the new media endangers diversity of thoughts, freedom of expression and ultimately the health of collective decision making processes.

Key challenges and questions

▪ **Challenge 11.3 — Digital divides.** New media technologies create digital divides with disparities in access, usage and influence, these divides driven by a variety of factors. These include sociodemographic and socioeconomic characteristics, personal elements, types of technology, degree of digital training, rights, infrastructure, among others [35]. Whereas existing divides have been explored, future divides may emerge from differences in algorithmic awareness or data inequalities. Identifying these new divides is a key research challenge, as is learning to identify the groups and collectives that will be most affected, and the likely consequences of group marginalisation. Other questions might be related to the mitigation measures: how can digital media literacy be improved? How can we best teach such literacy? What are the conditions for upskilling segments of the population that are currently most in need of education (e.g. seniors)?

▪ **Challenge 11.4 — Echo chambers and polarisation.** The Web 2.0 (social web) era has reinforced phenomena such as echo chambers and filter bubbles [14]. Individuals in echo chambers tend to consume and spread information aligned with their pre-existing beliefs. Yet algorithms in social media also favour equally aligned content, entering into vicious circles that lead to the polarisation of public opinion, with black-or-white opinions and beliefs. The possible perspectives and horizons of individual thought become limited [61]. Whereas this is a current issue, we need research on the long- term evolution of such dynamics. What are the effects of less randomness in encounters? How to prevent digital media bubble groups to evolve into digital ghettos?

▪ **Challenge 11.5 — Collective decision making.** New media exert a strong influence on collective decision making. The following questions remains open: Is there a way to maintain democratic control amidst ever-growing influential algorithms? Is it possible to avoid the creation of new digital feudalities in which platforms or individuals exert undue influence? Is there a way to preserve identification continuity in the process of collective action under the deployment of new technologies?

Topic 3. Social construction of risks in public opinion

Keywords: social construction; public opinion; hyper-reality.

The concepts populating the minds of the individuals of a society can be created or modified by social engineers through media technologies. Intersubjectivities can increasingly detach the collective imagination from the actual world. Something to explore is how this divergence between the real and the conceptualized becomes larger with the new media technologies, amplifying effects on human well-being. Digital media empowers agenda setting, or the ability to shape public opinion by determining which issues are given the most attention. It can be used in nefarious ways. Examples are the artificial construction of irrational fears in order to establish controversial and questionable policies, and the actions to foster moral panic or bellicose attitudes in international crisis situations.

Key challenges and questions

▪ **Challenge 11.6 — Moral panic.** The widespread feeling of irrational fear of a person, group or idea can be artificially triggered in online social networks. Online behaviour – violent attitudes, for example – can also spill over into the off-line world, and in this way, the shape of the public sphere can be manipulated. We have little fundamental understanding of the nature of the relevant dynamics, their relation to the topology of social networks; and whether moral

panic appears spontaneously or is induced by interested parties.

▪ **Challenge 11.7 — Human autonomy.** The application of artificial intelligence and machine learning algorithms to deliver personalized advertisements to individuals based on their digital footprint can become a form of manipulation. It may provoke threats to human autonomy, changing the course of elections or immersing individuals into echo chambers. The future consequences of increasingly powerful algorithms remains an open research area. How will AI-mediated communications influence human autonomy?

▪ **Challenge 11.8 — Erosion of individual identity.** For most of human history, individuals could take for granted the public validity of their individual identity. But emerging deepfake algorithms, AI-generated text and other AI techniques will democratise technologies for computers to impersonate humans or for people to impersonate others. The value of public identity could be partially lost. We must identify the immediate risks of impersonation, as well as their enabling technologies [90]. On the other hand, the implementation of measures to certify identity may lead to an excess of control threating again individual freedom. Finding a balance between these competing risks (for example in the metaverse) remains an open challenge. What are the legal and technological instruments to fight impersonation? How to empower individuals to control their own virtual identities and safeguard their biometric data?

Topic 4. Systemic, global and long-term risks

Keywords: systemic risks; global risks; long-term risks; digital preservation.

The global character of digital media, the lack of technodiversity, and the possibility of a doomsday invention [31] amplified by new digital media magnify the impact of global, systemic and long-term risks. The existential dependence of modern societies on electrical energy has been extended to telecommunications – a sustained internet failure would lead to catastrophic consequences. Whereas critical infrastructures will not depend on social media, the malfunction of the latter or their malicious operation may in the future lead to equally devastating results on society. The preservation of digital goods might be at risk, and so are democracy and public institutions.

Key challenges and questions

▪ **Challenge 11.9 — Digital preservation.** Proprietary social platforms are not open and are therefore beyond the scope of Web archive initiatives. Also, if efforts to decentralize the Web do not succeed, information archives might be in the hands of few entities, and so come to represent single points of failure. Web archives might in these cases be open to manipulation, biased or of limited diversity (socially, politically, geographically, etc.). A key challenge is to determine the means for granting web historians the digital preservation of the collective digital heritage. Should political, technological or legal instruments be deployed to grant the preservation of digital goods?

▪ **Challenge 11.10 — Democracy and public institutions.** Many challenges can be explored in relation to reinforcing democratic values and strengthening public institutions. These research avenues include methods to mitigate the risks that future digital media pose to democracy through generating unrest, instabilities, polarisation, or changing the relation between societies and their environment. Also, determining the extent of public intervention in the new digital media remains an open challenge, as well as determining the nature and extent of the control instruments (legal, technological, organisational).

▪ **Challenge 11.11 — Psychological well-being.** Psychological well-being is largely determined by self-perception. But self-perception has been dramatically reshaped in the online world, where different digital identities are created and evolve under the influence of external agents. A key challenge is to identify means to decelerate and to manage personal continuity and autonomy in ephemeral groups. How will future media technologies change our self-conception, our conception of reality and our interactions with reality? What is the precise extent of the onlife [76] in psychological well-being?

▪ **Challenge 11.12 — AI generated content.** Large Language Models and new image generation algorithms such as stable diffusion enable the generation of high-quality text and images. This AI generated content, possibly used for behavioural change purposes, may dominate the new media landscape and have incalculable consequences. This risk is particularly acute, if algorithms are iteratively refined – ChatGPT has been trained to improve by interacting with users by means of Reinforcement Learning techniques. An infosphere populated by interacting automata (inforgs of artificial nature) will irremediably imply the redefinition of what we are, repositioning the place of humans before the world [75] and creating a new sort of risks and opportunities that have to be researched.

12. Conclusions and recommendations

*"Science sans conscience n'est que ruine de l'âme"**
Technology without science is but ruin of the Humanity.
* "Science without conscience is but ruin of the soul", Rabelais [85]

The eleven chapters of this roadmap have identified more than one hundred research directions that should be supported to promote the development and use of digital media compatible with humanity well-being (see interactive map). We give below a summary which, without being exhaustive, highlights some general recommendations.

"The paradox of the human condition is that we can only become ourselves through the influence of others"[1]. The corollary of this claim, which has similar formulations by prominent researchers from sociology to psychology and neurosciences, is that a technology capable of mediating your interactions with your fellow citizens controls to some extent the development of your personality. A technology capable of mediating the interactions between the citizens of a country controls to some extent their collective behaviors, the culture they develop, and ultimately the democratic process itself.

This is what makes digital media a technology like no other. Their deployment at the scale of humanity (*e.g.*, >2.8 billion users for Facebook ; a penetration rate of social networking sites worldwide estimated at 70%), allows them to capture and organize an ever-increasing proportion of our relationships with others modulating, through their design and their recommendation

[1]Boris Cyrulnik, 1993 [72]

algorithms, the influence between each pair of citizen. As such, digital media touch the heart of our societies by acting as an interface between the citizens on the one hand and the collectives they form on the other, and ultimately the institutions. The consequences, both positive and negative, are innumerable both at the individual and the collective level.

The main conclusion of this roadmap is that the impact of digital media and their future developments on individuals and the collectives they form have not been sufficiently studied or anticipated from a scientific and technology point of view. Their contribution to the well-being of humanity has not been sufficiently explored and their damage not sufficiently anticipated (ex. Ch. 11). The very design of digital media platforms and the laws that regulate their use, especially for social networks, are still in their infancy and require interdisciplinary fundamental and applied research as well as systemic actions so that they can truly deliver their full potential without harming the individuals and the democracies. Social media have a systemic effect on societies, they thus should be addressed in a systemic ways, with actions at the individual, collective and institutional levels.

Here are some general recommendations steeming from this roadmap.

▪ **Recommandation 12.1 — Promote Digital Media and Information Literacy (DMIL - cf. Ch. 4)** Our human activities are increasingly embedded in digital environments which influence all aspects of modern life, including education, politics, work and leisure.

A growing research literature emphasizes the need for greater awareness of the impacts technology is having on an individuals' mental health, cognitive capacities, socialization and identity formation. This awareness should be developed as early as possible during scholarship. Vulnerable users and especially young people merit extra attention in learning to live safely in a digital world, managing screen time as well as their representations of self and others, and also facing other attendant risks to well-being.

Poor digital media and information literacy is also a factor of fake news belief and learners should be taught to develop fact-checking expertise and resist anxieties and fears caused by excessive algorithmic surveillance, systemic doubt and distrust about democratic institutions.

Actions. *Engage the entire educational institution community in tackling Digital Media and Information Literacy challenges and make both teachers and students more digitally literate. Better integrate DMIL in curriculum and make DMIL competences responsive and agile so as to equip learners with the mindset they will need for lifelong digital literacy updates. Make specific efforts to reach the most underrepresented and underserved communities such as low income people, migrants, rural area dwellers.*

▪ **Recommandation 12.2 — We must better understand the long-term impacts of ubiquitous digital media on individuals and their cognitive capacities (Ch. 2, Ch. 1).** How do digital media impact the formation of individual identities, the psyche and the construction of the self? How do they impact cognitive phenomena such as attention span or reduced judgmental capacity? How do they affect our behaviours and our moods? Our knowledge of the role of emotions in making choices is still very poor, and requires considerable further investigation. Does the online overexposure to emotional based contents such as images and videos change our decision making processes online and/or offline?

Actions. *We must promote both fundamental research and experimental/empirical research on the nature of cognitive processes and the disruptions/enhancement they undergo when a person is exposed to digital media. There is an urgent need for research in psychology, neurology and more generally in cognitive sciences to both rigorously define cognitive phenomena such as "attention" from a theoretical point of view, and then large longitudinal studies to measure how digital media*

interfere with these cognitive processes. Clinical research is also needed in order to identify vulnerable population and their specific risks and needs. New experimental protocols have to be invented to study these phenomena at scale and "in vivo".

▪ **Recommandation 12.3 — We must better understand the psychological and sociological mechanisms involved in the new forms of online socialization; and how they interfer with the formation of beliefs and opinions (Ch. 1, Ch. 7).** Social media are often the main source of information for a large part of the population and might also become one of the main source of social ties. What are the differences between online and offline mecanisms of socialization and information propagation? What kind of biases might be amplified or mitigated in these new environments and what are their effects at the collective level? Why the scaling effect of online media often lead to a situation where the truthfulness of a fact is less important than the interest and reactions that it generates? In particular, there is a need to understand the processes and mechanisms behind the creation of fake news, and its spread and virality as well as associated phenomenon such as deep fakes and echo chambers.

Actions. *We must promote interdisciplinary research at the interface between sociology, psychology, computational social sciences, complex systems sciences and data sciences to better characterize the specificity of online socialization and help formulate new guidelines and practices for technology deployment and for its regulation. In particular, traditional sociological surveys and online studies on behaviors and opinions should be conducted on the same group of people in order to be able to compare online socialization with offline socialization. Models of opinion dynamics and social networks formation should be developed and calibrated on these data to anticipate the long term effect of different implementation of recommender systems and social networking sites. This research is needed to design evidence-based regulation of digital media.*

▪ **Recommandation 12.4 — Mitigate the systemic effects of recommender systems, micro-targeting and their associated risks ; develop the legal tools to better regulate them (Ch. 1, Ch. 7, Ch. 9, Ch. 11).** Recommender systems or micro-targeting advertising are partially responsible for the toxification of our digital environments, spread of misinformation and social decohesion by mean of polarization and radicalization. Which technological measures (e.g. changes in recommender systems or advertising platforms) and what regulatory measures or policy interventions can help?

Actions. *Research that studies the systemic effect of the algorithmic layers that constrain our digital worlds should be supported, and steps should be taken to mitigate the associated risks. For example, there is a need to better understand the extent to which certain types of input data could be misused by self-learning recommender systems in order to eventually regulate their use.*
A better scientific understanding of the root causes of online hate speech and polarization with, among others, approaches from communication sciences, sociology, cognitive science and complex systems sciences should also be developped. It is important to develop collaborations with the tech industry, in order to find viable solutions.

▪ **Recommandation 12.5 — Promote access to data and develop digital observatories (cf. Ch. 1, 7, Ch. 9 & 10).** Access to data is one of the crucial and yet sensitive issues. The general public and government should become aware of the necessity for scientists to access data presently owned by the tech industry and stored in sealed silos. Such data would provide valuable insights on actual user behaviour for large and difficult-to-access populations. For example, the ability to have real-time access to aggregated data on collective behaviors, such as road traffic, rumor dissemination, or research on health issues, has proven to be of great value for decision-making and improving the daily lives of citizens.

The current system of hoarding data needs to change because society as a whole is being translated into databases. Opening the data produced by these companies is one of the key factors to free the full potential of digital media for humanity, enabling both research and application. This could be made without disclosing personal data since in most of these applications, aggregated data are enough. It is important to identify and recommend policies to guarantee better access to data, public or private, and to reconcile this with the privacy of data debate.

Actions. *1) Encourage companies to increase accessibility to data resulting from the use of their products and to be more transparent about the impact of their products. 2) Demand greater access to the technological black box of BigTech companies. 3) Facilitate data donation allowing users to share with third parties, and in particular academia, data about the content they are exposed to, and in particular recommended content; regulate obstruction to this sharing by platforms 4) Develop new legal frameworks, like digital commons, to bring any large aggregate social data set into the public domain.*

▪ **Recommandation 12.6 — Macroscope for a reflexive society immune to misinformation (Ch. 1)** Fighting against online disinformation and protecting online social networks users from manipulations require to understand the informational contexts in which they are influenced. In particular, misinformation campaigns are most effective when the contexts in which information is produced and the extent to which public opinion supports certain ideas are unclear.

As a reminder, a very common form of opinion manipulation consists in letting people believe that a problem is of concern to a large part of the population, when it is only the whim of a small group. Where does a rumor come from? We usually cannot know, because on networks, notions of space and borders are no longer valid. However, networks are not without a certain spatiality; it is simply invisible to a to a simple user. This "crisis of space" [27] is partly remedied by making an external point of view available to the public.

Actions. *4) Invest in the development of public digital macroscopes and observatories to provide citizens and policymakers with a reflexive insight on social dynamics from these digital platforms data. These macroscopes and observatories should help citizens to better contextualize the information and narratives they are exposed to. This would also help governments to take better-informed decisions thanks to a better anticipation of social dynamics.*

▪ **Recommandation 12.7 — Promote research and infrastructures for online trust and privacy (cf. Ch. 5 & Ch. 6 & Ch. 7).** How can we best deal with the critical issue of "trust", where an interdisciplinary view is most appropriate? How can we establish online trust and how is this different from offline trust? Digital trust depends on the result of experiences lived repeatedly in several digital spaces. Thus, the question of trust is linked to the question of how to make the digital experience "healthy" in any digital places, or in other words, how to make the digital space a "healthy space" whatever the place of experience. It is consequently closely related with the notion of online identities.

Actions. *Research on infrastructure that effectively protects privacy and promotes trust by design should be promoted. Research on new notions of identity and privacy should be developped in order to foster positive interactions and hinder the proliferation of online hate speech in current and future evolution of social media (e.g. metaverse). In particular, strong privacy policies and practices, such as encrypting personal data and limiting access, upon user's consent, to only those who need it should be more systematically implemented. This requires to develop research on privacy-enhancing technologies (PETs) such as zero-knowledge proofs and homomorphic encryption.*
Study of the mutual effects of on- and off-line behaviour should be conducted, to establish links

between dangerous behaviours offline and online.

▪ Recommandation 12.8 — Anticipate the future of social media (cf. Ch. 6 & Ch. 11). What is the future of social media – how will it look like in 10 years? Our societies may pass from a language base to a visual base and on to a virtual (multi-sensory) base. Will the impact of social media and social media content on individuals in a virtual environment be increasingly powerful? How will this influence the human brain? How will the interaction with AI, VR and AR modulated by AI/Machine Learning modify our everyday behaviours? If social media is driving the replacement of words with images, what does this mean for neuronal activities, brain physiology, and cognitive psychology? On the other hand, what will become of social media in a world where a very significant proportion of accounts will be operated by AIs indistinguishable from humans? What will happen when social interactions are mostly mediated by conversational agents that reduce the complexity of our social environment?

Actions. *Interdisciplinary research should be conducted on the impact of new forms of social interaction mediated by the virtual world and AI. A new generation of digital media should be developed that ensures that users can distinguish an AI from a human user in their online interactions. Research in psychology and neurology must be conducted on the effects on individuals of the large-scale deployment of immersive technologies such as the metaverse.*

▪ Recommandation 12.9 — Promote new data protocols for a free digital media ecosystem. Digital media are worlds of networks and reputation where Internet users can become captive users of a platform for fear of losing hard-won social capital over the years. From the perspective of economic actors, positive externalities and path dependencies exacerbate phenomena such as first-mover dominance and winner-take-all. In such environments, competition cannot play its role in selecting the best option for consumers and could lead to sub-optimal if not dangerous environments for individuals and collectives. In the same way that portability of cell phone numbers have once to be legally enforced, there are some issues about the portability of social media data, both with regard to content and to user to user connections.

Actions. *Europe should promote the research and legal implementation of protocols that would allow a user of a given social network to switch to another network while keeping the history of his or her posts and, to a large extent, social connections and social capital.*

Bibliography

Articles

[1] Mohammad Amir Anwar and Mark Graham. "The digital continent: placing Africa in planetary networks of work". In: (2022) (cited on page 82).

[2] Markus Appel. "Are heavy users of computer games and social media more computer literate?" In: *Computers & Education* 59.4 (2012), pages 1339–1349 (cited on pages 36, 40).

[3] Harald Atmanspacher. "Quantum approaches to brain and mind: an overview with representative examples". In: *The Blackwell companion to consciousness* (2017), pages 298–313 (cited on page 45).

[4] Albert Bandura. "Toward a psychology of human agency". In: *Perspectives on psychological science* 1.2 (2006), pages 164–180 (cited on pages 42, 43, 46).

[5] Katharina-Maria Behr et al. "Some practical considerations of ethical issues in VR research". In: *Presence* 14.6 (2005), pages 668–676 (cited on pages 66, 67).

[6] Manabendra Nath Bera et al. "Randomness in quantum mechanics: philosophy, physics and technology". In: *Reports on Progress in Physics* 80.12 (2017), page 124001 (cited on page 45).

[7] Mikołaj Biesaga, Szymon Talaga, and Andrzej Nowak. "The effect of context and individual differences in human-generated randomness". In: *Cognitive Science* 45.12 (2021), e13072 (cited on page 46).

[8] Manuel Castells. "The Rise of the network society Volume 1 With a new preface". In: *Book The Rise of the network society, Volume 1, With a new preface* (2009) (cited on page 91).

[9] Nina Cesare et al. "Promises and pitfalls of using digital traces for demographic research". In: *Demography* 55.5 (2018), pages 1979–1999 (cited on page 93).

[10] David Chavalarias and Jean-Philippe Cointet. "Bottom-up scientific field detection for dynamical and hierarchical science mapping, methodology and case study". In: *Scientometrics* 75.1 (Apr. 2008). 00030 bibtex: chavalariasBottomup2008, pages 37–50. ISSN: 0138-9130, 1588-2861. DOI: 10.1007/s11192-007-1825-6. URL: http://link.springer.com/10.1007/s11192-007-1825-6 (visited on 08/29/2015) (cited on page 20).

[11] David Chavalarias and Jean-Philippe Cointet. "Phylomemetic patterns in science evolution—the rise and fall of scientific fields". In: *PloS one* 8.2 (2013). 00000 bibtex: chavalariasPhylomemetic2013, e54847. URL: http://dx.plos.org/10.1371/journal.pone.0054847 (visited on 08/15/2015) (cited on page 20).

[12] David Chavalarias, Quentin Lobbé, and Alexandre Delanoë. "Draw me Science: Multi-level and multi-scale reconstruction of knowledge dynamics with phylomemies". en. In: *Scientometrics* (Nov. 2021). ISSN: 0138-9130, 1588-2861. DOI: 10.1007/s11192-021-04186-5. URL: https://link.springer.com/10.1007/s11192-021-04186-5 (visited on 11/24/2021) (cited on pages 20, 21, 23).

[13] David Chavalarias et al. "Les nouveaux fronts du dénialisme et du climato-scepticisme". fr. In: *HAL - https://hal.science/hal-03986798* (Feb. 2023). URL: https://hal.science/hal-03986798 (visited on 02/21/2023) (cited on page 33).

[14] Matteo Cinelli et al. "The echo chamber effect on social media". In: *Proceedings of the National Academy of Sciences* 118.9 (2021), e2023301118 (cited on page 98).

[15] Jean-Philippe Cointet and David Chavalarias. "Multi-level Science mapping with asymmetric co-occurrence analysis: Methodology and case study". In: *Networks and Heterogeneous Media* (2008). bibtex: chavalariasMultilevel2008, pages 267–276 (cited on page 20).

[16] Edward L Deci and Richard M Ryan. "Self-determination theory." In: (2012) (cited on pages 42, 44).

[17] Paul DiMaggio et al. "Social implications of the Internet". In: *Annual review of sociology* 27.1 (2001), pages 307–336 (cited on page 97).

[18] Matthew J Donald. "Quantum theory and the brain". In: *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 427.1872 (1990), pages 43–93 (cited on page 45).

[19] Yogesh K Dwivedi et al. "Climate change and COP26: Are digital technologies and information management part of the problem or the solution? An editorial reflection and call to action". In: *International Journal of Information Management* 63 (2022), page 102456 (cited on page 90).

[20] Dean Eckles. "Algorithmic transparency and assessing effects of algorithmic ranking". In: (2022) (cited on page 75).

[21] Abdulmotaleb El Saddik. "Digital twins: The convergence of multimedia technologies". In: *IEEE multimedia* 25.2 (2018), pages 87–92 (cited on page 97).

[22] "European Commission. Annexes to the proposal for a regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts". In: (2021) (cited on page 72).

[23] Christopher J Ferguson et al. "Like this meta-analysis: Screen media and mental health." In: *Professional Psychology: Research and Practice* 53.2 (2022), page 205 (cited on page 36).

[24] Luciano Floridi and Josh Cowls. "A unified framework of five principles for AI in society". In: *Machine learning and the city: Applications in architecture and urban design* (2022), pages 535–545 (cited on page 42).

[25] Luciano Floridi and Andrew Strait. "Ethical foresight analysis: what it is and why it is needed?" In: *The 2020 Yearbook of the Digital Ethics Lab* (2021), pages 173–194 (cited on page 97).

[26] Carl Benedikt Frey and Michael A Osborne. "The future of employment: How susceptible are jobs to computerisation?" In: *Technological forecasting and social change* 114 (2017), pages 254–280 (cited on page 78).

[27] Antoine Garapon and Jean Lassègue. "Le numérique contre le politique: crise de l'espace et reconfiguration des médiations sociales". In: *(No Title)* (2021) (cited on page 104).

[28] Mark Graham, Isis Hjorth, and Vili Lehdonvirta. "Digital labour and development: impacts of global digital labour platforms and the gig economy on worker livelihoods". In: *Transfer: European review of labour and research* 23.2 (2017), pages 135–162 (cited on page 82).

[29] Mathieu Jacomy et al. "ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software". en. In: *PLOS ONE* 9.6 (June 2014), e98679. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0098679. URL: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0098679 (visited on 06/24/2018) (cited on page 22).

[30] Christian Matthias Kerskens and David López Pérez. "Experimental indications of non-classical brain functions". In: *Journal of Physics Communications* 6.10 (2022), page 105001 (cited on page 45).

[31] Raffi Khatchadourian. "The doomsday invention". In: *The New Yorker* 23 (2015) (cited on page 99).

[32] Jonathan Liebers et al. "Identifying Users by Their Hand Tracking Data in Augmented and Virtual Reality". In: *International Journal of Human–Computer Interaction* (2022), pages 1–16 (cited on page 64).

[33] Quentin Lobbé, Alexandre Delanoë, and David Chavalarias. "Exploring, browsing and interacting with multi-level and multi-scale dynamics of knowledge". en. In: *Information Visualization* (Sept. 2021). Publisher: SAGE Publications, page 14738716211044829. ISSN: 1473-8716. DOI: 10.1177/14738716211044829. URL: https://doi.org/10.1177/14738716211044829 (visited on 09/24/2021) (cited on pages 20, 23).

[34] Federica Lucivero, Tsjalling Swierstra, and Marianne Boenink. "Assessing expectations: towards a toolbox for an ethics of emerging technologies". In: *NanoEthics* 5 (2011), pages 129–141 (cited on page 97).

[35] Sophie Lythreatis, Sanjay Kumar Singh, and Abdul-Nasser El-Kassar. "The digital divide: A review and future research agenda". In: *Technological Forecasting and Social Change* 175 (2022), page 121359 (cited on page 98).

[36] Michael Madary and Thomas K Metzinger. "Real virtuality: A code of ethical conduct. Recommendations for good scientific practice and the consumers of VR-technology". In: *Frontiers in Robotics and AI* (2016), page 3 (cited on page 67).

[37] Aniek F Markus, Jan A Kors, and Peter R Rijnbeek. "The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies". In: *Journal of Biomedical Informatics* 113 (2021), page 103655 (cited on page 73).

[38] Milagros Miceli and Julian Posada. "The Data-Production Dispositif". In: *arXiv preprint arXiv:2205.11963* (2022) (cited on page 82).

[39] Mark Roman Miller et al. "Personal identifiability of user tracking data during observation of 360-degree VR video". In: *Scientific Reports* 10.1 (2020), pages 1–10 (cited on page 64).

[40] Philippe Mongeon and Adele Paul-Hus. "The Journal Coverage of Web of Science and Scopus: a Comparative Analysis". In: *Scientometrics* 106.1 (Jan. 2016). arXiv:1511.08096 [cs], pages 213–228. ISSN: 0138-9130, 1588-2861. DOI: 10.1007/s11192-015-1765-5. URL: http://arxiv.org/abs/1511.08096 (visited on 03/24/2023) (cited on page 20).

[41] Shunsuke Nonaka, Tomoya Takeda, and Motohiro Sakai. "Who are hikikomori? Demographic and clinical features of hikikomori (prolonged social withdrawal): A systematic review". In: *Australian & New Zealand Journal of Psychiatry* 56.12 (2022), pages 1542–1554 (cited on page 68).

[42] Andrzej Nowak, Paul Lukowicz, and Pawel Horodecki. "Assessing artificial intelligence for humanity: Will AI be the our biggest ever advance? Or the biggest threat [Opinion]". In: *IEEE Technology and Society Magazine* 37.4 (2018), pages 26–34 (cited on page 44).

[43] Candice L Odgers and Michaeline R Jensen. "Annual research review: Adolescent mental health in the digital age: Facts, fears, and future directions". In: *Journal of Child Psychology and Psychiatry* 61.3 (2020), pages 336–348 (cited on page 36).

[44] Elisa Omodei et al. "Complex systems for the most vulnerable". In: *Journal of Physics: Complexity* 3.2 (2022), page 021001 (cited on page 93).

[45] Amy Orben and Andrew K Przybylski. "The association between adolescent well-being and digital technology use". In: *Nature human behaviour* 3.2 (2019), pages 173–182 (cited on page 36).

[46] Luiz Pessoa. "On the relationship between emotion and cognition". In: *Nature reviews neuroscience* 9.2 (2008), pages 148–158 (cited on page 35).

[47] Carina Prunkl. "Human autonomy in the age of artificial intelligence". In: *Nature Machine Intelligence* 4.2 (2022), pages 99–101 (cited on page 42).

[48] Hartmut Rosa. "Accélération. Une critique sociale du temps". In: *Lectures, les livres* (2010) (cited on page 92).

[49] Richard M Ryan and Edward L Deci. "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being." In: *American psychologist* 55.1 (2000), page 68 (cited on page 44).

[50] Richard M Ryan and Edward L Deci. "Self-regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will?" In: *Journal of personality* 74.6 (2006), pages 1557–1586 (cited on pages 42, 44).

[51] Juliet B Schor et al. "Dependence and precarity in the platform economy". In: *Theory and Society* 49 (2020), pages 833–861 (cited on page 78).

[52] Dean Keith Simonton. "Creativity as a Constrained Stochastic Process." In: (2004) (cited on page 45).

[53] Mel Slater and Maria V Sanchez-Vives. "Enhancing our lives with immersive virtual reality". In: *Frontiers in Robotics and AI* 3 (2016), page 74 (cited on page 64).

[54] Mel Slater and Maria V Sanchez-Vives. "Is consciousness first in virtual reality?" In: *Frontiers in Psychology* 13 (2022) (cited on page 66).

[55] Mel Slater et al. "The ethics of realism in virtual and augmented reality". In: *Frontiers in Virtual Reality* 1 (2020), page 1 (cited on page 67).

[56] Marco Tamietto and Beatrice De Gelder. "Neural bases of the non-conscious perception of emotional signals". In: *Nature Reviews Neuroscience* 11.10 (2010), pages 697–709 (cited on page 35).

[57] Astra Taylor. "The Automation Charade". In: *Logic Magazine* 5 (2018) (cited on page 78).

[58] Moshe Tennenholtz and Oren Kurland. "Rethinking search engines and recommendation systems: a game theoretic perspective". In: *Communications of the ACM* 62.12 (2019), pages 66–75 (cited on page 46).

[59] Marcos Treviso et al. "Efficient methods for natural language processing: a survey". In: *arXiv preprint arXiv:2209.00099* (2022) (cited on page 80).

[60] Paola Tubaro. "Learners in the loop: hidden human skills in machine intelligence". In: *Learners in the loop: hidden human skills in machine intelligence* (2022), pages 110–129 (cited on pages 80, 82).

[61] Joshua A Tucker et al. "Social media, political polarization, and political disinformation: A review of the scientific literature". In: *Political polarization, and political disinformation: a review of the scientific literature (March 19, 2018)* (2018) (cited on page 98).

[62] Patti M Valkenburg, Adrian Meier, and Ine Beyens. "Social media use and its impact on adolescent mental health: An umbrella review of the evidence". In: *Current opinion in psychology* 44 (2022), pages 58–68 (cited on page 36).

[63] Jay J Van Bavel and Andrea Pereira. "The partisan brain: An identity-based model of political belief". In: *Trends in cognitive sciences* 22.3 (2018), pages 213–224 (cited on page 35).

[64] Rachel E Warsaw et al. "Mobile technology use and its association with executive functioning in healthy young adults: A systematic review". In: *Frontiers in Psychology* 12 (2021), page 643542 (cited on page 40).

[65] Guy Williams. "Company Guidance: Worker Voice and Representation". In: *ShareAction* (2022) (cited on page 79).

[66] Zelda B Zabinsky and Zelda B Zabinsky. "Pure random search and pure adaptive search". In: *Stochastic Adaptive Search for Global Optimization* (2003), pages 25–54 (cited on page 46).

Books

[67] Anu Bradford. *The Brussels effect: How the European Union rules the world*. Oxford University Press, USA, 2020 (cited on page 72).

[68] Antonio A Casilli. *En attendant les robots-Enquête sur le travail du clic*. Média Diffusion, 2019 (cited on page 82).

[69] David J Chalmers. *Reality+: Virtual worlds and the problems of philosophy*. Penguin UK, 2022 (cited on page 66).

[70] David Chavalarias. *Toxic data*. Flammarion, 2023 (cited on pages 29, 34).

[71] David Chavalarias, Philippe Huneman, and Thibault Racovski. *Contribution of phylomemies to the understanding of the dynamics of science*. Edited by Grant Ramsey and Andreas de Block. Pittsburgh, Pa: University of Pittsburgh Press, 2022. ISBN: 978-0-8229-4737-0 (cited on page 23).

[72] Boris Cyrulnik. *Les nourritures affectives*. Odile Jacob, 1993 (cited on page 101).

[73] Antonio Damasio. *Feeling & knowing: Making minds conscious*. Pantheon, 2021 (cited on page 35).

[74] Hamid R Ekbia and Bonnie A Nardi. *Heteromation, and other stories of computing and capitalism*. MIT Press, 2017 (cited on page 78).

[75] Luciano Floridi. *The fourth revolution: How the infosphere is reshaping human reality*. OUP Oxford, 2014 (cited on page 100).

[76] Luciano Floridi. *The onlife manifesto: Being human in a hyperconnected era*. Springer Nature, 2015 (cited on page 100).

[77] Erich Fromm. *Escape from freedom*. Macmillan, 1994 (cited on page 42).

[78] Christian Fuchs. *Digital labour and Karl Marx*. Routledge, 2014 (cited on page 78).

[79] Mary L Gray and Siddharth Suri. *Ghost work: How to stop Silicon Valley from building a new global underclass*. Eamon Dolan Books, 2019 (cited on page 78).

[80] Yuval Noah Harari. *Homo Deus: A brief history of tomorrow*. random house, 2016 (cited on page 62).

[81] Jaron Lanier. *Dawn of the new everything: A journey through virtual reality*. Random House, 2017 (cited on page 67).

[82] Bruno Latour. *Où atterrir?: comment s' orienter en politique*. La découverte, 2017 (cited on page 91).

[83] Inès Leonarduzzi. *Réparer le futur: du numérique à l'écologie*. Éditions de l'Observatoire, 2021 (cited on page 91).

[84] Evgeny Morozov. *To save everything, click here: The folly of technological solutionism*. Public Affairs, 2013 (cited on page 92).

[85] François Rabelais. *La vie de Gargantua et de Pantagruel*. Volume 1. Chez Dalibon, Libraire, Palais-Royal, galerie de Nemours., 1823 (cited on page 101).

[86] Sarah T Roberts. *Behind the screen*. Yale University Press, 2019 (cited on pages 79, 82).

[87] Ronald Sandler. *Ethics and emerging technologies*. Springer, 2016 (cited on page 97).

Proceedings

[88] Lars Backstrom et al. "Group formation in large social networks: membership, growth, and evolution". In: *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2006, pages 44–54 (cited on page 97).

[89] M Bastian, S. Heymann, and M. Jacomy. "Gephi: an open source software for exploring and manipulating networks". In: *International AAAI Conference on Weblogs and Social Media*. 2009 (cited on page 22).

[90] Anne Bumiller et al. "Towards a Better Understanding of Impersonation Risks". In: *2022 15th International Conference on Security of Information and Networks (SIN)*. IEEE. 2022, pages 01–08 (cited on page 99).

[91] Maria Savona et al. "The Design of Digital Automation Technologies: Implications for the Future of Work". In: *EconPol Forum*. Volume 23. 5. 2022, pages 4–11 (cited on page 78).

[92] Angie Zhang et al. "Algorithmic management reimagined for workers and by workers: Centering worker well-being in gig work". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022, pages 1–20 (cited on page 78).

Softwares

[93] URL: https://influencermarketinghub.com/social-media-marketing-benchmark-report (cited on page 95).

[94] URL: https://incidentdatabase.ai (cited on page 97).

[95] Alexandre Delanoë and David Chavalarias. *GarganText, collaborative and decentralized LibreWare (V0.0.7)*. Paris, 2023. URL: https://gitlab.iscpif.fr/gargantext/main (cited on pages 16, 22).

[96] F Jarrige. *Technocritiques. Une histoire des résistances au progrès technique*. 2014 (cited on page 92).

[97] *Social media - Statistics on the use by enterprises*. 2021. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Social_media_-_statistics_on_the_use_by_enterprises#Social_media.2C_a_business_paradigm_shift (cited on page 95).

Index

A

Accountability 35
Advertising 64
Affect 35
Affective content 35
Agency 42, 47
Agenda-setting 35
AI 45, 74
AI regulation 73
Algorithm transparency 35
Algorithmic freedom 75
Algorithmic intelligibility 74
Algorithmic regulation 72
Annotation 81
Anonymity 57
Artificial Intelligence 77, 101
Authentication methods 57
Autonomy 42, 74, 91, 98
Axiological neutrality 59

B

Bias 67
Biases 104
Brain 35
Burnout 77

C

Citizen platforms 87
Civic participation 75
Clickfarmers 79
Code 73
Cognitive biases 29, 59, 75
Cognitive sciences 35
Collective action 34, 71
Collective dynamics 33
Collective intelligence 34
Commons 88
Confusion 67
Consent 74
Corporate transparency 80
Critical thinking 47

D

Data generation 81
Data protection 57, 89
Data reuse 89
Data security 59
Data work 77
Deep fakes 104
Democratic audit of platforms 74
Digital commons 85
Digital divide 98
Digital embedding 29
Digital frugality 93
Digital labour 77
Digital labour disclosure 84
Digital law 72
Digital literacy 47, 98
Digital platforms 77, 79
Digital preservation 100
Digital rights 85
Digital well-being 75
Disinformation 47
Diversity 75

E

E-commerce 57
Echo chambers 35, 98, 104
Education 47, 103
Egalitarian digital media 95
Emotion contagion 33
Emotional responses 34
Emotions 34, 35
Encryption algorithms 58
Environmental emergency 93
Equality 91, 94
Ethical design 75
Ethics 63

Explainability 74

F

Fake news 44
Fictional reality 67
Filter bubble 98
Fragmentation 51
Freedom 42
Freedom of choice 45
Freedom of speech 34

G

GDPR 57, 75
Global risks 100
Globalization 50
Green AI 78
Group formation 29, 34

H

Head-mounted displays 64
Homophily 29
Hyper-reality 99

I

Identity 57, 100
Identity formation 71
Identity protection 66
Immersive environment 67
Inclusive design 75
Inclusivity 92
Inclusivness 91
Infrastructure 59

L

Large scale on-line experiments 31
Legal framework 87
Legislation 73, 85

M

Macroscopes 33
Mass surveillance 33
Mental health 79
Metaverse 42, 63
Micro-targeting 44
Micro-workers 79
Misinformation 44
Moral panic 99
Multi-layer networks 31

N

Narratives 105
Non-verbal channels 31
norms
    moral 34
    social 34
Nudges 75

O

Offline discriminations 75
Open data 89
Opinion dynamics 33

P

Personalization 45
Polarization 51, 98
Political action 71
Prefrontal cortex 35
Privacy 57, 66
Pro-social values 75
Psychological well-being 101
Public institutions 100
Public opinion 99
Public participation 85
Public platforms 85

Q

Quantum computing 58

R

Radicalization 51
Randomness 42
Ranking algorithms 75
Rebound effects 93
Recommender systems 29, 31, 44

Reflexive social media 74
Reflexive tools 33
Regulation 35, 66, 89
Risk management 96
Risk mitigation 96

S

Selection 29
Self-determination 42
Self-governance 42
Self-reflexivity 44
Self-regulation
in humans 43
Shared norms 29
Shared reality 66
Smartness 91
Social activism 34
Social capital 31
Social cohesion 34
Social control 44
Social embedding 35
Social impact assessment 80
Social justice 34
Social movements 33
Social psychology 42
Social reflexivity 33
Socio-semantic networks 29
Sociological studies 31
Solutionism 75
Spatial embededdness 31
Subsectioning 50
Surveillance 44, 64
Sustainable development 91, 94

T

Temporal embedding 31
Transparency 58, 72, 78
Trust 72

V

Value-sensitive design 75
Virtual reality 42, 46, 63
Vulnerable populations 75, 87, 92

W

Web archives 87
Will 44

III Supplementary information

Appendices 117

.1 Queries used for the state-of-the-art 117

.2 Semantic perimeter of the research on *digital media and humanity well-being* 117

.1 Queries used for the state-of-the-art

We retreived from the WoS all the documents matching any AND conjunction between *AB= ((“social media platforms” OR “online communication” OR “Digital Media” OR “social media” OR “social networking site” OR “online social networks” OR “facebook” OR twitter OR “recommender systems” OR “screen time” OR “augmented reality” OR “virtual reality” OR “smartphone”)* and the following queries :

1. *(“digital natives” OR “digital heritage” OR “cultural diversity” OR minorities OR LGBT OR “community support”))*
2. *(“quantified self” OR “self-reporting” OR “health app” OR “tracking” OR “periods” OR fertility OR “health data” OR “sentitive data” OR “protected data” OR “confidential information”))*
3. *(“policy advocacy” OR “human rights” OR “sex workers” OR “sex rights” OR privacy OR “fair conditions” OR “digital labour” OR “virtual work” OR “online education”))*
4. *(“digital enclosure” OR “echo chamber” OR “filter buble” OR “exclusion” OR “digital commons” OR “peer production” OR “platform cooperativism”))*
5. *(“social justice” OR “political participation” OR “e-democracy” OR “access to information” OR “access to knowledge” OR “open data” OR “inclusivity” OR “open governance” OR “open governement” OR “open access” OR “open science” OR “decision fairness” OR “black box” OR “critical media” OR “digital capitalism” OR “cryptography” OR “human centric” OR “citizen centric” OR “employment” OR “freedom of speech” OR “surveillance” OR “ethics” OR “media regulation”))*

After removing duplicate, the corpora obtained had 58,266 documents up to Juily 2021.

.2 Semantic perimeter of the research on *digital media and humanity well-being*

These 1,650 expressions have been identified by text-mining with *GarganText* as the core vocabulary of the research on *digital media and humanity well-being*

- #metoo
- 3d
- 3d city models
- 3d data
- 3d environment
- 3d interaction
- 3d printing
- 3d simulation
- 3d space
- 4chan
- absence
- acceleration
- accelerometry
- acceptability
- accessibility
- accountability
- acquisition
- active participation
- activism
- activists strategies
- activity observation
- actor-network
- adaptation
- addiction
- adolescence
- adolescent,adolescents, teenagers, teens
- adolescent health
- adolescents
- adversarial
- adversarial testing
- advertising
- affective computing
- affective polarisation
- age-adjustment
- agency
- agent based models
- ai
- ai agents
- ai application
- ai content

- ai ethics,machine ethics
- ai-driven systems,ai-driven tools
- ai-driven tools
- alcohol
- alcohol use
- algorithmic bias
- algorithmic trading
- algorithms
- algorithms access
- alternative reality
- alzheimer
- alzheimer's disease
- amazon
- anger
- animosity
- anonymity
- anonymization
- anonymous
- anti-democratic
- anti-establishment
- anti-immigrant,anti-immigrant movement, anti-immigration, anti-migration
- anti-refugee
- anti-vaccination
- anti-vaccine
- anti-vaccine propaganda
- anticounterfeiting
- anxiety
- arab spring
- article analyses
- artifical intelligence
- artificial agents
- artificial intelligence,ai
- artificial moral agents
- artificial societies
- assortative
- astroturfer
- astroturfing
- asylum
- attention dynamics
- attention economy
- attention score
- audiences
- augmented reality,alternative reality, extended reality, hybrid reality
- augmented reality overlay
- augmented reality visualization
- authentication systems
- authenticity
- authoritarian
- authoritarian regimes
- authoritarian states
- authoritarianism
- autism spectrum disorder,autism spectrum disorders
- autocracies
- automated accounts
- automation
- autonomous ai
- autonomy
- availability
- avoidance behavior
- awareness
- awareness of biases,awareness of cognitive biases
- awareness of cognitive biases
- bayesian analysis
- behavioral neuroscience,behavioral neurosciences
- behavioral neurosciences
- belief in conspiracy theories
- belief system
- beliefs
- bias
- biases
- bibliometric
- big data
- big tech
- biodiversity
- biodiversity loss
- biomarkers
- biopolitics
- bipolar
- bipolar disorder
- bitcoin
- black boxes,black box, black-box
- blockchain
- blockchain-based
- blockchains
- blog analysis
- blogs
- blood pressure
- bmi
- body mass index
- borders
- bot
- botnet,botnets
- bots,bot, botnet, social botnet, socialbot, socialbots
- brain
- brain computer interfaces
- brain implants
- brand trust
- breast cancer
- brexit
- built-in smartphone sensors
- bullying
- bureaucracy
- buying choice,buying choices
- buzzfeed
- cambridge analytica
- cancel culture
- cancelculture
- cannabis

- capital
- capitalism
- care
- career choices
- carpal tunnel
- cartography
- causal modelling
- censorship
- cf-based recommender systems,cf-based recommender system
- childhood obesity,childhood overweight and obesity
- children
- children aged
- chinese social media
- cinahl
- citation count
- citizens participation
- citizenship
- civil societies
- civil society
- class struggle
- clean eating
- climate
- climate action
- climate change
- climate change mitigation
- climate denial
- climate denialist
- climate discourses
- climate disinformation
- climate justice
- clinical
- clinical centers
- clinical practice,clinical practices
- clinical psychology
- cloud computing
- cloud computing infrastructure
- cluster analysis
- co-design
- co-word
- co-word analysis
- cognition
- cognitive and affective trust
- cognitive behavioral therapy
- cognitive biais
- cognitive dissonance
- cognitive fatigue
- cognitive function
- cognitive loadawareness
- cognitive modelling
- cognitive overload
- cognitive science
- cohesion
- cohesion bi
- collaborative filtering
- collaborative plateforms
- collective action
- collective decision making
- collective dynamics
- collective freedoms
- collective intelligence
- collective organization
- college students
- common friends
- communication contract
- communication tactic
- communicative conventions
- community detection
- community development
- competition
- competitiveness
- complex identity
- complex systems approach
- compulsive use
- computer vision
- computerization
- conceptual engineering
- condom use
- confidentiality,confidential information
- confinement
- confirmation biases
- connected migrant
- connection
- consciousness
- consent
- conspiracy
- constraints
- construction of self
- consumer behavior
- consumer experience
- consumer trust
- consumers
- contagion
- content consumers
- continuous feedback
- convolutional neural networks
- cooperation
- coordinated campaigns
- coordination
- coordination patterns
- coronavirus disease
- corpus analysis
- correcting biases
- corruption
- counter narrative,counter-narratives
- counter-narrative
- counterfeit devices
- counterfeiting
- courts of justice
- covid-19
- covid-19 vaccines,covid-19 vaccine
- creativity
- crisis

- critical phenomena
- criticality
- crowdsourcing
- crowdwork
- cryptocurrency
- cryptography
- cryptomarket
- cultural differences
- cultural heritage
- cultural norms
- curation
- currency
- customer trust
- cyber capabilities
- cyber criminals
- cyber ethics
- cyber racism
- cyber security
- cyber-activity
- cyber-citizens
- cyber-ethnicity
- cyber-physical system
- cybercitizen,cyber-citizens
- cybersecurity
- cybersecurity systems
- cybersickness
- cyberspace,cyberspaces
- cyberspaces
- cybervetting
- dark pattern
- dark patterns
- data based models
- data donations
- data encryption
- data extraction
- data integrity
- data misuse
- data portability
- data protection
- data regulation
- data security
- data sharing
- data storage
- data visualization
- database quality
- dataism
- dataveillance
- debate
- debunking
- decentralisation
- decentralization
- decentralized identifiers
- decentralized infrastructure
- decentralized internet
- decentralized personal data management
- decentralized systems
- deception,deceptive intent
- decision support systems
- decision-making
- deep learning
- deep learning models
- deep networks
- deepfake
- deepfakes,deepfake
- deliberation
- deliberative
- deliberative arenas
- deliberative democracy
- deliberative processes
- deliberative theory
- demagogues
- democracies
- democracy,democracies
- democratic participation
- democratic societies
- democratic space
- denialist discourses
- depression
- depressive symptoms
- design
- design values
- developing countries
- development policies
- diabetes
- diabetes mellitus
- diaspora,diasporas
- diasporas
- diet
- differential privacy
- digital affordances
- digital age
- digital archives
- digital audiences
- digital biomarker
- digital book,digital books
- digital born
- digital bureaucracy
- digital citizen,digital citizens
- digital commons
- digital communities
- digital controls
- digital culture
- digital diaspora
- digital diasporas
- digital divide
- digital environments-augmented reality
- digital ethics
- digital exclusion
- digital extremism
- digital feminism
- digital footprints
- digital governance
- digital identity
- digital immigrants

- digital integration
- digital labor
- digital labour
- digital litteracy
- digital marketing
- digital media
- digital media literacy
- digital methodes
- digital narratives
- digital natives
- digital phenotyping
- digital piracy
- digital policing
- digital public sphere
- digital racism
- digital society
- digital sociology,digital sociologies
- digital sovereignty
- digital storytelling
- digital surveillance
- digital twin
- digital well-being
- digital work
- dignity
- dimentionality reduction
- direct interactions,direct interaction
- disassortative
- disaster management
- disbelief in institutions
- disclosure
- disclosure behavior
- disconnecting
- disconnection,disconnect
- discourse
- discrimination
- discriminatory treatment
- discursive strategies
- discussion groups,discussion group, group discussion, group discussions
- disease management
- disease surveillance
- disgust
- disinformation
- display technology,display technologies
- dissemination
- dissemination ethical approval
- distress
- distributed consensus
- distributed ledger technologies
- distrust
- diversity
- diving behavior
- drug administration
- dynamic networks
- dynamical systems
- e-commerce
- e-democracy
- e-government
- e-law
- e-learners
- e-learning
- early detection
- earth
- eating disorders
- echo chamber,echo chambers, echo-chamber
- echo chambers
- econometrics
- economic justice
- economic well-being
- ecosystem
- ecosystems
- edemocracy
- edge computing
- education
- education environments
- education level,education levels
- education systems,education system
- educational activities
- educational attainment
- educational material,educational materials
- educational practice,educational practices
- educational strategy,educational strategies
- educational tools,educational tool
- eeg
- ego network
- election
- election manipulation
- elections
- electrodermal activity
- electromagnetic tracking
- electronic freedom frontier
- embedding-based representations
- emergence
- emotion,emotions
- emotion contagion
- emotional anxiety
- emotional communication
- emotional exhaustion
- emotional experience
- emotional expression,emotional expressions
- emotional state
- emotions
- empathic responses
- empathy
- employee
- employees
- employment
- empowering users
- empowerment,empowering
- encouragement to violence
- encryption
- endometriosis
- energy consumption
- enhancement of the public discourse

- enrollment
- environment
- environmental protection
- ephemeral environnment
- epidemiology
- epistemic bubbles
- epistemic networks
- equality
- equilibrium
- equity
- espionage
- ethereum
- ethereum-based
- ethical approval
- ethical assessment
- ethical considerations
- ethical decision-making
- ethical framework
- ethical implications
- ethical institutions
- ethical issues
- ethical principles
- ethical recommendations
- ethical standards,ethical framework
- ethical use
- ethical values
- ethics,ethical principles
- ethics approval,ethical approval
- ethnic composition
- ethnic minorities
- ethnography
- european union
- event-centric datasets
- evidence-based policy
- evolution
- excessive use
- exclusion
- experience escapes
- explainability
- explainable recommendation
- exploitation
- exploratory data analysis
- extended reality
- extremism
- eye tracking
- face-to-face interaction
- facebook
- facebook advertising,facebook marketing
- facebook behavior
- facebook friends
- facebook information
- facebook marketing
- facebook messenger
- facebook usage
- facebook users
- facial expression
- facial expressions
- facial recognition
- fact-checking
- fair
- fair trade
- fairness
- fake images
- fake narratives
- fake news,“fake news”
- fake news detection system
- fake profiles,fake profile
- fake publishers
- fake users
- false information
- falsehood
- families
- family identity
- far-right
- fast algorithms
- fatigue
- fear
- fear of missing out
- feature representation
- feedback
- femicide
- feminism
- field experiment
- filter bubble
- filter bubbles
- first-person cameras
- fmri
- focus group
- focus on important things
- focusgroup
- fomo
- food safety
- formal language
- fragmentation
- free labour
- free press
- free software
- free will
- freedom of expression
- freedom of speech
- friends
- friendship
- friendships
- functional programming
- fundamental rights
- gambling
- game streams
- gameplay
- gaming
- gargantext
- gatekeepers
- gaza war
- gaze
- gdpr

- gender
- gender diverse
- gender divisions
- gender gap
- gender identity
- gender minority
- generation z
- geo-information
- geomatics technology
- gesture
- gesture tracking sensors
- gis
- global climate movement
- global connectivity
- global integration
- global positioning system
- global warming
- glucose monitoring
- google
- google scholar
- gossip
- governance
- government
- government power
- governmental institutions
- governmental regulations
- governments
- gps
- gpus
- graph embeddings
- graph mining
- graph neural networks
- grassroots
- grassroots mobilization
- green transition
- greta thunberg
- group attitudes
- group messaging
- group-based
- growth-cruve analysis
- hını
- habermas
- habermas's
- habituation
- hacker groups
- hacktivism
- hactivism
- hand gesture,hand gestures
- happy
- harassment
- harmful practices
- harms
- hashtag
- hashtags
- hate
- hate speech
- hate speech reduction
- hateful comments
- hateful content
- head-mounted displays
- health
- health authorities
- health care
- health care professionals
- health care providers
- health care system,health care systems
- health data,health information, health-related information
- health disparities
- health education
- health information
- health issues
- health literacy
- health outcomes
- health professionals,health care professionals, health care providers, health professional
- health promotion
- health research
- health-related information
- health-related quality
- healthcare,health care
- healthy behaviors
- healthy individuals
- heterophily
- high risk
- historical events
- history
- hiv
- hiv prevention
- hiv testing
- hoax,hoaxes
- homelessness
- homophily
- homotopy
- hospitality
- hospitals
- human autonomy
- human communication
- human in-the-loop
- human inforgs
- human papillomavirus
- human rights
- humanity
- humiliation
- hybrid methods
- hybrid reality
- hypermedia
- hypernets analysis
- icts
- idealization
- identities
- identity
- identity thievery
- illiberal

- illiberalism
- image verification
- image-recognition
- immersive
- immersive technologies
- immersive virtual environments
- immersive virtual reality
- immigrant,immigrants
- impact of the social media
- improved mental health
- in-person
- incel
- incels
- inclusion
- inclusivity
- incorrect-information
- individual freedoms
- individualization
- indoor environment,indoor environments
- inequalities
- inequality,inequalities
- inequality reduction
- infectious diseases,infectious disease
- infiltration
- infiltration campaign
- influence
- influence online
- influencing behaviors
- infodemic
- inforgs
- information age
- information consumption
- information credibility
- information diffusion
- information diffussion
- information disorders
- information dissemination
- information flow
- information infrastructure
- information misuse
- information overload
- information sharing,information shared, share information
- information society
- informed consent
- infosphere
- infrastructures
- instability
- instagram
- instant messaging
- institutional safeguards
- insulin
- integration policy
- integrity
- intellectual property
- interaction
- internet abuse
- internet access
- internet addiction
- internet culture
- internet monitoring
- internet of things,iot
- internet-related behaviors
- interpersonal trust
- intersubjective
- interviews
- ip addresses
- isolation
- issue engagement
- job seekers
- journalism
- journalists
- joy
- judges
- jurisdictions
- justice
- key challenge,key challenges
- knowledge
- knowledge democratization
- knowledge discovery
- knowledge dissemination
- knowledge mapping
- labeling machine
- labor
- labour
- language
- lapses
- large-scale experiments
- law enforcement
- legal analysis
- legal issues
- leisure
- lgbtq
- lifestyle
- lifestyle changes,lifestyle change
- link analysis
- linkedin
- llogic
- location recommendation,location recommendations
- location-based social networks,location-based social network
- lockdown
- locked in
- log data
- loneliness
- longitudinal studies
- loss of privacy
- low energy
- low tech,low technology
- low technology
- lower personal perceptions
- machine ethics
- machine learning

- malicious
- malicious actors
- malicious attacks
- malicious profiles
- malware
- manipulate public opinion
- manipulation
- manipulative
- marital status
- marker occlusions,marker occlusion
- market segments
- marketing
- marx
- mass media
- media consumption
- media manipulation
- media sharing
- mediated social structure
- medical education
- medical information
- medication adherence
- memory biases
- mental disorders
- mental fatigue
- mental health,mental health care, mental health services, mental healthcare
- mental health care
- mental health services
- mental health treatment
- mental healthcare
- mental obesity
- metadata
- metaverse
- metoo
- mhealth
- mhealth apps
- micro to macro
- microblogging,micro-blogging
- microblogging platform,microblogging platforms
- microblogs
- microwork
- migrant,migrants
- migrant crisis
- migration
- migration flows
- military power
- mimic
- mindfulness
- minority group,minority groups
- mirror illusion
- mirror neurons
- mirror therapy
- mis/disinformation
- misinformation
- misinformation online,online misinformation
- misjudgments
- mistrust
- mixed methods
- mobile and social media,social and mobile media
- mobile app,mobile apps
- mobile app developpement
- mobile augmented reality,mobile augmented reality application
- mobile augmented reality application,mobile augmented reality applications
- mobile computing
- mobile devices
- mobile health,mhealth
- mobile health apps
- mobile internet
- mobile learning
- mobile log data
- mobile media
- mobile phone applications
- mobile phones
- mobile robot,mobile robots
- mobile sensing
- mobile social networks
- mobile terminal,mobile terminals
- mobile vr
- mobility
- moderators,moderator
- money
- mood
- moods
- moral
- moral attributions
- moral character
- moral machines
- moral outrage
- moral theory
- morality
- mortality
- motion tracking sensor
- motion-tracking device,motion tracking device, motion tracking devices
- motor recovery
- msm
- multi-agents modelling
- multi-level interactions
- multi-user interactions
- multimedia course recommendation
- multimedia courses
- multiscale
- multitasking
- museums
- music
- mutual trust
- myspace
- narrative
- narrative synthesis

- narratives,narrative, stories
- narratives analysis
- national identity
- nationalism
- natural language processing
- navigation systems,navigation system
- negative emotions
- negative feedback
- negative health behaviors
- neoliberalism
- netizens
- netizenship
- network analysis
- network interaction,interaction network, interaction networks
- network providers,network provider
- network science
- networks of websites
- neural network,graph neural networks
- neuroimaging
- neuronal assembly
- neuroplasticity
- neurorehabilitation
- neuroscience
- new media
- news
- news content
- news coverage
- news literacy
- news websites
- nirs
- nlp
- non-binary
- non-invasive brain stimulation
- non-state actor
- note books
- novel mobile application,novel mobile applications
- novel virtual reality paradigm
- obesity,overweight and obesity, overweight or obesity
- objectivity
- off-the-shelf smartphones,off-the-shelf smartphone
- offline interactions
- offline political participation
- offline social networks
- on-field observations
- online brain
- online communications,online communication
- online communities
- online courses
- online debate,online debates
- online education
- online environment
- online experimentation platforms
- online harms
- online learning
- online political expression
- online reputation
- online risk
- online self,online self-presentation
- online self-presentation
- online social media
- online social movement
- online social network
- online social networks
- online textual data
- online trading
- online vulnerability
- open access
- open data
- open government
- open government data
- open innovation
- open source,free software
- open-ended questions
- open-source
- openness
- opinion diffusion
- opinion dynamics
- opinion mining
- opinions
- opposing views
- outdoor augmented reality
- outgroup
- overlay network
- overload
- overuse
- ownership
- p2p
- p2p network
- pain
- pandemic,pandemics
- parental control
- parents
- participation
- participative deliberation
- passive sensing
- passive sensor
- patient care
- patient education
- patient engagement
- patient privacy
- patients
- peace
- peer-to-peer,p2p
- people living
- performance expectancy
- person centered care
- personal data
- personal information
- personal privacy

- personalised treatments
- personality
- personalization
- personalization algorithms
- personalized recommendation
- personalized recommender system,personalized recommender systems
- personally identifiable information
- pharmacy
- phase transition
- phone
- phylomemy
- physical activity
- physical architecture
- piracy
- plastic waste
- platform
- platform capitalism
- platform policy
- platform-goverance
- poi recommendation
- polarisation
- polarization
- polarized
- police violence
- policy makers
- policymakers
- political advertising
- political and social changes,social and political change
- political astroturfing
- political attitudes
- political battlefield
- political communication
- political conflicts,political conflict
- political content
- political conversation,political conversations
- political corruption
- political debate
- political disaffection
- political discourse
- political dissent
- political economy
- political efficacy
- political elites
- political engagement
- political expression
- political extremism
- political fake news
- political groups
- political hoax
- political identity,political identities
- political inequality
- political influence
- political information
- political institutions
- political interactions
- political interests,political interest
- political issues
- political knowledge
- political mobilizations
- political participation
- political parties
- political polarization
- political power
- political preferences
- political processes,political process
- political purposes
- political science
- political stability
- political strategy
- political system,political systems
- political tool
- political views
- political weapon
- politicization
- politics
- polls
- poor decisions
- popularity
- populism,populist
- populist
- positive effect,positive effects
- positive feedback
- post-truth
- postfeminism
- potential benefits,potential benefit
- poverty
- predictive power
- pregnancy,pregnant women
- preoperative
- prescriptive debates
- presidential campaign
- privacy
- privacy behavior
- privacy breach
- privacy breaches
- privacy issues,privacy concerns
- privacy leakage
- privacy legislation
- privacy management
- privacy perceptions
- privacy preservation,privacy-preserving
- privacy rights
- privacy risk perceptions
- privacy risks
- privacy settings
- privacy violation
- privacy violation detection
- privacy violations
- privacy-friendly
- privacy-intrusive
- privacy-preserving
- private data,private information

- private information
- pro-democracy
- pro-vaccine
- problematic internet use
- procedural justice
- producers
- profile injection
- profiling systems
- propaganda
- protest,protests
- protest repression
- protests
- psychiatric disorder
- psychoeducation
- psychological distress
- psychological well-being
- psychology
- psychosis risk
- psychotherapy
- psycinfo
- public agenda
- public attention
- public awareness
- public debate,public debates
- public debates
- public discourse
- public health
- public health surveillance
- public networks
- public opinion
- public perception,public perceptions
- public relations
- public sphere
- public trust
- purchase intention
- qr codes
- quality of governance
- quantum computing
- racial
- racial disparities
- racism
- randomized controlled trials
-
- real-time augmented reality
- real-time tracking
- real-world datasets,real world dataset, real world datasets, real-world dataset
- reciprocity
- recommendation,recommendations
- recommendation algorithm,recommender algorithm
- recommendation performance
- recommendation processes,recommendation process
- recommendation quality,quality recommendations, recommendations quality
- recommendation strategies
- recommendation system,recommendation systems
- recommendations
- recommender algorithm, recommender system, recommender systems,recommendation, recommendation algorithm, recommendation system, recommender system
- recovery
- recruiters
- recruitment
- reddit
- redpilling
- referendum
- reflectivity
- reflexivity
- refugee
- refugees
- regulating information
- regulation
- regulations
- rehabilitation
- reinforcement learning
- rejection
- rejection of election results
- relationships
- religiosity
- religious moralism
- renormalisation
- representativity
- repression
- repressive tools
- reputable source
- reputation
- reputation rankings
- resilience
- resilient democracies
- responsibility
- responsibility-based
- responsible consumption
- reward mechanism
- risk
- risk perception
- risks
- risky drinking
- rituals
- robotics
- rule of law
- rumour
- rural communities
- rural life
- russia
- sad moods
- sadness
- safety
- scalability
- schizophrenia
- schoolchildren

- science mapping
- scientometrics analysis
- screen time
- search engines,search engine
- sectarian politics
- security
- sedentary
- sedentary time
- segregation
- segregationist
- selection of interactions
- self
- self awareness
- self consistency
- self construction,self-construction
- self-assessments
- self-awareness,self awareness
- self-blame
- self-censorship
- self-construction
- self-control
- self-determination
- self-efficacy
- self-esteem
- self-expression
- self-management
- self-organization
- self-perception
- self-presentation
- self-regulating
- self-report
- self-reports
- self-monitoring
- semantic analysis
- semantic approach
- semantic information
- semantic mapping
- semantic segmentation
- semantic web
- semi-structured interviews,semi-structured interview
- sense of self
- sensitive data
- sensitive information
- sensor technologies
- sentiment
- sentiment analysis
- sentiment dynamics
- sentiment indices
- sentiments
- serendipitous recommendations
- serendipity
- serious gaming
- service providers,service provider
- service recommendation,recommendation services, service recommendations
- sex differences
- sexual abuse
- sexual assault
- sexual health
- sexual identity
- sexual minorities
- sexual trauma
- share content,content shared, content sharing
- share resources,resource sharing, resources shared
- sharing
- side effects
- silence
- simulation
- sina weibo
- single sign-on
- skype
- sleep
- sleep disturbance
- sleep disturbances
- sleep duration
- sleep onset
- sleep patterns
- sleep quality
- smart cities
- smart city
- smart contract
- smart environment
- smart glasses
- smart refugees
- smartphone
- smartphone addiction
- smartphone app,mobile app, mobile phone applications, smartphone apps
- smartphone apps
- smartphone ownership
- smartphone security
- smartphone sensors,built-in smartphone sensors
- smartphone usage,smartphone use
- smartphone usage data
- smartphone use
- smartphone user,smartphone ownership, smartphone users, user smartphone
- smartphone-based self-monitoring
- smartphones,smart phone, smart phones, smartphone
- smoking
- smoking cessation
- smugglers
- snapchat
- snss
- social attention
- social behavior
- social behaviors
- social botnet
- social capital
- social change

- social cohesion
- social commerce
- social concerns
- social connectedness
- social context
- social data
- social distancing
- social diversification
- social dynamics
- social exclusion
- social experiments
- social fragmentation
- social graph
- social groups
- social impacts
- social implications
- social information
- social integration
- social interaction
- social isolation
- social issues
- social justice
- social media addiction
- social media campaigning
- social media communications,social media communication
- social media content,social media posts
- social media data
- social media engagement
- social media fatigue
- social media marketing
- social media mining
- social media mobilisation
- social media monitoring
- social media networks,online social networks, social networking sites
- social media platform,online social media, social media platforms, social media sites
- social media platforms
- social media posts
- social media provider
- social media regulation
- social media screening
- social media sites
- social media tools
- social media use,social media used, social media using, use social media
- social media users
- social medias,social media
- social messaging
- social movements
- social network analysis,social networks analysis
- social network datasets
- social network service
- social network sites
- social network users
- social networking sites,social network service, social network sites
- social networks
- social networks analysis
- social neuroscience
- social norms
- social outcomes
- social pain
- social platform,social platforms
- social psychology
- social recommender systems
- social rejection
- social relationships,social relationship
- social smart contract
- social stability
- social structure
- social tensions
- social ties
- social trust
- social unrest
- social work
- social-cyber maneuvers
- socialbot
- socialbots
- socially-aware systems
- societal cohesion
- societal sustainability
- socio-demographic variables
- socio-economic inequality
- sociodemographic data
- socioeconomic status
- sociotechnical devices
- solidarity
- source-checking
- sovereignity
- sovereignty
- spatial location
- spatial representation
- speech
- spiral of silence
- spirituality
- stability
- state-society divide
- statistical field theory
- stereotypicality
- stock market
- stock price crash risk
- stock returns
- stock trading
- stories
- storytelling
- streamer
- stressors
- stressors-strains-outcomes
- stroke patients
- structural position
- students

- subjective experiences
- subjective norms
- subjective sleep quality
- subjective well-being
- subjective wellbeing
- substance use
- substitution behavior
- subversion
- suicidal ideation
- suicide
- suicides,suicide
- surveillance
- surveillance and control
- survey
- survey experiments
- survival of societies
- survivors
- sustainability
- sustainable development
- symbolic approach
- tablets
- teachers
- technology acceptance model
- technology mediated interaction
- technosocial environment
- teenagers
- teens
- tele-immersion
- teleoperation
- telerobotics
- television
- temporal graph analysis
- term enrichment
- term extraction
- text messaging
- tgdnb
- thematic evolution
- threats to individual freedoms
- tiktok
- token,tokens
- tolerance
- tolerance behavior
- topic analysis
- topic modeling
- topic-level trust-based recommendation algorithm
- topology of interaction
- tourism
- traceability
- tracking accuracy
- tracking of bodies
- tracking systems
- trading platforms
- traditional journalism
- traditional media
- transgender
- transnational
- transnational networks
- transparency
- transparent
- traumatic issues
- treatment
- trolls
- trust
- trust evaluation
- trust in news
- trust inference
- trust information
- trust management
- trust mechanisms,trust mechanism
- trust prediction
- trust score,trust scores
- trust values,trust value
- trust-aware collaborative filtering
- trustworthiness
- trustworthy
- truth
- tweet
- tweets
- twitch
- twitter
- twitter data,twitter data access plan
- twitter data access plan
- twitter messages
- type theory
- ubiquitous computing
- undecided voters
- undergraduate
- undergraduate students
- underrepresented minorities
- unvaccinated
- urban environments
- usability
- user awareness
- user data
- user engagement
- user privacy
- user private data,private user data
- user profiling
- user trust
- user-generated content
- user-generated information
- users experience,user experience, user experiences
- users identity,user identities, user identity
- utilitarian
- vaccination,vaccinations
- vaccine
- vaccine hesitancy
- value-sensitive design
- veracity
- vice
- victimization
- video games,video game

- video streaming
- video-sharing
- violence
- virtual 3d world,3d virtual world
- virtual and augmented reality technologies,virtual and augmented reality technology
- virtual buttons,virtual button
- virtual classroom
- virtual communities,virtual community
- virtual elements
- virtual environment,virtual environments
- virtual information
- virtual keyboard
- virtual museums,virtual museum
- virtual object,virtual objects
- virtual planes,virtual plane
- virtual reality,vr, vr technologies, vr technology
- virtual reality exposure
- virtual reality head-mounted display,virtual reality head-mounted displays
- virtual reality laboratory
- virtual reality-based interventions
- virtual scenarios,virtual scenario
- virtual scenes,virtual scene
- virtual task,virtual tasks
- virtual technologies
- visual feedback
- visualization,data visualization
- vr
- vr interaction
- vulnerabilities
- waist circumference
- walkability
- warfare
- wealth
- wearable devices
- wearable sensors
- wearable technologies
- web 2 0
- web archives
- web cartography
- web ecosystem
- web-ethnography
- webometrics
- wechat
- well-being,wellbeing
- whatsapp
- wikipedia
- wildlife
- withdrawal behavior
- women
- worklife balance
- workplace
- world health organization
- wrist
- xenophobia
- xenophobic
- xenophobic right-wing
- young adults
- young people
- young women
- youtube
- “bothsidesism”
- “dark patterns ”
- “echo chambers ”
- “fake news”
- “filter bubbles”
- “redpilling”