Manifesto from Dagstuhl Perspectives Workshop 18262

Web Futures: Inclusive, Intelligent, Sustainable
The 2020 Manifesto for Web Science

Bettina Berendt[1], Fabien Gandon[2], Susan Halford*[3], Wendy Hall[4], Jim Hendler*[5], Katharina E. Kinder-Kurlanda[6], Eirini Ntoutsi*[7], and Steffen Staab*[8]

1 **TU Berlin and Weizenbaum Institute for the Networked Society, Germany; KU Leuven, Belgium.** berendt@tu-berlin.be
2 **Inria, Université Côte d'Azur, CNRS, I3S, France.** fabien.gandon@inria.fr
3 **University of Bristol, UK.** susan.halford@bristol.ac.uk
4 **University of Southampton, UK.** wh@ecs.soton.ac.uk
5 **Rensselaer Polytechnic Institute, USA.** hendler@cs.rpi.edu
6 **University of Klagenfurt, Austria.** katharina.kinder-kurlanda@aau.at
7 **Freie Universität Berlin, Germany.** eirini.ntoutsi@fu-berlin.de
8 **Universität Stuttgart, Germany, and University of Southampton, UK.** s.r.staab@soton.ac.uk

Abstract

This Manifesto was produced from the Perspectives Workshop 18262 entitled "10 Years of Web Science" that took place at Schloss Dagstuhl from June 24 – 29, 2018. At the Workshop, we revisited the origins of Web Science, explored the challenges and opportunities of the Web, and looked ahead to potential futures for both the Web and Web Science.

We explain issues that society faces in the Web by the ambivalences that are inherent in the Web. All the enormous benefits that the Web offers – for information sharing, collective organization and distributed activity, social inclusion and economic growth – will always carry along negative consequences, too, and 30 years after its creation negative consequences of the Web are only too apparent.

The Web continues to evolve and its next major step will involve Artificial Intelligence (AI) at large. AI has the potential to amplify positive and negative outcomes, and we explore these possibilities, situating them within the wider debate about the future of regulation and governance for the Web. Finally, we outline the need to extend Web Science as the science that is devoted to the analysis and engineering of the Web, to strengthen our role in shaping the future of the Web and present five key directions for capacity building that are necessary to achieve this: *(i)*, supporting interdisciplinarity, *(ii)*, supporting collaboration, *(iii)*, supporting the sustainable Web, *(iv)*, supporting the Intelligent Web, and *(v)*, supporting the Inclusive Web.

Our writing reflects our background in several disciplines of the social and technical sciences and that these disciplines emphasize topics to various extents. We are acutely aware that our observations occupy a particular point in time and are skewed towards our experience as Western scholars – a limitation that Web Science will need to overcome.

Perspectives Workshop June 24–29, 2018 – www.dagstuhl.de/18262
2012 ACM Subject Classification Information systems → World Wide Web; Human-centered computing → Collaborative and social computing theory, concepts and paradigms; Computing methodologies → Philosophical/theoretical foundations of artificial intelligence; Social and professional topics → History of computing; Social and professional topics → Censorship; Social and professional topics → Surveillance; Social and professional topics → Network access control
Keywords and phrases Web Science, Artificial Intelligence, Web Governance, Capacity Building
Digital Object Identifier 10.4230/DagMan.9.1.1

* Editor / Organizer

Web Futures: Inclusive, Intelligent, Sustainable, *Dagstuhl Manifestos*, Vol. 9, Issue 1, pp. 1–42
Editors: Bettina Berendt, Fabien Gandon, Susan Halford, Wendy Hall, Jim Hendler, Katharina E. Kinder-Kurlanda, Eirini Ntoutsi, and Steffen Staab
Dagstuhl Manifestos
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Funding *Steffen Staab*: The work performed by Steffen Staab has been partially supported by EU-Project Co-Inform which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 770302.

Acknowledgements The authors of this Manifesto are greatly indebted to the co-participants of the Dagstuhl Perspectives Workshop. We also appreciate the support and patience that Dr. Kunz and the team at Dagstuhl had with us when we took much longer than planned in putting together this Manifesto.

Executive Summary

The Web has changed the world, and the world has changed the Web. From relatively humble origins at CERN in 1989, the Web has become the central nervous system of our planet, used by over 4 billion people across all spheres of life and reaching into the lives of those who do not even know it exists, shaping global markets, climate science, political conflict and war.

The possible futures of the Web concern us all.

In this Manifesto, we address potential futures arising from key ambivalences that lie at the heart of the Web. For all the enormous benefits on offer – for information sharing, collective organization and distributed activity, social inclusion and economic growth – the negative consequences are only too apparent. Examples include surveillance and privacy concerns, fake news and election interference, and the massive centralisation of data and infrastructure in the hands of a few near-monopolistic private companies.

Artificial Intelligence (AI) has the potential to amplify these outcomes – for better and for worse. AI might radically extend capabilities for propagation of misinformation, biased decision making, social polarization and new forms of exploitation. AI also offers opportunities to mobilise against these practices and trajectories, to support a web that is more inclusive and collaborative than ever before. We explore these possibilities, situating them within the wider debate about the future of regulation and governance for the Web, the future of Web Science and how this can be supported to shape a Web and a world that we would prefer to see.

In Section 2, we outline the ambivalences that sit at the heart of the Web, and its future:

- Information freedom vs. information quality
- Personalisation vs. privacy
- Mass participation vs. manipulation of the masses
- Inclusiveness and fairness vs. exploitation
- Sustainability vs. Growth

These tensions will never be fixed once and for all. Our question is: what can be done to meet the challenges that they pose and drive the Web towards the most positive outcomes possible?

Section 3 explores this question at the cutting edge of today's Web, outlining how AI extends and deepens the tensions described above, as the Web becomes an ever more complex network of human and technical agents. We develop an agenda for harnessing AI towards progressive ends, for example, to mitigate the activities of regressive and malign actors and towards a benevolent, inclusive, empowering and sustainable Web for the future.

In Section 4, we consider the wider landscape of governance and regulation that will be required in order to support the future Web. This requires interdisciplinary analysis, technical innovation and regulatory systems to govern humans, AIs and their interactions online. We

propose that the concept of "internet governance" conceived to manage IP addresses and Domain Name Servers should be expanded to "Web governance" as a the field of practice through which we can shape the future of the Web. Beyond the value of governance-as-usual, for example on hate-speech or fraud, Web-specific forms of governance are required to tackle new problems and interdependencies online, including the governance of artificial agents. This will only be effective if developed through global collaboration and enforcement.

The interdisciplinary capacity of Web Science is critical to all the endeavours described in Sections 1–4. Without well-developed interdisciplinary collaboration that draws together all the necessary forms of expertise, we cannot expect this to succeed. We have learnt a great deal over the past decade about how difficult this is to do and what it takes to do this effectively. In Sections 5 and 6 we explore how to build on what has been learnt, to support research practices that integrate expertise across the disciplines to deliver the next decade of Web Science. Section 5 explores how to support and develop Web Science as a field of research and collaborative practice, through joint research and shared infrastructures, shared methods to grow interdisciplinary communication and create combined expertise, harnessing AI for Web Science research and through the use of participatory methods for widespread engagement.

In Section 6, we turn attention to how this capacity can be developed through policy in academic institutions, with funding authorities, educators, governments and external partners in industry, business and the not-for-profit sector. We postulate that all these stakeholders must individually and jointly support *(i)*, interdisciplinarity, *(ii)*, collaboration, *(iii)*, the Sustainable Web, *(iv)*, the Intelligent Web, and *(v)*, the Inclusive Web.

Our views were informed by the Perspectives Workshop 18262 on "10 Years of Web Science" that took place at Schloss Dagstuhl from June 24 – 29, 2018. Web Science was established in 2008 to develop new knowledge and understanding of the past, present and future of the Web. This demanded collaboration across computational and social sciences, the humanities and beyond. The Web cannot be understood only as a technical system or only as a social system. It is both at once. This manifesto combines the expertise of workshop participants and includes a range of writing conventions, crafted as a shared path, with generous compromises in all disciplinary directions. At the workshop, we reviewed what we have learnt from the past decade of Web Science and looked ahead to potential futures for Web Science, and for the Web.

Participants in the workshop included sociologists and computer scientists, philosophers, political scientists, anthropologists, media scholars and industry partners. In spite of such breadth, we are acutely aware that our observations occupy a particular point in time and are skewed towards our experience as Western scholars. The key challenge for Web Science is to broaden this base as we move forwards towards a promising future of Web Science and a beneficial future Web.

Table of Contents

1 Introduction

Thirty years have passed since Tim Berners-Lee proposed a new architecture for information sharing for Physicists working with the high energy Physics laboratories at CERN. Shortly afterwards, this came to be known as the World Wide Web. And the rest is history. The Web has evolved from a draft proposal to a global infrastructure with over 4 billion regular users, and it shapes the lives of those who do not even know of its existence.[1]

The Web has become the nervous system for our planet.

This has been an unexpected journey. What was originally imagined as a technical system is now deeply embedded in and shaped by the social world.

Web architecture and infrastructure are technical, but the Web cannot be understood as only technical. The Web is shaped by social activity – by individuals, communities, governments and businesses – but it cannot be understood as only social.

In 2006, Berners-Lee et al. [8] proposed a new interdisciplinary field of study "Web Science", to research how the Web was evolving and what might be done to shape its future. Even then, it was clear that the Web was already implicated in some fundamental social and economic transformations – and that it posed some unexpected challenges, for example cybercrime, hate-speech and the increasing centralisation of content and infrastructure.

Over the past decade, Web Science has grown and spread across the world[2], integrating expertise from across the engineering sciences, social sciences, humanities and beyond (for definitions and state-of-the-art surveys, see [44, 72]). Web Scientists have generated new knowledge and understanding of how the Web has changed the world, and how the world has changed the Web.

The World ⇌ The Web

Web Scientists have traced how the Web generates new business models [38], enabled the effective use of big data and Web infrastructures for peace-building in fragile states [81] and built tools to analyse Web data at scale and speed for a wide variety of purposes [76]. On the dark side of the Web, Web Scientists have documented atrocities in information warfare [66, 67], identified how fake news is influencing our political landscapes [57, 59] and developed methods to identify individual cybercriminals on the Dark Web [79].

And still the Web evolves. In the years since Web Science was founded, we have seen the emergence of social media, a fully fledged data economy and revelations of mass surveillance and interference in democratic elections. And now, a new wave of Artificial Intelligence, spurred on by the phenomenal data resources created (in large part) by the Web, has begun a new round of transformations that mark a step-change in the Web of the future.

The Web ⇌ Artificial Intelligence

The Web has both facilitated the current and rapid growth of Artificial Intelligence as a field and – in turn – will be increasingly shaped by a variety of artificial intelligences, developed by researchers in academia, in business and elsewhere. Looking towards the future, the Web will be comprised of a broader range of actors than ever before, as humans, artificial agents and other technologies interact shaping each-other and shaping outcomes across the globe.

---

[1] For example, influencing commodity prices, global politics and climate change research.

[2] `http://wstnet.webscience.org`

Exactly how this will happen is, as yet, undetermined. On the one hand, there are very real concerns. For example, AI is radically extending the capabilities for propagation of misinformation, of biases in information distribution and decision making, and it further extends social polarization, harassment and hate speech. On the other hand, there are also opportunities for AI to mobilise against these activities and support a Web that is more inclusive and collaborative than ever before.

The future of the Web is a deeply sociotechnical question [72]. Web Science is essential if we are to address these concerns and harness the opportunities for an AI-enabled Web.

Web Science ⇌ The Web of Humans and Artificial Intelligences

This Manifesto for Web Science addresses the future of the Web, and arises from our Dagstuhl Workshop in June 2018. Participants in this workshop included sociologists and computer scientists, philosophers, political scientists, anthropologists and media scholars. Staying true to the interdisciplinary commitments of Web Science and the diverse disciplinary expertise of its authors, this report is following a range of writing conventions. It has been crafted as a shared path, with generous compromises in all disciplinary directions.

We opened our manifesto with the observation that the Web is now the nervous system for our planet. All our futures are entangled with the future of the Web. So "Web Science" is not just for researchers. It is for everyone. It is for entrepreneurs with an eye for a new business proposition, corporations seeking to innovate. It is for governments wanting to make the most of the Web for services and engagement, whilst assuring their citizens that they have robust regulation and governance in place to manage the Web. And Web Science is for individuals and communities, whose everyday activities make up so much of the content online, whose data are driving the data economy, and whose lives are increasingly unimaginable without the Web. Web Science seeks to engage all the interests above, to create new dialogues and ensure that we build a comprehensive and actionable understanding of the Web as we seek to shape its future.

In Section 2 below, we outline the current challenges facing the Web, in general and with particular reference to Artificial Intelligence. In Section 3, we outline how Web Science can address these challenges, by shaping the development of AIs for the Web in a socially responsible way. In Section 4, we broaden the view to consider the wider landscape of governance and regulation that will be required in order to support a fair and inclusive Web. The interdisciplinary capacity of Web Science is critical to these endeavours. Without well-developed collaboration that draws together all the necessary forms of expertise, we cannot expect this to succeed. We have learnt a great deal over the past decade about how difficult this is to do and what it takes to do this effectively. In Section 5 we explore how we can build on what has been learnt, to support research practices that integrate expertise across the disciplines to deliver the next decade of Web Science. A key challenge for Web Science is to broaden our base as we move forwards in time. In Section 6, we turn attention to how this capacity can be developed by academia, enterprises, funding authorities and governments. In doing so, we are acutely aware that our observations occupy a particular point in time and are skewed towards our experience as Western scholars. Addressing this will also be critical for the future of Web Science and the future of the Web.

2 Ambivalences of The Web

Whilst the Web is a mighty instrument for individuals, institutions and society, as an instrument it can be put both to beneficial or to detrimental use. For all the enormous benefits that the Web offers for information sharing, collective organization and distributed activity, social inclusion and economic growth, the negative consequences are only too apparent. It is a core task for Web Science to analyse, understand and navigate these ambivalences of the Web, ideally amplifying the good and countering the bad.

2.1 Information Freedom vs. Information Quality

There is a very long history behind the idea that the sharing of ideas and of evidence-based information is positive for social inclusion, productivity and well-being in society.[3] The printing press greatly reduced the costs of information dissemination and allowed the sharing of ideas and evidence that supported the development of modern science. The invention of the Web further reduced these costs by many orders of magnitudes democratizing access to formal knowledge (open science, free textbooks, historical sources) and informal knowledge (recipes, reviews, etc.). Both the production and consumption of information such as political news and scientific results has been greatly widened, accelerated, which is *prima facie* a positive outcome.

In the pre-Web era, publishers were gatekeepers. Their role could be ambivalent, too, as they could exert quality control or censorship on content. In the era of the Web, it is inexpensive to become a publisher. While it is easy to publish facts and judgments, it is equally possible to widely disseminate mis- and disinformation, unfounded beliefs and prejudices.

Mis- and disinformation existed before the Web, but the Web has facilitated its distribution, at scale and speed. Artificial Intelligence further amplifies the problem. It has been a core principle of modernity to question printed information, but to believe in original sources, such as photography, audio or video. Artificial Intelligence allows for the creation of *deep fakes*, e. g., video or audio faking the voice and the appearance of anybody while being indistinguishable from true video and audio.

2.2 Personalization vs. Privacy

Core to the success of the Web is the convenience that it offers. Using our personal devices, information consumption and production has become frictionless. One or two clicks are sufficient to buy, to like or to share, and service providers know about our preferences facilitating selection of music, audio, restaurants or other products from swathes of offerings. Such *personalization* is made possible through *online behavioural tracking mechanisms* able to harvest minute details of online activity. A variety of actors aggregate online activity across populations in order to analyze, profile and serve targeted information. The more detailed the tracking, the smoother our experience of the Web, and the higher the revenues for the corporates.

While every small improvement derived from closer tracking and more precise analytics turns into a sizable business benefit for the trackers motivating them to ever more comprehensive monitoring of our lives, there are diminishing returns yielded to the individual. The

[3] `http://www.rogerclarke.com/II/IWtbF.html`

consequence is a deep intrusion into the privacy of individuals, what they do and like, what their political beliefs are, whom they do business with, and whom they spend their days and nights with. At best this intrudes people's privacy and exploits it without fair remuneration, at worst it may lead to misuse manipulating individuals or deriving damaging conclusions about them.

Consider a case of manipulation in 2014. Facebook faced a storm of protest after it was revealed that via an *online psychological experiment* they were analyzing if they could make users feel happier or sadder by controlling their feeds. This incident provoked widespread public concern regarding the effect of such experiments and interventions and the lack of agreement on expertise and ethics knowledge about how to do experimental Web research. Consider the topic of possibly damaging conclusions as demonstrated in early experiences with prototypes of a planned Chinese social scoring system. There have been reports that individual's political opinion and even the political opinion of their friends influence the score and hence impact freedom of movement, ability to work, or chances to date other people.[4]

Artificial Intelligence techniques, such as text and video understanding, intelligent emotion recognition, or face recognition, may fully leverage these developments. Then, companies or governments might have almost complete knowledge about the private life of citizens. One might have a complete account of traffic rules violations or of CO2 emission with incentives to avoid both. But one would also incur far-ranging possibilities for misuse of such knowledge and severe implications on the resulting social behavior of citizens.

2.3 Influence by the Masses vs. Manipulation of the Masses

The Web offers the opportunity for countless people to connect across geographical boundaries in ways that were difficult to imagine in the context of previous media systems. Since the inception of the Web, the resulting network allows for a scale and speed of information dissemination that has been seen as a huge opportunity for democratization, for example by allowing the wide publication of injustices. A recent example was the German video by Rezo on Youtube which allowed a previously unpolitical social media influencer to affect the German political discussion on climate change and political (in-)activity.[5]

The Web also allows to tap into the "wisdom of the crowd" by sourcing its seemingly infinite variety of connected entities. However, this network has from its beginning always held surprises and developed in unexpected ways. While we have made huge inroads into understanding phenomena of how and when people mutually influence each other, there is still a lot to learn about how this affects the behaviour of crowds online in the large. This is especially the case if we look at the less positive aspects of the Web as a public network. Crowd sourced judgment can be good or bad, depending on how you want to see the world [54].

The extent to which crowd behaviour may be manipulable or determining at what point a righteous public outcry may turn into undesirable and vindictive pursuit are not well understood. And both the positive and the negative sides of the Web may be amplified by

[4] Cf., e. g., `https://www.wired.co.uk/article/china-social-credit-system-explained` on the ambivalences of the Chinese social scoring systems.

[5] `https://www.spiegel.de/wissenschaft/mensch/rezo-video-die-politik-lullt-ein-und-die-jugend-wehrt-sich-a-1269173.html`

the possibilities that algorithmic analyses and AI bring. For example, in 2018 it was revealed that Cambridge Analytica had harvested the personal data of millions of people's Facebook profiles without their consent and used it for political purposes (*the Facebook–Cambridge Analytica data scandal*), like targeting voters in the US presidential election.[6] This was a turning point in the public understanding of personal data and caused many calls for tighter regulation of tech companies' use of data. It also showed that there is a potential to influence large amounts of people – and uncertainty as to how exactly such influencing works, may already be at work and who may be trying to influence people.

These uncertainties also became clear in connection with the Brexit referendum when it turned out that bot accounts may have been used to boost follower numbers for parties on Twitter.[7] This example highlights the uncertainty about the influence of bots: their actual impact is uncertain as is their power to influence opinions – there certainly seems to be large potential for it.

Thus, it is clear now that interconnected networks of the Web may influence people for the better or the worse, but it is neither clear when a purpose turns out to be "good" or "bad" and who, if anyone, should have the authority and power to make such a judgement.

2.4 Inclusiveness and Fairness vs. Exploitation

The Web has been designed to be inclusive. This means that all individuals, but also all institutions – companies, associations, governmental or non-governmental organizations – can make best use of it, contributing their content and data and being able to access it.

While this may sound simple enough, it is not. Even low costs of internet access might be too high for some to bear. In this context, Facebook's `http://internet.org` initiative proposes to offer for free a limited version of the Web with access to specific Web sites like Wikipedia as well as to Facebook and Facebook partner businesses. The Indian government has banned this initiative. The underlying rationale was that this initiative would turn Facebook into a quasi-monopolistic gatekeeper to the internet in India, such that only businesses who partnered with Facebook would be able to thrive there. At the same time, this initiative would have given many people access to *some* internet opportunities they seek and any alternative will be more expensive and therefore more exclusive for many of them.

The Web as a system of people, platforms, and economic affairs naturally tends towards such a monopoly as that `internet.org` wants to create. Indeed, other monopolies or oligopolies exist in the Web, e. g., for mobile phone operating systems, search, social networking, business networking, shopping, or video hosting, where companies have created free offers that attracted a majority of people rendering further competition economically unfeasible. While it seems fair enough that these companies draw a profit from their investments, once these monopolies/oligopolies have grown large enough, they are the gatekeepers that can exploit their customers as well as the existing infrastructure and business partners almost at will.

There are fundamental problems in defining what constitutes fair behaviour and what does not, and the fairness of economic behavior cannot be disentangled from the other questions raised before, which were the ambivalences of personalization versus privacy violations or wisdom of the crowd versus mass manipulation.

---

[6] `https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election`

[7] Researchers Say Many Of The Brexit Party's Twitter Followers Aren't Behaving Like Genuine Voters `https://www.buzzfeed.com/alexspence/nigel-farages-brexit-party-twitter-following`

Artificial Intelligence seems to amplify this problem. Progress in AI depends on the availability of trained researchers and availability of data, both of which are found at the Web behemoths more than anywhere else, and AI allows for those value-added services that are likely to attract even more customers thus nurturing the spiral that strengthens existing monopolies.

In addition to these challenges to economic fairness and the risk of exploitation, AI and its uses in the Web also present challenges to fairness in the sense of equalities (of opportunities, outcomes, etc.) and the risk of discrimination. There is an active discussion as to when AI systems are fair or when they are biased [62].[8]

Regulations that aim at creating fairness for all, thus should not only consider the individual, but they should consider inclusiveness and fairness for all groups. Small companies and social groups may suffer most from regulations that large companies can deal with. One of the authors was briefly involved in running a truly free social networking site, Metalcon, with more than 10,000 users. That social network could not be maintained in spite of its attraction to new visitors and in spite of free voluntary work, as legal compensations for copyright breeches provoked by its users far exceeded its minor income.

Further discussions concerning fairness versus exploitation are ongoing considering issues like *net neutrality* (especially vivid in USA) or *upload filters* (especially vivid in Europe). This should not come as a surprise. While our societies may have improved fairness over the last centuries at large, there is an ongoing need for improving understanding and discussing as to what constitutes (un-)fairness – also in the Web.

2.5 Sustainability vs. Growth

The Web happened during what has come to be called the anthropocene, the current geological age during which human activity has been strongly influencing the climate and the environment. As a global artefact the Web continues to grow to cover and connect ever more people, usages, cultures and resources. But leading the Web to its full potential requires us to strike a balance between its growth and its costs of all sorts: the economical cost of the infrastructure it requires (e. g., Internet, telco networks) and the maintenance of its resources (e. g., maintaining and evolving Wikipedia), the ecological cost of its deployment and use (e. g., Google's post on powering a search[9]), the educational cost of the experts it requires (technical aspects, content production, legal framework, etc.)

And as we will see in Section 3, the encounter of the Web with AI can both amplify or alleviate this problem [80] as some AI techniques can be used to downscale the resources needed to operate Web applications [41] and at the same time some of the AI techniques are consuming a lot of energy [73]. This is also the case with other technology the Web is linking to such as blockchain [78], the internet of things [39], or mobile and ubiquitous computing, to name a few.

The emergence and scaling of new capabilities and usages of the Web are not only raising questions about our ability to sustain the evolution of the Web in terms of energy and money but in other domains as well. In the educational domain the Web curricula keep growing with new topics, standards, frameworks, tools, etc. In the legal domain each evolution requires

[8] See also the ACM FAccT conference series on Fairness, Accountability, and Transparency in socio-technical systems. `https://facctconference.org/`

[9] `https://googleblog.blogspot.com/2009/01/powering-google-search.html`

the legislator to monitor and control it closely. In other words it is important to not only consider the full potential of the Web, but also the one we can afford.[10]

We will therefore use the term "sustainability" in two of its many meanings. The first is *ecological sustainability*, i. e., questions regarding the long-term viability of our planet and the life it supports. The second meaning is the *sustainability of the Web* itself, i. e., the long-term viability of the Web as the complex socio-technical system that is evolved by society and that Web Science co-develops and studies. In addition, in Section 4.4 we will briefly sketch the connections with the UN's Sustainable Development Goals, in order to highlight the close relationship of our notions of sustainability with human rights and social goals. We understand *the Sustainable Web* to mean a Web that is long-term viable in this sense – by considering the goals of ecological sustainability as well as other goals.

2.6 Ambivalences without an End

We have discussed ambivalences of the Web today. It is germane to the Web as a powerful instrument that such ambivalences that affect human well-being for the better or the worse will never disappear. However, we, using Web Science as a discipline, must ask ourselves how to deal proactively with these ambivalences. In the following sections, we will try to address this question starting with the challenges that the encounter between AI and the Web is bringing.

3 The Web linking all Forms of Intelligence

Half of humanity uses the Web directly, and nearly all people are affected by it. But this is only a fraction of the agents who contribute to the Web. More and more human and artificial forms of intelligence are connected to the Web and participate in its content, structure, services and all the activities it supports. This section focuses on the opportunities and challenges of a Web that links all forms of intelligence and the role Web Science and AI have to play in providing a multidisciplinary approach to enhance the opportunities and to contain the problems such as the ones we identified in the previous section.

The first section introduces this evolution as a continuation of the historical trend toward a Web of everything and on everything. We then divide the challenges in evolving Web Science to study a Web that includes AIs into three sections: challenges at the individual level (e. g., a user, a software agent), challenges at the level of collaborations of different forms of intelligence (e. g., hybrid societies over social media), and cross-domain challenges to highlight the inter-disciplinary nature of the research needed. Before closing this section we stress the specific potential and challenges in using AI to study and manage the Web and to help us conduct Web Science research at Web scale.

[10] https://www.w3.org/community/wwca/

3.1 A Historical Trend toward a Web of Everything on Everything

When the Web started to spread at the beginnings of the 90s, it was perceived – from a technical point of view – as a hypertext system distributed over an open set of servers, in particular over the Internet and described using metaphors of a universal library formed by linked HTML documents called Web pages and primarily for human consumption. Social Scientists paid less attention to the technical specifications, but recognised the emergent social phenomena that these enabled and drew on a range of metaphors to describe this, including the Web "as culture" and the Web "as community". In fact, as soon as it started to spread, the Web also started to reveal to everyone the full potential of the distributed hypermedia software architecture it had globally standardized and that goes far beyond the image of a universally shared library [35]. Although the Web was initially open to contribution (HTTP PUT and POST) this was not widely implemented in the first browsers and until the mid-90s, with wikis, blogs and forums, for the Web to be re-opened in read-write mode. This paved the way to numerous new social media applications and a rapid increase in the number of people using the Web as well as the scale of its content. In parallel, the initial use of programs to generate pages on the fly opened the way to a second major evolution of complex Web applications and dynamic pages leading to a deep Web far exceeding the librarian metaphor and effectively linking programs, as well as people [30]. As programs became more and more connected to the Web the need to exchange other data than pages became clear [7], and the Web first separated form (CSS) and content (HTML) to then be able to exchange different types of content (XML, RDF, JSON) and ultimately became a place for publishing and linking datasets. Both trends made the Web more machine-friendly and supported the development of Web-connected devices starting with mobile access to the Web and leading to the more recent notion of a Web of Things where the Web is now in position to provide a universal software architecture and framework to access the Internet of Things (IoT) and program its applications and interfaces.

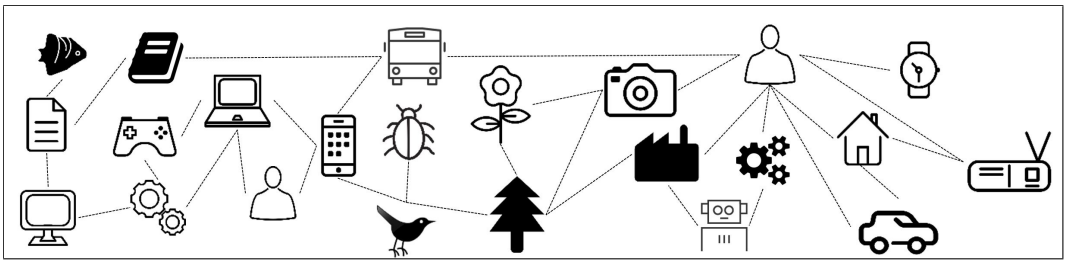

Figure 1 A Web of everything and on everything.

Despite these different evolution trends, it is remarkable to see that the Web architecture remains largely one and unique. One Web is now linking documents, people, data, programs, and virtually anything that can be identified and one Web is accessed via multiple means: computers, smartphones, watches and a growing heterogeneity of connected things. Together "Web of everything" and "Web on everything" are two viral and synergistic evolution trends of the Web: every new device connected to the Web provides more data and more resources to access and to link and, in return, every new piece of data linked to the Web supports new applications and provides resources for the connected devices.

In Section 2.1, we have identified the tension between the freedom to contribute information and the need to ensure the quality of the information we find on the Web. We can now extend that challenge to everything the Web will connect raising the problems of searching and finding relevant resources on the Web while ensuring security, certification and authentication in these connections we make.

As a result the Web has become a universal collaborative space for actors that consume and produce information including all different forms of natural and artificial intelligences. Web Science was set up to understand the Web and evolve it for the benefit of society, including the generation of opportunities and the protection of the vulnerable. The linkage of diverse forms of intelligences necessitates accelerated development of Web Science as a discipline. As Artificial Intelligence augments the means of the rich and powerful, Web Science must be developed to leverage opportunities for society as a whole and to protect weaker member of societies against negative ramifications of a Web of linked intelligences.

3.2 Challenges at the Individual Level: A Variety of Intelligences on the Web

Figure 2 Connecting and linking a variety of intelligence forms on the Web.

A first challenge we can identify is the one of designing artificial Web intelligence, i. e., the study of the specific problems in designing AI agents that are connected to the Web – its resources and users – with all the heterogeneity, different scales and speeds this involves. AIs connected to the Web must be robust and benevolent by design as they can potentially be in contact with billions of users.

Challenge – Benevolence: Ensure that AI brings no harm to Web users and to make AI Web bots benevolent by design.

A short term objective is the study of new AI approaches to traditional Web tasks that can benefit from AI, such as indexing, searching, or browsing but also more complex tasks such as fact checking [5], fraud detection, or protection enforcement to support privacy.

To the technical challenges identified in this section we should add upfront the extension of the tension identified in Section 2.2 between the ability to adapt and the need to respect privacy and confidentiality: The more AI techniques are deployed on the Web and its data the more important it becomes to monitor and control the usage of data they consume and produce. Inversely, but still on a short term perspective, the Web has the opportunity to become the most important source of data sets to feed data-hungry AI and the challenge here is to ensure the Web can provide high quality data sets with the metadata required to ensure trust and proper usages.

Challenge – Web Intelligence: Include more intelligence in classical Web tasks and more Web resources in classical AI methods.

The diversity of AI techniques in existence is an asset as it allows for addressing many types of diversity we find on the Web in terms of content, users, contexts, tasks, usages, resources, etc. Doing so will require the adaptation and extension to the Web of different forms of AI – machine learning, knowledge representation and reasoning, constraint solving, etc. All the domains of AI – symbolic AI to neural AI – should be considered in relation to the Web: The next revolution may come from a currently very quiet field and the broad

picture of the Web requires all types of expertise. The heterogeneity of the Web resources and actors in general requires a diversity of methods to manage them. Clearly, supporting different AI techniques is not enough to cater to a diversity-aware Web but the ability to combine different techniques might help in providing measures to contain problems. For instance, some statistical methods could be more robust to noise, some others could be able to detect malevolent behaviors or a bias [62], and other symbolic approaches could be useful in certifying the provenance or in explaining the process that produced a dataset or a result [58]. All of them combined could contribute to address the problem of quality data as input and output of agents on the Web.

Challenge – Web Resilience: Design a variety of AIs resilient to the Web, its heterogeneity, different scales and speeds, its noise, open-world nature and uncertainty.

Another challenge is to harness the multidisciplinary nature of Web Science to identify new ways of simulating, reproducing or engaging intelligence including emotional intelligence, communication skills, or imagination. A diversity of intelligent skills and new ways to simulate them in the open-world context of the Web are needed. On this point the Web is both raising questions and providing first answers by proposing a universal space to link different forms of intelligent processing such as knowledge-based systems, statistical learning, learning by reading, wisdom of crowd, etc. Many different sources of intelligence can be found or connected to the Web.

Challenge – Intelligence Diversity: Understand and harness the diversity of intelligent skills and behaviours that we need on the Web.

The advent of AI on the Web requires Web Science to evolve toward the studying of a Web that includes AI agents as first class citizens. The first challenge is to make the Web "AI-friendly" at the core of its architecture by providing knowledge level means to connect an AI to the Web. Beyond semantic Web, linked data and knowledge graphs [45], formalization and models are needed that support the publication and exchange of the many types of data an AI may want to obtain or provide on the Web including: input data, training data, output data/results, parameters, configurations, schemata, embeddings, etc. All kinds of semantic information networks can potentially be exchanged by AIs on the Web including complete neural networks, layers or sub-networks.

Challenge – AI Friendliness: Augment the Web architecture and standards to turn the Web into an environment fully supporting the hosting of AI agents.

Compared to other AI domains, an important specificity of AI on the Web is that it will systematically be in contact with a huge variety of human users. Human-AI interaction on the Web is an important challenge for interaction design and HCI. In Web-based interaction with AIs, an increasing number of users can interact with many different AI agents using and combining multi-modal means such as written chat, voice recognition, gesture recognition, multi-touch, etc. Moreover, the Web also holds the potential to augment interactions in particular by providing background knowledge and contextual data that can feed AI, assisting smarter interactions and supporting, for instance, context awareness [37]. In other words, the challenges are both to use the Web for better AI interaction and to use AI for better interactions with the Web.

Challenge – Positive User Experience: Design advanced human-AI interaction on the Web and assist human-Web interactions with AI support.

Finally a special case of interaction with AI is the generation of explanations and justifications to support trustable human-centered AI on the Web. The challenge is not only to generate single explanations but to weave a Web of explanations and justifications allowing users to follow their noses in linked traces possibly distributed over the Web (distributed data sources, distributed processing) to understand and take ownership of the results and to achieve decision making.

Challenge – Trust Building: Provide user-friendly explanations, justifications, and traces of AI activities and results to foster trust and support decision making.

Design for user-friendliness and understandability of explanations should also respond to diverse abilities and needs (accessibility), and designers must not forget the interests of indirect stakeholders, who are not necessarily using the Web but are still affected by it.

3.3 Challenges at Group Level: Collaborations of Different Intelligences on the Web

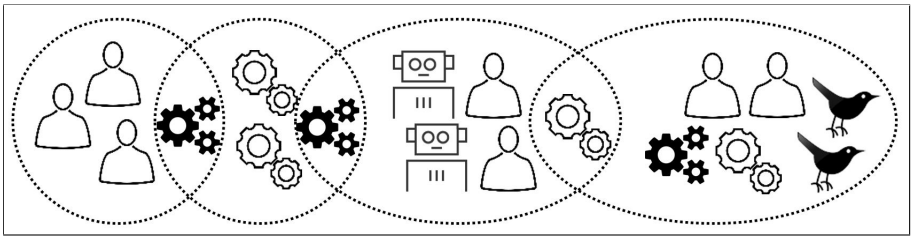

Figure 3 Shaping hybrid societies of intelligence types on the Web.

The simple fact that complex tasks combine several intelligent abilities – learning, reasoning, planning, solving, analyzing, extracting, communicating and collaborating – shows that Web Science will have to consider the Web evolution toward a multi-AI system. This evolution is both required and allowed by the distributed Web architecture leading us to the challenges of building and managing collective intelligence on the Web.

The long term challenge for Web Science is the study and design of human-agent collectives and social machines including AIs that form hybrid Web societies which include very different types of intelligence: people, connected animals, connected plants, Artificial Intelligence that reasons, that learns, or that induces. Consequently we need to study on the Web various combinations of possible collectives, interactions and their combinations: human-human, human-AI and AI-AI. Web Science must consider the different types of groups that will be formed, from swarms to complex societies with their normative rules, roles, and social constructs.

We have already mentioned that these hybrid communities will require new means to bridge the semantics used by machines and the semantics used by humans in their interactions [36]. The challenge will also be to study the interactions of Web AIs with the resources of the Web (linked pages, linked data, connected objects, etc.) forming the environment of these forms of intelligence. Moreover, beyond individual interactions, a challenge will be to study and design the orchestration of these multiple interactions. Web Science will have to face the problem of this massive interaction design with the Web and everything it links and AI will have to face the problem of engaging in very different types of interactions with different forms of intelligence including different kinds of AIs.

Challenge – Hybrid Communities: Support the formation of groups of human and artificial agents and the interactions within these groups.

In Section 2.3 we mentioned the tension between the wisdom we can gain from the social activity of the Web and the risk of mass manipulation. This perspective of hybrid communities is amplifying that challenge of fostering the beneficial social phenomenon and preventing the dangerous ones as we are putting together billions of users and millions of software agents.

For instance, one may study the role AI can play in human-agent teaming and the role of a personal assistance to team members. Moreover, the notion of norms raises the question of having "laws of Web AI" since a Web AI is at least as dangerous as a robot in terms, for instance, of harming users by learning and disclosing sensitive knowledge. A complex question for Web science would then be to incentivize the development of benevolent Web AI through ethical principles of Web Science research and the study of normative systems for hybrid societies. Going even further, one could envision the design of Web AI to observe other Web AI, watchdogs checking the behaviours of other AIs to detect and protect against unwanted (e. g., unauthorized face recognition), malicious (e. g., forbidden fingerprinting) or defective AI (e. g., AIs with a bias).

Challenge – Social Contracts: Support the establishment and enforcement of social rules governing human and artificial agents behaviours within their hybrid communities.

To help communities life cycle and routines and to help users in participating in large scale groups with large collections of resources, Web Science should study Web-dedicated AI taking part in online social activities. These goal-driven agents could actively participate in the online activity and, for instance, foster linkage, interactions and convergence, bridge, translate, check, or augment our posts and maintain for us an overview of our social context and activity. They could also prevent or report problems such as bullying, harassment and polarization. Each of these tasks requires dedicated intelligent processing that would need to be resilient to the variety of users, (mis)usages and content one can find on the Web. Ultimately, a very ambitious goal would be to have AIs with skills about human social interactions and psychological characteristics to help humans face their own humanity on the Web and help us, as individuals, reach the World-Wide-Web scale.

Challenge – Online Intelligence Augmentation: Support human users online by helping them understand the complexity of the Web, its (mis)usages and content, and face humanity in all its complexity on the Web.

This macro perspective also shows that Distributed AI (DAI), in particular, has a *rendez-vous* with the Web and its sciences because the distributed nature of the Web calls for distributed AI approaches. Multi-agent systems and distributed AI blackboards are examples of distributed AI architectures which, if merged with the Web architecture would allow for many different kinds of AIs to collaborate worldwide to the benefit of the Web. The AIs and the multi-agent systems would also in return benefit from the Web, its resources, its standards, its users and its methods. This trend has already started in some domains (e. g., B2B) but is at a very early stage and a conceptual challenge is to study the merging of these architectures to make the Web DAI-friendly and vice-versa as, so far, these approaches are more often juxtaposed than merged [13].

Challenge – Distributed Intelligence: Bring together the architectures of distributed AI and the architecture of the Web.

In Section 2.5 we have seen that we need to address the sustainability of the growth of the Web and of AI techniques [80] and we will have to add to the need for new social contracts a further need for the Web to respect the limits of growth establishing a natural contract [68] with our environment.

4 Governance and Regulation of People and AIs in the Web

The ambivalences (Section 2) and challenges (Section 3) of the Web require more than interdisciplinary analyses and technical progress. They also require systems that regulate the interactions of players on the Web, both human and artificial – i. e., governance. Governance comprises all of the processes of governing, whether by a government or other actors, over a social system, through laws, norms, or power (based on [9]). We include code (or, more generally, software and hardware) in these categories in line with the adage that "code is law" [51]. Web governance is complex because moral values, jurisdictions and even fundamental rights differ across the globe and also change over time. Different values have to be balanced against each other. What is more, governance itself is a complex undertaking, especially as much that happens on the Web is not governed by laws alone.

4.1 The Scope of Web Governance

To investigate *Web governance*, one first needs to delineate what falls inside the scope of the term. It is closely related to "Internet Governance", a term that, in the early years of the Internet, used to refer to the technical core of the Internet: standards for data transmission, the management of IP addresses and Domain Name Service, and the operation of the root servers. The concept soon grew in scope. According to the definition published under the auspices of the UN in 2005, Internet Governance comprises "the development and application by governments, the private sector and civil society, in their respective roles, of shared principles, norms, rules, decision-making procedures, and programmes that shape the evolution and use of the Internet".[11] We propose to delineate "Web governance" analogously, by substituting "Web" for "Internet" in this definition.

In spite of the continuing importance of the governance of the technical basis (such as Web protocols), attention is thus shifting to a wide array of social governance questions. Questions center on how much influence each sector of society should have on the development of the Web and how issues such as freedom of speech and censorship, attempts to influence elections, misinformation, hate speech and cyber-bullying, data protection, privacy violations and surveillance should be approached. In addition, with the rise of collaborations of different forms of intelligence, the laws, norms, powers and languages governing AI are more and more relevant for how the Web evolves.

With the Web becoming increasingly interwoven into all fabrics of life, it could be argued that all societal norms, rules etc. co-shape the evolution and use of the Web. This could render the term "Web governance" overly wide or even redundant. However, it still remains possible and important to delineate it. In the following section, we will illustrate such

[11] `https://publicadministration.un.org/en/internetgovernance`

attempts at delineation with the help of examples, concentrating on legal approaches that show how the non-trivial interactions between different realms of governance play out in practice.

4.2 Balancing Ambivalences through Legal Web Governance

Laws have been and are being implemented by states and supra-national bodies such as the European Union. These laws govern what is legal (or not) on the Web and how violations of these rules can be sanctioned.

"The use of the Web" (as one domain of Web governance) can comprise, in principle, many domains of human activity. For example, people may use the Web to plan a murder. The associated communicative acts may constitute a conspiracy to commit a crime, which in itself is sanctioned by the criminal law of many jurisdictions. The legal status of the acts may not be affected by their happening online, and in this sense this "use of the Web" is not an object of Web governance. However, states increasingly pass laws and take actions to intercept such communications (in attempts to prevent crimes from happening), and it is *these* laws and actions, which do fall under the heading of "Internet/Web governance", that have manifold implications also on other communications. For example, attempts at targeted surveillance of suspicious behaviour and individuals often results in dragnet surveillance of unrelated individuals. Dragnet surveillance is not only unethical, but in many jurisdictions also illegal (cf. the arguments around NSA surveillance brought forward in connection with the Snowden revelations). Similarly, the push by some governments for backdoors into encryption schemes would not only allow law enforcement to intercept communications that they may legally intercept, but it would break the confidentiality of all communications, thereby infringing on the rights of innocent individuals. Thus, privacy violations can result from the mathematical-technical aspects of the Web as a communication medium.

Challenge – New Interdependencies: Mathematical and technical aspects of digital communications can lead to mission creep and collateral damage of governance measures, and Web governance needs to take these interdependencies into account.

In addition and as discussed in Sections 2.2 and 3.2, privacy violations can be a flipside of personalisation and other data analyses. Data protection legislation is one form of governance addressing these risks of "uses of the Web". We will explain the interactions with Web governance using the example of the General Data Protection Regulation (GDPR), the European Union's (EU) data protection law, which regulates also non-EU actors if they process data of EU citizens (this includes search and social-network giants). The GDPR regulates the processing of personal data, and its first top-level aim is to protect the fundamental rights and freedoms of natural persons. The law therefore covers the processing of personal data in any form, ranging from hand-written data to big data on the Web. At the same time, much of today's personal-data processing *is* Web-based, and a second top-level aim of the GDPR is to "ensure the free flow of personal data between [EU] Member States", a type of processing that *typically* involves Internet/Web structures. Core GDPR principles and more specifically regulated concepts include general issues around personal data such as purpose limitation, data minimisation, and transparency and accountability, but they also include issues that *typically* (even if not exclusively) arise in Web-based and Web-scale data processing, such as profiling. In addition, there are uses of the Web and other data sources that have developed through Internet/Web-level-scale processing of personal data (such as group profiling) which are *not* regulated well through this law, which focusses on protecting

individuals. In addition, some of its mechanisms (such as the large reliance on individual consent as a lawful ground for data processing) are increasingly regarded as ineffective in practice. As a consequence, many – but not all – characteristics of data processing that emerge from the technical and applications infrastructure of the Web combined with the way people use the Web, become the subject of the GDPR, giving the law an important role in Web governance.

Profiling based on personal data can not only violate rights to data protection and privacy, it can also lead to unfairness and discrimination (Section 2.4). The GDPR recognises this threat in a recital, thus clarifying that discrimination must not result from data processing, but it does not specify concrete rights or duties. Since discrimination is illegal anyway, and Web-related laws thus as of now refer to general anti-discrimination laws, we expect that this will be an area in which the *case law* side of governance will play an important role in regulating discrimination via networked AI (cf. the court case against the personalised delivery-assignment algorithm in the app of a food delivery company, the first of its kind in Europe, in which the plaintiff trade union claims discrimination [23]).

Challenge – Limitations to Coverage: Laws need a certain degree of technology agnosticism in order to protect effectively and for extended periods of time. On the other hand, they need to respond to technology-typical effects. This will usually lead to some threats not being covered. Case law, and other forms of governance, are needed to compensate for these gaps.

Some threats of mis- and disinformation (see Section 2.1) are addressed by laws. Being untruthful does violate laws in some circumstances: when being a witness in a court of law, when reporting business results as a public limited company, with respect to information given out as a public official, etc. Being untruthful about a person is sanctioned by laws in many jurisdictions when the untruths can have severe impacts on a person, e. g., in cases of defamation. Other laws outlaw types of speech such as threats of criminal acts against the person or incitement to hatred.

On the Web, the lines between such hurtful speech become blurred, and the increasingly used term "hate speech" captures many different phenomena. Some forms of hate speech therefore fall under the governance of general laws. However, Web-specific aspects including speed and range of distribution as well as business models of social media platforms, have led to a proliferation of such forms of speech and to platforms arguable neglect of their duties to counteract such content.

Some countries have therefore created specific laws that constitute specific *Web* governance. An example is the German Network Enforcement Act (NetzDG), designed to enforce a functioning complaints management system in large social-network platforms, imposing heavy fines if these do not remove content that has been flagged to them as violating specific illegal forms of speech. This approach – to refer to laws outside the narrow scope of Web governance, coupled with Web-governance-specific legal measures – however fails to cover forms of hate speech that are not currently covered by laws (such as doxing[12] and other threats). The boundaries between accepted and non-accepted speech are changing, and therefore it can be expected that legal regulation will change as well.

Other countries and the EU itself have implemented or are debating laws specifically against misinformation, hate speech, or cyber-bullying [25]. In the wake of allegations concerning election manipulation, some countries have decided to regulate speech in pre-

[12] Broadcasting private or identifying information, for example people's home address.

election times, and to address specifically influence by foreign-national actors.[13] In general, the balancing between free speech and the protection of individuals who may fall victim to the free speech of others, is handled differently in different countries, as witnessed by the US view of freedom of speech as a near-absolute fundamental right (cf.,) e. g.,, [32]).

However, there are no *general* "laws against lying", and there probably should not be such laws, given the fine line, in many contexts, between mis-/disinformation, bias, and freedom of speech, as well as the positive roles that omissions and even white lies can have for social cohesion. This can become a serious problem if (as with climate-change denial speech) lies affect and endanger all of humankind (and the planet as a whole). Our legal systems generally focus on protecting individuals and are therefore ill-equipped to handle such threats. The question is whether specific lies will come to be recognised as illegal (such as holocaust denial is in Germany). Similarly, there are and probably cannot/should not be general "laws against expressing hatred", and the balancing of the fundamental rights of the haters and the hated remains an ongoing challenge.

Restrictions on which content may be distributed via the Web may also be based on a balancing against other interests and rights, such as intellectual property (IP) and copyright. Laws that have created wide-scale protests such as the EU Directive on Copyright in the Digital Single Market (which needs to be implemented in national laws in the Member States by June 2021) attempt to enforce proportionality in these balancing acts. Arguably, these protections too can contribute to mis- and disinformation (Section 2.1), as when IP-based restrictions of access deprive the public of relevant information (e. g., scientific publications not in open access, or copyright laws used to combat critical reporting online [50]).

These examples of phenomena that are in need of, and increasingly addressed by, forms of law-based Web governance also demonstrate that conflicts are inevitable. Typical conflicts between criminal justice and fundamental rights such as the privacy of communications have already been described in the first example. The protections afforded by data protection and protections from hate speech and related phenomena, as well as the protection of intellectual property, on the other hand, often conflict with interests in the freedom of information or freedom of speech. Openness and freedom of speech, and also decentralisation, have been founding principles and values of the Web community. Not only can this tradition conflict with the restrictions on speech imposed in different jurisdictions (whether Web-specific or general, see [22] for an overview of how restrictions on speech are key elements of most jurisdictions), but these restrictions also differ between countries and cannot easily be enforced globally.

Challenge – The Need to Balance Rights, and to Recognise the Effects of the Context of *Web* Speech: Web ambivalences are often also conflicts between different rights (often fundamental rights of natural persons), and these always have to be balanced in governance systems that consider themselves fair and just. At the same time, laws and case law have to recognise the specific effects of speech on the Web, such as its huge speed, reach, and therefore possible impact.

4.3 Governance of Artificial Agents

The examples above all centered on "uses of the Web" by natural persons as well as private and state organisations. With the role of AI increasing and hybrid societies becoming Web

[13] France's law concerning the fight against the manipulation of information (loi relative à la lutte contre la manipulation de l'information) is an example of this.

users (see Section 3.3), Web governance will also have to regulate the "uses of the Web" by artificial agents. One example are autonomous agents on the Web, such as bots, which can play major roles both in spreading information and misinformation via the Web (see Section 2.3). What should these agents be allowed to do or not do? Who will be liable when things go wrong? Lawmakers may in the future draw on current activities such as regulations concerning what autonomous vehicles are allowed to do and not do on public streets, and how questions such as liability for accidents are handled [46].

Challenge – Governance of Artificial Agents: Should artificial agents become first-class citizens on the Web also in legal terms? What rights and duties do they have, and how can responsibility, liability, and sanctions be regulated?

4.4 Beyond Laws

Norms as another approach to governance are developed and enacted at different levels. One is corporate governance and self-regulation. For example, with the help of human fact-checkers and automated classification, social media companies have tried to combat the spread of misinformation across their platforms, sometimes in efforts to obviate legislation.

One advantage of non-legal governance processes is that ethics-related desiderata that are difficult or impossible to formalise into law can be addressed. This may not be instigated by management, but also or only by workers and employees. Recent examples are the protests by Google employees against the company's participation in weapons development [69], or the price surges created by Uber drivers in an attempt to raise their earnings to acceptable levels and raise public awareness of their working conditions [42].

NGOs and not-for-profit organisations are further types of governance actors. The focus may be on how to govern the organisation itself (an example is the Wikimedia Foundation Governance system) or on monitoring other actors (an example is AlgorithmWatch).

Ethics-based desiderata and governance norms may also arise via codes of conduct that professional associations give themselves (relevant examples for Web Science include the AoIR Recommendations for Ethical Decision-Making and Internet Research [63] and the IEEE Ethically Aligned Design Guidelines [31], the ACM's Statement on Algorithmic Transparency and Accountability [16]) or that national or supra-national bodies create (such as the EU's 2019 Ethics Guidelines for Trustworthy AI [14]).

Professional norms will remain key to addressing some of the ambivalences that cannot be addressed by laws alone. For example, journalistic codes of conduct mandate truthfulness and objectivity, and media regulation measures and attempts to foster diversity to both allow for desired biases and combat undesired biases. For these measures to succeed in mitigating mis- and disinformation, they however need to be combined with regulation of platform providers and changes in the consumption and sharing behaviour of end users.

Code itself is an instrument of governance [51, 65]. For example, the music industry has witnessed many attempts to technically block behaviours that were seen by the industry and some creators as detrimental to their interests: the sharing of media files between users. The code-based approaches include Digital Rights Management (to control access and copying), streaming (to not even create a media file on the user's device) and content inspection and upload filtering algorithms (to detect unwanted storage of copyrighted content).

Last but not least, infrastructures are instruments of governance. This fact has been thrown into stark relief by recent debates over who owns and controls the infrastructures of connected mobile devices (one of the most important physical platforms for Web usage). Early in 2020, the Covid-19 pandemic spurred a worldwide push to develop smartphone

apps for contact tracing and exposure notification, designed to break infection chains. A debate evolved around the question of whether software protocols should be centralised or decentralised. Many countries and regions opted for decentralisation, arguably mirroring (among other things) decentralisation as a fundamental principle of the Web. Code and tracing apps built on either decentralised or centralised protocols can comply with the GDPR [27, 60].[14] However, both types of systems are limited by restrictions posed by the devices, their operating systems and overall infrastructure, which are built and controlled by Apple and Google, the quasi-monopolists of the global smartphone market [70].

Power relations and perceptions change over time, and this impacts possibilities and impact of governance. For example, the GDPR is now, after much initial industry lobbying and claims to the effect that it would hamper innovation, hailed by many as an example for legislation protecting personal data and privacy. Conversely, after more than a year of NetzDG application in Germany, there appears to be no evidence of the over-blocking by social-network platforms that had been feared by many civil-society advocates. As the Snowden revelations have illustrated, laws on national security passed by dominant nations govern how individuals around the world can (or cannot) effectively exercise their fundamental rights and freedoms, including those afforded to them by their own countries' laws. Tech giants wield enormous power over citizens and conceivably even over elections and thus the fabric of democracy.

It has also been argued that the blocking of content that violates economically powerful companies' copyright, enters laws and is enforced faster and more thoroughly than the blocking of content that violates vulnerable individuals' and groups' personality rights, even though the technical processes to detect and block such content may be similar. Within companies, ethics-related desiderata formulated by management become binding rules for governance more easily than those formulated by workers and employees.

Power is not in itself bad or good, and it exists and will exist in one way or other in all communities and societies. What governance can and should do is control power and ensure that the powerful are held to account. Importantly, governance has to create structures and processes that ensure that control and accountability will *continue* to function and to make sure that human rights are and remain protected. The global agreement on human rights and social goals is best described by the UN's Declaration of Universal Human Rights and the UN Sustainable Development Goals (SDGs). The discussions around the SDGs show the interdependencies between human rights, sustainability in the ecological sense, and peace, justice and strong institutions (the latter specifically named as SDG 17). While this holds for economic and social rights such as food and health and for justice against violent crimes, similar interdependencies hold for political and civil rights such as those discussed in the current section, and for justice against violations of these rights. We therefore posit a challenge modelled on SDG 17.

Challenge – The Need for Justice and Strong Institutions: All governance systems are faced with (often economic) power and interests. The ambivalences and conflicts between these have to be addressed, and power has to be controlled and held to account, by strong institutions focussed on justice.

[14] The European Data Protection Board adds that "[i]n general, the decentralised solution is more in line with the [data] minimisation principle" [27, p. 9].

4.5 Global Web Governance

It appears likely that any global, Web-scale agreement on rights and on the need for weighing different rights against each other in governance processes will only be possible at the level of the UN. The ongoing quest for a declaration or even bill of digital human rights [40] may realise the ROAM Principles, endorsed in 2015 by UNESCO's General Conference, which state that the internet should be "(i) Human Rights based (ii) Open, (iii) Accessible to all, and (iv) Nurtured by Multistakeholder participation".

As discussed in Section 2.5, it is becoming increasingly clear that in addition to human rights, we also need to think of the rights of other beings, nature and ultimately the planet. Towards sustainability in this sense, local, national and global governance activities are needed. Towards this end, Web users and Web science will have to ask themselves the same questions that need to be asked about non-digital resource use. The Web in its current form and development consumes large amounts of energy and resources [48, 80], and it is poised for growth: streaming content, higher bandwidth, the push to ever more documentation and (over?)sharing. Is it desirable and realistic to reduce these activities and thereby the Web's resource consumption? Or is the only possibility "green growth" based on renewable energies and materials? These questions and decisions will contribute to the already well-known local and global conflicts concerning climate accords. In addition, they will interact with Web-specific conflicts of interest. For example, should the trend to ever more streaming of content be stopped by less energy-hungry text formats? How would this affect accessibility or the delicate balance around intellectual property rights (see Section 4.4)?

Challenge – Global Collaboration on Web Governance: For governance to succeed and to contribute to sustainable outcomes, many stakeholders and forms of governance have to collaborate, across sectors and across the globe. Conflicts are unavoidable. Structures and processes for effecting such collaboration need to be (re-)negotiated and maintained continuously.

5 Extending Web Science to Address Ambivalences of the Web

Understanding the Web, its current and future trajectories through AI, demands that we craft new approaches at the intersections of traditional disciplinary practice. As the object of our study evolves and the necessity for collaboration across disciplines and with researchers in industry and government becomes ever-greater, it is urgent that we innovate our methods and research infrastructures accordingly.

Key challenges for Web Science at present are as follows:

- How can we unlock the promise of interdisciplinarity inherent in the vision of Web Science?
- How can we engage with AI for Web Science research?
- How to share data, algorithms and tools between researchers and with industry?
- How to create inclusive and robust public debate about the Web and its futures?

In the following we explore how these issues might be addressed.

5.1 Collaborative Research and Shared Infrastructures

Drawing together strong social research questions with engineering approaches allows Web Scientists to accomplish valid and ethical social science research with Web data towards

understanding and impact. Three clear benefits of such a collaboration are (i) that both communities can unlock new sources of data for Web Science research (ii) that social science training includes a strong focus on critical analysis of ethics, methods, and effects (iii) and that engineering has a more intuitive knowledge and understanding of industry practice (by inference at least). There are various examples of how such collaborations are starting to manifest in shared infrastructures. New models for access to data and tool sharing are starting to emerge, from Web observatories [77] to online tools specifically aiming to Open Web data to interdisciplinary communities[15]. Examples include:

- building Web Science data archives, data management tools (for documenting research processes throughout a project's lifetime) and infrastructures
- sharing algorithms across the academy and through dialogue with industry partners to develop innovative practices and mechanisms for collaboration
- actively seeking collaborations across domains and keeping the door open for including new disciplines in Web Science

Such infrastructures must benefit from the latest AI techniques to support mixed research methods at a Web scale: For example, AI techniques can help researchers (and data archivists, curators and others) perform a number of tasks required for collecting and compiling datasets that adhere to scientific best practice and are of shareable quality, such as detecting bias or enforcing anonymity in the data.

Challenge – Collaboration and infrastructures: Collaborative work and shared infrastructures such as observatories and repositories allow for leveraging the strengths from different communities for everyone.

5.2 Boundary Objects for Interdisciplinarity

Academic disciplines work with a variety of objectives and thus have developed often incommensurable epistemologies. Especially, there are very different ideas within the social sciences and computer sciences as to what constitutes a method or how knowledge should be derived. Often there is a misalignment of interests, for example, if social scientists would like computer scientists to become "epistemic partners" [12] in the collaborative production of theories and knowledge while computer scientists would like social scientists to inform the design and building of content and infrastructure and thus expect a particular type of results from the research [49]. Web Science has been calling for interdisciplinarity since its beginning and has been providing platforms for interdisciplinary work for a long time. There is a large potential in integrating a comprehensive understanding of the data with sophisticated computational methods [43]. However, transcending established knowledge frameworks to build new understandings is difficult and demands not only to take risks and to deal with uncertainty but also requires patience and indulgence.

Finding common points of interaction and discussion can be beneficial to an interdisciplinary process. For example, visualisation offers particular potential for interdisciplinary collaboration, to draw together technical methods for data exploration associated with Data Science with domain expertise from other disciplines. Whilst data scientists can identify patterns, they do not (usually) have the expertise to interpret the results substantively. And, whilst social scientists and humanities scholars understand their field of practice in detail,

[15] http://mediamonitoring.gesis.org/

they do not (usually) have the expertise to engage with data science, at least not to the depth of the possibilities that are emerging with AI. Visualisations may operate as "boundary objects" [10] where data science methods can display results in a way that is intelligible by domain experts, and domain experts can question the techniques and methods that are used for data interrogation, e. g., by ruling out irrelevant patterns or suggesting changes to the parameters of data interrogation.

Visualisation is thus not only about communicating results and presenting findings: It is required to make results understandable, but it needs to be a concern from the beginning, not just at the end of a research project [33].

Visualization can enable an interdisciplinary "interviewing" of data – of enabling a dialogue between data, method and theory – using iterative visualisations to support abductive reasoning. The complexity of data makes formulating hypotheses very difficult, and one way of allowing such formulating could be through a dialogue where we query the data and refine ideas in an iterative process. This could allow computer scientists and social scientists to work together, to bring theory and data into dialogue much earlier in the process, by looking at the same data, asking questions, visualizing the data in different ways. An example is Visual Analytics / Visual Data Mining, where questions are derived from the data, to derive new hypotheses.

Questions could then become how we might be able to map interviewing methods used in the social sciences (e. g., expert interviews, semi-structured interviews employing an interview guide, biographical interviews) to how we might interview data; can we map social science methods of interviewing to different computation techniques? Can visualization allow us to annotate quantitative data with qualitative information?

Collaborating in crafting and interrogating visualizations would allow us to innovate in interdisciplinarity, to find new ways to collaborate on an eye-to-eye level, especially between computer scientists and social scientists.

Challenge – Action: Joint "interrogation" of visualizations is a way to bring different methods into dialogue from the beginning and to unlock the potential of interdisciplinarity for understanding the Web and its intelligences.

5.3 Engaging with AI for Web Science Research

Over the coming years, Web Science will focus closely on the emergence of Artificial Intelligence and its implications for the Web. However, AI tools also hold various potential for Web science research itself, enabling us to observe, analyse and intervene in the evolution of the Web and may even serve as boundary objects for interdisciplinary processes.

In line with the focus detailed above (i. e., accessing the strengths of interdisciplinary collaboration by building shared and collaborative infrastructures) one interest of Web Science in the endeavour of both understanding and making use of AI is in the role of data and their management. The open and linked data facet of the Web is of particular importance when considering the links between AI and Web (science) data. Intelligent agents can help us produce, curate, share and maintain corpora and datasets. For instance AI techniques can be designed to check certain formal quality criteria of a dataset, look for bias in it, or address issues such as loss of context. AI can be also used for simulations producing data and to generate synthetic data, thus addressing privacy concerns. AI can thus be used to produce

datasets that are meaningful to Web Scientists from various disciplinary and methodological backgrounds (and in particular to social science inquiry). The study and design of high quality datasets allows their usage beyond Web Science, e. g., for the training of AIs.

Inversely, Web Science can produce multidisciplinary methods and tools to certify and characterize training sets to improve the quality of the learning and conclusion made by AIs using them. To support such an endeavour reproducible research and dataset sharing need to become a core feature in Web Science research. They are required to support the use of provenance to understand how content is generated. Data sharing architectures therefore are required to agree on standards for vocabularies, metadata (of various kinds) and criteria for establishing trust in the architecture. Importantly, this trust not only concerns the three dimensions of confidentiality, integrity and availability of data, but their more complex quality criteria such as transparency of data provenance, support for ethical scrutiny of data processes, support for assessing its validity, etc. to allow for reproducibility and (multidisciplinary) secondary use of the data.

In its endeavour to understand and engage with AI, Web Science can be seen as building a research program as a joint effort with two research fields born in the 50s: "AI" for Artificial Intelligence [56] and "IA" for Intelligence Amplification [4] and Intelligence Augmentation [28]. While AI focuses on methods to simulate and scale automated intelligent behaviours, IA focuses on methods to amplify intelligence with systems where human intelligence is at the center from the start. As a universal platform linking all kinds of intelligence, the Web has the potential of being the meeting point for AI and IA. But many open questions will have to be addressed in Web Science to reconcile these two historical trends.

For example, the Web already is populated by Web bots but they usually are restricted to certain realms while they could be generalized. How could we generalize the bots as the ones of Wikipedia to bots on the open Web designed to monitor and preserve certain characteristics of the Web? AI might be used to detect and counter-attack some undesirable network effects. Starting with defining the Web we want, how can AI help? For instance, could we design agents that would derive high-quality recommendations without having centralized data repositories?

As a first step, and from the infrastructure point of view, we could imagine Web farms for Web AIs hosting autonomous agents that would study, monitor and report on the Web. Then benevolent AI for the Web (e. g., watchdogs for the Web) could be used to regulate, as AI hold the promise of supporting diversity on the Web and of preventing unwanted behaviours such as bullying. Problems that could be targeted by Web bots include the detection of metrics manipulation (e. g., link farms), cross-language plagiarisms, centralization or digital divides, vandalism or spamming etc. The bot agents would need to be based on policies and values important to a philosophy of the Web that seeks decentralization and equality of access to improve the Web's resilience.

Web agents working to improve users' experience, understanding, awareness and control of their participation and contributions to the Web would also support such a philosophy. For instance, educational AI could help educate Web users in many domains including Web literacy or ethical thinking. Agents could also provide customized descriptions of the context in which a user is, including security, neutrality and privacy notices or his human-computing participation when it occurs. As mentioned before, AI could also help users burst filter bubbles and foster serendipity. In the long term, benevolent AIs could actively help enforce (human) rights on the Web and be scrutiny agents for important values of the Web.

At the start of this section we stressed that workflows and data are key drivers of Web Science; we then went on to explain how the multidisciplinary nature of Web Science also puts it in an ideal position to explore and expand the forms of intelligence on the Web. First, both Web Science and AI are highly multidisciplinary [34] and the multiple disciplines that

are common to both fields act as bridges to make them interact. AI could also be used to operationalize the expertise from each domain into agents that help us provide assistance, reporting or training from the domains they represent. These agents could help us find and support a massively multidisciplinary method and allow us to scale to the multi-disciplinary interactions required by the design and study of the Web. One possibility, for instance, would be for these AI agents to produce and maintain boundary artifacts at the frontiers of disciplines.

Challenge – Tooling: Leverage AI tools for Web Science.

5.4 Participatory Methods: a Web "For Everyone"

In thinking about the future of the Web, Web Science must engage with multiple and often competing stakeholders. The Web does not belong to anyone and it must meet the needs of everyone. As Web Scientists it is our responsibility to ensure that the full range of voices is heard as we build our understanding of the Web and seek to shape its future. In order to achieve this, we are committed to working with participatory methods that allow us to gain insight of diverse perspectives and build dialogues between these. Questions of ethics and responsible innovation have become pressing questions for Web Science. How do we ensure that Web technologies are ethical? That they are designed and implemented in ways that take into account potential harms? This is not simply a question of "ethical approval" for Web Science research but a wider "ethics of care" towards the kind of society we are making [55, 19], as we work to shape the Web towards the future. This means working with citizens as well as governments, community groups as well as businesses, to ensure that the future of the Web takes into account the diverse views, feelings and needs of all its stakeholders.

There are already some good models for collaboration with government and business – co-funding of research, knowledge exchange partnerships, even joint-research – and it is important that we extend these in the coming years. However we also need to extend collaboration with citizens and communities to enable meaningful participation and effective voice in shaping the future of the Web. These methods may include citizen science, whereby individuals participate in distributed models of data collection and analysis [11].

Similarly citizen social science can be used to encourage people to participate in research on areas of particular concern or interest to their lives and communities[16]. We should also consider how new and emerging forms of data can be used for community purposes, not just for recommender systems and targeted advertisements. For example, Web Science research has already developed environmental pollution apps to inform asthmatics which parts of cities to avoid [74], and built platforms for open data sharing and community-relevant analytics [1]. Beyond these excellent examples of "tech for good" we should also think about empowering citizens to be active in debates about the future of the Web, through education in schools, and through concerted efforts to build "anticipatory capacity" [3, 53, 2]; and thorough specific online mechanisms that encourage deliberative democracy [21] and actively seek to build widespread participation [75].

[16] See `https://www.socialsciences.manchester.ac.uk/social-statistics/research/projects/citizen-social-science-methods-for-social-research/` or `http://letters1916.maynoothuniversity.ie/`

Challenge – Action: Continuously working towards finding innovative, caring forms of citizen participation allows to build democratizing facets of the Web.

5.5 The Role of Web Science in Governing the Web

The complex governance processes described in Section 4 open up possibilities for intervention for individual researchers, platforms, other private and public actors, governments, and supra-national bodies. Example methods for intervention include, e. g., *(i)*, embedding of ethical concerns and values in the design and revision processes (such as value-sensitive design and reflective methods [17, 18]), *(ii)*, affirmative/corrective action, and, *(iii)*, the raising of awareness.

An important role for Web scientists is to study, describe and analyse the socio-technical system of the Web, for example to contribute a scientific view to supplant media and policy discourse that is often carried out in ideological and emotional terms. The gathering and analysis of a strong evidence base is an important task for Web Science. Principled evaluations of the processes and successes of legal as well as self-regulation efforts are another important area for Web-Science research. A further key contribution are critical analyses of the ideals invoked by laws and other professional, national and international agreements, such as the investigation of multi-stakeholderism in Internet Governance by [24]. Especially given the pervasive and dynamic power structures across the Web, *critical* research on governance and its evolution is another key task for Web-Science research and policy advice.

In hybrid human-agent environments, often unexpected situations arise that may require affirmative or corrective actions. Risk analysis during the design phase and before deployment may help preview some possible risks and to be pro-active as well as reactive. Existing laws may provide effective means against some risks, and unexpected means against other risks (cf. the proposal to use copyright law to combat revenge porn [52]). Emergencies can arise from feedback loops, and when these involve also or only artificial agents, they easily become fast and large-scale runaway feedback loops (as observed with Tay, a bot designed to learn jokes but ending up learning to be a racist from the human Twitter users it interacted with [61]; or the stock-market trading algorithms that created the 2010 Flashcrash [29]). Important questions about human and non-human agency arise that need to be addressed in order to ensure that Web content and services remain transparent and controllable and responsive to advice and, where possible, also draw on research findings. A positive example is the reaction of the social network platform nextdoor.com to the occurrence of racial profiling of neighbours by neighbours: Users who referred to "race" when posting to "Crime & Safety" forums were prompted to provide additional information, such as hair, clothing, and shoes, which helped reduce the occurrence of racial profiling [71].

In different roles, we are also responsible for informing users, the public, and other stakeholders. We can and should raise awareness about technical possibilities and facts (such as the presence and activities of bots [14]) and about what is regulated by laws (e. g., due diligence to clarify liability questions, bot content production in Wikipedia as a case of public speech). Information activities can aim at raising awareness, and they can also aim at changing behaviours in certain directions, from nudging to forcing. Nudging towards behaviours may operate via information about psycho-social effects and nudging towards attitudes such as empathy. Well-known examples from the Web include differences between online and offline behaviour (in terms of social checks and balances that are missing online), and the reductions in empathy often associated with cultural or spatial distance.

However, most informing and certainly all nudging or forcing rests on values and choices, and neither can be assumed to be "good" in and of itself [20, 6]. Self-reflection about these values and choices, and about whether and how to integrate which other parties' values and choices, is as important as the attempt to be "ethical" itself. This has impacts on the definition and implementation of the concept of "benevolence" that was introduced above.

Challenge – Action: Web science can and should leverage its interdisciplinary potential for shaping and analysing the Web along entire life cycles of Web-based systems: for analysing and demonstrating what works and what does not, for designing systems that take these insights into account, and for recognising when it hits the limits of academia and needs to collaborate with non-academic stakeholders.

5.6 Futures of Web Science

As Web Science evolves with the Web, so too it becomes apparent, and indeed pressing, that our methods evolve. At the beginning of Web Science this meant integrating robust social science methodologies with the engineering methods of computer science. This has proven both more difficult and more rewarding than originally imagined, and still we continue to learn about how to do it better. As we look towards Web Futures, we must innovate again. This will include engaging with AI for Web Science research. And it will also mean adjusting Web Science research to the emergent AI Web. As the Web becomes ever more heterogeneous, to include animals, plants, ships, rocks (and so it goes on) we need to think about how we research these new networks, their activity end effects, the opportunities and the challenges that they present, not only for Web Science but for the future of the Web. And we must do so in a context where the continued growth of the Web is a pressing question for the sustainability of the planet. Environmental Web Science is as yet in its infancy and we will need to draw in new disciplines, from the life sciences in particular, and new approaches that help us to research, understand and intervene in this truly inclusive Web.

6 Capacity Building

The Web is not fixed or finished. It remains open to possibilities, of all kinds. The very properties of openness, flexibility and dynamism that drove its growth now threaten its future (Section 2). As we move towards a Web of AI (Section 3), both the challenges and opportunities are amplified. We are already all too familiar with hate speech, cybercrime, violations of privacy with data and algorithmic bias, for example. These social problems and risks may become worse as AI drives new methods of information proliferation, individual profiling and targeting, and the means to override security systems. And there will be challenges that we cannot even imagine as yet – this much we know from the history of the Web. Yet, at the same time, AI brings opportunities. Opportunities to harness the Web in positive and progressive ways, to counter the morally unacceptable and criminal activities more efficiently and effectively than ever before and to provide resources for human prosperity and well-being.

It is a matter of pressing global concern that we support these opportunities and tackle the challenges, in the present and with a clear focus on the future societies that we are making. This is no simple matter. It raises

- *political questions* of competing interests, across sectors and between nations;
- *philosophical questions* about moral standpoints;
- *social questions* about inclusion, inequality and power;
- *legal, ethical and practical questions* about regulation, sanctions and enforceability; and
- *technical questions* about the architecture, operation and accountability of Web technologies.

These are questions for us all, and demand that we build the capacity to answer them as a matter of urgency. Getting to the answers demands a radical acceleration of the new forms of knowledge – across the established disciplines – that have been pioneered by Web Science. And it demands robust long-term participatory mechanisms to link research with the wider education system, with industry, policy makers and citizens, to create an ecosystem that can build ongoing socio-technical and inclusive capacity to face the challenges and shape the opportunities.

In this section, we outline the actions that can be taken by different stakeholders in the Web ecosystem to build this capacity and bring forward a future Web that is more inclusive, more intelligent, more sustainable, more collaborative.

6.1 Actions to be Taken by Researchers

Web Science researchers have made great strides forward in integrating social and technical understandings of the Web. They need to and will continue to apply the best methods from individual disciplines and to develop these methods further. However, more remains to be done to achieve the capacity required to face current challenges effectively.

1: **Supporting Interdisciplinarity:** The plethora of questions raised cannot be answered by any discipline on its own. Hence, researchers must resist "research silos". Researchers must avoid academic wars between schools of thoughts and rather work on bridging, benchmarking, and hybridizing. They must leverage or create recognized and highly ranked venues that award recognition for interdisciplinary results.

2: **Supporting Collaboration:** In spite of a momentum towards fragmentation [64], the Web is a global phenomenon that cannot be understood from a national or regional point of view. Hence, researchers must collaborate at global scales, they must actively share their problems as well as their resources. This includes the sharing of data sets beyond the boundaries of domains, cultures, or disciplines [72], and the dealing with cultural, legal and other challenges that may arise from such intentions to share.

3: **Supporting the Sustainable Web:** Research must provide facts and raise awareness about the implications of our digital lives on the environment and the climate informing organizations and government to support behaviour change. Research might support the individual, too. For example, research on interaction design might show which interfaces best reconcile computing objectives with environmental strains. In addition, researchers can contribute by developing and using more sustainable technologies and practices, and by highlighting rebound effects in which increases in consumption/deployment neutralise or even reverse such progress, as illustrated in Section 2.5.

4: **Supporting the Intelligent Web:** The Web is a showroom and testbed for AI research. With increasing power of AI research, there is a growing potential for AI research on the Web to harm people calling out for the protection of people when AI takes its infant steps on the Web. AI researchers need to adapt review boards as known from medicine and the social sciences to protect their work from overpromising the good and overdelivering the bad – for the sake of AI and the Web.

5: Supporting the Inclusive Web: The influence of research on politics is more limited than some researchers wish. To influence policy making, researchers must actively collaborate with stakeholder groups and recognize their contributions to society. Many stakeholder groups share the objective of shaping the Web for a better future. Positive influence in the Web comes from many stakeholder groups: NGOs (like International Fact Checking Network[17] or Freedom House[18]), customer protection agencies, labour unions[19], and many more aim to make the Web a better place. Researchers can build greater collaboration with partners in the Web ecosystem, to ensure more widespread and effective knowledge-exchange with businesses, governments, state agencies and civil society in order to accelerate understanding and innovation for an inclusive approach to the evolution of the Web.

6.2 Actions to be Taken by Academic Institutions and Funding Agencies

Academic institutions as well as funding agencies achieve effects by setting up or funding organizational structures and rewarding people. Rewards are issued along a range of well-understood incentives imposed by academia and funding agencies.

The cross-cutting concerns of the Web run counter to the organizational structures imposed by academic institutions and funding agencies as well as to major parts of the reward structures (e. g., promotions). To support the Web, academic institutions and funding agencies need to change. As most academic institutions and funding agencies depend on the political will that funds them, suggested actions must not be considered in isolation from political support for structural changes as indicated in Section 6.5 below.

1: Supporting Interdisciplinarity: Interdisciplinary courses are most effective if co-lectured by researchers from two disciplines, which increases the price per teaching unit. Interdisciplinary research is harder to publish and interdisciplinary basic research projects are scarcer and harder to get than disciplinary ones. Implied costs cannot be borne by the individual academic, as a further increased workload would jeopardize both their career and their health. Hence, reward structures must be modified towards supporting the novelty of approaches and the diversity of research and other outcomes in order to support interdisciplinary approaches.

2: Supporting Collaboration: A global understanding of the Web improves the prospects of all, students, researchers, academic institutions, and society. Organizational measures of sharing beyond institutional boundaries and funding initiatives may improve overall knowledge about the Web enhancing effectiveness and efficiency of academia. Mundane examples include information and resource/data set sharing, such as websites that list courses, introductory texts, or example data.

3: Supporting the Sustainable Web: Beyond sustainability in terms of overconsumption of physical and environmental resources, such as energy or atmosphere capacity for CO2, the existence of the Web is threatened by its ambivalences. Uncertainty, vagueness and conflicting objectives will require an evolving, but also never-ending academic dialogue. Academic institutions must respond to this need supporting a constant vigilance, scrutiny, and reflection with discussions on public platforms, in teaching programmes and through continued support for Web Science research initiatives, even at low tides of funding cycles. Funding agencies must provide support for understanding the Web without pauses

[17] https://www.poynter.org/ifcn/
[18] https://freedomhouse.org/report-types/freedom-net
[19] http://faircrowd.work/unions-for-crowdworkers/

that disrupt and destroy academic knowledge. In addition, institutions and funders can incentivise research on sustainability as well as professional practices that save resources.

4: Supporting the Intelligent Web: At this point in time, there is a proliferation of new AI institutions. Few of them, however, embrace the Web as a topic of its own and tackle AI on the Web. Thus, a broader and interdisciplinary perspective is lacking, although the ethics, laws, and politics concerning AI in the Web are issues *right now*.

5: Supporting the Inclusive Web: Academic institutions are biased towards people from specific regions of the worlds, social classes and gender. Beyond the generic problem of the 'leaky pipeline', the effect that gender parity becomes less and less balanced over academic career progression, the Web itself is a filter that may reward Web-extroverts, but deter those that shy away from Web activity, e. g., to avoid misogyny online [47]. Academic institutions must promote diversity programmes and research applied to reducing all types of divides that threaten the universal accessibility of the Web, whether it is by being physically or mentally challenged, economically limited, or socially intimidated.

6.3 Actions to be Taken by Educators

Guiding the Web into the future requires a long-term perspective that carries all its stakeholders into the future. It is essential that all citizens develop Web literacy, i. e., the competence and the capacity to engage knowledgeably in public debate and discourse on and about the Web. Education and educators play a pivotal role to achieving this goal.

Supporting the Sustainable Web: Many of the Web's problems cannot be solved once and for all. To ensure the long-term viability of the Web, we need a *process of civilizing the Web* that ultimately builds on mutual understanding and respect as can only be developed with the help of education at large.

Supporting the Intelligent Web: Intelligent behavior of, in and by the Web is not given as a fact, but emerges from actions and actors. As teachers, we need to educate how intelligences can behave for the better or worse, how they can be used or how they might misguide us.

Supporting the Inclusive Web: As a society, we must aim for and increase inclusiveness. As teachers we must strive to recognize *exclusion* and open our eyes and the eyes of others, young and old, to where and when people are ostracized on the Web.

Supporting Interdisciplinarity and Collaboration: Teaching about the Web is more interdisciplinary than what most teachers can or will be able to accomplish. Hence, teachers need to collaborate, including across disciplines and with their students, to achieve a joint understanding across disciplines, generations and social groups.

Educators must not be left alone with solving the issue of teaching Web literacy. Programs must be developed that educate about consumption in and of the Web (searching skills, critical reading, fact checking, Web usage and privacy), about contribution to the Web (Wikipedia, social media), and about related aspects from technology to data analytics to social science studies.

6.4 Actions to be Taken by Enterprises

The business models of all large and most small- end medium-sized enterprises revolve around or are deeply affected by the digital economy. Hence, all these enterprises rely on a Web that

is reasonably secure and protects their interests as well as those of their customers. This implies that enterprises must take action to safeguard this ecosystem.

Supporting Interdisciplinarity: In the past, engineering companies might have built cars and sold them to customers. In the future, mobility companies need to revise their thinking about when and for which life situation a customer needs which mobility service. Why and when would the customer feel safe and secure and trust this service, and when not? How would the customer contribute to a digital ecosphere that makes the enterprise's platform more valuable than the one of others and that provides her with a tailor-made, optimized solution?

Such questions are asked now, and not only in car companies, and they will become even more important in the future. Answers require the collaboration of people from a multitude of disciplines: engineers, computer scientists, economists, sociologists, psychologists and others. To ask and to answer these questions enterprises need models and expertise as provided by Web Science.

Supporting the Sustainable Web: Enterprises rely on the long-term viability of their business models, which includes evolution, adaptations to changing circumstances, and sometimes revolutionary changes to strategy.

Increasingly, business models that are not ecologically sustainable are coming under attack from various sides, including customers, investors and regulatory bodies and governments. Enterprises also need to ensure the long-term viability of their business networks aligned to help customers fulfill their objectives (aka "customer ecosystem"). Attention to such developments, innovation and self-regulation can answer some of these demands and, even if they come at the expense of short-term revenue, increase the long-term viability of business models and enterprises, and thereby also make the parts of the Web that these enterprises support more sustainable.[20] Laws and regulations may stifle some business models, and they may also drive innovation and new services/products and markets with GDPR being a recent example.

Supporting the Intelligent Web: Commercial AI advances are not only propelled by the Web, they are also more early visible in the Web than anywhere else. Downsides range from the merely annoying, such as incapable chatbots, to phenomena that affect the governing of nations, e. g., when information platforms regulate or do not regulate information diffusion. While in the discussion of AI and ethics, future versions of the trolley problem are discussed at length, AI and the Web has received little attention.

For example, AI contributes to automatically deleting billions of social media posts and accounts per year[21], and the decision is often wrong, because AI is nowhere near understanding the nuances of harassment, hate speech or misinformation.[22]

---

[20] Examples include Tumblr forbidding pornographic content, Instagram limiting the viewing of number of likes, and Twitter denying political advertisements.

[21] Cf., e. g., `https://about.fb.com/news/2018/05/enforcement-numbers/`.

[22] AI-based filters by themselves are unable to protect fundamental rights. For example, it has been observed that they block legitimate speech while at the same time tolerating non-legitimate speech in the same domain: Many antisemitic German tweets remain unfiltered whereas the account of the only German-Jewish weekly newspaper was automatically blocked ("Zensur im Netz: Twitter blockiert 'Jüdische Allgemeine'", 2019, `https://www.tagesschau.de/inland/twitter-juedische-allgemeine-101.html`); Islamophobic tweets run free while a Muslim politician was blocked after mentioning Muslims ("Berliner Politikerin Chebli: Twitter-Account von SPD-Staatssekretärin zeitweise gesperrt," 2019, `https://www.faz.net/aktuell/wirtschaft/unternehmen/twitter-begruendet-sperrung-mit-falschinformation-ueber-wahlen-16173579.html`).

Supporting Collaboration: To support the intelligent Web that enterprises seek to increase their profits, they must embrace the collaboration with the intelligent in the Web. Enterprises must seek more proactive interaction with the public, with academia, with NGOs, making good use of participatory methods and being open to regulation. While this may hurt business in the short term, it may be unavoidable to achieve long-term goals such as efficiency, customer trust, and even innovation.

Supporting the Inclusive Web: Datafication of businesses may make unfair practices of enterprises more visible. Society is less willing to accept these and expects companies to take advantage of available data and make their businesses more inclusive.

The presence of regulation on inclusion not only leads to obstacles on the way to profit making, it can also help create new products and business models and to include previously excluded individuals and groups into the workforce.

Based on ubiquity, inclusion, intelligence, and collaboration, the Web allows for sharing products, services, and events (transport, accommodation, repair, experiences etc.) that would never have been shared before. This sharing economy of the Web has enabled novel enterprise opportunities for value creation. Now, we need to become creative to use the sharing economy for reducing, not worsening[23], physical and environmental footprints.

6.5 Actions to be Taken by Policy Makers

Policy makers and state agencies play a double role when they interact with Web Science. At one end, they impose conditions that constrain how the Web, and Web Science, operate and, thus, are critical to shaping the future of both. Also, they determine support for academia and funding agencies at large, nurturing or impeding Web Science research.

At the other end, policy makers and state agencies might benefit from knowledge provided by Web Science, researchers, practitioners and think tanks – helping to bring forward laws and regulations that would be based on more profound knowledge and hence better serve society.

Supporting Interdisciplinarity: As yet, we do not have a "joined up" approach to Web related policies across governments, e. g., linking education policy with policy supporting technical innovation and policy on the environment. Developing problem-focussed research for policy making, drawing in necessary expertise, is key to achieving such a cross-cutting approach.

Supporting the Sustainable Web: The technical infrastructure of the Web consumes enormous amounts of energy and creates hazardous waste, while the systems that operate on the Web (e. g., cryptocurrencies) have taken energy consumption to the level of individual nation states.[24] It is imperative that innovative solutions are found to incentivize stakeholders and regulate future technologies supporting innovation for climate and environment protection.

Supporting the Intelligent Web: Artificial Intelligence is an area that is in the focus of governments, e. g., the European Commision [15]. As Dignum et al. [26] review: "Commission takes up a rule setting role... In our opinion, this is a good first step. If we were

[23] https://www.transportenvironment.org/news/uber-and-lyft-increase-pollution-and-undermine-public-transport-study-shows

[24] https://www.economist.com/the-economist-explains/2018/07/09/why-bitcoin-uses-so-much-energy

to draw the analogy with a game, independently of who is playing the game, without rules no one wins." The same applies for the intersection of AI and the Web, which is so omnipresent in our digital lives that it warrants dedicated attention on its own – which it has not received, yet.

Supporting the Inclusive Web: It is a difficult task for politicians to formulate the laws that bring forward a civilized Web [64]. To date, policy makers have occupied a largely reactive position, chasing the Web as it has evolved and rarely equipped with the cutting-edge knowledge to make timely interventions in this domain.

Laws for inclusion exist, but often they are not enforced. At the technical level, many concepts and tools exist that enhance Web accessibility[25], but few Web sites actually implement these ideas. At the social level, datafication provides the basis for monitoring fairness, but no legal consequences are drawn from such data. At the societal level, the digital divide must be reduced and intimidation of benevolent users must be prevented. Thus, laws that target the inclusion and the safety of all the must be redrawn to include evaluation and continuous improvement of their effectiveness.

Supporting Collaboration: The Web is global and its phenomena are crossing sectors of life. Hence, policy makers must organize and enable collaboration at all scales. The Web is an issue for our global society and it must be addressed internationally. This requires action comparable to other large transcontinental efforts, such as the Square Kilometre Array[26], the fusion reactor ITER[27], or the International Space Station[28].

The careful, scientific sharing of Web data must be made possible by laws. There are models for sharing medical or social science data, which must be transferred to Web Science such that scientific investigations are made possible. The individual must be protected, but also the societal interest of learning about social cohesion, democracy and public security must not be neglected, even when the data is owned by private corporations. Public discussion of these laws will require time in order to achieve a broad consensus on the models and their purposes.

Taking such support measures will not be possible by individual policy actions. Rather, policy makers will need to create organizational structures that allow them to provide such support: *(i)*, the Web is a fundamentally international affair, but regulatory bodies for the Web are virtually non-existent. *(ii)*, policy makers should and must rely on Web expertise, but no comprehensive expert structure exists. Industry takes a view that is biased in several ways. And academia alone cannot provide the expertise. It finds itself between the rock of over-competition, which makes service to society at large ungratifying if not impossible, and the hard place of struggling for interdisciplinarity without having the means to really do so. Policy makers must create spaces for bringing together all Web stakeholders, researchers, industrial practitioners, interest groups, and politicians, but only in this way government will have the most profound expertise to guide their policy making – especially at time when the Web is moving so fast. These issues are deeply challenging and urgent.

[25] https://www.w3.org/WAI/
[26] https://www.skatelescope.org
[27] https://www.iter.org/
[28] https://www.nasa.gov/mission_pages/station/main/index.html

7 Conclusion

The Web has evolved from a Web of documents to a Web that is so omnipresent that societies almost take it for granted and spend little thoughts about it. At the same time the Web affects individual lives, economies as well as national and international politics. The next wave, which is already visible now, will be a Web that links all forms of Intelligence and that continues to evolve at accelerating speed. A new wave of AI, powered to a great extent by the phenomenal data resources largely created by the Web and by the global infrastructure of the Web, has begun a new round of transformations with unforeseen impacts on the Web futures and the future of our world.

The "Web is a never ending project" [35] and perpetual changes are at the very core of the Web. Change is neither good nor bad and as with every change, there is the potential to bring as much disruption as benefit. We cannot stop change, therefore the question is how we deal with change and in particular, how we embrace it as an opportunity for reflection on the past and to guide us in the future while staying loyal to Web's founding values of information sharing and inclusiveness.

As our societies are faced with increasing levels of complexity and uncertainty, and as the future, deeply impacted by new technologies, is an open space of possible developments that we cannot foresee, we cannot hope that the problems of the Web will magically disappear. We have therefore to accept its ambivalences and to build resilient processes with the capacity to provide all the required safeguard mechanisms. We must find out how to prevent negative consequences of the technology, how to detect emerging problems and how to react via timely and effective interventions. The role of Web Science with its different disciplines is to reduce the time between a problem arising and (intermediate) solutions being found.

The Web is for everyone and everyone is also responsible for the Web and its "healthy" operation. Individual responsibility can range from protecting your private information to reporting misuse of information and respecting others online. Paraphrasing John F. Kennedy's historic words:

Ask not what the Web can do for you – ask what you can do for the Web.

Individual actions are desperately needed to incur collective actions and systemic change. By the latter we do not want to reduce the responsibilities borne by stakeholders and other institutions, rather we want to highligh the importance of both individual and collective actions.

8 Participants

- Robert Ackland (Australian National University – Canberra, AU)
- Ricardo A. Baeza-Yates (NTENT – Carlsbad, US & Northeastern Univ., US)
- Bettina Berendt (TU Berlin and Weizenbaum Institute for the Networked Society, DE & KU Leuven, BE)
- Noshir S. Contractor (Northwestern University – Evanston, US)
- David De Roure (University of Oxford, GB)
- Kemal A. Delic (Hewlett Packard – Grenoble, FR)
- Nikolaus Forgó (Universität Wien, AT)
- Fabien Gandon (INRIA Sophia Antipolis, FR)
- Susan Halford (University of Bristol, GB)
- Wendy Hall (University of Southampton, GB)
- Lynda Hardman (CWI – Amsterdam, NL)
- Andreas Hotho (Universität Würzburg, DE)
- Katharina E. Kinder-Kurlanda (University of Klagenfurt, AT)
- Claudia Müller-Birn (FU Berlin, DE)
- Wolfgang Nejdl (Leibniz Universität Hannover, DE)
- Eirini Ntoutsi (Freie Universität Berlin, DE)
- Paolo Parigi (Airbnb and Stanford University – San Francisco, US)
- Evaggelia Pitoura (University of Ioannina, GR)
- Oshani Seneviratne (Rensselaer Polytechnic, US)
- Elena Simperl (University of Southampton, GB)
- Steffen Staab (Universität Stuttgart, DE & University of Southampton, UK)
- Guglielmo Tamburrini (University of Naples, IT)
- Pinelopi Troullinou (The Open University – Milton Keynes, GB)

References

1 Giovanni Allegretti and Sofia Antunes. The Lisbon participatory budget: results and perspectives on an experience in slow but continuous transformation. *Field Actions Science Reports. The journal of field actions*, Special Issue 11, 2014.

2 S. Amsler and K. Facer. Contesting anticipatory regimes in education: exploring alternative educational orientations towards the future. *Futures*, 94:6–14, 2017.

3 A. Appadurai. *The Future as Cultural Fact*. Verso, London, 2013.

4 Ross Ashby. Design for an intelligence-amplifier. *Automata studies*, 400:215–233, 1956.

5 Ipek Baris, Lukas Schmelzeisen, and Steffen Staab. CLEARumor at SemEval-2019 task 7: Convolving ELMo against rumors. In *Proceedings of the 13th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2019, Minneapolis, MN, USA, June 6-7, 2019*, pages 1105–1109, 2019. URL: `https://www.aclweb.org/anthology/S19-2193/`, `doi:10.18653/v1/s19-2193`.

6 Bettina Berendt. AI for the common good?! Pitfalls, challenges, and ethics pen-testing. *Paladyn*, 10(1):44–65, 2019. `doi:10.1515/pjbr-2019-0004`.

7 Tim Berners-Lee, Robert Cailliau, Ari Luotonen, Henrik Frystyk Nielsen, and Arthur Secret. The World-Wide Web. *Commun. ACM*, 37(8):76–82, 1994. `doi:10.1145/179606.179671`.

8 Tim Berners-Lee, Wendy Hall, James Hendler, Nigel Shadbolt, and Daniel J Weitzner. Creating a science of the Web. *Science*, 313(5788):769–771, 2006.

9 M. Bevir. *Governance: A very short introduction*. Oxford University Press, Oxford, UK, 2012.

10 J. Bowker and S.L. Starr. *Sorting Things Out: Classification and its' consequences*. MIT Press, Cambridge, MA., 1999.

11 Irene Celino, Oscar Corcho, Franz Hölker, and Elena Simperl. Citizen science: design and engagement (dagstuhl seminar 17272). In *Dagstuhl reports*, volume 7. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

12 Karin Knorr Cetina. *Epistemic cultures: How the sciences make knowledge*. Harvard University Press, 2009.

13 Andrei Ciortea, Simon Mayer, Fabien Gandon, Olivier Boissier, Alessandro Ricci, and Antoine Zimmermann. A Decade in Hindsight: The Missing Bridge Between Multi-Agent Systems and the World Wide Web. In *AAMAS 2019 - 18th International Conference on*

Autonomous Agents and Multiagent Systems, page 5p, Montréal, Canada, 2019. URL: https://hal-emse.ccsd.cnrs.fr/emse-02070625.

14 European Commission. Ethics guidelines for trustworthy ai, 2019. URL: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.

15 European Commission. On Artificial Intelligence — a European approach to excellence and trust. Technical report, European Commission, Brussels, BE, February 2020. Whitepaper. URL: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

16 ACM US Public Policy Council, 2017. URL: https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf.

17 Loy D. Watley. Training in ethical judgment with a modified Potter box. *Business Ethics: A European Review*, 23(1):1–14, 2014. doi:10.1111/beer.12034.

18 Mathieu d'Aquin, Pinelopi Troullinou, Noel E. O'Connor, Aindrias Cullen, Gráinne Faller, and Louise Holden. Towards an "ethics by design" methodology for AI research projects. In Jason Furman, Gary E. Marchant, Huw Price, and Francesca Rossi, editors, *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, AIES 2018, New Orleans, LA, USA, February 02-03, 2018*, pages 54–59. ACM, 2018. doi:10.1145/3278721.3278765.

19 María Puig de La Bellacasa. *Matters of care: Speculative ethics in more than human worlds*, volume 41. U of Minnesota Press, 2017.

20 R. De Wolf, E. Vanderhoven, B. Berendt, J. Pierson, and T. Schellens. Self-reflection on privacy research in social networking sites. *Behaviour & Information Technology*, 36(5):459–469, 2017.

21 R. Dean, J. Boswell, and G. Smith. Designing democratic innovations as deliberative systems: The ambitious case of NHS citizen. *Political Studies*, 2019. Advanced online publication. doi:10.1177/0032321719866002.

22 Ronald J. Deibert, John G. Palfrey, Rafal Rohozinski, and Jonathan Zittrain. *Access Denied: The Practice and Policy of Global Internet Filtering (Information Revolution and Global Politics)*. MIT Press, 2008.

23 CGIL (Confederazione Generale Italiana del Lavoro [Italian General Confederation of Labour]. Riders: CGIL, causa contro multinazionale food delivery per condotta discriminatoria, algoritmo penalizza diritti di sciopero e malattia, 2019. URL: http://www.cgil.it/riders-cgil-causa-contro-multinazionale-food-delivery-per-condotta-discriminatoria-algoritmo-penalizza-diritti-di-sciopero-e-malattia/.

24 L. DeNardis and M. Raymond. Thinking clearly about multistakeholder internet governance. In *GigaNet: Global Internet Governance Academic Network, Annual Symposium 2013*, 2013. Available at SSRN: https://ssrn.com/abstract=2354377 or http://dx.doi.org/10.2139/ssrn.2354377.

25 Deutscher Bundestag – Wissenschaftliche Dienste. Regulierung von Hate Speech und Fake News in sozialen Netzwerken durch EU-Mitgliedstaaten [regulation of hate speech and fake news in social networks by eu member states], 2018. WD 10-3000 -032/18, https://www.bundestag.de/resource/blob/566942/a5eb997872bbe5dbca3f47112eb04c46/wd-10-032-18-pdf-data.pdf.

26 Virginia Dignum, Catelijne Muller, and Andreas Theodorou. First analysis of the EU whitepaper on ai. Technical report, ALLAI, February 2020. URL: http://allai.nl/first-analysis-of-the-eu-whitepaper-on-ai/.

27 European Data Protection Board (EDPD). Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak, 21 April 2020. URL: https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_20200420_contact_tracing_covid_with_annex_en.pdf.

28 Douglas C Engelbart. Augmenting human intellect: A conceptual framework. SRI Summary Report AFOSR-3223, Stanford Research Institute, 1962.

29 Omer Farooq, Salman Ahmed Khan, and Sadaf Khalid. Financial ethics: A review of 2010 flash crash. *International Journal of Computer and Information Engineering*, 8(6):1652–1654, 2014.

30 Roy T Fielding and Richard N Taylor. Principled design of the modern Web architecture. *ACM Transactions on Internet Technology (TOIT)*, 2(2):115–150, 2002.

31 The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems, 2016. URL: `http://standards.ieee.org/develop/indconn/ec/ead_v1.pdf`.

32 B. Fox. Campaigners lock horns with EU justice chief on 'hate speech', 2018. EURACTIV, `https://www.euractiv.com/section/freedom-of-thought/news/campaigners-lock-horns-with-eu-justice-chief-on-hate-speech/`.

33 Peter Fox and James Hendler. Changing the equation on scientific data visualization. *Science*, 331(6018):705–708, 2011.

34 Fabien Gandon. The three 'W' of the World Wide Web call for the three 'M' of a Massively Multidisciplinary Methodology. In Valérie Monfort and Karl-Heinz Krempels, editors, *WEBIST 2014 - 10th International Conference*, volume 226 of *Web Information Systems and Technologies*, Barcelona, Spain, 2014. Springer International Publishing. `doi:10.1007/978-3-319-27030-2`.

35 Fabien Gandon. For everything: Tim Berners-Lee, winner of the 2016 Turing award for having invented... the Web. *1024 : Bulletin de la Société Informatique de France*, 11:21, 2017. URL: `https://hal.inria.fr/hal-01843967`.

36 Fabien Gandon, Michel Buffa, Elena Cabrio, Olivier Corby, Catherine Faron Zucker, Alain Giboin, Nhan Le Thanh, Isabelle Mirbel, Peter Sander, Andrea G. B. Tettamanzi, and Serena Villata. Challenges in Bridging Social Semantics and Formal Semantics on the Web. In Hammoudi, S., Cordeiro, J., Maciaszek, L.A., Filipe, and J., editors, *5h International Conference, ICEIS 2013*, volume 190 of *Lecture Notes in Business Information Processing*, pages 3–15, Angers, France, 2013. Springer. URL: `https://hal.inria.fr/hal-01059273`.

37 Fabien Gandon and Alain Giboin. Paving the WAI: Defining Web-Augmented Interactions. In *Web Science 2017 (WebSci17)*, pages 381–382, Troy, NY, United States, 2017. `doi:10.1038/nbt0609-508`.

38 Paul Gaskell. *Forecasting financial markets with online information.* PhD thesis, University of Southampton, Faculty of Business, Law and Art, 2015. URL: `https://eprints.soton.ac.uk/393779/`.

39 K. Georgiou, S. Xavier-de-Souza, and K. Eder. The IoT energy challenge: A software perspective. *IEEE Embedded Systems Letters*, 10(3):53–56, September 2018. `doi:10.1109/LES.2017.2741419`.

40 L. Gill, D. Redeker, and U. Gasser. Towards digital constitutionalism? Mapping attempts to craft an internet bill of rights. Technical report, Berkman Center Research Publication No. 2015-15, 2015. Available at SSRN: `https://ssrn.com/abstract=2687120` or `http://dx.doi.org/10.2139/ssrn.2687120`.

41 Christophe Guéret, Victor de Boer, and Stefan Schlobach. Let's "downscale" linked data. *IEEE Internet Computing*, 18(2):70–73, 2014.

42 Seda Gurses, Rebekah Overdorf, and Ero Balsa. POTs: The revolution will not be optimized? *CoRR*, abs/1806.02711, 2018. `arXiv:1806.02711`.

43 Susan Halford and M. Savage. Speaking sociologically with Big Data: Symphonic social science and the future for Big Data research. *Sociology*, 51(6):1132–1148, 2017.

44 James A. Hendler, Wendy Hall, and Noshir Contractor. Web science: Now more than ever. *IEEE Computer*, 51(6):12–17, 2018. `doi:10.1109/MC.2018.2701637`.

45 Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard de Melo, Claudio Gutierrez, José Emilio Labra Gayo, Sabrina Kirrane, Sebastian Neumaier, Axel Polleres, Roberto Navigli, Axel-Cyrille Ngonga Ngomo, Sabbir M. Rashid, Anisa Rula, Lukas Schmelzeisen, Juan Sequeda, Steffen Staab, and Antoine Zimmermann. Knowledge graphs, 2020. arXiv:2003.02320.

46 Viktória Ilková and Adrian Ilka. Legal aspects of autonomous vehicles – an overview. In *Proceedings of the 2017 21st International Conference on Process Control (PC), Štrbské Pleso, Slovakia, June 6 – 9*, pages 428–433, 2017. doi:10.1109/PC.2017.7976252.

47 Emma A. Jane. Online misogyny and feminist digilantism. *Continuum*, 30(3):284–297, 2016. doi:10.1080/10304312.2016.1166560.

48 N. Jones. How to stop data centres from gobbling up the world's electricity. *Nature*, 561:163–166, 2018. doi:10.1038/d41586-018-06610-y.

49 Katharina E Kinder-Kurlanda. Ethnography in a computer science centered project. In *ICC'14 Workshop, ACM WebSci'14 Conference, Bloomington. Online: http://www. icc. ecs. soton. ac. uk/papers/icc14_submission_2. pdf*, 2014.

50 Dave Lee. How scientology changed the internet. *BBC News*, 2013. URL: https://www.bbc.com/news/technology-23273109.

51 L. Lessig. *Code and Other Laws of Cyberspace*. Basic Books, 1999.

52 Amanda Levendowski. Using copyright to combat revenge porn. *3 N.Y.U. J. INTELL. PROP. & ENT. L.*, pages 422–446, 2014. Available at SSRN: https://ssrn.com/abstract=2374119 or http://dx.doi.org/10.2139/ssrn.2374119.

53 R. Levitas. *Utopia as Method: the imaginary reconstitution of society*. Palgrave Macmillan, London, 2013.

54 Annette Markham. Fabrication as ethical practice: Qualitative inquiry in ambiguous internet contexts. *Information, Communication & Society*, 15(3):334–353, 2012.

55 Aryn Martin, Natasha Myers, and Ana Viseu. The politics of care in technoscience. *Social Studies of Science*, 45(5):625–641, 2015.

56 John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon. A proposal for the Dartmouth summer research project on Artificial Intelligence, August 1955. URL: http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf.

57 Panagiotis Takis Metaxas and Eni Mustafaraj. From obscurity to prominence in minutes: Political speech and real-time search. In *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line*, April 2010.

58 Luc Moreau, Paul T. Groth, James Cheney, Timothy Lebo, and Simon Miles. The rationale of PROV. *J. Web Semant.*, 35:235–257, 2015. doi:10.1016/j.websem.2015.04.001.

59 Eni Mustafaraj and Panagiotis Takis Metaxas. The fake news spreading plague: Was it preventable? In *Proceedings of the 2017 ACM on Web Science Conference, WebSci 2017, Troy, NY, USA, June 25 - 28, 2017*, pages 235–239, 2017. doi:10.1145/3091478.3091523.

60 CNIL (Commission nationale de l'informatique et des libertés). La CNIL publie son deuxième avis adressé au parlement sur les conditions de mise en œuvre de SI-DEP, Contact Covid, Vaccin COVID et TousAntiCovid [The CNIL publishes its second advice to parliament about the conditions of implementation of SI-DEP, Contact Covid, Vaccin COVID and TousAntiCovid], 21 January 2021. URL: https://www.cnil.fr/fr/la-cnil-publie-son-deuxieme-avis-adresse-au-parlement-sur-les-conditions-de-mise-en-oeuvre-de-si-dep.

61 Gina Neff and Peter Nagy. Automation, algorithms, and politics — talking to bots: Symbiotic agency and the case of Tay. *International Journal of Communication*, 10:17, 2016.

62 Eirini Ntoutsi, Pavlos Fafalios, Ujwal Gadiraju, Vasileios Iosifidis, Wolfgang Nejdl, Maria-Esther Vidal, Salvatore Ruggieri, Franco Turini, Symeon Papadopoulos, Emmanouil Krasanakise, Ioannis Kompatsiaris, Katharina Kinder-Kurlanda, Claudia Wagner, Fariba Karimi, Miriam Fernandez, Harith Alani, Bettina Berendt, Tina Kruegel, Christian Heinze, Klaus Broelemann, Gjergji Kasneci, Thanassis Tiropanis, and Steffen Staab. Bias in data-driven

AI systems — an introductory survey. *WIREs Data Mining and Knowledge Discovery*, 2020. doi:10.1002/widm.1356.

63 AoIR (Association of Internet Researchers), 2012. URL: https://aoir.org/reports/ethics2.pdf.

64 Kieron O'Hara and Wendy Hall. Four internets: The geopolitics of digital governance. *Center for International Governance Innovation*, 206, December 2018. URL: https://www.cigionline.org/publications/four-internets-geopolitics-digital-governance.

65 J.R. Reidenberg. Lex informatica: The formulation of information policy rules through technology. *Texas Law Review*, 76:553–593, 1998.

66 S. Saad, S. Bazan, and C. Varin. Asymmetric cyberwarfare between Israel and Hezbollah: The Web as a new strategic battlefield. In *Poster Proceedings WebSci'11, June 14-17 2011, Koblenz, Germany*, pages 1–4, 2011.

67 Sabrine Saad, Muriel Chamoun, and Stéphane B. Bazan. Infowar on the web: When the caliphate goes online. In *Proceedings of the ACM Web Science Conference, WebSci 2015, Oxford, United Kingdom, June 28 - July 1, 2015*, pages 53:1–53:3, 2015. doi:10.1145/2786451.2786487.

68 Michel Serres. *The natural contract*. University of Michigan Press, 1995.

69 Scott Shane and Daisuke Wakabayashi. 'the business of war': Google employees protest work for the Pentagon. *The New York Times*, 2018. URL: https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html.

70 Tamar Sharon. Blind-sided by privacy? digital contact tracing, the apple/google api and big tech's newfound role as global health policy makers. *Ethics and Information Technology*, 2020. doi:10.1007/s10676-020-09547-x.

71 Phil Simon. How nextdoor addressed racial profiling on its platform. *Harvard Business Review*, 2018. URL: https://hbr.org/2018/05/how-nextdoor-addressed-racial-profiling-on-its-platform.

72 Steffen Staab, Susan Halford, and Wendy Hall. Web Science in Europe: beyond boundaries. *Commun. ACM*, 62(4):74, 2019. doi:10.1145/3312569.

73 Emma Strubell, Ananya Ganesh, and Andrew McCallum. Energy and Policy Considerations for Deep Learning in NLP. *arXiv e-prints*, page arXiv:1906.02243, June 2019. arXiv:1906.02243.

74 Mengfan Tang, Pranav Agrawal, and Ramesh Jain. Habits vs environment: What really causes asthma? In *Proceedings of the ACM Web Science Conference*, pages 1–5, 2015.

75 The speakers commission on digital democracy:, 2015. URL: http://www.parliament.uk/business/commons/the-speaker/speakers-commission-on-digital-democracy/.

76 Ramin Tinati, Susan Halford, Les Carr, and Catherine Pope. Big Data: methodological challenges and approaches for sociological analysis. *Sociology*, 48(4):663–681, 2014.

77 Thanassis Tiropanis, Wendy Hall, Nigel Shadbolt, David De Roure, Noshir Contractor, and Jim Hendler. The Web Science observatory. *IEEE Intelligent Systems*, 28(2):100–104, 2013.

78 Jon Truby. Decarbonizing bitcoin: Law and policy choices for reducing the energy consumption of blockchain technologies and digital currencies. *Energy Research & Social Science*, 44:399–410, 2018. doi:10.1016/j.erss.2018.06.009.

79 Gert Jan van Hardeveld, Craig Webber, and Kieron O'Hara. Deviating from the cybercriminal script: Exploring tools of anonymity (mis)used by carders on cryptomarkets. *American Behavioral Scientist*, 61:1244–1266, October 2017.

80 Ricardo Vinuesa, Hossein Azizpour, Iolanda Leite, Madeline Balaam, Virginia Dignum, Sami Domisch, Anna Felländer, Simone Daniela Langhans, Max Tegmark, and Francesco Fuso Nerini. The role of Artificial Intelligence in achieving the Sustainable Development Goals. *Nature Communications*, 11(1):233, 2020. doi:10.1038/s41467-019-14108-y.

81 Jennifer Welch, Susan Halford, and Mark Weal. Conceptualising the Web for post-conflict governance building. *Peacebuilding*, 3(1):58–74, 2014.