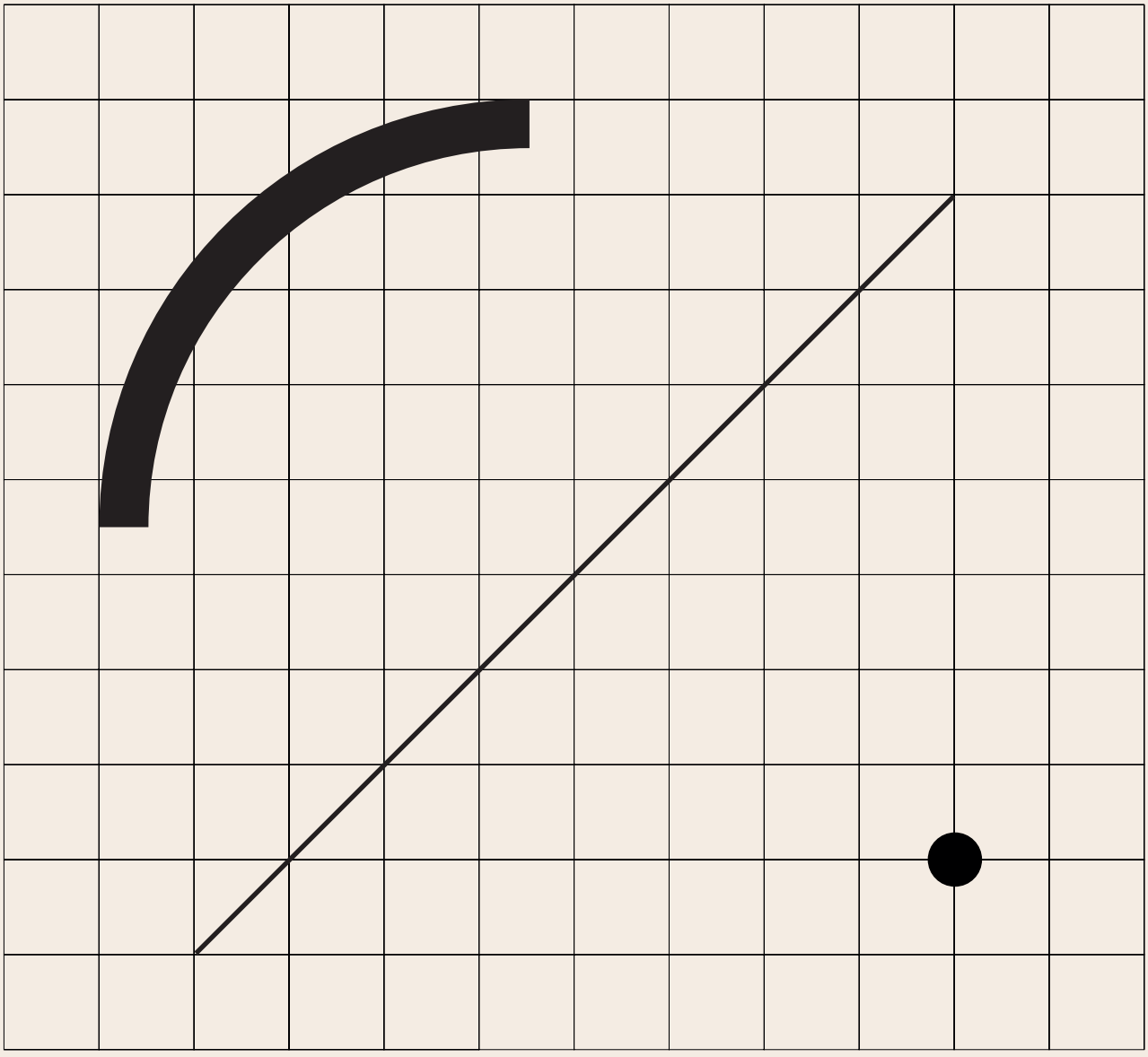
For a Decentralised Global Disinformation Research Cloud

POLICY PAPER - OCTOBER 2021



groupe d'études geopolitiques

groupe d'études geopolitiques

For a Decentralised Global Disinformation Research Cloud

AUTHOR
SAIPH SAVAGE, ISHAN SHARMA, CRISTINA MARTINEZ PINTO,VICTOR STORCHAN

FOREWORD
PATRICE GEOFFRON

45, RUE D'ULM 75005 PARIS
LEGRANDCONTINENT.EU
GEG@ENS.FR

REFERENCES
SAIPH SAVAGE, ISHAN SHARMA, CRISTINA MARTINEZ PINTO,VICTOR STORCHAN, FOR A DECENTRALISED GLOBAL DISINFORMATION RESEARCH CLOUD, POLICY PAPER, GROUPE D'ÉTUDES GÉOPOLITIQUES, OCTOBER 2021.

Foreword

PATRICE GEOFFRON • PROFESSOR OF ECONOMICS AND DIRECTOR OF THE CENTRE FOR GEOPOLITICS OF ENERGY AND RAW MATERIALS AT PARIS-DAUPHINE

Let's start by acknowledging the fact that the work proposed here is first and foremost a source of amazement: unfortunately, there is little doubt that the phenomena listed in the preamble of the document have indeed occurred, especially in terms of the integrity of electoral processes. The count of 639 distinct disinformation campaigns related to the 2020 American elections (by the Electoral Integrity Project) projects us directly into a dystopia...

And it is not the almost two years of health crisis that will have blurred this sensation, with 'infodemic' among the neologisms of this period.

How did we get here? Is it possible to identify a rise in disinformation operations and a tipping point from which an exponential runaway effect can be observed?

Tracing the path of the protean threats of disinformation is a complex task. This explains why the volume of research devoted to disinformation is growing strongly. The ScienceDirect bibliographic database (consulted on October 9, 2021) indicates that 126 articles have included the keyword «Disinformation» since 1987, when it first appeared (an article on AIDS in Africa). If, for several decades, this database has recorded only one article per year (and very often none), 21 were published in 2020 and, already, 54 in 2021 (current year...). Of the 2021 publications, 33 mention Twitter, 28 Facebook, 23 China, 14 Russia, which gives a panoramic view of a few nodes of the problem. Some titles among the most recent papers give a light on the issues raised: 'Examining characteristics of prebunking strategies to overcome public relations disinformation attacks', 'Misinformation, disinformation, and fake news: Cyber risks to business', 'Relating foreign disinformation through social media, domestic online media fractionalization', 'CoVerifi: A COVID-19 news verification system',... Let's also note that, as early as 2013, the World Economic Forum's Global Risk Report had highlighted the risk of digital disinformation, directly related to the cyber attacks

Undoubtedly, we need research that analyzes the motives behind misinformation and its impact. In this perspective, the first merit of the paper by S. Savage, I.Sharma, C.M. Pinto and V. Storchan is to analyze the 'costs' of disinformation analysis: the combination of massive computing power, access to relevant data, the availability of sophisticated AI models, ... And these conditions are best met by large technological platforms, even though they are among the channels through which disinformation circulates, so that a threat of 'pervasive bias in the large-scale disinformation research that is conducted' (in the words of the authors) emerges.

Beyond these observations, and since the robustness of very mature democracies has been tested by disinformation operations, the authors call for a reduction in the asymmetries of means observed in order to conduct relevant research and build a space for cooperation between universities, NGOs, public institutions and companies. The strength of their proposal, beyond the «petition of principle», is to define in great detail the arcana of the system, both in technical terms (a global and decentralized cloud, progressive feeding of a relevant database) and in terms of governance (so as to avoid a «curse of the commons» and to allow for the indispensable multidisciplinary interactions). Their proposal is probably only the beginning of a long process, but is consistent with the initiatives of research clouds and public computing infrastructures with which it is intended to form an ecosystem.

For a Decentralised Global Disinformation Research Cloud

SAIPH SAVAGE • DIRECTOR NORTHEASTERN CIVIC A.I. LAB AND ASSISTANT PROFESSOR AT NORTHEASTERN UNIVERSITY
ISHAN SHARMA • FELLOW AT THE FEDERATION OF AMERICAN SCIENTISTS AND POLICY ANALYST AT THE DAY ONE PROJECT
CRISTINA MARTINEZ PINTO • FOUNDER AT PIT POLICY LAB AND SENIOR RESEARCH SCIENTIST AT THE HUMAN-COMPUTER INTERACTION LAB AT WEST VIRGINIA UNIVERSITY **VICTOR STORCHAN** • STANFORD ALUMNUS, AI ENGINEER IN FINANCIAL INDUSTRY

Disinformation is increasingly fomenting public distrust and impacting national security, election integrity, public health, and other critical U.S. sectors [8, 21]. The 7.2 million mentions of rumored antifa violence and 6.2 million mentions of the QAnon conspiracy theory on social media are believed to have contributed to the January 6th attack on the U.S. Capitol[1]. The Election Integrity Partnership (EIP) identified 639 distinct misinformation and disinformation campaigns related to the 2020 U.S election across 15 social media platforms, 72% of which explicitly aimed to delegitimise the election[2].

Note that disinformation campaigns are not an isolated incident that only affects the United States. Following Russia's interference in the 2016 US presidential campaign, "disinformation" has become a global problem. European democracies have been threatened by disinformation campaigns [26, 27]. Several of these campaigns have aimed to sow distrust in European elections, and promote social divisions to create chaos across Europe and strengthen Kremlin influence [3, 16].

Bridging the resource gap between industry, multilateral open science, & academic initiatives

The newest technological capabilities, like large AI language models, are powerful tools that can speed up the deployment of hostile influence operations and increase their scale [6]. As many democracies are trying to protect themselves against disinformation through their legislative bodies [1], they need to rely on flexible, multidisciplinary, and plurilateral coalitions to bring expertise to rapidly and continuously evolving technology. Beyond technology itself, deeper understanding of the mechanisms, scope, and impacts of disinformation campaigns and dynamics are needed to address this pressing issue [7,31]. However, in order to conduct large-scale disinformation research to understand how these complex networks and relationships are playing out in the wild, it is necessary to have both massive and expensive amounts of computational power and access to relevant data [24,25]. The only entities capable of such research at scale are large technology companies (e.g., Google, Facebook, Twitter, etc), which have substantial computing resources and the latest hardware to train billion parameter models (OpenAI's GPT-3 or Google's Switch-C). They serve as the gatekeeper for research to user-generated data on popular internet platforms [2,11,24]. The result is pervasive bias in the large-scale disinformation research that is conducted. In particular, there has been a tendency to focus on research that matches the interests and agenda of large technology companies [2, 37] (especially as they are the ones who provide the data and computation to conduct the studies).

Research falling outside corporate interests has tended to find itself with limited access to the computation and data resources needed to conduct their investigations [2, 13, 30], such as the recent incident between Facebook and researchers at New York University (NYU) whose access was revoked upon exposing how Facebook profited from manipulative political ads[3]. Even more recently, internal Facebook emails revealed that in a Facebook academic research initiative, where the company would share data with academics, only half of the Facebook user interactions data was shared, not all, as the company had

1 — Misinformation 2020: What the Data Tells Us About Election-Related Falsehoods, Defense One, November 2020.

2 — Center for an Informed Public, Digital Forensic Research Lab, Graphika, & Stanford Internet Observatory (2021). The Long Fuse: Misinformation and the 2020 Election. Stanford Digital Repository: Election Integrity Partnership. v1.3.0.

3 — NYU Researchers Were Studying Disinformation On Facebook. The Company Cut Them Off, NPR, August 2021.

originally claimed[4]. This incident resulted in the undermining of the power and accuracy of research claims made over the last three years that used Facebook's data[5].

Limited access to computation and data for disinformation investigations not only makes knowledge progress difficult, it also inhibits federal governments' abilities to create national defense strategies to combat the phenomenon at a national level as the race to master this technology has become a geopolitical issue[6]. Overall, it is important to highlight that AI models used to combat disinformation and, more generally, technologies used across the globe, require more scrutiny. Open science initiatives (like the Big Science Workshop or the EleutherAI group) as well as academia, should be empowered to provide impartial and transparent audits of technology, especially if these technologies are having a critical effect on public dialogue and civic engagement [5, 14].

Designing a Decentralised Global Disinformation Research Cloud

To address these pressing problems, we propose the creation of a Decentralised Global Disinformation Research Cloud, a tool that will enable a larger and more diverse number of stakeholders (i.e., academics, NGOs, governments, startups) to collaborate and securely conduct large scale disinformation research. Part of the challenge inevitably involves answering questions over how best to share disinformation data and computational resources with stakeholders without compromising privacy, security, or proprietary interests. For this reason, we propose the need for a cloud governance model that facilitates collaborations across sectors around disinformation research. Such governance includes coordinating how the stakeholders will access and share resources between each other to conduct disinformation research, as well as coordinating how they will start collaborating with actors outside their immediate sector. We expect that by powering these multi-disciplinary collaborations, we will be able to derive more actionable real world solutions to address the complex phenomena of disinformation which impact the security of nations, the safeguarding of our democracies, and public health. Some first sketches of solutions exist such as the concept of "middleware", a disintermediation software produced by a third party as an independent entry point which is integrated into platforms (Google API, Facebook, Twitter, etc.) to moderate content, regulate feeds, and filter or classify information[7]. However, this should be the subject of a more systematic large-scale impact study. Likewise, the collaborative project between NYU and the Poynter Institute for Media Studies to train researchers and policy makers on the use of technology for fact-checking could thus be scaled up and sustained[8]. Notice that ensuring multi-stakeholder governance not only removes a large coordination burden from large communication platforms (e.g., Facebook), it can also limit conspiracy theories about censorship towards certain research directions, and even voices, who are allegedly not aligned with the interests of the large communication platforms [38]. Recently, for example, British politicians working directly in the European Parliament have pushed conspiracies about Facebook censoring Conservative voices, citing a heavy drop in engagement on their own political Facebook pages [15, 18, 39].

Our proposed tool provides multiple stakeholders with the data and computational power that they need to conduct disinformation research. Our tool, which we have already designed with the Federation of American Scientists [29], consists of three main parts to facilitate multidisciplinary and decentralised collaborations around disinformation: (1) a decentralised network of servers that will be readily available for stakeholders to conduct large scale disinformation research; a (2) "data library" that holds data to help researchers to conduct their disinformation research; (3) a governance model to coordinate decentralised stakeholders to share the computational resources and data, while also facilitating multidisciplinary disinformation research collaborations. Our Research Cloud illustrates two core benefits relative to traditional approaches for conducting disinformation research: 1) scale, the ability to create and coordinate decentralised research teams dynamically in response to research interests; 2) diversity, the ease with which stakeholders from diverse sectors and disciplines can be brought together to conduct disinformation research

4 — Facebook made big mistake in data it provided to researchers, undermining academic work, The Washington Post, September 2021.

5 — Facebook admits it messed up again, Insider, September 2021.

6 — Facebook reportedly provided inaccurate data to misinformation researchers, The Verge, September 2021.

7 — Francis Fukuyama, Barak Richman, Ashish Goel, Roberta R. Katz, A. Douglas Melamed, Marietje Schaake, Middleware for Dominant Digital Platforms: A Technological Solution to a Threat to Democracy, Standford University.

8 — Institute for Data, Democracy and Politics, The George Washington University.

while having access to the computation and data that they need. We contend that by providing a larger and more diverse number of stakeholders with access to computational capacity, data, and decentralised governance mechanisms, we will democratise the study of disinformation. This will lead to more diverse studies around disinformation, helping us to better understand how disinformation targets minorities across platforms. Our tool will also help us to design more effective socio-technical solutions to address the problem of disinformation, as well as facilitate open research and cross-country disinformation collaborations [23–25, 28].

Data Policy

Notice that for our tool to function, it is necessary to have access to data to conduct the disinformation research. For this purpose, our International Cloud holds a data library for storing disinformation data that stakeholders can use for their studies. To obtain initial access to data from communication platforms for disinformation research, our tool operates with subscription based API access. Here we work with the APIs that Facebook and Twitter [17,19,22,34] are already offering to scientists, as well as with the APIs that different companies, such as Meltwater [32], have made available. Meltwater is the world's first online media monitoring company, and it offers a vast range of APIs for collecting data from different sources. Specifically, Meltwater offers data from 15 different communication platforms channels (including access to all of Twitter's content), blogs, the comments from different online posts, news articles, and product reviews. Meltwater can also provide data for over 270,000 news sources globally, over 25,000 podcasts, as well as offering several different options for collecting data about TV and radio shows. Through this, we can fuel the initial data that is fed into our Research Cloud.

Data Privacy and Ethics

Our Cloud will be closed, meaning that only institutions and individuals approved by the governance team will be able to access its data and computational resources. The governance team itself would be the result of the cooperation between major research organisations such as the Federally Funded Research and Development Center (FFRDC) in the United States or the European Research Center (ERC) in Europe. Similar to practices in data collaboratives and data archives for sensitive data [33, 35], we will follow a philosophy that providing data access is not a pretext for privacy violations or erosion. The disinformation data that is shared will be aggregated and anonymised, following strict rules to ensure privacy. The governance team will work with Institutional Review Boards and ethicists on a set of operating principles ensuring that all disinformation research produced on the cloud respects human rights and privacy law.

One point that is critical to consider is how to enable data sharing between Europe and other continents. It is important to note that this has become increasingly complex because the European Court of Justice broke the privacy shield which was the principal mechanism that allowed for easy data sharing between Europe and the American continent[9]. The European court ruled that, due to the US Cloud Act and related laws within the American continent, the continent was not able to guarantee to European citizens that their data would be treated within a framework as strong as GDPR[10]. We therefore must study new mechanisms that can enable efficient data sharing between Europe and the U.S.

European Research Clouds & Public Computational Infrastructure in the US and Latin America

Considering the severe threats disinformation poses to the longevity of democracy, it is crucial for democratic nations and their shared democratic processes to coordinate research and devise cohesive strategies. We believe that together, Europe, the United States, and Latin America should establish a strong trilateral partnership on technology to put shared democratic values and the rule of law at the top of the agenda. Our proposal for a National Cloud for Disinformation Research builds on existing transnational Research Clouds such as the European Cloud Initiatives' Open Science Cloud (EOSC) and the Franco-German GAIA-X initiative [10, 12, 20, 36]. The latter is promoted as a "federated ecosystem of cooperation" to manage public problems from a multi-stakeholder perspective and enable computing power to develop better policies. Its Technical Architecture Report provides a framework for the core architecture, governance, operating ecosystem, security elements, and data-protection provisions that could serve our

9 — The European Court of Justice has ruled that Privacy Shield is invalid, Wired, July 2020.

10 — EU-U.S. Privacy Shield Program Update, Privacy Shield Framework.

proposed Research Cloud[11]. The Franco-German GAIA-X initiative argues that through the creation of this type of European-based public data infrastructure, the EU will become more competitive and promote innovative processes, products, and services [12].

We believe that our Global Disinformation Research Cloud could work within Europe's digital strategy which includes the Gaia-X project, VIGINUM[12], and Italy's new national hub to fight disinformation, the Italian Digital Media Observatory[13]. Together, we will enable a flourishing ecosystem where academics, companies, government, and NGOs work together to address the problem of disinformation. Our proposed governance model will facilitate a harmonised regulatory and standard framework through which stakeholders from different European countries will be able to access and use open and public disinformation data to understand how these campaigns operate within a wide range of different scenarios. Collaborating closely with these European initiatives will also help to unite fragmented research efforts around disinformation and help to understand how foreign disinformation campaigns can potentially play countries against each other [4, 9]. Connecting with the Gaia-X project and the Italian Digital Media Observatory will also help European countries to achieve digital sovereignty around the large-scale study of disinformation. We also believe there will be value in connecting with the European Commission, which launched a public consultation on its Digital Services Act package (DSA or IA Act)[14]. Working with the European Commission will likely help us to better identify the market imbalances that exist within different sectors to study disinformation at scale, as well as to create strategies for better supporting these sectors.

In the U.S. context, our proposal draws on the efforts of the Stanford Institute of Human-Centered Artificial Intelligence (HAI) to create a National Research Cloud for AI, including spawning a National AI Research Resource Task Force to investigate the feasibility and advisability of a cloud-based platform for large-scale AI research. Within the context of Latin America we will build off of the public computational distributed infrastructure we have designed previously with Mexico's presidency[15][16]. Should a National Research Cloud for AI be created in the United States or Latin America, there will undoubtedly be lessons to share across the Atlantic Ocean. These lessons may be codified in a unified research partnership to counter disinformation, a persistently destabilising force for an inherently fragile democracy. In short, if we are to make our democratic norms more resilient, we must learn and implement solutions that work, together.

11 — GAIA-X: A Franco-German pitch towards a European data infrastructure, Ministerial talk and GAIA-X virtual expert forum 4 June 2020, Livestream.

12 — LE SERVICE DE VIGILANCE ET DE PROTECTION CONTRE LES INGÉRENCES NUMÉRIQUES ÉTRANGÈRES (VIGINUM).

13 — The Digital Media Observatory Against Fake News and Disinformation is Born, Teller Report, September 2021.

14 — The Digital Services Act package, European Commission.

15 — BlockChain Hack MX, Gobierno de México.

16 — Modelo de Gobernanza para implementar la RedBlockchain México, Gobierno de México.

REFERENCES

[1] Law n° 2018-1202 of december 22 2018 on fighting disinformation. 2018.

[2] M. Abdalla and M. Abdalla. The grey hoodie project: Big tobacco, big tech, and the threat on academic integrity. In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, pages 287–297, 2021.

[3] M. T. Bastos and D. Mercea. The brexit botnet and user-generated hyperpartisan news. Social science computer review, 37(1):38–54, 2019.

[4] M. Baumann. 'propaganda fights' and 'disinformation campaigns': the discourse on information warfare in russia-west relations. Contemporary politics, 26(3):288–307, 2020.

[5] H. P. Beck and P. Charitos. The Economics of Big Science: Essays by Leading Scientists and Policymakers. Springer Nature, 2021.

[6] M. M. Ben Buchanan, Andrew Lohn and K. Sedova. Truth, lies, and automation: How language models could change disinformation. 2021.

[7] N. Bliss, E. Bradley, J. Garland, F. Menczer, S. W. Ruston, K. Starbird, and C. Wiggins. An agenda for disinformation research. arXiv preprint arXiv:2012.08572, 2020.

[8] A. Bovet and H. A. Makse. Influence of fake news in twitter during the 2016 us presidential election. Nature communications, 10(1):1–14, 2019.

[9] S. Bradshaw and P. N. Howard. The global organization of social media disinformation campaigns. Journal of International Affairs, 71(1.5):23–32, 2018.

[10] P. Budroni, J. Claude-Burgelman, and M. Schouppe. Architectures of knowledge: the european open science cloud. ABI Technik, 39(2):130–141, 2019.

[11] A. Carstens. Big tech in finance and new challenges for public policy. speech to FT Banking Summit, 2, 2018.

[12] E. Celeste. Digital sovereignty in the eu: challenges and future perspectives. Data Protection Beyond Borders: Transatlantic Perspectives on Extrater ritoriality and Sovereignty, pages 211–228, 2021.

[13] A. Esthose Suresh. People over profit: saving social media from big tech. LSE Business Review, 2021.

[14] S. Fachard, S. C. Murray, A. R. Knodell, and K. Papangeli. The fortress of eleutherai: New insights from survey, architecture, and epigraphy. Hesperia: The Journal of the American School of Classical Studies at Athens, 89(3):475–549, 2020.

[15] K. Fangen and C. R. Holter. The battle for truth: How online newspaper commenters defend their censored expressions. Poetics, 80:101423, 2020.

[16] Y. Gorodnichenko, T. Pham, and O. Talavera. Social media, sentiment and public opinions: Evidence from# brexit and# uselection. European Economic Review, page 103772, 2021.

[17] D. Jacobson, G. Brail, and D. Woods. APIs: A strategy guide. « O'Reilly Media, Inc.», 2012.

[18] A. Juhász and P. Szicherle. The political effects of migration-related fake news, disinformation and conspiracy theories in europe. Friedrich Ebert Stiftung, Political Capital, Budapest, 2017.

[19] R. Kimmons, J. Rosenberg, and B. Allman. Trends in educational technology: What facebook, twitter, and scopus can tell us about current research and practice. TechTrends, pages 1–12, 2021.

[20] K. Komaitis. Europe's ambition for digital sovereignty must not undermine the internet's values. Computer Fraud & Security, 2021(1):11–13, 2021.

[21] S. Kumar, R. West, and J. Leskovec. Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. In Proceedings of the 25th international conference on World Wide Web, pages 591–602, 2016.

[22] S. Lomborg and A. Bechmann. Using apis for data collection on social media. The Information Society, 30(4):256–265, 2014.

[23] I. V. Pasquetto, B. Swire-Thompson, M. A. Amazeen, F. Benevenuto, N. M. Brashier, R. M. Bond, L. C. Bozarth, C. Budak, U. K. Ecker, L. K. Fazio, et al. Tackling misinformation: What researchers could do with social media data. The Harvard Kennedy School Misinformation Review, 2020.

[24] N. Persily and J. A. Tucker. Conclusion: The challenges and opportunities for social media research. Social Media and Democracy: The State of the Field, Prospects for Reform, page 313, 2020.

[25] N. Persily and J. A. Tucker. Social Media and

Democracy: The State of the Field, Prospects for Reform. Cambridge University Press, 2020.

[26] F. Pierri, A. Artoni, and S. Ceri. Investigating italian disinformation spreading on twitter in the context of 2019 european elections. PloS one, 15(1):e0227821, 2020.

[27] O. Pollicino and E. Bietti. Truth and deception across the atlantic: a roadmap of disinformation in the us and europe. Italian J. Pub. L., 11:43, 2019.

[28] M. Saad, A. Ahmad, and A. Mohaisen. Fighting fake news propagation with blockchains. In 2019 IEEE Conference on Communications and Network Security (CNS), pages 1–4. IEEE, 2019.

[29] S. Savage. A national cloud for conducting disinformation research at scale. Policy Brief for the Federation of American Scientists 2021., page 1, 2021.

[30] J. Schoenfeld. Big tech, cloud computing, and accounting. Tuck School of Business Working Paper, (3596065), 2021.

[31] K. Starbird. Online rumors, misinformation and disinformation: The perfect storm of covid-19 and election2020. 2021.

[32] I. Stavrakantonakis, A.-E. Gagiu, H. Kasper, I. Toma, and A. Thalhammer. An approach for evaluation of social media monitoring tools. Common Value Management, 52(1):52–64, 2012.

[33] I. Susha, M. Janssen, and S. Verhulst. Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: Taxonomy development. 2017.

[34] R. Tromble. Where have all the data gone? a critical reflection on academic digital research in the post-api age. Social Media+ Society, 7(1):2056305121988929, 2021.

[35] S. Verhulst, A. Young, and P. Srinivasan. An introduction to data collaboratives. Creating Public Value by Exchanging Data. The GovLab, UNICEF, Omidyar Network, New York, 2017.

[36] I. I. Veršić and J. Ausserhofer. Social sciences, humanities and their interoperability with the european open science cloud: What is sshoc? Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare, 72(2):383–391, 2019.

[37] Z. Vorhies and K. Heckenlively. Google Leaks: A Whistleblower's Exposé of Big Tech Censorship. Simon and Schuster, 2021.

[38] R. Willis. Observations online: Finding the ethical boundaries of facebook research. Research Ethics, 15(1):1–17, 2019.

[39] D. Yoldas. Eu-russia information war, human right, and democracy-fake news, fact-checking, conspiracy theories and hate-speech in post-truth and illiberal democracies age.