



Received: 20 June 2021
Accepted: 16 August 2021

*Corresponding author: Katalin Parti,
Department of Sociology, Virginia
Polytechnic Institute and State
University 225 Stanger St, Blacksburg,
Virginia, 24061, United States
E-mail: kparti@vt.edu

Reviewing editor:
Sandro Serpa, University of the
Azores, Ponta Delgada, Portugal

Additional information is available at
the end of the article

SOCIOLOGY | RESEARCH ARTICLE

The Future of Interdisciplinary Research in the Digital Era: Obstacles and Perspectives of Collaboration in Social and Data Sciences - An Empirical Study

Katalin Parti^{1*} and Akos Szigeti²

Abstract: In the last decade, a transition in research design and methodology is identified in social research methodology; however, the high entry threshold (i.e., technical knowledge) to utilize computational methods and the ethical concerns seem to slow down the process. A possible way out is that social sciences collaborate with computational or data scientists in interdisciplinary research projects to rely on each other's skills and to develop jointly accepted ethical principles. In this exploratory study, we collected data from researchers with a variety of academic backgrounds to find out their views of interdisciplinary projects and related methodological or ethical issues. Our findings derived from one-on-one interviews ($n = 22$) reinforce the importance of interdisciplinary collaboration and highlight the significance of "interpreters," i.e., individuals able to communicate with and connect various areas of science, education, and academic institutions' role in enhancing interdisciplinary collaborations of sciences. Additional concerns of participants emerged in terms of research methodology applied in the digital world (i.e., data

ABOUT THE AUTHOR

Katalin Parti is an MA in Law and Political Science (Hungary, 1998), MA in Sociology (Hungary, 2013), a PhD in Criminal Justice (Hungary, 2008), and MSc in Criminal Justice with a Concentration on Cybercrime Investigation and Cybersecurity (Boston University, 2021). She is an Assistant Professor of Sociology, Virginia Polytechnic Institute and State University. A Fulbright and Max Planck Research scholar, she studies interpersonal cyber violence, gender, crime, sexuality, and the digital transition of social sciences.

Akos Szigeti is a sociologist and a PhD student at the Doctoral School of Law Enforcement at the National University of Public Service, Hungary. In his PhD work, he investigates the opportunities of adapting text mining in cyber criminology, specifically in the exploration of textual data on the darknet. This study helped to build the foundations of his PhD dissertation by exploring scholarly attitudes towards innovative research methods in the intersection of computer and social sciences.

PUBLIC INTEREST STATEMENT

We interviewed researchers in the intersection of social science and data sciences to share their experiences about teaming with other disciplines. Among the themes, educational challenges, interdisciplinary communications barriers, academic gatekeeping, the reliability of digital data, and ethical challenges concerning digital sampling, data collection, and analysis came to the forefront. The current paper can assist the reader in figuring out whether it is worth becoming a social scientist in the digital world. What would aspiring academics and graduate students face with research goals to examine socialization in the digital era? What are the realities to assess digital footprints without inter and transdisciplinary teams bridging across sciences? How helpful is the current academic environment to the publication and dissemination of research results in the intersection of social and data sciences? The paper also sheds light on the importance of transparency and open science in research, emphasizing that making data, analysis, and methods visible helps repeat and evaluate the findings.

validity, credibility and research ethics). Finally, participants identified open science and the transparency of research as the key to the future development of social sciences.

Subjects: Computer Science (General); Education - Social Sciences; Information Technology

Keywords: mixed methodology; interdisciplinarity; transdisciplinarity; social science; data science; collaboration

1. Introduction

As society is increasingly digitized, the number of interconnected users grows year by year (ITU, 2019). Citizens use social media networks and the personal and business services of these with increasing regularity (Digital, 2020). The role of sociologists in this situation is twofold. On the one hand, they must follow and analyze the latest processes in society, and on the other, they must explore the future trends of societal problems. Social science can only make use of the opportunities of digitization “if they are capable of renewing their research methodology while maintaining their critical reflection” (Németh & Barna, 2019, p. 121). This self-reflexion is a prerequisite of the adaptation of digital methods to rethink sampling procedures and the interpretation of data (boyd & Crawford, 2012; Karpf, 2012) to strike a bridge between the strict research ethics rules of social sciences and the ad hoc approach of data science, and develop these further (Salganik, 2018). It also represents a challenge to acquire the technical expertise required for applying data science methods, for which reason social scientists typically conduct research involving large data sets in interdisciplinary cooperation (Metzler et al., 2016). Teamwork involving multiple disciplines is increasingly emphasized in research, education, and policy (Guimaraes et al., 2019; Hadorn et al., 2008; Heinzmann et al., 2019; Schmalz et al., 2019). Choi and Pak (2006) used the words additive, interactive, and holistic to define multiple disciplines’ collaborative efforts. Collaboration can be multi-, inter-, and transdisciplinary, according to varying degrees and levels of collaboration: “Multidisciplinarity draws on knowledge from different disciplines, it stays within their boundaries. Interdisciplinarity analyses, synthesizes, and harmonizes links between disciplines into a coordinated and coherent whole. Transdisciplinarity integrates the natural, social, and health sciences in a humanities context, and transcends their traditional boundaries.” (Choi & Pak, 2006, p. 1). Since our research explored the intersections of social and data sciences, “interdisciplinarity” and “transdisciplinarity” are used interchangeably in this paper. For our research, we wish not to make a strict distinction between the two since we suggest an integrated approach involving social and data sciences in the humanities context, not excluding other natural sciences from the collaboration. In the paper, we also use “multiple disciplinarity,” an umbrella term suggested by Choi and Pak (2006) that conceives all three concepts and provides different perspectives on problems and offers comprehensive services.

In this study, we explored the possibilities and obstacles of cultivating social science based on collaboration by teaming with data sciences in research projects. In addition, we examined the willingness and tendencies of researchers to use methods in the intersection of social and data sciences to study digital social life and tools capable of collecting digital footprints. Our purpose was to serve the development of social science research by formulating questions about the adaptation of digital methodological tools, uncovering the methodological pitfalls of using digital transactional data, and identifying the questions of research ethics in the digital era.

2. Theoretical background

2.1. The digitization of social life as a subject of research

The research of the digital landscape of the 21st century has new opportunities in store for social scientists. Social media websites, video sharing sites, online games, and immersive worlds have become places where young people socialize (Ito et al., 2008). The digital world created new opportunities for the Y and Z generations to learn and practice social norms and behavioral patterns, to find areas of interest, acquire new skills, experiment with new forms of self-

expression, test their independence, and set out on their journey to adulthood (Sparks & Honey, 2014). Socialization in the digital age challenges families to handle the risks accompanying the above-mentioned opportunities (Livingstone & Blum-Ross, 2020). Digital platforms also represent domain for adults to realize their professional and private goals for self-development and entertainment (Madden, 2010; Villanti, 2017). The mass presence of users in the digital space enables sociology, anthropology, philosophy, criminology and other social sciences to examine human behavior and societal phenomena.

The continued digitization of society required a research approach with interdisciplinary foundations capable of integrating social sciences and information technology knowledge (Loader & Dutton, 2012). We can find examples of computer science and social sciences working together from ethnographic field study (Goulden et al., 2016) through cyber criminology (Lavorgna, 2021) to research on cybersecurity (Jacob et al., 2019). Computational Social Science and Social Computing came to scene as adequate fields of blending the knowledge of computer and social sciences. The former area relies on the cooperation of two sciences: while social scientists lay the foundations for the collaboration by providing the context required for interpretation, the methodological approach and research questions, statisticians and IT experts have the knowledge of mathematical modeling, calculation tools and collecting data sources. In contrast, social computing deals with the interaction of people with computer systems. Researchers in this field are looking for answers to questions such as why and how users generate content and how they are aided in this by the design and development of systems (Mason et al., 2014). The overlap of these two areas has already led to a possible fusion: Evans (2020) proposed to reimagine Computational Social Science as Social Computing, based on “recognizing societies as emergent computers of more or less collective intelligence, innovation and flourishing” (Evans, 2020, p. 1). We assume that social scientists know and use the tools of Computational Social Science and/or Social Computing. Yet, we still find relatively few examples for the cooperation of social sciences and IT to develop research methodology tools. Golder and Macy (2014) said that the “virtual laboratory” created by online communication in social media can only be researched in possession of complementary technical skills, which extend to the collection, storage, manipulation (systematic organization), analysis, interpretation and validation of data.

2.2. Digital footprints as assessment opportunities

As Salganik said, “Researchers are in the process of making a change akin to the transition from photography to cinematography” (Salganik, 2018, p. 5). This transition needs time and requires an innovative mindset of the researchers. There is an emerging trend (Zyoud et al., 2018) for using user data readily available in social media and adapting data mining and automated analytical tools to understand the latest societal phenomena and explore future challenges in society. According to Savage and Burrows (2007), web-based transaction data may show the way out of the “empirical crisis” of sociology (Savage & Burrows, 2007, 2009) caused by the hollowing out of the traditional research methods of social sciences in the digital world. Commercial sociology may represent a new lease of life for sociological research (Burrows & Savage, 2014), i.e., behavioral research based on big data developed by for-profit companies for market research and focusing on digital user communication. Commercial sociology uses traditional social research methods but adjusts those radically to web-based transactional data, i.e., the nature of large transactional databases generated as a by-product during the use of public, industrial or digital services (Savage & Burrows, 2007). Starting with the emergence of web 2.0 and social media, users no longer generate datasets by using services but continuously communicating and producing their own content (Burrows & Savage, 2014). Osborne et al. (2008) go as far as saying that the statisticians, journalists, economists, educators, communication analysts, activists, and policymakers, who make use of the digital capital and systematically analyze digital data, using a combination of research methods, do better social research than professional sociologists, who only use the acknowledged (or pre-digital) scientific methods of sociology (Osborne et al., 2008). According to Burrows and Savage (2014), this list should also include data scientists beginning 2013, since that

was the first year when according to Google Trends, there were more searches for “data scientist” than “statistician” (Burrows & Savage, 2014; McKie & Ryan, 2012).

2.3. The increasing difficulties of conventional assessment

Although the digitization of society made data collection simpler in many aspects (online survey tools, rapid data processing, cheap data storage) (Tourangeau, 2004), there have been numerous arguments against the use of web-based data for social sciences. Specific authors (e.g., Golder & Macy, 2014; Uprichard, 2012) considered the ahistoric nature of web-based data, i.e., the fact that they are only suitable for assessing a snapshot in time (this is what Erickson called “persistent conversations”, Erickson, 1999); whether the data used are ideal for estimating future trends does not only depend on methodology but also the nature of data sources. According to Golder and Macy (2014), the results of network analyses are difficult to interpret due to potential distortions, since what and how many other persons participants consider as members of their network greatly depends on how the questions are formulated (see also: Butts, 2009). We face a similar dilemma when deciding where the users’ sampling threshold to be involved should lie (Preece & Shneiderman, 2009). Researchers also face the challenges of generalizing online behavior since online interactions differ from offline ones (Golder & Macy, 2014). Tinati et al. (2013) saw in digital communities a heterogeneous, uneven, constantly changing, and intangible global stream of humans, information, objects, money, images, and at the same time risks, which would also demand that we rethink how we interpret sampling procedures and results (boyd & Crawford, 2012; Karpf, 2012).

One of the problems is the validity of data appearing on digital platforms. The online data mass does not represent the diversity of society (*platform bias*: Ruths & Pfeffer, 2014). The distortion related to the nature of the platform arises from the preferences of the persons visiting it and, understandably, determines the answers received on the platform. Popular culture is no longer homogeneous due to the diversity of the internet. It is a rather fragmented and colorful culture, segmented by the advertisers of online platforms and programmers, who direct users into various “lifestyle clusters” (Weiss, 2000). *Sampling bias* is a common phenomenon in online sociology research; it occurs if a researcher uses only one platform to collect data (Tufekci, 2014). The concept of *data availability bias* (Ruths & Pfeffer, 2014) is related to this and means that we can only use the data of registered, communicating users on the given platform. Another problem is *data authenticity*: although data may be distorted even collected by a paper-based questionnaire, the image displayed in social media may be oblique in many more ways (Xiang et al., 2018). *Survey fatigue*, i.e., the validity problem caused by the loss of interest of users, results from the fact that technology has made online opinion polls easier and cheaper (Massey & Tourangeau, 2013). Several researchers discussed the methodological and ethical concerns regarding big data (Frické, 2015; Massey & Tourangeau, 2013; Ruths & Pfeffer, 2014), for which, however, there is no general solution.

Research ethics should be given even more attention in the case of big data analyses: the social science approach based on strict rules and the data science’s ad hoc approach has to be developed further into a principle-driven solution, where observing such fundamental ethical principles as the respect of individuals, the principle of beneficence and fairness, and the respect of the law and public interests is a priority (Salganik, 2018).

At the same time, it is challenging to enforce ethical norms in the digital research space, where large amounts of data are available. Users did not initially upload their data—or joined an online community—to participate in research. Users can be participants in research projects and may make their data available for analysis without recognizing that they agreed to this by joining a digital community or downloading an application. Digital data is itself a commodity. *Data capitalism* (Mayer-Schönberger & Ramge, 2018; Sadowski, 2019) is a new driving force of the economy in the era of big data when the traditional function of money as an asset is replaced by user data uploaded to digital platforms. *Surveillance capitalism* (Zuboff, 2019) is

only one step forward from this. It arose from data capitalism, and its goal is to observe and analyze user data by data oligarchs as Google, Facebook, Amazon, or the Chinese company Baidu. Tech giants glean user data and use it for their economic purposes (see also: *platform capitalism*: Srnicek, 2017). One such goal, in particular, is predicting or influencing consumer behavior. Total surveillance, the concept on which Jeremy Bentham based his idea of the modern prison, Panopticon (Bentham, [1813] 2008), and about which Foucault said that mass media is a perfect instrument for influencing behavior (Foucault, [1975] 1995), has reached its peak in the internet era with the data collection, data harvesting, and data mining activity of digital platforms (Parti, 2008). Although Bentham's surveillance concept only tackled monitoring of prison inmates, where prisoners were watched knowingly and directly by a human being (the guard positioned in the center of the Panopticon, the modern prison), Foucault's surveillance concept, in the dawn of television as mass medium, criticizes mass media for dispossessing viewers of free thinking entirely by penetrating homes every night through the evening news, where viewers do not even know that they are "controlled." Overall, early philosophical studies (for an overview, see Anders, 1956) already suggest a grim picture of mass media, positing that, with information society coming into age, surveillance is getting voluntary (by simply using social media), and highly manipulative of desires, realities, and behaviors of users. This original concept has just become more articulated in the age of the internet and social media (e.g., Zuboff, 2019).

A further problem is the scientific grant system, which only supports interdisciplinary collaboration involving humanities in rare, exceptional cases. Although interdisciplinarity may shed light on the critical societal and economic impacts and the mechanisms of digitization, social scientists can only collaborate with related humanities (such as philosophy, education, and communication science, history, archaeology, and tourism), as the grants intended to encourage cooperation between different sciences are announced for other sciences (Khan et al., 2019). The findings of scientific research are best published at conferences and academic journals for two reasons: to promote the evaluation, replication, and further development of the results, and also because the academic promotion and tenure system reward such publications. However, journal editors and reviewers may reject studies merely because they present unusual findings or are realized as uncommon cooperation between disciplines. This phenomenon is called *science-metrics gatekeeping*, the existence of which is supported by empirical data mainly in the field of medical publications (Siler et al., 2015), but the phenomenon is widely and long known, also in the field of social sciences (Gieryn, 1983; Lindsey, 1988; Peters & Ceci, 1982). The concept of gatekeeping appears in our study concerning interdisciplinary collaborations.

The novel, computer-aided analysis (machine learning, sentiment analysis, topic modeling) of data available on digital platforms may involve another, separate set of problems in the context of analyzing data for the purpose of social science, which we do not describe in detail here (for a summary analysis see Golder & Macy, 2014). All in all, computerized analytical methods have not spread widely in social sciences because—since they are atheoretical applications—predicting linguistic observations has greater value for them than assessments based on conventional, theory-based hypothesis testing (Golder & Macy, 2014).

2.4. The links between open science and IT development

The transparency and reproducibility of scientific knowledge are characteristics of scientific research, which academic journals increasingly require. Still, we know of very few research projects where the available documentation covers everything from the idea to the raw data through results (Toelch & Ostwald, 2018). Womack (2015) believes that only 13% of published studies contain raw data, and even less contain the codes needed to analyze the data (Morin et al., 2012). This may have several reasons, such as data privacy regulations, pending patents, lack of technical know-how, or platform capitalism itself (Srnicek, 2017). Platform capitalism means that the *sharing economy* also called *gig economy* makes use of the digitization of communication and tries to generate profit by collecting and selling user data. The majority

of user data of online platforms is in the hands of private companies, which only allow restricted access to the data. The lack of transparency, however, limits not only accessibility but also reproducibility. This prompted academic journals (Stodden et al., 2013) and research financing agencies (Baker, 2016a; Collins & Tabak, 2014) to introduce strict rules for publishing hypotheses, data, and analytical processes, collectively referred to as *open science*. Open science increases the accessibility of research: studies accessible free of charge (without a subscription) are more often cited, in particular, if they also publish raw data and codes. According to a report published by the European Union in 2017 (O'Carroll et al., 2017), there is a great need for education that extends to the practice of open science, but little such educational effort is known so far (Schmidt et al., 2016; Teal et al., 2015). Authors of scientific publications have also become aware of the *reproducibility crisis*: according to a 2016 survey of *Nature*, 90% of the authors of scientific publications think the crisis has become significant (Baker, 2016b, 452).

Our research aimed to identify the obstacles in interdisciplinary research collaborations in the intersection of social sciences and data science. Although many studies are dealing with the impact of big data on sociology, they do so in a purely theoretical manner. Our project is novel because we directly asked those involved about the opportunities of using big data for social sciences while throwing some light on the concerns and preferences related to allowing data capitalism in sociology. In the following, we conceptualize interdisciplinarity, describe our sample and methodology, introduce our findings, and then interpret them. Finally, we provide a summary of the conclusions of the study important for interdisciplinary collaborations.

3. Methods

Our exploratory research was based on expert interviews, which were aimed at an in-depth exploration of the willingness of social and data sciences teaming up in interdisciplinary teams. We also conducted an online survey based on purposive sampling, where we tried to assess which digital data-oriented research methods and tools (e.g., data analysis software) are being used by social scientists. Although the survey results are not included in this study due to the small sample size, we consider it important to present our full research design in the Methods section. The research plan was approved by the Institutional Review Board of Virginia Tech under #19-882.

3.1. Survey

We asked professionals working in the field of social research about the following in an online questionnaire (survey): (1) What data collection approaches and methods they used based on SAGE's research methodology map, which is a comprehensive collection of research designs and methods utilized in contemporary social sciences (SAGE, n.d.); (2) What are the research methodology tools they used; (3) What digital tools and software they used to analyze the results; and (4) What significance inter- and transdisciplinarity have in today's social research.

We disseminated the questionnaire online, through the following public channels and professional forums of LinkedIn and Facebook: LinkedIn Research Methodology group, LinkedIn Criminal Law, Criminology and Cybercrime community, LinkedIn Research Methods and Data Science community, LinkedIn Qualitative Research Using Social Media, LinkedIn Social Innovation Research, Socio-economic Sciences and Humanities, LinkedIn Applied Sociology, LinkedIn American Society of Criminology Group, LinkedIn Law Enforcement Crime Center, LinkedIn Research Methodology, Facebook Qualitative Research Group, Facebook Research Scholars Group, Facebook Data Science. We also sent targeted emails to mailing lists of professionals, such as the Cybercrime and Women & Crime¹ working groups of the European and the American Society of Criminology, and the mailing list of the American Society of Sociology. We also disseminated the questionnaire through personal contact, e.g., at Virginia Tech's Data and Decision conference and the in-person conferences of the American Society of Criminology in November 2019. The questionnaire could be filled in between 1 November 2019, and 31 December 2020. Participants could join from all over the world, the only prerequisites being 18 years of age or more, command of the English language, and

professional involvement in social science research at an academic level which were indicated in specific questions. Despite considerable efforts for dissemination, as a result of which 813,131 potential participants received the questionnaire, only 126 filled it out. After filtering out incomplete items, we could include only the answers of 84 participants in the analysis. This sample size does not allow us to include the survey results in this study.

3.2. Expert interviews

We approached those participants who provided their email addresses on the questionnaire and thereby agreed to be interviewed as experts. We also reached out to social researchers and data scientists with research intersecting data science and social sciences, working at universities in the United Kingdom and the United States. Our aim with the interviews was to (1) gain a deeper insight into the characteristics of social research methods applied in the digital world and (2) the possibilities and obstacles of their use. In order to obtain this information, we composed a semi-structured interview guideline, which also let us ask specific questions related to the field of study or methodological approach of the particular interviewee. Seventeen participants who filled in the questionnaire let us know that they would do an expert interview, but we managed to interview only 5 of them. In addition, we conducted 17 expert interviews with researchers in the field of social and data science using purposive sampling. In sum, we conducted 22 interviews. Interview participants had degrees from various fields, such as in criminology, criminal law, criminal justice, philosophy, psychology, anthropology, computing science, computational social science, biological engineering, business analytics, and mathematics. Since we have disseminated the online survey tool mostly on conferences and online platforms for researchers of crime, criminology, and crime science, it is not surprising that most of the interview participants had research related to these areas. Consequently, and also a limitation, our results should be interpreted with caution as they mainly reflect ideas from these particular sciences. Most of the interviews took place online due to the global pandemic or the interview participants' remote location (17 interviews). We transcribed the audio recordings of the interviews with the speech-to-text transcription solution of otter.ai. Team members coded all interviews separately, and then all interviews were matched for code differences. The research team used Atlas.ti cloud version for coding; thus, the process of code matching had been performed online, in the cloud, by team members discussing discrepancies and deciding about recoding differently coded text segments (Armstrong et al., 1997; Bryman & Burgess, 1994) upon need. We identified 44 codes in the interviews, but we only mention the most important of these codes in this study. The interviews were conducted in English. Findings are supported throughout the study by interview excerpts.

4. Results

4.1. Educational challenges

When we asked participants about their willingness to create interdisciplinary research teams and applying methodologies and tools for data harvesting and the study of digital lifestyles, one of the most clearly identifiable observations of the participants was related to education. Only a marginal number of social sciences curricula currently contain modules, which provide guidance for navigating the complexities of the digital, data-based approach, digital sampling, data recording and analysis. The fact that higher education is slow to catch on with digital developments also shows in the preparedness of social sciences. When social scientists do not make use of the opportunities in the digital world, it is not the attitude of the scientists that is to blame, but rather the lack of relevant training and preparedness. In countries where the education of social sciences does not contain a sufficient number of classes in research methodology for digital platforms, students will graduate without being able to use these methods or without knowing the opportunities on digital platforms, such as social media research, data mining or mass collaboration.

Understanding an IT and programming-based approach and sampling represents a significant challenge for social scientists educated in traditional methodology used in offline, pre-digital environments. Problems primarily arise concerning planning and data collecting since the

experiment-based planning used on digital platforms limits the representativeness of data and the reliability and transparency of the research. While the mathematical approach involves purely logical analysis, social science relies on context. However, this does not mean that the digital approach and sampling cannot, by definition, be acceptable for social sciences, only that compared to the ad hoc nature of data science, social science is overly rigorous as regards planning and sampling.

“The rigor of data collection that social scientists apply is difficult to translate to the digital age.” (Participant #6)

“The verbal rigor is much more developed in social sciences in general than it is in computer science. We [in computer science] are much more ad hoc.” (Participant #16)

4.2. The need for “interpreters”

Additional problems attribute to the difficulties of cooperation between sciences. Representatives of data sciences and social sciences were socialized in very different research environments and use different terminology. As recommended by interview participants, transdisciplinary research groups can only work successfully with the help of “liaison officers” or “interpreters,” capable of transdisciplinary conceptualization and can translate the various concepts and terminology to the language of the others.

“It’s very hard for a researcher to have all of the skills in practicing different methods. It’s also really hard to assemble a team of people who have different skills and also good at talking to each other. And you sort of have to do one or the other. Either I have to be able to do statistics and philosophy and a systematic review [...] or I have to find other people who want to collaborate on that project, and we have to be able to talk past disciplinary boundaries.” (Participant #7)

“There are difficulties along the way because these are two very separate disciplines. And there has to be a translator between them to facilitate that communication ... but because they don’t understand each other, the collaboration terminates itself at some point.” (Participant #14)

4.3. Academic gatekeeping

At the same time, scientific cooperation should be transdisciplinary rather than interdisciplinary. Since computer science is not sufficient to establish causality, it is the analytical approach of mathematicians that is suited for the purpose. This gives rise to further complications, which could only be resolved if rigid university structures are bypassed. However, university departments prefer research methods, accepted by the given science, tried, tested, and based on doctrines. It is also a prerequisite of obtaining a doctorate to prove that they can apply the accepted methods and produce results using them relevant to science, based on a conventional, accepted methodology. Another obstacle in inter- and transdisciplinarity is the gatekeeping activity of scientific publication outlets, which do not let methods diverging from doctrines pass. Scientific publication outlets tend to reject analyses that test known theories in an unconventional (digital) environment, with unconventional methods (data mining, mass collaboration).

“A transdisciplinary approach is one of those things that everyone endorses, and almost no one does. ... Universities remain set up in disciplines [...]. Journals remain set up in disciplines ... you know, the *Journal of Psychology*, the *Journal of Geography* and research councils ... as set up in disciplines and we’ve applied for grants, and we’ve been told it’s not social science enough.” (Participant #11)

“I don’t think academia values this collaboration. They say they do. But actually, fellow reviewers at conferences or journals make it difficult to get to perceive interdisciplinary research the same quality as single-discipline research.” (Participant #6)

4.4. Digital data reliability

If we meet the requirement of precision for the methodology, data reliability still remains a question. The validity of results is at risk due to the lack of reliable data. The members of online communities may present an image of themselves on digital platforms different to reality, which casts doubt on the truthfulness of the results, the quality of data, and the general applicability of the findings.

“If you’re going to deal with a database using digital platforms to collect data, the quality of data is crucial because if you fail with the quality of data, you are failing with the conclusion.” (Participant #4)

“One of the problems with these datasets in comparison to traditional social science methods is we know very little about the sample.” (Participant #11)

“When you’re developing a questionnaire, you do all these tests of validity. Whereas here, it’s like, oh, there’s this data set. I think it represents this. But it is vital to understand what that data represents.” (Participant #13)

“Big data has really provided us with many challenges, computing challenges because there’s so much data but actually ... getting it integrated in ways that allow for analysis is really challenging.” (Participant #19)

4.5. Ethical challenges

In addition to the reliability of data, the question of ethical research also arises on digital platforms. Is it ethical to collect data on a platform where users do not know that the data they publish for public access will also be used for research purposes? Has the user automatically agreed to participate in the research by registering on the given online platform? Is it ethical to link various databases, which do not contain personal data, but user identity can be established from the combined data? Is it ethical to use data mining in crime prevention projects where citizens were not informed that they would participate in such a project? The question of data capitalism and surveillance capitalism also arises concerning research ethics.

“Should we actually understand humans better than they understand themselves? Should we have access to so much data about an individual?” (Participant #12)

Digital data collection has its pitfalls, although it is more convenient in many ways than data collection based on conventional social science sampling. Data on digital platforms (e.g., social media) are not uploaded by users for specific envisaged research. We do not know precisely how behavior-influencing mechanisms work, what influences—messages, personalized advertising, nudges, and triggers—users were exposed to, which prompted them to make the given decision. Social scientists cannot factor all these underlying processes into the analysis formula, so research planning cannot be precise. It is impossible to select an analytical method suited for the characteristics of the data. One interview participant used the example of decision processes in cyberattacks based on social engineering to illustrate the complexity of thought processes scientists need to understand. However, user data alone do not allow for drawing too sophisticated deductions about how the actual cyberattack happened.

“How do we formulate questions that capture the mental decision process that people using technology go through when faced with social engineered cyber-attacks?” (Participant #2)

“I think the challenge is to make sure that people really understand the data they’re using, and any particular shortfalls within it, and the potential for biases within the data.” (Participant #15)

4.6. Challenges of open science

Although the publication of data and databases promotes the verifiability, transparency, and reproducibility of scientific findings, it also raises ethical questions. The sharing of codes and codebooks developed by scientists increases transparency and reproducibility of results and

provides an opportunity for the continued development of open-source software used for scientific data analysis.

“I think there is really an overdue change in the field to the use of open-source software, and have reproducible, basically script-based software. Not only because there is a lot more, in some cases, more sophisticated work to be done, but also because they provide more rigor and reproducibility.” (Participant #8)

Transparency is a prerequisite of reproducibility and reliability and may promote cooperation between scientists. There are numerous good practices for achieving transparency, such as the publicly accessible publication databases of scientific journals (e.g., the COVID databases created during the pandemic²) or institutional research seminars. However, the crisis in academic trust is still preventing cooperative processes, according to interview participants.

“I think the key in legitimizing science is how much trust we can place in published findings. On the one hand we have to trust that the data has been processed in the correct way and that the person has not made an error in the process of performing analysis on the other. These are the kinds of things which lead to a reproducibility crisis.” (Participant #8)

“A drive towards transparency and research. One thing to achieve this is running a seminar series on open research and moving people towards open tools to understand open science. And yeah, pushing that transparency and clarity.” (Participant #15)

Transparency is an essential criterion of future social research: research using new methods or combining known ones in new ways will be seen as legitimate and an example to follow if the findings are verifiable and reproducible.

“People need to be more cognizant of their metadata. People need to be a lot more transparent about what exactly they’re doing to their data, so that it can be more replicable across the board. Those data, any data and metadata should be published with the article, much like in the hard sciences.” (Participant #17)

5. Discussion

In this research, we explored scientists’ attitudes towards social research methods for the exploration of digital footprints, and data harvesting, during one-on-one interviews ($n = 22$). The goal was to map scientists’ willingness to work together in interdisciplinary teams and map the obstacles of interdisciplinary collaboration. Although several sciences are dealing with the impact of data capitalism and big data on sociology on a theoretical level, our project fills a gap in that we asked scientists to tell us what *they* think about the points of connection between social and data science and the obstacles of their cooperation.

The continued digitization of society and the changes in response to the pandemic are putting pressure on social researchers and force them to adapt and digitize rapidly. The digitization of social sciences more than ever requires interdisciplinary planning and a preference for methodologies, sampling, data collection tools, and online dissemination techniques suitable for digital platforms. This makes it necessary to develop ethical guidelines and practices, which take big data and social media characteristics and the challenges these pose into account. Representative sampling is one such challenge in the digital space, as is the analysis of digital data based on a rigorous methodology, highlighting causal links, and respecting the privacy of the research subjects.

All this requires a contribution of data sciences, since social scientists, no matter how open to computer programming and mathematical analysis, will not by themselves be able to cope with the above hurdles while remaining competitive in the ever-powerful market of data capitalism (Mayer-Schönberger & Ramge, 2018; Sadowski, 2019). Scientific collaborations play an essential role in this process. Thus, the theory-oriented social sciences approach is to be coupled with data

sciences' digital tool (application, software) developer approach when focusing on big data analysis.

Interview participants were open to scientific collaborations and innovative solutions, but—when we asked—also listed the problems that slow down the digital shift in social sciences, making forming multiple disciplinary research groups difficult. They considered the most severe issues as follows: education is lagging behind the requirements of research into the digital, interdisciplinarity teams need “interpreters” for bridging between terminologies, academic gate-keeping is real, digital data suffers from validity and reliability issues and poses ethical challenges, and the lack of academic trust hinders open science.

The recommendations formulated during our study may serve as important guidance for the future of social sciences. Such is the idea that imparting scientific research methodology, research planning, data collecting, data analysis in the digital world should be given more emphasis in higher education. Cooperation between various branches of science and between different sciences should also be promoted in academic institutions. These recommendations are absolutely compatible; what is more, Golder and Macy (2014) believe that multidisciplinary collaborations are only suited short-term to make up for the lack of skills. Specific modules, which in the age of big data prepare scientists for online research (e.g., programming interfaces, use of APIs, manipulating unstructured data, building online survey databases, machine learning, topic modeling) will have to become part of the postgraduate programs of universities in the long run.

Participants agreed that IT, mathematics, other classic sciences, and social sciences need to rely on each other's resources in data collection and data analysis in the digital world. At the same time, collaborative group work involving humanities and natural sciences has no tradition and usually does not make it past the “academic gatekeepers” (i.e., publication outlets and academic departments). Although doctrine-based science does contribute to further assessing existing findings, it only tolerates interdisciplinary results in exceptional cases. Scientific journals established to publish scientific findings and scientific conferences will have to stop methodically excluding other sciences. This would allow projects to be realized as collaborative efforts of data sciences and social sciences, which use uncommon methods and tools to arrive at their scientific findings. The fact that several areas of science need to cooperate in a given project does not automatically mean that the findings will be unreliable. Both the organizers of conferences and the editorial teams of journals will have to accept the risk of publishing interdisciplinary results since the emergence of novelty or innovation can be specifically expected due to such collaborations.

In this context, another recommendation of the research participants would be to have more active communication between sciences and science branches. For interdisciplinary research to be possible, there is a need for, as participants put it, “liaison officers,” who understand the language of the various sciences, know the terminology and act as a kind of interpreter-translator in the communication between the teams of researchers or within interdisciplinary research groups. Data sciences and social sciences may have different sets of terminology, but that does not mean that they are incapable of cooperation. People are needed who take responsibility for the coordination and for resolving the differences in vocabulary.

When we asked interview participants to tell us in their own words what directions social research should have in the digital world, it was not the use of digital tools and smart devices, but rather inter- and transdisciplinarity, a combination of existing methods, and interoperability between areas of research that they highlighted. They did not obviously consider applying new methods, not even the application of digital technologies. Instead, they suggested combining existing methods, the cooperation of sciences, and selecting the best method for the research. Although large number of users on digital platforms increases statistical power (Erickson, 1999; Golder & Macy, 2014), interview participants also confirmed that digital sampling might compromise the representativeness and reliability of data and the validity of results (Golder & Macy, 2014).

Regarding ethical concerns, they mentioned the lack of consent of subjects, the dilemma of combining databases, and the confidentiality and publication of databases. They suggested that researchers should make sure that before using new methods, they are relevant and fitting to the project goals, and data reliability should be verified with yet another method.

Undoubtedly, the participants recognized the potentials of open science: ensuring the reproducibility of findings, the transparency of science, sharing and publishing data and databases, rigorous description of research methods, and the verifiability of research findings. All the above promote the transparency of science and help resolve the *reproducibility crisis* (Baker, 2016b; Toelch & Ostwald, 2018). The forced digitization as a result of the pandemic will likely have a positive effect on sciences, both as regards making scientific findings public (e.g., through broader accessibility of conferences taking place in the digital space), and the establishment of inter- and transdisciplinary research groups. Several leading scientific publishers made their website freely accessible (called Science 2.0 by Mirowski, 2018), and journal editors encourage cooperation between the sciences (Jamali et al., 2020). As open databases become widespread, the time needed for review is also shortened. Further research groups based on the cooperation of humanities and natural sciences may be established to examine the societal impacts of the pandemic (Holmes et al., 2020) if their creation is promoted by new grant application procedures supporting collaboration. One example is Scholarly Communications Networks (SCNs), enabling the establishment of private, invitation-based research groups for individual research projects, which become particularly significant in times of large-scale social and economic changes, like those that started in 2020 (Beeny et al., 2020). These include groups researching the pandemic's societal, medical, psychological, and neurobiological impacts (Holmes et al., 2020).

Overall, our findings highlight researchers' openness to establishing interdisciplinary collaborations between social and data science that itself is a suggested way to develop social science in the future. Our conclusions confirm the statements of previous research about multiple disciplinarity. Still, their validity should not be considered absolute since we could not support our qualitative results with analyzable quantity of survey data, and most of our participants are from the field of criminology, crime research, and sociology. Nevertheless, we are convinced that future development in social sciences can be encouraged by establishing interdisciplinary research teams which can provide a better understanding of and more successfully manage the problems arising in a digital society.

Funding

The study was supported by the Hungarian Ministry of Innovation and Technology. Grant number: KDP-2020/1967.

Author details

Katalin Parti¹

E-mail: kparti@vt.edu

ORCID ID: <http://orcid.org/0000-0002-8484-3237>

Akos Szigeti²

ORCID ID: <http://orcid.org/0000-0003-3641-1648>

¹ Department of Sociology, Virginia Polytechnic Institute and State University 225 Stanger St, Blacksburg, Virginia, 24061, United States.

² Doctoral School of Law Enforcement Sciences, National University of Public Service 1441 Budapest, Hungary.

Citation information

Cite this article as: The Future of Interdisciplinary Research in the Digital Era: Obstacles and Perspectives of Collaboration in Social and Data Sciences - An Empirical Study, Katalin Parti & Akos Szigeti, *Cogent Social Sciences* (2021), 7: 1970880.

Notes

1. The Women & Crime working group of the American Society of Criminology consists of scientists using a feminist research method, who characteristically rely on a combination of quantitative and qualitative data collection mechanisms and a mixed methodology in their research projects.
2. See for example, LitCovid: <https://www.ncbi.nlm.nih.gov/research/coronavirus/>; or the database of the National Institute of Health about the latest medical research: <https://www.nih.gov/coronavirus>, but bigger social science journals (e.g., Wiley Online Library) also made their databases publicly accessible with respect to the coronavirus pandemic.

Disclosure statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Anders, G. (1956). The world as a phantom and as matrix. *Dissent*, 3(1), 14–24. Retrieved from August 14, 2020 <https://marcuse.faculty.history.ucsb.edu/projects/>

- anders/
 Anders1956DissentWorldPhantomAndMatrixOCR.pdf
- Armstrong, D., Gosling, A., Weinman, J., & Marteau, T. (1997). The place of inter-rater reliability in qualitative research: An empirical study. *Sociology*, 31(3), 597–606. <https://doi.org/10.1177/0038038597031003015>
- Baker, M. (2016a). Dutch agency launches first grants programme dedicated to replication. *Nature News*. July, 20(2016). <https://doi.org/10.1038/nature.2016.20287>
- Baker, M. (2016b). Is there a reproducibility crisis? *Nature*, 533(7604), 452–453. <https://doi.org/https://doi.org/https://doi.org/10.1038/533452a>
- Beeny, M., Hotchkiss, V., & Spence, C. N. (2020). The Impact of COVID-19 on Scholarly Communications. August 26, 2020, *H-Net: Humanities and Social Sciences Online*. Retrieved from September 28, 2020 <https://networks.h-net.org/node/1883/discussions/6367139/13-impact-covid-19-scholarly-communications>
- Bentham, J. ([1813] 2008). *Panopticon: Or, the Inspection House*. Dodo Press.
- Bryman, A., & Burgess, R. G. (Eds.). (1994). *Analyzing Qualitative Data*. Routledge. <https://doi.org/10.4324/9780203413081>
- Burrows, R., & Savage, M. (2014). After the crisis? Big Data and the methodological challenges of empirical sociology. *Big Data & Society*, 1(1), 205395171454028. <https://doi.org/10.1177/2F2053951714540280>
- Butts, C. T. (2009). Revisiting the foundations of network analysis. *Science*, 325(5939), 414–416. <https://doi.org/10.1126/science.1171022>
- Choi, B. C., & Pak, A. W. (2006). Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: 1. Definitions, objectives, and evidence of effectiveness. *Clinical and Investigative Medicine*, 29(6), 351–364. <https://doi.org/10.25011/cim.v31i1.3140>
- Collins, F. S., & Tabak, L. A. (2014). NIH plans to enhance reproducibility. *Nature*, 505(7485), 612–613. <https://doi.org/10.1038/505612a>
- Digital. (2020). Digital 2020: 3.8 billion people use social media. *Digital 2020 Reports*. <https://wearesocial.com/digital-2020> Retrieved from May 13, 2020
- Erickson, T. (1999). Persistent conversation: An introduction. *Journal of Computer-Mediated Communication*, 4(4), 447–450. <https://doi.org/10.1111/j.1083-6101.1999.tb00105.x>
- Evans, J. (2020). Social Computing Unhinged. *Journal of Social Computing*, 1(1), 1–13. <https://doi.org/10.23919/JSC.2020.0002>
- Foucault, M. ([1975] 1995). *Discipline & Punish: The Birth of the Prison* (2nd Edition ed.). Random House.
- Frické, M. (2015). Big data and its epistemology. *Journal of the Association for Information Science and Technology*, 66(4), 651–661. <https://doi.org/10.1002/asi.23212>
- Gieryn, T. F. (1983). Boundary-work and the demarcation of science. *American Sociological Review*, 48(6), 781–795. <https://doi.org/10.2307/2095325>
- Golder, S. A., & Macy, M. W. (2014). Digital footprints: Opportunities and challenges for online social research. *Annual Review of Sociology*, 40(1), 129–152. <https://doi.org/10.1146/annurev-soc-071913-043145>
- Goulden, M., Greiffenhagen, C., Crowcroft, J., McAuley, D., Mortier, R., Radenkovic, M., & Sathiaselan, A. (2016). Wild interdisciplinarity: Ethnography and computer science. *International Journal of Social Research Methodology*, 20(2), 137–150. <https://doi.org/10.1080/13645579.2016.1152022>
- Guimaraes, M. H., Pohl, C., Bina, O., & Varanda, M. (2019). Who is doing inter- and transdisciplinary research, and why? An empirical study of motivations, attitudes, skills, and behaviours. *Futures*, 112, 102441. <https://doi.org/10.1016/j.futures.2019.102441>
- Hadorn, G. H., Biber-Klemm, S., Grossenbacher-Mansuy, W., Hoffmann-Riem, H., Joye, D., Pohl, C., Wiesmann, U., & Zemp, E. (2008). The Emergence of Transdisciplinarity as a Form of Research. In G. H. Hadorn, H. Hoffmann-Riem, S. Biber-Klemm, W. Grossenbacher-Mansuy, D. Joye, C. Pohl, U. Wiesmann, & E. Zemp (Eds.), *Handbook of Transdisciplinary Research* (pp. 19–39). Springer. https://doi.org/10.1007/978-1-4020-6699-3_2
- Heinzmann, J., Simonson, A., & Kenyon, D. B. (2019). A Transdisciplinary Approach is Essential to Community-based Research with American Indian Populations. *American Indian and Alaska Native Mental Health Research* (Online), 26(2), 15–41. <https://doi.org/10.5820/aian.2602.2019.15>
- Holmes, E. A., O'Connor, R. C., Perry, V. H., Tracey, I., Wessely, S., Arseneault, L., Ballard, C., Christensen, H., Cohen Silver, R., Everall, I., Ford, T., John, A., Kabir, T., King, K., Madan, I., Michie, S., Przybylski, A. K., Shafran, R., Sweeney, A., Worthman, C. M., ... Bullmore, E. (2020). Multidisciplinary research priorities for the COVID-19 pandemic: A call for action for mental health science. *Lancet Psychiatry*, 7(6), 547–560. [https://doi.org/10.1016/S2215-0366\(20\)30168-1](https://doi.org/10.1016/S2215-0366(20)30168-1)
- International Telecommunication Union (ITU) (2019). *ICT Facts and Figures*, 2019. Retrieved from May 13, 2020 <https://www.itu.int/en/ITU-D/Statistics/Documents/facts/FactsFigures2019.pdf>
- Ito, M., Horst, H., Bittanti, M., Boyd, D., Herr-Stephenson, B., Lange, P. G., Pascoe, C. J., Robinson, L., Baumer, S., Cody, R., Mahendran, D., Martínez, K., Perkel, D., Sims, C., & Tripp, L. (2008). *Living and Learning with New Media: Summary of Findings from the Digital Youth Project*. The John D. and Catherine MacArthur Foundation Reports on Digital Media and Learning, November 2008. Groupe de Recherche en Littérature Média Multimodale Hybridité. Retrieved from September 26, 2020 <https://digitalyouth.ischool.berkeley.edu/files/report/digitalyouth-WhitePaper.pdf>
- Jacob, J., Peters, M., & Yang, T. A. (2019). Interdisciplinary Cybersecurity: Rethinking the Approach and the Process. In -K.-K. R. Choo, T. Morris, & G. Peterson (Eds.), *Advances in Intelligent Systems and Computing* (pp. 61–74). Springer International Publishing. https://doi.org/10.1007/978-3-030-31239-8_6
- Jamali, D., Barkemeyer, R., Leigh, J., & Samara, G. (2020). Open Access, Open Science, and Coronavirus: Mega trends with historical proportions. *Business Ethics: A European Review*, 29(3), 419–421. <https://doi.org/10.1111/beer.12289>
- Karpf, D. (2012). Social science research methods in internet time. *Information, Communication and Society*, 15(5), 639–661. <https://doi.org/10.1080/1369118X.2012.665468>
- Khan, A., Choudhury, N., & Uddin, S. (2019). Few research fields play major role in interdisciplinary grant success. *Scientometrics*, 119(1), 237–246. <https://doi.org/10.1007/s11192-019-03043-w>
- Lavorgna, A. (2021). Epistemologies of Cyberspace: Notes for Interdisciplinary Research. In A. Lavorgna & T. J. Holt (Eds.), *Researching Cybercrimes* (pp. 5–22).

- Palgrave Macmillan. https://doi.org/10.1007/978-3-030-74837-1_1
- Lindsey, D. (1988). Assessing precision in the manuscript review process: A little better than a dice roll. *Scientometrics*, 14(1–2), 75–82. <https://doi.org/10.1007/BF02020243>
- Livingstone, S., & Blum-Ross, A. (2020). *Parenting for a Digital Future*. Oxford University Press. <https://doi.org/10.1093/oso/9780190874698.001.0001>
- Loader, B. D., & Dutton, W. H. (2012). A decade in internet time. The dynamics of the Internet and society. *Information, Communication and Society*, 15(5), 609–615. <https://doi.org/10.1080/1369118X.2012.677053>
- Madden, M. (2010). Older Adults and Social Media, August 27, 2010. Pew Research Center Internet & Technology. Retrieved from September 13, 2020. <https://www.pewresearch.org/internet/2010/08/27/older-adults-and-social-media/>
- Mason, W., Vaughan, J. W., & Wallach, H. (2014). Computational social science and social computing. *Machine Learning*, 95(3), 257–260. <https://doi.org/10.1007/s10994-013-5426-8>
- Massey, D. S., & Tourangeau, R. (2013). New challenges to social measurement. *The Annals of the American Academy of Political and Social Science*, 645(1), 6–22. <https://doi.org/10.1177/2F0002716212463314>
- Mayer-Schönberger, V., & Ramge, T. (2018). *Reinventing Capitalism in the Age of Big Data*. Basic Books.
- McKie, L., & Ryan, L. (2012). Exploring trends and challenges in sociological research. *Sociology*, 46(6), 1–7. <https://doi.org/10.1177/0038038512452356>
- Metzler, K., Kim, D. A., Allum, N., & Denman, A. (2016). *Who is doing computational social science? Trends in big data research* (White Paper, pp. 1–23). London, UK: SAGE Publishing. <https://doi.org/10.4135/wp160926>
- Mirowski, P. (2018). The future(s) of open science. *Social Studies of Science*, 48(2), 171–203. <https://doi.org/10.1177/0306312718772086>
- Morin, A., Urban, J., Adams, P. D., Foster, I., Sali, A., Baker, D., & Sliz, P. (2012). Shining Light into Black Boxes. *Science*, 336(6078), 159–160. <https://doi.org/10.1126/science.1218263>
- Németh, R., & Barna, I. (2019). Digitális társadalomkutatások az ELTE-n: Beszámoló A társadalom kutatásának új útjai című workshop digitális társadalomkutatással foglalkozó szekcióról [Digital Social Research at Eötvös Loránd University: Report on the Digital Social Research Sections of the Workshop on New Ways of Social Research]. *Szociológiai Szemle*, 29(4), 115–121. <http://real.mtak.hu/id/eprint/107508>
- O’Carroll, C., Kamerlin, C. L., Brennan, N., Hyllseth, B., Kohl, U., O’Neill, G., & Van Den Berg, R. (2017). Providing researchers with the skills and competencies they need to practise Open Science. *Report of the Working Group on Education and Skills under Open Science* (pp. 1–36). Directorate-General for Research and Innovation. European Commission. <https://doi.org/10.2777/121253>
- Osborne, T., Rose, N., & Savage, M. (2008). Reinscribing British sociology: Some critical reflections. *Sociological Review*, 56(4), 519–534. <https://doi.org/10.1111/j.1467-954X.2008.00803.x>
- Parti, K. (2008). Deviances in the Virtual Reality or the Character-Altering Power of Virtual Communities. *European Journal of Crime, Criminal Law and Criminal Justice*, 16(3), 325–343. <https://doi.org/10.1163/157181708X333373>
- Peters, D. P., & Ceci, S. J. (1982). Peer review practices of psychological journals: The fate of published articles, submitted again. *Behavioral and Brain Sciences*, 5(2), 187–195. <https://doi.org/10.1017/S0140525X00011183>
- Preece, J., & Shneiderman, B. (2009). The reader-to-leader framework: Motivating technology-mediated social participation. *AIS Transaction on Human-Computer Interaction*, 1(1), 13–32. Retrieved from August 12, 2021 <https://aisel.aisnet.org/thci/vol1/iss1/5>
- Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, 346(6213), 1063–1064. Retrieved from October 4, 2020 <https://people.cs.umass.edu/~brenocon/smacss2015/papers/Science-2014-Ruths-1063-4.pdf>
- Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society*, 6(1), 1–12. <https://doi.org/10.1177/2053951718820549>
- SAGE (n.d.). *Methods Map*. Retrieved from August 31, 2020 <https://methods.sagepub.com/methods-map>
- Salganik, M. J. (2018). *Bit by Bit. Social Research in the Digital Age*. Princeton University Press.
- Savage, M., & Burrows, R. (2007). The coming crisis of empirical sociology. *Sociology*, 41(5), 885–899. <https://doi.org/10.1177/0038038507080443>
- Savage, M., & Burrows, R. (2009). Some further reflections on the coming crisis of empirical sociology. *Sociology*, 43(4), 762–772. <https://doi.org/10.1177/0038038509105420>
- Schmalz, D. L., Janke, M. C., & Payne, L. L. (2019). Multi-, inter-, and transdisciplinary research: Leisure studies past, present, and future. *Journal of Leisure Research*, 50(5), 389–393. <https://doi.org/10.1080/00222216.2019.1647751>
- Schmidt, B., Orth, A., Franck, G., Kuchma, I., Knoth, P., & Carvalho, J. (2016). Stepping up open science training for European research. *MDPI Publications*, 4(2), 1–10. <https://doi.org/10.3390/publications4020016>
- Siler, K., Lee, K., & Bero, L. (2015). Measuring the effectiveness of scientific gatekeeping. *Proceedings of the National Academy of Sciences*, 112(2), 360–365. <https://doi.org/10.1073/pnas.1418218112>
- Sparks & Honey (2014). *Meet Gen Z: Forget everything you learned about Millennials*. Retrieved from September 13, 2020 www.slideshare.net/sparksandhoney/gen-eration-z-final-june-17
- Srnicek, N. (2017). *Platform Capitalism*. Polity Press.
- Stodden, V., Guo, P., Ma, Z., & Zaykin, D. (2013). Toward reproducible computational research: An empirical analysis of data and code policy adoption by journals. *PLoS ONE*, 8(6), e67111. <https://doi.org/10.1371/journal.pone.0067111>
- Teal, T. K., Cranston, K. A., Lapp, H., White, E., Wilson, G., Ram, K., & Pawlik, A. (2015). Data carpentry: Workshops to increase data literacy for researchers. *International Journal of Digital Curation*, 10(1), 135–143. <https://doi.org/10.2218/ijdc.v10i1.351>
- Tinati, R., Halford, S., Carr, L., & Pope, C. (2013). *The Promise of Big Data: New Methods for Sociological Analysis*. SOTON Eprints.
- Toelch, U., & Ostwald, D. (2018). Digital open science—Teaching digital tools for reproducible and transparent research. *PLoS Biology*, 16(7), e2006022. <https://doi.org/10.1371/journal.pbio.2006022>
- Tourangeau, R. (2004). Survey research and societal change. *Annual Review of Psychology*, 55(1), 775–801. <https://doi.org/10.1146/annurev.psych.55.090902.142040>
- Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. In *ICWSM ‘14: Proceedings of the 8th International AAAI Conference on Weblogs and*

- Social Media*, 2014, Cornell University. Retrieved from October 4, 2020 <https://arxiv.org/pdf/1403.7400.pdf>
- Uprichard, E. (2012). Being stuck in (live) time: The sticky sociological imagination. *The Sociological Review*, 60(S1), 124–138. <https://doi.org/10.1111/j.1467-954X.2012.002120.x>
- Villanti, A. C., Johnson, A. L., Rath, J. M., Jacobs, M. A., Graham, A. L., & Rath, J. M. (2017). Social media use and access to digital technology in US young adults in 2016. *Journal of Medical Internet Research*, 19(6), e196. <https://doi.org/10.2196/jmir.7303>
- Weiss, M. J. (2000). *The Clustered World: How We Live, What We Buy, and What It All Means About Who We Are*. Little, Brown & Co.
- Womack, R. P. (2015). Research Data in Core Journals in Biology, Chemistry, Mathematics, and Physics. *PLoS ONE*, 10(12), e0143460. <https://doi.org/10.1371/journal.pone.0143460>
- Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2018). Assessing reliability of social media data: Lessons from mining TripAdvisor hotel reviews. *Information Technology & Tourism*, 18(1–4), 43–59. <https://doi.org/10.1007/s40558-017-0098-z>
- Zuboff, S. (2019). *The Age of Surveillance Capitalism*. Public Affairs.
- Zyoud, S. H., Sweileh, W. M., Awang, R., & Al-Jabi, S. W. (2018). Global trends in research related to social media in psychology: Mapping and bibliometric analysis. *International Journal of Mental Health Systems*, 12(1), 4. <https://doi.org/10.1186/s13033-018-0182-6>



© 2021 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format.

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



Cogent Social Sciences (ISSN: 2331-1886) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

