Quaderni di Comunità

Persone, Educazione e Welfare nella società 5.0

Community Notebook

People, Education, and Welfare in society 5.0

n. 3/2023

REINVENTING UNIVERSITY.

THE DIGITAL CHALLENGE IN HIGHER EDUCATION

Edited by

Stefania Capogna, Ligita Šimanskienė, Erika Župerkienė

eurilink

University Press

Iscrizione presso il Registro Stampa del Tribunale di Roma
al n. 172/2021 del 20 ottobre 2021

Eurilink University Press Srl
Via Gregorio VII, 601 - 00165 Roma
www.eurilink.it - ufficiostampa@eurilink.it
ISBN: 979 12 80164 71 1
ISSN: 2785-7697

Prima edizione, febbraio 2024
Progetto grafico di Eurilink

INDICE

5. THE IMPACT OF DIGITALISATION IN SCIENTIFIC RESEARCH: RISKS, OPPORTUNITIES AND ETHICAL CHALLENGES[1]

by Maria Chiara De Angelis*

Abstract: *Internet of Things, robotics, biometrics, persuasive technology, Big Data, virtual, augmented reality, and digital platforms are pervading the world of scientific research, influencing how everyday scientific work practices are organized and conducted. In this emergent context, the essay wants to offer a theoretical reflection on the purposes and functions of digital technology in scientific research processes, enhancing the dimension of ethical choice and crucial judgment in governance processes and researchers' professional development.*

Keywords: digital transformation, scientific research, digitalisation of science, social and ethical challenges.

Abstract: *Internet of Things, robotica, biometria, tecnologia persuasiva, Big Data, realtà virtuale e aumentata e piattaforme digitali stanno pervadendo il mondo della ricerca scientifica, influenzando l'organizzazione e la conduzione delle pratiche quotidiane del lavoro scientifico. In questo contesto nascente, il saggio vuole offrire una riflessione teorica sugli scopi e le funzioni del digitale nei processi di ricerca scientifica, valorizzando la dimensione della scelta etica e del giudizio cruciale nei processi di governance e nello sviluppo professionale dei ricercatori.*

[1] Accepted June 2023 - Published December 2023.
* Research Fellow at Link Campus University, mc.deangelis@unilink.it.

Parole chiave: trasformazione digitale, ricerca scientifica, digitalizzazione della scienza, sfide sociali ed etiche.

Introduction

The widespread digital transformation and the subsequent development of data science, both in public and private life dimensions, characterise the epochal change we are going through (Floridi & Taddeo, 2016). Internet of Things, robotics, biometrics, persuasive technology, Big Data, virtual augmented reality, and digital platforms are pervading the world of scientific research, influencing how everyday scientific work practices are organised and conducted. The current scenarios open to reflections that affect scientific research at various levels: from the organisation of the researcher's work to the research design, from the relationship with colleagues to the dissemination of research results, from the evaluation practices to the ethical dimension.

Digitalisation in academic research shows many faces. The OECD (Bello & Galindo-Rueda, 2020a) underlines four specific areas from the results of the *OECD International Survey of Scientific Authors (ISSA*): the adoption of digital scientific collaboration and productivity tools throughout the different stages of the scientific process; the digitally enabled diffusion of, and access to, data and code; the use of advanced and data-intensive digital tools to gain insights and develop predictions and, the development of digital identities and online communication of scientific work.

The ISSA results reveal the following challenges:

- *Access to data:* Scientists across nearly all fields consider data collection and curation skills the most important challenge.

- *Skills:* Advanced programming skills appear to be more crucial in computer science, earth and planetary science, and mathematics,
- *Infrastructure:* While legal knowledge of intellectual property, privacy and confidentiality seem key for authors in medicine and health.

This transformation has concrete impacts on three aspects of open science: dissemination of scientific information, access to research data and engagement with stakeholders outside of research (Bello & Galindo-Rueda, 2020a).

Due to its complexity, the theme can be approached from various points of view. In this paper, we have chosen to focus on the social and ethical implications of digital innovation on the so-called productive research routines. Based on those premises, this work intends to explore social and ethical digital transformation issues in scientific research to stress the main effects on practices, considering the 'generative' and a 'transformative' dimension of digital technologies on environments (Floridi, 2016).

The questions that drive our work are: How do research practices change the way of doing science with digitalisation? What are the main ethical impacts of digitalisation in academic research, and what role does the researcher's subjectivity play in governing the digital transition we are going through?

To this end, we will outline the most significant elements with which academic research must measure itself in the ongoing change. We will endeavour to examine how automation and digitisation affect scientific work and the research professions (§ 1), the figure of e-scholar and its implications in the research process (§ 2), and the emergent ethical issues of the digital revolution in academic research (§ 3).

1. Digitalisation and research in practice

In recent years, we have definitively entered the digital age. A new epochal chapter in human history and civilisation (Arendt, 1958) directly involves science and scientific research as creators and promoters of the change. What scenarios are foreseen for scientific research? How is the way of doing research changing? What kind of research and what researcher can be expected in a future that is already present? How has the pandemic accelerated the adoption of digital tools and techniques and changed the openness and inclusiveness of research and innovation ecosystems?

To answer these questions, we deem it appropriate first to underline the tangible consequences of the advent of digital technologies in our lives and then move on to examine the direct impacts on the daily practices of scientific work. The new technological revolution based on cybernetic-physical systems is moving in the direction of integrating the Internet and the Internet of Everything (IoE) into complex systems based on the pervasive convergence between information technologies, multi-level computing and communication, such as cloud computing and fog and dew computing[2]. This evolution has led to the rapid transformation of production processes, including the automation of industrial and factory processes, accelerated by the advent of ever more dynamic, fast and intelligent forms of exchange and interaction between machines and between them and the environment, promoting the development of increasingly complex adjustment and maintenance systems capable of learning from experience.

[2] The basic idea of fog computing and edge computing is to bring cloud servers closer to end users at the network's edge, reducing the code/data transferring and exchanging time between mobile devices and servers.

Consequently, digitisation in the so-called Industry 4.0 produces a radical impact on the labour market in terms of professional figures required for the governance of the emerging complexity, together with important changes in production processes and the allocation of resources in people's daily lives. If we then consider the current communication flows stimulated by the advent of the Internet, these have generated an artificial space of a cognitive type where ubiquity and instantaneity are affirmed, which is the disengagement of space and time from material ties. The state of connection is the physical and cognitive prerequisite of current socialisation processes. We are immersed in a meta-technology in which complex systems of a technological nature are integrated to create environments for the synthesis and hybridisation of technologies (Boccia Artieri, 2012, p. 26; Boccia Artieri *et al.*, 2022).

The current communication flows stimulated by the advent of the Internet have created an artificial cognitive space where ubiquity and instantaneity assert themselves, which is the release of space and time from material ties. In the network society, the state of permanent connection creates a new ontological space that requires new interpretative and epistemological models. The infosphere, states Floridi, is therefore not a virtual environment supported by a genuinely "material" world behind it; rather, the world itself will increasingly be interpreted and understood informationally as part of the *infosphere.* At the end of this shift, the infosphere will have gone from referring to the space of information to being synonymous with Being. We will find this informational metaphysics increasingly easier to embrace (Floridi, 2007, p. 61).

The squares become virtual squares but not less real. The Internet quickly became an organising principle that gave an organisational form even to less organised entities. An example of

this is the practice of smart working, which has forced us to rethink the role of managers in coordinating work groups and has posed dilemmas on the trade union level and the rights/duties of workers, etc. The network society re-designs the network social relationships, the times and spaces of the relationship, and the organisational forms of culture.

Considering the literature on the topic, there are recurring elements that characterise the digital revolution in terms of impact on the bio-psycho-economic-social and cultural environment, which we try to summarise in three keywords and then move on to understand more closely how these changes are transforming the way we do academic research:

Access to information and breadth of applications (e.g., among the databases used for research, for example, ScienceDirect database provides 1.4 million open access articles; JSTOR provides access to more than 12 million journal articles, books, images, and primary sources in 75 disciplines, and we can continue to list).

Access to a virtually infinite amount of data, while on the one hand, broadening the possibilities of knowing and exercising active and aware citizenship (EE UU, 2020). On the other hand, it produces noise. It creates significant problems in selecting relevant and essential information: an accumulation of information which stimulates fatigue and inhibits the critical ability of judgment from a certain point onwards. There is a breaking point beyond which value is no longer created but even disvalued and lost (Han, 2015, p. 76).

The infinite connection of informational reality translates into a constant interaction between material and concrete reality and virtual and interactive reality (Castells, 2009). This recursivity describes the expansive potential of technologies and the relationship between them and existential and organisational environments. This pervasiveness operationalises the online and

offline multi-life dimension that we all live in today, producing a cognitive and emotional re-design and opening to new relational possibilities and new connected solitudes (Turkle, 2011).

Disintermediation shortens the distribution and information chains and simultaneously introduces a regime of equivalence and horizontalisation of communication (Boccia Artieri, 2012, p. 26; Boccia Artieri *et al.*, 2022). As Han points out, the digital medium is a medium of presence that abolishes the action of the mediating instance in the name of the search for transparency and efficiency (2015:29). De-mediatisation abolishes the representation that characterised the classic electronic media, giving way to an extemporaneous presence, which is exhausted in the here and now, effectively inhibiting the political action that characterised the age of representation and fuel, together with the possibility of exchange and collaboration, subtle forms of conformism.

These three dimensions of digital transformation can also be expressed in scientific work, where the phases of research development are re-designed, the role assumed by the researcher, the setting of the scientific work and its diffusion, the "publics" to which it is addressed, the relations with subjects inside and outside the academic world.

Already fifteen years ago, Jankowski summarised the growing affirmation of e-research by exhaustively underlining its main expressions/declinations (2009:6):

- Increasing computerisation of the research process, often involving high-speed, large-capacity machines configured in a networked environment;
- Reliance on network-based virtual organisational structures for conducting research increasingly involves distant collaboration among researchers, often international in scope;

- Development of Internet-based tools facilitating many phases of the research process, including communication, research management, data collection and analysis, and publication;
- Experimentation with new forms of data visualisation, such as social network and hyperlink analysis, and multimedia and dynamic representations;
- Publication, distribution and preservation of scholarship via the Internet, utilising traditional and formal avenues (e.g., publishing houses, digital libraries) as well as those less formal and less institutionalised (e.g., social networking sites, personal Web sites).

Access. The development of technologies has always gone hand in hand with the emerging needs of science to equip itself with tools and processes capable of accelerating the exchange of information and data between researchers overcoming space-time dislocation. ICT-based infrastructures and services, the use of the cloud for large-scale storage and data storage and retrieval systems, and social media have radically changed how knowledge is created, exchanged and disseminated, fuelling the development of polycentric and decentralised organisational systems**.**

A "widespread" research model is established based on demanding investments at a system level and on interaction networks between multiple subjects (universities, research centres, companies, institutions, foundations, and public and private entities). Open access to research data, new citation tools and interoperable databases has made access to large-scale research results possible, favouring a more democratic and less elitist approach. However, this questions us about *the problem of information overexposure,* which inevitably involves the researcher during all phases of the research process, particularly in its initial exploratory phase, which consists of identifying the sources for framing the research problem and in their evaluation and

selection. The researcher constantly deals with search engines, digital archives, and bibliographic and citation management platforms. Moreover, if, on the one hand, digital technologies favour access to sources and their management, on the other, it becomes urgent for the researcher to implement search and filtering strategies for information and refine his critical thinking concerning the evaluation of sources and their re-use (Fedeli, 2017).

Martin Weller, coining the term *digital scholar,* underlines how being researchers in the digital age implies a change in attitudes, which is reified in the acceptance of values and ideologies inherent in the so-called "open access" movement and the recognition of collaborative work and networking between peers (Weller, 2011).

Connection/Networking. Daily social networking activities and the proliferation of platforms capable of supporting collaborative research also allow us to build connections with distant contexts in both the space-time and cultural dimensions. International collaboration networks allow the definition of exchange spaces to be unimaginable until a few years ago through infrastructures and innovations in digital science.

Disintermediation. Added to the attention relating to the dimension of data care, which also includes access, organisation, management, and re-use, is the question of legitimation, dissemination of research results and so-called data care (Borgman, 2007, p. 824). With the use of open-access platforms and social media, the legitimation of the final product of the research no longer ends with the post-referee publication. However, it becomes a recursive process, with multiple voices along a continuum that extends over time. Even the drafts become objects of exchange reflection and fully enter the scientific production process, according to a logic of greater transparency

than in the past. The research products partly participate in disintermediation and horizontalisation of information flows peculiar to digital media. The researcher can also use non-institutional or formal channels to contribute to the visibility of his work and use accessible archives, citation databases (e.g., Scopus, Web of Science), and scientific search engines (e.g. Google Scholar) for the preservation of his research work and its dissemination/ visibility. The radical change, however, does not only concern the number of available resources or the processes that affect their management and communication but also the legitimation of these resources by the academic world and, last but not least, the attribution of an economic value to the dissemination of scientific work products by platforms and publishers who manage their dissemination and indexing across the network. Concerning the dimension of work organisation, then, on the one hand, artificial intelligence and robotics have favoured the automation and standardisation of processes; on the other hand, what some authors call "work of mundane knowledge," including practices of control, data sharing and standardisation, and the preparation, repair and supervision of laboratory robots. These subsidiary practices, often invisible compared to the scientific results disseminated through scientific papers, require routine knowledge work, defined as "banal", which constitutes a fundamental part of the researcher's work and only the latter can carry out. Contrary to what one might think, automation and digitalisation have produced an amplification and diversification of research work, which has an impact in terms of quantity and diversification of activities, contributing to increasing the complexity, number and diversity of the researcher's tasks, impacting unevenly in the scientific hierarchy (Ribeiro *et al.*, 2022). This is due to the peculiar characteristics of scientific work, which concerns open, complex, and continually evolving "knowledge" objects that require a

structured, non-obvious capacity for analysis and synthesis. Routine in science takes place in continuous change such that automation and digitalisation slide from their presumed status of technical and stable objects to that of epistemic objects. The routine knowledge work performed by highly skilled researchers is different from the routine work performed by other types of professionals because the objects engaged in scientific work include objects that "are typically open-ended, question-generating, and complex" (Knorr-Cetina, 2001, p.190; 1997). Recently, the focus has also shifted to implementing e-infrastructures for research from an institutional, organisational unit and service perspective. Holewa *et al.* (2015) integrates the capability approach and the socio-technical model in the *HWDM maturity model* and examine a) what services the researcher/research group requires to undertake research in a highly productive technology-enabled way leveraging available data/ information and b) what core integral service functions are also required to enable the researcher/research group to improve performance.

The HWMD model was created as a self-assessment model for research units and is configured as a *six-dimensional framework* (Fig 1) that is relative to effective e-research delivery: a) governance and leadership, b) research information and data management, c) technological infrastructure, d) collaboration and community engagement, e) workforce education, training and development and, f) service delivery and management.

Figure 1: The HWMD maturity model

Source: Holewa et al.*, 2015*

In the socio-technical model, technology denotes research and helps draw boundaries by re-designing the relationships between the subjects involved.

Digital transformation changes the epistemological status of research, drawing new epistemic boundaries. E-research can then be read through the definition of social relationship as a symbolic and intentional reference that connects social subjects as it actualises or generates a link between them, generatively reconfiguring purposes, means, norms and values according to the three semantics proposed by Donati (2004)[3].

[3] Through the AGIL relational model, Donati goes beyond the classic social and post-modern categories of society by proposing a representation based no longer on

Therefore, the researcher's identities and practices are constantly constructed, starting from the relationships between the parties involved, designing new systems of alliances and meanings (Callon, 1984; Latour, 2005). The digitalisation of scientific research then involves translation work on the part of the actors involved (Callon & Law, 2005), which affects communication flows, the quantity/quality of relationships, and the sphere of production of material/economic and immaterial/knowledge-based value: how scholars communicate, the types of results they produce, and the networks within which they operate.

2. *Digital transformation and scientific work: from scholar to e-scholar*

The relationship between technology and scholars is changing, along with the definition of technology (information technology, ICT, instructional technology) and its applications in academia and research (Lei *et al.*, 2021; Sun *et al.*, 2020; Zhang, 2022).

The way academic community members develop their research activities, access information resources, and

networks of objects or individuals but on networks of relationships. According to the sociologist, the relationship is expressed according to three semantics:
1) Referential: the relationship as a referent, as referring something to something else within a framework of symbolic meanings with different types and degrees of intentionality and more or less shared between the actors in the field
2) Structural: the relationship as religion, as bond, connection, constraint, mutual conditioning, structure
3) Generative: It shows how the different components and acting subjects that enter into relationships produce an effect that cannot be explained based on the properties of these social components and actors but takes on its quantum-qualitative connotations. The relationship becomes an emergent effect rather than an aggregate effect.

communicate with each other has dramatically changed with the irruption of the web. Nevertheless, the tools provided by today's web aren't efficient enough to satisfy many of the specific requirements of this new generation of e-scholars. In 1990, using data gathered from more than 5,000 faculty members, Ernest L. Boyer classified the types of activities scholars regularly engaged in.

In Ernest L. Boyer's definition of scholarship, there are *four components*, each of which, he suggests, should be considered equal value by universities and government policy.

Discovery – This is creating new knowledge in a specific area or discipline. This is often taken to be synonymous with research. This is probably closest to the public conception of scholarship, as universities are often the site of significant breakthroughs.

Integration – This is focused on interpretation and interdisciplinary work. It is moving away from the pure 'genesis' research of discovery. Boyer states that it is 'making connections across the disciplines, placing the specialities in a larger context, illuminating data in a revealing way, often educating non-specialists.

Application – This is related to the concept of service. Still, Ernest L. Boyer distinguishes between citizenship and scholarly types of service, and for the latter, it needs to build on the scholar's area of expertise. It can be seen as an engagement with the wider world outside academia, which might include public engagement activities and input into policy and general media discussions. This can also include the time spent peer-reviewing journal articles and grant applications and sitting on various committees.

Teaching – Much of the interpretation of Boyer can be seen as an attempt to raise the teaching profile. He argues that 'the

professor's work becomes consequential only as others understand it. Yet, today, teaching is often viewed as a routine function, tacked on.

According to Martin Weller's contribution (2011), how these main dimensions are changed in the digital age?

Discovery – The creation of new knowledge. An open, digital, networked approach to discovery could relate to data sharing (e.g., generating and analysing unprecedented amounts of data). The access of a great amount of data originates some implications: a) the application of grid computing or crowdsourcing analysis, b) Unexpected applications, c) Data visualisation, and d) combination

Integration – This is focused on interpretation and interdisciplinary work. The question relating to data care, which also includes access, organisation, management and re-use, is added to the question relating to legitimation, dissemination of research results and so-called data care (Borgman, 2007, p. 824). With open-access platforms and social media, the legitimacy of the final research product no longer ends with the post-referee publication. Still, it becomes a recursive process, with many voices along a continuum that extends over time. Even the drafts become an object of exchange reflection and fully enter the scientific production process, according to a logic of greater transparency than in the past.

Application – This is related to the concept of service. It can be seen as engagement with the wider world outside academia. We are seeing the development of a 'personal brand' amongst academics as new technologies allows them to establish an audience complementary to their institutional one. These new channels are also beginning to compete with traditional means of public engagement in terms of influence. New technologies are

facilitating access to a new audience, which is disintermediating many of the conventional channels.

Teaching – the ability to be understood by others referring to a specific area of knowledge. With the advent of a wide variety and high quality of freely available academic content online, the individual student is no longer limited by the physical resources they can locate, and the lecturer is therefore no longer regarded as the sole source of knowledge, as the learner can pick and choose elements from a variety of courses provided by any number of diverse institutions and individuals.

Starting from Martin Weller's analysis, the digitisation of scientific research refers to how scholars communicate outputs/results, the networks they operate within and the outputs/results released. In Weller's definition, the networking dimension actualises the researcher's reality in the digital age: in a state of permanent connection.

Wim Van Petegem *et al.* (2021) present a digital scholar framework and its key components, considering five specific dimensions of analysis, referring to the scholars as human beings about the Self about the digital identity, the team they belong to, the HE institution, the Local Community and global word that the e-scholar inhabits.

Figure 2: The Digital Scholar Framework

Source: Van Petegem et al.*, 2021*

The e-scholar moves then into research, teaching and service, and dialoguing with the academy and the community. Based on the dimensions described above, the figure of the e-scholar is acted upon from time to time by taking on different roles and positions. The roles assumed by the e-scholars depend strictly on the level of digital competencies needed to play in the specific situations. The e-scholar can be, at the same time or with a different level of expertise, an author, a storyteller, a content creator, an integrator among different media and languages, and a

networker who involves a large audience and can create connections and intersections among researchers and disciplines (Van Petegem *et al.*, 2021, pp. 28-29). The complexity of the researcher's work has increased, which entails the need to develop adequate skills to face the challenges of the frontier profession. The development of data gathering and research analytical skills to comprehend and evaluate data is undoubtedly necessary, but more skills are needed. In the so-called semantic web, the e-scholar is increasingly configured as an active subject in the process of thematization, systematization and problematization of data in terms of exploration of meanings "in situation", incorporated into the contexts and experiences in which individuals interact with data, considering also possible known and subconscious biases (Špiranec *et al.*, 2019; Capogna, 2022).

From a critical data literacy perspective, the e-scholar must be able to read and make the data speak, combining the ability to identify reliable data sources, evaluate information collected following a solid research design and methodologies, guarantee the reliability, replicability, and the dissemination of research's results.

According to the critical perspective, the emerging themes that concern the researcher's professional development affect the ontological treatment of data, their epistemological status, the emerging social critical issues directly related to data processing, and the pedagogical and ethical aspects of their usage (Špiranec *et al.*, 2019). A more holistic, epistemological and relational approach to academic development is required to face the incoming challenges.

The complexities associated with living and working in a digital world as a scholar have implications for considering how we create and share knowledge. Data literacy skills find meaning and concrete application in research if, at the same time, the

researcher grows more and more in the discipline of thinking, that reflection *in action* and *on action* that makes it possible to explore reality by questioning preconceived knowledge, with the awareness of acting in a scientific community and a defined cultural fabric, in the predisposition to think about experience and thought (Schön, 1993). The researcher's activity, therefore, far from being merely technical, requires the development of meta-reflective skills, open to questioning the ontological and epistemological nature of the observed reality, political-strategic and ethical, as we will have the opportunity to explore in greater depth in the next paragraph.

3. Research in the digital era: Open questions for ethical challenges

Digital technologies transform the surrounding environment and create new ontological spaces. In these environments, technologies can interact, and sometimes without our intermediation. Social media, Big Data (BD) [4], and e-health technologies are all examples of ICTs that are believed to raise new ethical problems or dilemmas. To be adequately addressed, the ethical issues raised by ICT need a different conceptual framework, particularly an information approach (Floridi & Taddeo, 2016).

In the last decade, ethical reflection has shifted attention from technological means to contents (information) that can be created, recorded, processed and shared through such means,

[4] The word *Big Data, which* refers to the data set in digital format that is collected, archived, and managed through large datasets, which cannot be processed using the software and hardware systems traditionally used. The contents published include the so-called transactional data (information collected in the context of exchanges between citizens and administrations and between consumers and companies) and the digital by-product data, data created and inserted by users through the platforms typical of the 2.0 era of the web (Ganz & Reinsel, 2011).

moving from a *computer-centric approach* to a *data-centric approach*. The data-centric approach thus focuses on the different moral dimensions of information according to the whole cycle of information creation, sharing, storage, protection, usage, and possible destruction (Floridi & Taddei, 2016). The pervasiveness of big data is not only due to its volume but also to the variety and speed with which it is processed and disseminated (Laney, 2001), changing people's experience online in real-time (Salganik, 2018). Data can be defined as the *oil of the 21st century* and its analysis, a powerful means of extraction and dissemination on the market (Ganz & Reinsel, 2011). In combination with new statistical modelling techniques, BD may enable advances in many practically important areas, with benefits in the health field and in the market where the transition costs are significantly reduced (Weinhardt, 2021). However, at the same time, there are no fewer potential risks and negative outcomes to consider: the improper creation, use, and dissemination of data without consent, the problem of controlling data, the dissemination of sensitive information and violation of privacy, the counterfeiting of data, and its dissemination for illicit purposes are some of the open questions to face. In scientific research of every discipline, these risks are evident and raise questions about how to protect people from improper use of data, starting from the people involved in the research activities as participants and ending with the researchers and the products of their scientific work (e.g., plagiarism, dissemination without consent, improper use of AI). Generative artificial intelligence technologies (e.g., ChatGPT), for example, on the one hand, can be used by students and teachers alike to develop new learning experiences and promote creativity; on the other hand, they question the originality of the final product on the choice and reliability of the sources, questioning the entire research process. In the digital era, what are the rights and duties

of scholars and researchers who use e-methods across the digital projects of social research? What are the problems in terms, for example, of exchange, handling, and "cannibalisation" of data and analysis?

The phenomenon of datafication (Van Dijck, 2014; Van Dijck *et al.*, 2018) has given rise to a new branch of ethical studies called Internet Research Ethics (IRE). IRE is defined as the analysis of ethical issues and application of research ethics principles about research conducted on and on the Internet. The IRE studies focus on data (including generation, recording, curation, processing, dissemination, sharing and use), algorithms (including artificial intelligence, artificial agents, machine learning and robots) and corresponding practices (including responsible innovation, programming, hacking and professional codes), to formulate and support morally good solutions (e.g., right conducts or right values). Internet-based research, broadly defined, is research that utilises the Internet to collect information through an online tool, such as an online survey; studies about how people use the Internet, e.g., through collecting data and examining activities in or on any online environments; and, uses of online datasets, databases, or repositories (Buchanan & Ess, 2008, 2009). Various scientific associations have been concerned with drawing up guidelines that respond to the new frontiers of digital transformation. For example, the British Psychological Society's (2017) guidelines provide a concise chart for researchers and ethics boards based on four main ethical principles (Fig.3):

1. Respect for the autonomy, privacy and dignity of individuals and community
2. Social responsibility
3. Maximising benefits and reducing risks
4. Scientific integrity

In addition to following these broad ethical principles, for researchers seeking more direct best practices, Gabrielle Samuel and Elisabeth Buchanan (2020) have suggested the following when preparing an Internet research protocol:

- Consider data in use, at rest, in transit, and in deletion: different ethical considerations and security measures; describe procedures (including safeguards for collecting, storing, processing subject data, and data destruction) for minimising potential risks to the subject's confidentiality;
- Learn the nuances between and among data management practices, including de- and re-identification; anonymised, coded, aggregated;
- Data sharing and data use agreements: It is important for researchers to work with the research ethics review process (RECs/IRBs) in planning for data sharing;
- Specify where and under what conditions individuals will have access to the data, what will be available and to whom (air gap, clean rooms, data access levels);
- Address uncertainty in data longevity in more open-ended terms: "Data may exist on backups or server logs beyond the time frame of this research project";
- Clarify that one's consent to use, for example, Facebook, is not the same as consent to participate in research;
- Ensure research is not in violation of terms of service, user standards, or norms;
- Disclose what third-party sites may be used for collection, storage, dissemination and that access by third parties is possible;
- Confirm if the research will NOT involve merging any of the data sets in such a way that individuals might be identified;
- Confirm if the researcher will NOT enhance the public data set with identifiable or potentially identifiable data.

Emphasis is placed on researcher responsibility, digital torts, information harms, informed consent, information privacy theory, data linking practices and web data anonymisation/ encryption techniques.

In parallel to this approach focused on normative and data literacy skills aspects in creating, using, and sharing data, we advise developing more extensive, inter-and-multi-disciplinary reflections based on human-centred research, where the focus returns on *human being*: what "image of man" (Jonas, 1979) we desire to promote and develop? How do digital technologies affect the nature of acting? More specifically, who are the researchers in the digital age and how/can you contribute positively to building a *good* society? The human sciences and social sciences have the arduous task of taking up these challenges and defining the epistemological statutes that can guide the change underway (Capogna, 2023). Scientific research, like any human activity, must be read as a whole as an *ethical practice*. Following, we proposed some questions to open a scientific debate on the issue.

- *Research freedom.* Every day, we read articles in the press about how, in a data-driven economy, everything that is digital or moves in the digital sphere can be monetised. In such a world of large digital platforms and technology companies constantly looking for new avenues of profit, academia offers an interesting way to diversify the business. The development of large digital platforms and technological companies constantly looking for new ways of profit has made the academic world fertile ground for diversifying its business. While digitalisation of the subject of science and of science itself can hold the promise of better and more thoughtful research, digitalisation also opens the door to greater commercialisation (Jansen, 2021; UNESCO, 2017). In this situation, the academic research system must preserve its

digital sovereignty through the possibility of making research results and data open through Open Science. However, it must also work to control the data generated by its various activities. Big tech companies can and are using this data for commercial services, sometimes in competition with universities and potentially undermining university values (Bello & Galindo-Rueda, 2020a).

- *Research creativity.* Digital allows massive and potentially unlimited access to data and information, but when does data/information become knowledge? Which inferential processes are supported by new technologies, and which, on the contrary, are hindered? How does digital impact creativity? In the so-called *infosphere*, one of the greatest challenges is to continue cultivating the personal character of research so that originality and methodological rigour are not lost, encouraging the intersection of different disciplines and preparing for the unexpected that may emerge from the initial question. It is more necessary than ever to reflect on which new or "old" *epistemological statutes* are necessary for science on a human scale, which is at the service of the community and does not become scientific dogma. This work of reflection on the epistemé of an era that Foucault calls *archaeology* (1966, 1969) continues to prove necessary to define and rethink the space of possibilities within which the characteristic knowledge of that era is constituted and operates. Furthermore, given the possibility of creating and processing a potentially infinite amount of data, it is necessary to ask ourselves what spaces this opens up for prediction and what limits they may have. The AI potentiality can fuel the Promethean delirium of omnipotence that blocks thought, reducing the ability to deal with complexity. For example, the thorny issue of sources and plagiarism has forced many universities to equip themselves

with anti-plagiarism systems, which will have to deal with increasingly advanced artificial intelligence in their self-learning processes.

- *Cultivate community.* Digitalisation allows the network to be quickly fed and opportunities for exchange and knowledge to grow. On the other hand, an element to consider is the quality of the exchange and the depth of scientific reflection, which cannot elude the *body*. The body is the primary place of human perception and experience (Merleau-Ponty, 1945), the most natural technical object (Mauss, 1934) and the existential foundation of culture and the Self (Csordas, 1994). The body builds culture and knowledge through the experience generated by the embodiment of the subjects (Csordas, 1994). The quality of knowledge also depends on embodied engagement and sharing in the scientific community of reference and requires a new epistemological approach that considers the complex interaction among human beings, artificial intelligence and bio-social environments.

Conclusions

The effects of ICT on an epistemological level are profound and pervasive. Human beings, machines, institutions or environments contribute in a systemic way to creating new knowledge through a complex situated, embodied, distributed and relational process. This requires an effort on the part of the scientific community to question the nature of information organisms to provide a methodological framework (ontological and epistemic) that allows us to better face the ongoing ethical and socio-political challenge (Russo, 2018, p. 665; Floridi, 2015). HEIs must promote multi-disciplinary and multi-level reflection to steer

the ongoing challenges and to stimulate the awareness of teachers, staff, researchers and students, questioning the means in relation to the aim that is intended to be achieved and to the ethical implications that digital transformation affects in all phases of scientific research. The universities and the entire educational system are called to rethink and promote spaces of social criticism in which it is possible to freely question the purposes and functions of digital technology in scientific research processes, enhancing the dimension of ethical choice and critical judgement in governance processes and researchers' professional development, promoting a more mature style of data literacy and a more critical understanding of the reality and new environments that ICTs contribute to creating.

References

Arendt, H. (1958). *The human condition*. University of Chicago Press.

Bello and Galindo-Rueda (2020a). "Charting the digital transformation of science: Findings from the 2018 OECD International Survey of Scientific Authors (ISSA2)", https://doi.org/10.1787/18151965, based on OECD ISSA 2018, http://oe.cd/issa.

Boccia Artieri, G. (2012). *Pubblici, cittadini e consumatori nella (Social) Network Society,* Milano: FrancoAngeli.

Boccia Artieri, G., Colombo, F., Gili, G. (2022). *Comunicare. Persone, relazioni, media*. Bari: Laterza.

Boyer, E. L. (1990), *Scholarship reconsidered: Priorities of the professoriate.* Carnegie Foundation for the Advancement of Teaching.

Buchanan, E. A., & Ess, C. (2008). *Internet research ethics: The field and its critical issues*. In K. Himma & H. Tavani (eds.), The Handbook of Computer and Information Ethics (pp. 272–92). Boston: Wiley.

Buchanan E. A., Ess C. M. (2009). *Internet research ethics and the institutional review board: Current practices and issues*. Computers and Society, 39(3), 43–49.

Callon, M. (1984). *Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay.* The Sociological Review, 32(1_suppl), pp. 196–233.

Callon, M. & Law, J. (2005). *On qualculation, agency and otherness*. Environment and Planning D: Space and Society 23: pp. 717–733.

Capogna S. *Sociology between big data and research frontiers, a challenge for educational policies and skills*. Qual Quant. 2023;57(1):193-212. doi: 10.1007/s11135-022-01351-7. Epub 2022 Mar 5. PMID: 35283540; PMCID: PMC8897736.

Castells, M. (2009) *The Rise of the Network Society*. John Wiley and Sons Ltd.

Csordas Thomas J. (1994) *Embodiment and experience. The existential ground of culture and self*. Cambridge: Cambridge University Press.

Donati, P. (2004) *Introduzione alla sociologia relazionale*, Milano: FrancoAngeli Milano.

EE. UU. (2020) *Strategic plan 2020-2024 – Research and Innovation*. At https://commission.europa.eu/publications/strategic- plan-2020-2024-research-and-innovation_en#details Last Access: 08/09/2023.

Fedeli, L. (2017), *La ricerca scientifica al tempo dei social media.* Milano, Franco Angeli; pp. 11 - 134 (ISBN: 9788891757494.

Floridi, L. (2007). *A Look into the Future Impact of ICT on Our Lives*,The Information Society, 23:1, 59-64, DOI: 10.1080/ 01972240601059094.

Floridi, L. (2013). *The ethics of information.* Oxford University Press.

Floridi, L. (Ed.) (2015*). The onlife manifesto. Being human in a hyperconnected era*. Springer.

Floridi, L. (2016). *The 4th revolution: How the infosphere is reshaping human reality*. Oxford University Press.

Floridi L, Taddeo M. (2016). *What is data ethics?* Phil. Trans. R. Soc. A 374: 20160360. http://dx.doi.org/10.1098/rsta.2016.0360.

Foucault, M. (1966) *Le parole e le cose: un'archeologia delle scienze umane*, (trad. Emilio Panaitescu, Milano: Rizzoli, 1967.

Gantz, J. and Reinsel, D. 2011. *Extracting Value from Chaos*. IDC iView, International Data Corporation (IDC).

Han, B.C. (2015) *Nello sciame. Visioni del digitale.* Nottetempo editore.

Holewa, H., Wolski, M., Dallest, K., & McAvaney, C. (2015). *The HWMD maturity model: A foundational framework to measure the effectiveness of institutional research infrastructures*. Proceedings of the 2015 Higher Education Technology Agenda Conference (TETHA 2015), 1–17. Gold Coast, Australia. Retrieved from http://dro. deakin.edu.au/view/DU:30073806.

Laney, D. (2001). "*3D Data Management: Controlling Data Volume*, Velocity, and Variety", Gartner, file No.949. Feb 6 2001, At https://studylib.net/doc/8647594/3d-data-management--controlling-data-volume--velocity--an... Last access: 05/10/2023.

Jankowski, N.W. (Ed.). (2009). *E-Research: Transformation in Scholarly Practice* (1st ed.). Routledge. https://doi.org/10.4324/9780203875049.

Jansen, D. (2021). *Safeguarding academic and digital sovereignty: a model for action.* EUA Expert Voices https://eua.eu/resources/expert-voices/250-safeguarding-academic-and-digital-sovereignty-a-model-for-action.html.

Jonas, H. (1979). *The imperative of responsibility: In search of an ethics for the technological age*. Chicago (Ill.): The University of Chicago Press.

Knorr Cetina, K. D. (2001) 'Postsocial Relations: Theorizing Sociality in a Post-Social Environment', G. Ritzer and B. Smart (ed.) *Handbook of Social Theory* (pp. 520-37), London and Thousands Oaks, CA: Sage.

Latour B., *Reassembling the Social. An Introduction to Actor-Network Theory,* Oxford University Press, Oxford, 2005.

Law, J., (Ed.) (1986) *Power, action belief: A new sociology of knowledge?* London: Routledge and Kegan Paul.

Lei, H., Xiong, Y., Chiu, M. M., Zhang, J., & Cai, Z. (2021). *The relationship between ICT literacy and academic achievement among students: A meta-analysis*. Children and Youth Services Review, p. 127, 106123.

Mauss, M. (1934) *Body Techniques.* Article Originally Published Journal of Psychology, XXXII, ne, 3-4, 15 March-15 April 1936. Paper presented at the Society of Psychology on May 17 1934.

Merleau-Ponty, M., (1945) *Phénoménologie de la perception*. Paris: La Librairie Gallimard.

Ribeiro, B., Meckin, R., Balmer, A., & Shapira, P. (2022). *The digitalisation paradox of everyday scientific labour: How mundane knowledge work is amplified and diversified in the biosciences.* Research Policy, 52(1), 104607. https://doi.org/10.1016/j.respol.2022.104607.

Russo, F. *Digital Technologies, Ethical Questions, and the Need of an Informational Framework.* Philos. Technol. 31, 655–667 (2018). https://doi.org/10.1007/s13347-018-0326-2.

Salganik M.J. (2018), *Bit by bit. Social research in the digital age*. Princeton: Princeton University Press.

Samuel, G. Buchanan, E. (2020). *Guest Editorial: Ethical Issues in Social Media Research.* Journal of Empirical Research on Human Research Ethics. 15(1-2):3-11. doi:10.1177/1556264619901215.

Schon Donald A. (1993) *Il professionista riflessivo. Per una nuova epistemologia della pratica professionale.* Edizioni Dedalo.

Špiranec, S., Kos, D. & George, M. (2019). Searching for critical dimensions in data literacy. In Proceedings of CoLIS, the Tenth International Conference on Conceptions of Library and Information Science, Ljubljana, Slovenia, June 16-19, 2019. Information Research, 24(4), paper colis1922. Retrieved from http://InformationR.net/ir/24-4/colis/colis1922.html.

Sun, Y., Zhang, C., & Kok, R. A. (2020). *The role of research outcome quality in the relationship between university research collaboration and technology transfer: empirical results from China.* Scientometrics, 122, 1003-1026.

Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other.* Basic Books.

UNESCO (2017). *Recommendation on Science and Scientific Researchers.* https://en.unesco.org/themes/ethics-science-and-technology/recommendation_science.

Van Dijck J. (2014), *Datafication, dataism and dataveillance. Big data between scientific paradigm and ideology.* Surveillance & Society 12 (2): 197-208.

Van Dijck J., Poell T., de Waal M. (2018). *The Platform Society: Public Values in a Connective World,* New York, NY, Oxford University Press.

Van Petegem, W., Bosman, J., De Klerk, M., & Strydom, S. (2021). *Evolving as a Digital Scholar: Teaching and Researching in a Digital World.* Leuven University Press. https://doi.org/10.2307/j.ctv20zbkk0.

Zhang, Y., Adams, D., & Lee, K. C. S. (2022). *Assessing academic staff's ICT integration in higher education: instrument development and validation*. Journal of Computers in Education, pp. 1–20.

Weinhardt, M. (2021). *Big Data: Some Ethical Concerns for the Social Sciences*, Soc. Sci. 2021, 10(2), 36; https://doi.org/10.3390/ socsci10020036.

Weller, M. (2018). *The Digital Scholar Revisited. The Digital Scholar*: Philosopher's Lab, 1(2) 52–71. DOI: https://doi.org/10.5840/dspl20181218.