SOCIOLOGICAL ARTIFICIAL INTELLIGENCE RESEARCH (SAIR). A REVIEW AND CONCEPTUAL FRAMEWORK

DENIS IORGA[*]
BOGDAN BOCȘE[**]
DIANA OLAR[***1]

ABSTRACT

The phrase "Artificial Intelligence (AI)" appears more and more often in discussions about technology and society. Consequently, the sociological community needs a conceptual framework to study the phenomenon and its symbols. Although there are academic efforts in this sense, they tend to accentuate the scientific fragmentation that emerged due to the interpretative flexibility of AI and the lack of an accessible interdisciplinary language. The current work advances a conceptual framework for a sociological AI research by disentangling the interpretative flexibility of this technology, defining it, and reviewing existing sociological approaches on the subject. The framework was developed from a pragmatic point of view in which the aims of easing scientific progress monitoring and orientation of new researchers in the field of Sociological Artificial Intelligence Research (SAIR) guided the process.

Keywords: sociology, Artificial Intelligence, AI, conceptual framework, SAIR.

INTRODUCTION

Digital technologies offer a large number of benefits, for instance, faster communication, access to new opportunities, and more efficient ways to control the

[*] Interdisciplinary School of Doctoral Studies, University of Bucharest, e-mail: *denis.iorga@drd.unibuc.ro*.

[**] Knosis.AI, e-mail: *bogdan@knosis.ai*; DeepVISS.ORG, e-mail: *bogdan@deepviss.org*.

[***] Interdisciplinary School of Doctoral Studies, University of Bucharest, e-mail: *diana.a.olar@gmail.com*.

[1] As the main author of this article, Denis Iorga was responsible for designing and writing the manuscript using input from all authors. Both the secondary authors provided critical feedback and helped shape the manuscript. Bogdan Bocșe provided the main conceptual ideas discussed in the conclusion and discussion section and critically reviewed the manuscript; Diana Olar performed part of the literature review concerning sociological approaches to AI, contributed to the development of the framework, and critically reviewed the manuscript.

„Revista română de sociologie", serie nouă, anul XXXIV, nr. 1–2, p. 5–19, Bucureşti, 2023

environment. However, they also pose challenges related to ethics, privacy, power asymmetries, and inequalities, to name but a few. As a response to the diffusion of digital technologies across the social sphere, sociology adapted itself by trying to use the affordances brought forth by these tools and to understand their relationship with society.

The use of digital technologies in the sociological research practice led to new and better ways of gathering and analyzing data, building research networks, and instructing students (Lupton, 2014). Likewise, attempts to examine the relationship between digital technologies and society have led to a new understanding of the social order, as well as of current and future possibilities, risks, and challenges posed by these tools. The most illustrative example is provided by Zuboff (2019) who described the development and consequences on human freedom of a “surveillance capitalism” built on recent technological advancements (e.g., new types of sensors, social networks, predictive algorithms).

Nonetheless, the integration of some digital technologies into the sociological research practice can present challenges, given their interpretative flexibility (Pinch and Bijker, 1984): under the same technology definition, one finds different and often divergent meanings and designs. An illustrative example in this sense is *Artificial Intelligence (AI)*: different sociological research groups may have a different interpretation of what constitutes AI. These inconsistent interpretations of AI can lead to problems such as overestimation and underestimation of the impact of AI on society, overlooking potential risks and possibilities, naively investing resources in analyzing artifacts that are falsely labelled as AI, and providing conflicting recommendations to policymakers.

In other words, the interpretative flexibility of AI induces scientific fragmentation with respect to its integration into the research practice: having multiple definitions of AI can lead to a lack of clarity and consistency in defining and understanding both the problem spaces and collaboration opportunities related to AI, making it difficult for researchers to communicate and work across disciplinary boundaries and for readers to understand and evaluate research findings. In turn, scientific fragmentation hinders scientific progress (Balietti et al., 2015). This is particularly evident in the case of new researchers interested in the sociological investigation of AI.

Faced with an elusive subject matter, new researchers to the field of *Sociological Artificial Intelligence Research (SAIR)* often spend considerable time and resources in making sense of AI before producing added-value research. The difficulty for new researchers in navigating the SAIR literature was already signaled by Liu (2021). Moreover, the interpretative flexibility of AI makes communication of new researchers with peers from other scientific communities strenuous. Even more, the interpretative flexibility of AI hinders scientific progress monitoring due to the numerous definitions associated with this technology. In an attempt to address these challenges, *the current work focuses on developing a* *conceptual framework for AI research in sociology*, anchored in a narrative literature review.

The conceptual framework advanced in the current work aims to ease the orientation of new researchers interested in sociologically researching AI, provide a reference point in scientific cross-community communication, and also lessen the effort required for scientific progress monitoring on the topic. Its purpose is to act as a heuristic for new sociological researchers interested to navigate and contribute to the SAIR field. As such, the first section deals with disentangling the meaning of AI. It does so by distinguishing between intelligent machines and AI as the computer instructions driving their behavior. The subsequent section delineates the current sociological perspectives on AI, which encompass two distinct approaches: one that regards AI as a topic of investigation and another that perceives AI as a research instrument. Finally, the third section advances a novel conceptual framework for SAIR. The framework was developed through the integration of the findings presented in the initial and subsequent sections of this work, in addition to the authors' accumulated expertise gained from diverse AI endeavors.

1. AI AS COMPUTER INSTRUCTIONS

AI was originally defined by John McCarthy, in 1955, as the "science and engineering of making machines intelligent" (Manning, 2020). However, AI is often equated with the intelligent machines themselves in the public imagination (Cave et al., 2020), rather than with the science and engineering behind them. In turn, this entanglement opens room for the constellation of myths and metaphors surrounding intelligent machines, often amplified through science fiction media (Hermann, 2021), to obscure the economic, social, and power relationships involved in the making of AI[2]. Therefore, the current section will attempt to distinguish between intelligent machines (*e.g.*, robots, chat-bots, self-driving cars) as an embodiment of AI and AI as a set of computer instructions.

A definition of intelligent machines is first required to distinguish between intelligent machines and AI. However, Woolgar (1985) asserted the difficulty of this task due to the elusiveness of the term "intelligence". For this purpose, the following paragraphs will review a series of existing attempts to define what makes a machine intelligent. Afterward, the distinction between AI as a set of computer instructions and intelligent machines will be expanded.

Squibb (1973) argued that, from a sociological point of view, intelligence is an abstraction made from *patterns of behavior that are socially accepted as being*

[2] Plato's cave allegory is here illustrative for supporting the argument: the shadows (in this case intelligent machines) are mistaken for the fire (in this case the set of computer instructions defined as AI) and the source of the shadows (the network of humans and non-humans responsible for their implementation).

intelligent. As such, the question of "what makes a machine intelligent?" should be replaced with that of "how we tell a machine is intelligent?" In line with this view, further sociological efforts to define intelligent machines attempted to answer the latter question by focusing on the ability of such devices to perform in social settings in a manner accepted as being intelligent.

The sociological approach to defining intelligent machines draws attention to their capability of performing social roles (Schwartz, 1989). However, it can be argued that attention should be placed on the way the social role is performed rather than simply on whether a machine performs a social role or not. For example, both an android that closes a door and Latour's hinge-pin (Johnson, 1988) perform a similar social role: they assist humans in closing a door. Yet, an android closing the door is more connected to the notion of "intelligent machines" than a hinge-pin due to its way of performing the respective social role.

The focus on the social performance that allows machines to be regarded as intelligent was also taken by Alan Turing when he dismissed the question of "Can machines think?" (Turing, 2009). He then proceeded to work on the epistemology of defining intelligent machines by arguing that we say a machine is intelligent if we can't differentiate it from a human in a conversation. That is, an intelligent machine is a machine capable of *successfully performing the social role of a human conversation partner*.

Another attempt to define the behavioral characteristics that make a machine intelligent was provided by Russell and Norvig (2020). The authors used the bounded rationality of humans as their reference point. They argued that for a machine to be considered intelligent, it has to either *think humanly, think rationally, act humanly,* or *act rationally*.

Both examples illustrate anthropocentric attempts to define the patterns of behaviors that a machine has to display for it to be generally accepted as intelligent. Other efforts provide a more species-neutral approach. For example, Coelho Mollo (2022) argued that a machine can be regarded as intelligent if it *can perform appropriate aimed behavior* (goal-directed) *in different contexts* (general) *and under changing circumstances* (flexible)*, by taking into account previous interactions with the world* (adaptive). However, this approach re-opens a discussion that took place at the turn of the 21st century regarding the question of whether a pocket calculator can be regarded as an intelligent machine (Hauser and Rapaport, 2005).

Following the discussion of Hauser and Rapaport (2005) and the species-neutral approach presented by Coelho Mollo (2022), a pocket calculator can solve mathematical problems (goal-directed) in various types of changing environments (flexible and general) by computing numbers previously inserted by a human (adaptive). This leads to the counter-intuitive conclusion that a pocket calculator is an intelligent machine. In this regard, the scholarly literature proposes at least two alternative approaches that avert this inference.

The first solution refers to adopting a libertarian approach in describing intelligence. As the argument between Hauser (1993) and Rapaport (1993) goes, there is a minimum and a maximum of intelligence, with humans being placed near the maximum through their *sensory-perceptual* (*e.g.*: hearing, seeing, identifying), *cognitive* (*e.g.*: calculating, knowing, believing), and *conative behaviors* (*e.g.*: needing, wanting, seeking) (Hauser and Rapaport, 2005). As a pocket calculator partly displays the patterns of behaviors accepted as intelligent it can therefore be regarded as being minimally intelligent.

The second solution is to use a conservative approach to defining intelligence. Here, Coelho Mollo (2022) argued for a set of empirical standards, *a threshold in terms of goal-directedness, generalizability, flexibility, and adaptivity of behaviors*. Despite no mention of specific empirical standards, the author suggested it as a way to distinguish between entities capable of performing computational processes (a pocket calculator, for example), entities capable of self-organization and self-maintenance (plants, bacteria, fungi), and entities capable of our common-sense understanding of intelligent behaviors (animals, humans, and machines that mimic animal and human behaviors).

Based on the perspectives presented above, intelligent machines seem to represent any ensemble of non-natural objects that can perform one or more of the behaviors below:

- Successfully perform the social role of a conversation partner.
- Think or / and act humanly or /and rationally.
- Display enough sensory-perceptual (see), cognitive (calculate), and/or conative (want) behaviors to be regarded as intelligent by humans.
- Display enough goal-directed, context-agnostic, change-responsive, and past-related adaptable behaviors, above a yet-to-be-defined empirical threshold.

Notwithstanding the definition chosen for intelligent machines, a distinction between intelligent machines and AI can be made. AI refers to the intelligence of intelligent machines, rather than to the intelligent machines themselves. In other words, current intelligent machines (*e.g.*, self-driving cars, chatbots, robots) are able to manifest one or more of the behaviors listed above due to AI. Yet, if AI refers to the intelligence of intelligent machines, what exactly is AI?

As presented at the beginning of this section, the founder of the field, John McCarthy, defines AI as an application of science and engineering. For a better understanding, The European Commission's AI Act refers to this application of science and engineering in terms of software built through a set of *specific techniques* and capable of generating content, decisions, predictions, and recommendations that influence the environment (Council of the European Union, 2021).

Furthermore, Annex I of the AI Act provides an updatable list of techniques specific to software defined as AI: "(***a***) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning, (***b***) Logic – and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference, and deductive engines, (symbolic) reasoning and expert systems; and (***c***) Statistical approaches, Bayesian estimation, search and optimization methods" (Council of the European Union, 2021).

Citing Marcus Tomalin, Cave *et al.* (2020) argued that the commonality of these techniques lies in the automation of functions of the human brain. A less anthropomorphizing and accurate description (considering advancements in swarm intelligence, for example) would assert that the commonality lies in the automation of the function of biological systems. Nonetheless, the founder of the term "AI" seems to suggest that the set of techniques involved in the development of AI does not have to be confined to methods that are observable in biological systems (McCarthy, n.d).

Based on the arguments presented so far, AI refers to a *specific set of computer instructions* (or techniques) capable of generating content, decisions, predictions, and recommendations that influence the environment. Adopting this approach in order to define AI has two main advantages. Firstly, it manages to steer the attention away from anthropomorphisation, fetishism, and yet-to-be-defined thresholds and spectrums of intelligence, toward a discussion concerning the development, use, and consequences of these instructions. Secondly, it enhances the capability of explaining AI, as compared to other more nuanced and extended definitions.

However, it is important to notice that this approach to defining AI also has a major drawback. For example, can the computer instructions of a proximity sensor-based door closer be regarded as AI? After all, these instructions are implemented using a logic-based technique: if the sensor detects movement, it decides to open the door. To address this challenge, a possible refinement of the definition would refer to acknowledging AI *as a source of promises of computation*.

Unlike the promises of computation made by an AI system, the promises of computation made by a proximity sensor door-closer tend to realize as expected, or, in Ellul's (1990) terms, with a low degree of unpredictability. That is, the computational promises made by AI distinguish themselves in terms of output *volatility*. The promises of computation made by AI, as compared to other computer instructions, are subject to faster and more unpredictable change. This approach is in line with the libertarian approach in defining intelligent machines by appealing to a difference of degree to distinguish between AI and other computer instructions. Nonetheless, this approach has the benefit of a more established and grounded empirical tradition in measuring unpredictability, compared to the measurement of intelligence.

Another possible approach to distinguish AI from other computer instructions refers to acknowledging that both are responsible for turning *extelligence*[3] (here defined as interpretable and relevant ex-corpore knowledge) into *intelligence* (here defined as interpretable and relevant in-corpore knowledge). In this sense, AI differs from other computer instructions in terms of the *extelligence space* they use. This view follows the conservative approach in defining intelligent machines by allowing for a difference of type to be established between AI and other computer instructions.

As regards the characteristics of the extelligence space, the concept of "the network" here becomes relevant to illustrate the difference of type between AI and other computer instructions. More precisely, aspects related to the topology of the human and non-human network responsible for the production of the extelligence used by a set of computer instructions represent good candidates to decide whether it is AI or not. However, this aspect is yet to be explored. Until then, characterizing AI as a specific set of computer instructions facilitates the disassociation of intelligent machines from the overarching notion of "AI".

2. SOCIOLOGICAL AI RESEARCH (SAIR)

The current section focuses on existing academic efforts that aimed to conceptualize SAIR. The purpose of this section is to highlight the existence of two complementary approaches described by Mlynář *et al*. (2018): the humanist approach (*that treats AI as an object of study*) and the computationalist approach (*that treats AI as a tool for research*). As it will be presented in the following lines, both approaches seem to focus on one or more of the following three objects of study: inner workings, applications, or / and narratives.

Liu (2021), for example, showed that approaches for sociologically researching AI as an object of study can be grouped into three analytical perspectives. The "scientific AI" perspective focuses on AI as a scientific field and looks at the social context in which AI is researched and developed. The "technical AI" perspective focuses on AI as a meta-technology and seeks to understand the social ramifications of specific AI applications and technologies. The "cultural AI" perspective focuses on AI as a phenomenon with consequences on the social, economic, and cultural environment.

Another view was presented by Sartori and Theodorou (2022) who identified two main directions for the study of AI as a sociological object of study. Unlike the work of Liu (2021), which used different understandings of AI to define possible research approaches, Sartori and Theodorou (2022) focused on the purpose of AI research. As such, they identified approaches that seek to enhance human control

[3] The term "extelligence" was adapted from Stewart and Cohen (1997).

over AI and approaches that seek to understand narratives used to socially construct AI.

An earlier work developed by Bailey and Barley (2020) pointed toward two main approaches when researching AI as a sociological object of study. The first approach focuses on how AI is designed. The second approach focuses on how AI is used. However, the authors argued that both these approaches tend to ignore issues of social institutions, power, and ideology. Thus, they suggested the existence of a third approach that considers the aforementioned issues.

Joyce *et al.* (2021) outlined a research agenda for the sociology of AI by stressing the important role of concepts such as inequalities and structural change. In their view, there are two potential approaches for sociologists concerning AI as an object of study. The first approach is to focus on analyzing the politics of data, algorithms, and code. The second approach is to focus on how AI is socially shaped in practice.

Approaches that take AI as a tool for research look at AI from a pragmatic point of view and equate it with certain computational techniques that sociologists can use to enhance their research practice. For example, Carley (1996) discussed different approaches within sociology that use AI as a tool for research. The author argued that AI can be understood as a tool for text analysis, a tool for network analysis, and a tool for theory development and evaluation.

A more recent perspective on using AI as a tool of research focused on the applications of machine learning in sociology (Molina and Garip, 2019). The authors distinguished between the applications of unsupervised and supervised machine learning algorithms. They argued that AI can be used for policy prediction, causal inference, data augmentation, phenomena measurement, population heterogeneity characterization, and model checking.

Chen *et al.* (2021) focused on how social sciences can harness the predictive power of machine learning algorithms. The authors argued that AI can be used to find latent indicators, augment data, generate theoretical hypotheses, and make causal inferences. Additionally, the authors emphasized the potential of AI to be used for interventional research aimed at helping disadvantaged social groups and improving social governance in various fields such as criminology, health, and international politics[4].

Yet another approach was described by Briscoe (2022) at a conference regarding the emergence of automated social sciences. The host of the conference illustrated how AI can facilitate and improve: the collection of relevant literature

[4] Here, it should be noted that the use-case presented by the authors for AI as an interventional research tool (classifying disadvantaged social groups in order to provide targeted social assistance) seems to be developed on a utilitarian moral compass, which, in conjunction with the limitations of AI, carries the risk of sacrificing the individual in favor of the collective. While the potential of AI to help disadvantaged groups exist, our view is that its implementation should be made based on a Kantian moral compass.

through natural language analysis and bibliometrics, study design (through identifying scientific gaps, running meta-analysis, creating nomological networks, and summarizing related studies), as well as measuring scientific integrity and validating scientific claims.

Two observations can be made based on the review of existing approaches in SAIR. First, none of the authors illustrate the use of existing AI applications to enhance other laboratory and academic tasks beyond data analysis (for instance, writing or teaching), despite the increasing adoption of such tools by academics. Second, regardless of the approach, each of the presented authors tends to define AI as either a computer technique, an application, or a concept. As will be presented in the following section, this allows for the development of a conceptual framework for SAIR that is compatible with both the computationalist view and the humanist view.

3. A CONCEPTUAL FRAMEWORK FOR SAIR

The review presented in the previous section suggested the existence of three objects of inquiry in SAIR. Based on this inference, the current section will attempt to advance a novel conceptual framework for SAIR that considers AI as an object of research and as a tool for research. The aim of the framework is to act as a heuristic for sociological researchers interested to navigate and contribute to the SAIR.

The first object of study in SAIR refers to *AI inner workings*. Approaching AI inner workings as an object of study means examining the design and implementation of the specific set of computer instructions defined as AI. Following the work of Woolgar (1985), this can be translated into a sociology of AI researchers.

AI inner workings also represent a central object of study for sociologists that seek to use AI as a tool for research. However, this approach can be translated into a form of applied statistics and computer programming, rather than a sociology of AI researchers. Sociologists that are interested in using AI study its inner workings in order to apply them for different research purposes (*e.g.*: data analysis). Sociologists interested in taking AI as an object of research study its inner workings to understand its social construction.

The second object of study refers to *AI applications*. Taking AI applications as an object of study means analyzing the behavior of hardware and software apparatuses powered by the specific set of computer instructions defined as AI. In other words, researching AI applications refers to studying how intelligent machines interact with the social world. The aim of this inquiry can range from identifying implementation challenges to understanding the impact of AI applications on social groups.

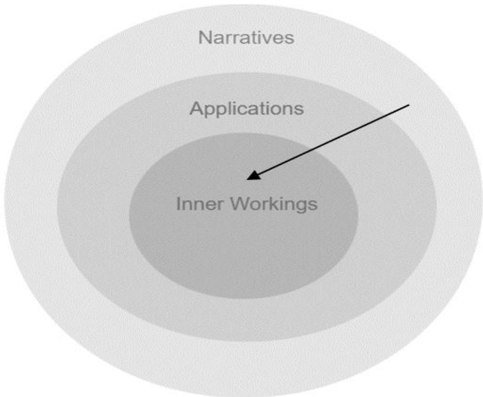
Sociologists interested in using AI as a tool for research can also study the behavior of AI applications to identify opportunities for enhancing the research practice. For example, a sociologist may focus on understanding how AI-powered text analysis software can be used, in order to apply it for data analysis. In this case, the difference between using AI applications and AI inner working lies in the degree of technical knowledge required to obtain the desired result. In other words, using AI applications as a tool for research requires a “black box” understanding, as opposed to the “white box” understanding required for using AI inner workings.

The third object of study refers to what Cave *et al*. (2020) defined as *AI narratives*. Studying AI narratives means investigating stories circulated through various media regarding AI applications and inner workings. While both the study of AI inner workings and AI applications involves working with such stories, studying AI narratives does not acknowledge a single definition of AI and seeks to understand the rules of discourse production. Alternatively, those that seek to use AI as a tool for research can use AI narratives to identify appropriate AI applications and inner workings for their purpose.

Figure 1 illustrates the relationships between AI inner workings, AI applications, and AI narratives. The illustration suggests, through the black arrow, that identifying AI in practice can be achieved by starting with AI narratives. AI narratives can be used as rhetorical devices to identify claims of AI applications and AI inner workings. It is important to notice that the inquiry can focus specifically on AI narratives, with no requirement to investigate the artifacts on which the narratives were built. When this is not the case, AI applications and inner workings can be identified through AI narratives and be approached as tools for research or/and as objects of research.

Figure 1

Three objects of study when researching AI.



With a working definition of AI (Section 1) and clearly defined objects of study at hand (AI narratives, AI applications, and AI inner workings), this section now introduces a conceptual framework for SAIR. The development of the conceptual framework presented in *Figure 2* was conducted from a pragmatic philosophical standpoint (Creswell, 2009) in which the goals of the framework guided the process, namely effortless orientation for new researchers to the field and scientific progress monitoring. Thus, the reader should remember that the framework aims to guide new researchers interested in studying AI from a sociological point of view, and also allow the classification of existing and future sociological research regarding AI.

Figure 2

Conceptual framework for Sociological Artificial Intelligence Research (SAIR).

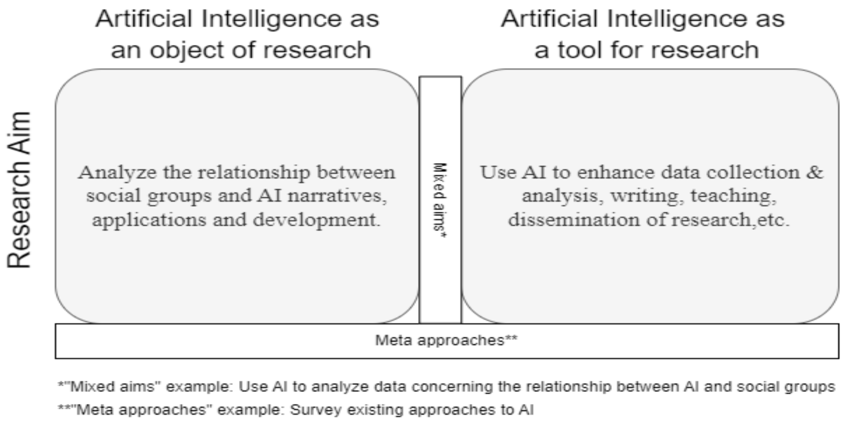


Figure 2 distinguishes between the "AI as an object of research" and the "AI as a tool for research" perspectives. Moreover, it distinguishes between different research aims. Meta approaches and mixed aims were included as categories. The latter two categories account for studies such as those that, for example, use AI to study the relationship between AI and social groups (mixed aims that merge the two perspectives), and studies such as the present one (meta approaches).

The first step taken to develop the framework was to consider both the "AI as an object of research" and "AI as a tool for research" perspectives. Thus, the conceptual framework presented in *Figure 2* distinguishes between approaches that focus on understanding AI in relationship with the social world ("AI as an object of study") and approaches that focus on using AI to understand the social world ("AI as a tool for research"). This perspective thus illustrates both the humanist and the computationalist approaches.

The second step taken was to identify possible research aims that a new researcher can have for both approaches. Based on the analysis of existing efforts to sociologically conceptualize SAIR, the framework presented in *Figure 2* identifies *the study of the relationship between AI and social groups* and *the use of AI to enhance research practice* as the main research interests in SAIR. It is worth noticing that the framework also includes the use of AI to assist in writing and teaching.

The use of AI to enhance research practice involves using AI to collect, augment, or analyze data, improve writing, help reading and teaching, disseminate results, and other related research activities. It involves various use cases, from theory generation to identifying latent indicators. Important to notice here is that the use of AI as an interventional research tool was not included due to its ambivalent potential consequences.

New researchers interested in the "AI as a tool for research" perspective can start by using various information sources to find products or techniques with the AI label (AI narratives) that seem useful in their research practice. For example, a sociologist may come across an AI-based PowerPoint generator by searching for the "AI PowerPoint generator" keywords in a search engine. Following the identification of a product or technique that may seem helpful for research practice, the sociologist can apply it. Going back to the example presented above, the sociologist may choose to use the PowerPoint generator to create materials for research results dissemination.

It is important to mention that using AI for research practice also opens room for interdisciplinary collaboration. Sociologists can go beyond using AI and choose to contribute to the development of AI (AI inner workings). For example, they may get involved in a project that aims to use AI to automatically analyze online newspapers. Their expertise in analyzing public opinion can be used in every stage of the project, from establishing functional requirements to evaluation.

The study of the relationship between AI and social groups refers to the examination of the mutual influence between AI and social groups. Thus, sociologists can focus on understanding how social groups influence AI narratives, applications, and inner workings. Some examples here would be examining why and how certain groups use cultural resources to tell stories concerning AI, how groups hinder or promote the use of AI, and how developers embed bias in AI. On the other hand, sociologists can focus on understanding how AI narratives, applications, and inner workings influence social groups. Some examples here would be analyzing how different stories of AI influence beliefs regarding other spheres of life, how AI applications have negative or positive consequences on various groups, and how the development of an AI application leads to various resource/knowledge mobilization between groups.

CONCLUSION

Inspired by the European Commission's AI ACT, the current work asserted that AI refers to a specific set of instructions that computers use to generate content, predictions, decisions, and recommendations that influence the environment. The present paper advocated for a distinction between AI and intelligent machines. It was argued that equating AI with intelligent machines, and not with the computer instructions behind them, amplifies the entanglement of AI with terms such as "robots", "self-driving cars", "chatbots" and others. In turn, this entanglement steers the attention away from the social dynamics behind these computer instructions. In other words, the present work suggested that the interpretative flexibility of AI, as well as the problems emerging from it, is amplified through the lack of a clear-cut distinction between AI and intelligent machines as an embodiment of AI.

By defining AI in terms of computer instructions, its interpretative flexibility is displaced from the term "intelligent machines" to that of "computer instructions". Nonetheless, the volatility of the output of computer instructions and characteristics of the used extelligence space were advanced as potential solutions to differentiate AI from other types of computer instructions. With this in mind, the current work then reviewed existing attempts to sociologically conceptualize AI. The insights resulting from this endeavor suggested the existence of two sociological approaches to AI: "the humanist" perspective and "the computationalist" perspective. The examination of the two approaches created the grounds for developing a conceptual framework designed to assist in the sociological investigation of AI.

In an attempt to provide a heuristic for new researchers to the field of SAIR, the present work advanced a conceptual framework for SAIR that illustrates three possible objects of study (narratives, applications, inner workings), four main approaches (object of study, tool for research, mixed aims, and meta approaches), and two potential research aims (understand the relationship between AI and the social world and use AI to understand the social world). The framework aims to assist those interested in navigating and contributing to the SAIR field and also in scientific progress monitoring. Moreover, the proposed framework can also act as a reference point in interdisciplinary discussions related to AI, thus enhancing the potential for scientific collaboration between different research communities. For this purpose, further research should focus on evaluating the capability of the framework to achieve its intended purpose.

REFERENCES

1. BAILEY, D.E. & BARLEY, S.R. (2020). Beyond design and use: How scholars should study intelligent technologies. *Information and Organization,* 30(2), 100286. *DOI: https://doi.org/10.1016/j.infoandorg.2019.100286*.
2. BALIETTI, S., MÄS, M. & HELBING, D. (2015). On disciplinary fragmentation and scientific progress. *PloS one,* 10(3), e0118747. *DOI*: *https://doi.org/10.1371/journal.pone.0118747*.
3. BRISCOE, E. Computational Social Science. FY (2021) for the Fulbright Specialist Program (2022). University Politehnica of Bucharest.
4. CARLEY, K.M. (1996). Artificial intelligence within sociology. *Sociological methods & research,* 25(1), p. 3–30. DOI: *https://doi.org/10.1177/0049124196025001001*.

5. CAVE, S., DIHAL, K. & DILLON, S. (2020). *AI narratives: A history of imaginative thinking about intelligent machines,* England, Oxford University Press. ISBN: 0192586041.
6. CHEN, Y., WU, X., HU, A., HE, G. & JU, G. (2021). Social prediction: a new research paradigm based on machine learning. *The Journal of Chinese Sociology,* 8(1), p. 1–21. *DOI: https://doi.org/10.1186/s40711-021-00152-z*.
7. COELHO MOLLO, D. (2022). Intelligent Behaviour. *Erkenntnis*, p. 1–17. *DOI: https://doi.org/10.1007/s10670-022-00552-8*.
8. COUNCIL OF THE EUROPEAN UNION (2021). *Regulation of the European Parliament and of the Council Laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts* [Online]. Available: *https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206* [Accessed October, 2022].
9. CRESWELL, J. (2009). Philosophical Worldviews. *Research design: Qualitative, quantiative and mixed method approaches.* 3rd Edition ed. California: SAGE Publications. pp. 5–10. ISBN: 9781412965576.
10. ELLUL, J. (1990). Unpredictability. *The Technological Bluff.* Grand Rapids, Michigan: William B. Eerdmans. ISBN: 0-8028-3678-X.
11. HAUSER, L. (1993). Why isn't my pocket calculator a thinking thing? *Minds and Machines,* 3(1), p. 3–10. *DOI: https://doi.org/10.1007/BF00974301*.
12. HAUSER, L. & RAPAPORT, W.J. (2005). *Why isn't my pocket calculator a thinking thing?* [Online]. Available: https://cse.buffalo.edu/~rapaport/Papers/joint.pdf [Accessed September, 2022].
13. HERMANN, I. (2021). Artificial intelligence in fiction: between narratives and metaphors. *AI & society*, p. 1–11. *DOI: https://doi.org/10.1007/s00146-021-01299-6*.
14. JOHNSON, J. (1988). Mixing humans and nonhumans together: The sociology of a door-closer. *Social problems,* 35(3), p. 298–310. *DOI: https://doi.org/10.2307/800624*.
15. JOYCE, K., SMITH-DOERR, L., ALEGRIA, S., BELL, S., CRUZ, T., HOFFMAN, S. G., NOBLE, S. U. & SHESTAKOFSKY, B. (2021). Toward a sociology of artificial intelligence: A call for research on inequalities and structural change. *Socius,* 7. *DOI: https://doi.org/10.1177/2378023121999581*.
16. LIU, Z. (2021). Sociological perspectives on artificial intelligence: A typological reading. *Sociology Compass,* 15(3), e12851. *DOI: https://doi.org/10.1111/soc4.12851*.
17. LUPTON, D. (2014). Scoping Digital Sociology. *Digital sociology.* Oxfordshire: Routledge. pp. 12–19. ISBN: 9781315776880.
18. MANNING, C. (2020). *Artificial Intelligence Definitions* [Online]. Standford University: Standford University Human-Centered Artificial Intelligence. Available: *https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf* [Accessed September, 2022].
19. MCCARTHY, J. n.d. *What is AI? / Basic Questions* [Online]. Standford University. Available: *http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html* [Accessed October, 2022].
20. MLYNÁŘ, J., ALAVI, H.S., VERMA, H. & CANTONI, L. (2018). Towards a sociological conception of artificial intelligence. *Artificial General Intelligence: 11th International Conference.* Prague: Springer. pp. 130–139. *DOI: https://doi.org/10.1007/978-3-319-97676-1_13*.
21. MOLINA, M. & GARIP, F. (2019). Machine learning for sociology. *Annual Review of Sociology,* 45, p. 27–45. *DOI: https://doi.org/10.1146/annurev-soc-073117-041106*.
22. PINCH, T.J. & BIJKER, W.E. (1984). The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social studies of science,* 14(3), p. 399–441. *DOI: https://doi.org/10.1177/030631284014003004*.
23. RAPAPORT, W.J. (1993). Because mere calculating isn't thinking: Comments on Hauser's Why Isn't My Pocket Calculator a Thinking Thing? *Minds and Machines,* 3(1). *DOI: https://doi.org/10.1007/BF00974302*.
24. RUSSELL, S. & NORVIG, P. (2020). What is AI? *Artificial intelligence a modern approach 4th Edition.* New York: Pearson Education, Inc. pp. 1–5. ISBN: 0136042597.

25. SARTORI, L. & THEODOROU, A. (2022). A sociotechnical perspective for the future of AI: narratives, inequalities, and human control. *Ethics and Information Technology,* 24(1), p. 1–11. *DOI: https://doi.org/10.1007/s10676-022-09624-3.*
26. SCHWARTZ, R.D. 1989. Artificial intelligence as a sociological phenomenon. *Canadian Journal of Sociology/Cahiers canadiens de sociologie*, p. 179–202. *DOI: https://doi.org/10.2307/3341290*.
27. SQUIBB, P. (1973). The Concept fo Intelligence–A Sociological Perspective. *The Sociological Review*, 21(1), p. 57–75. *DOI: https://doi.org/10.1111/j.1467-954X.1973.tb00479.x.*
28. STEWART, I. & COHEN, J. (1997). *Figments of reality: The evolution of the curious mind,* England, Cambridge University Press. ISBN: 1139425293.
29. TURING, A. M. (2009). Computing machinery and intelligence. *Parsing the turing test.* Dordrecht: Springer. pp. 23–65. *DOI: https://doi.org/10.1007/978-1-4020-6710-5_3*.
30. WOOLGAR, S. (1985). Why not a sociology of machines? The case of sociology and artificial intelligence. *Sociology,* 19(4), p. 557–572. *DOI: https://doi.org/10.1177/003803858501900400.*
31. ZUBOFF, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power,* London, Profile Books. ISBN: 1782832742.