You have downloaded a document from

The Central and Eastern European Online Library

The joined archive of hundreds of Central-, East- and South-East-European publishers, research institutes, and various content providers

Source: Bulletin of the Transilvania University of Braşov, Series VII: Social Sciences and Law

Bulletin of the Transilvania University of Braşov, Series VII: Social Sciences and Law

Location: Romania

Author(s): Livia Dana Pogan, Radu-Ioan Popa

Title: A Theoretical and Practical Comprehensive Framework for Artificial Intelligence and Socio-Psychological Challenges

A Theoretical and Practical Comprehensive Framework for Artificial Intelligence and Socio-Psychological Challenges

Issue: 1/2020

Citation style: Livia Dana Pogan, Radu-Ioan Popa. "A Theoretical and Practical Comprehensive Framework for Artificial Intelligence and Socio-Psychological Challenges". Bulletin of the Transilvania University of Braşov, Series VII: Social Sciences and Law 1:33-40.
https://www.ceeol.com/search/article-detail?id=882815

CEEOL copyright 2023

Bulletin of the *Transilvania* University of Braşov
Series VII: Social Sciences • Law • Vol. 13(62) No. 1 – 2020
https://doi.org/10.31926/but.ssl.2020.13.62.1.3

A THEORETICAL AND PRACTICAL COMPREHENSIVE FRAMEWORK FOR ARTIFICIAL INTELLIGENCE AND SOCIO-PSYCHOLOGICAL CHALLENGES

L. POGAN[1] R. I. POPA[2]

Abstract: *The present paper aims to highlight the major up to date findings and research trends on the topic of artificial intelligence (AI) from a theoretical perspective, analysed in association with the human-machine interaction challenges. The existing research and scholarly literature display quite heterogenous approaches on these topics, and only a few of them tackle the social sciences view on the matter. In conclusion, there is a need for a more specific approach from the socio-psychological perspective, especially in topics concerning human centricity, human training at work, personality variables and other indicators when using AI devices.*

Key words: *artificial intelligence, human centricity, personality.*

1. The context

Artificial intelligence (AI) is undoubtedly already present in our lives, impacting steadily more and more domains. If at the beginning, innovations brought by Industry 4.0 were implemented in factories, their spread involves, nowadays, individuals` functioning. Thus, smart devices, intelligent apparatus or assistive robots are a part of peoples` daily existence.

These transformations are seen as emergent of the Fourth Industrial Revolution, firstly mentioned by Klaus Schwab, the founder and executive chairman of the World Economic Forum. He describes the numerous challenges faced now-a-days, understanding and adapting to the technology revolution being one of the most important ones, and impacting everyone and humanity itself (Schwab, 2018). Furthermore, he notices the echoes of this revolution in all life sectors, as work, relations or entertainment and considers that the amplitude of these changes cannot be compared to anything that humanity has ever experienced before (Schwab, 2018).

The history of humanity has always been channelled by transformations and change

[1] Lucian Blaga University of Sibiu, Romania, livia.pogan@ulbsibiu.ro, corresponding author
[2] Lucian Blaga University of Sibiu, Romania

CEEOL copyright 2023

CEEOL copyright 2023

has accompanied and fostered progress. Each industrial revolution reshaped peoples` existence, living, working and interacting. From a technology revolution point of view, the first step was established in the 18th century, when production moved outside the house, due to the steam-power engine, an age considered to be the First Industrial Revolution. Furthermore, electricity started to be used in factories and manufacturing units, making mass production possible. Thus, the first years of the 20th century were marked by the Second Industrial Revolution. The Third Industrial Revolution, situated in the 1970`s, again changed the way people worked, reshaping several industrial domains, not only production. The previous two revolutions had already brought workers a greater independence from nature and many activities were no longer undertaken by men. Automatic production chains have made jobs both easier and less dangerous, while making affordable goods available for large categories of consumers.

Each technological jump was translated into an industrial revolution, fundamentally changing lives for many people. The actual global context makes the spread of novelty, innovations and technology, at a faster peace (Pogan, 2019). Therefore, the impact of the Fourth Industrial Revolution is greater than any of the previous ones.

2. State of the Art Framework concerning AI use

In the last years, the research interest has been developing side to side with the implementing process when it comes to AI platforms, devices and solutions from various sectors of activity. The most prominent fields that invested both in research and implementation, as presented in the coming sections, are related to the industry, medical, transportation and IT domains.

In the following sub-chapters, we will present several up to date target fields of interest and research directions, the social sciences` contribution, concluding with possible implications and applications for the future.

2.1. Target fields of interest and research directions

The industrial sector is one of the major contributors when it comes to investing in the development of smart solutions and performance concerning human-machine interaction. Despite the fear that machines replace more and more the human manual work, a study by Ganschar et al. (cited in Bertram et al., 2018, p. 170) showed that companies still rely on human workers and manual work in a proportion around 43% to get the job done. From this point of view, it can be observed that there is still a certain flexibility and freedom inside the human work and task response, that machines cannot cover yet. At this point, the industry is highly interested in supporting the human work with intelligent assistive devices, in order to meet the growing demands of the customer, to bring out new technologies to train and help workers act faster, more prompt, avoid accidents and errors, thus improving product quality and performance, while making the production more flexible and smart orientated.

In terms of trends for industry and research, Bertram et al. (2018) described the assistive systems projects and prototypes for manual working stations known so far to

CEEOL copyright 2023

CEEOL copyright 2023

the public. The study concludes that only some of these existing products manage to offer automated work plans, having flexible and adaptable integration in the production or possessing the autonomous learning ability, as required for AI equipment.

Another focus in the industrial area consists of the interest in managing artificial-intelligent devices and virtual realities (VR) at work. This experience was extended to education, training, entertainment, culture, sports and transportation also. VR application integrates multimedia, sensors, various displays, human-machine interaction, ergonomics, simulations, computer graphics and of course AI technologies (Zhu, Fan & Zhang, 2019). In a review conducted by Zhu, Fan and Zhang (2019), data showed that the major research trends and applications concerning the digital human models and the use of VR circulate mainly in the fields of automotive industry (41.30%), followed by aerospace industry (19.57%) and the industrial plants (19.57%), then the energy industry (10.87%) and military industry (8.70%). According to the same authors, the application type focuses as follows on assembly maintenance/simulation and evaluation (41.3%), workplace design and optimization (17.4%), automotive interior assessment (17.4%), human-robot collaboration simulation (10.9%), education/training (6.5%), human behaviour (6.5%) (Zhu, Fan & Zhang, 2019, p. 563). From this point of view, the social and psychological dimensions are represented and studied at the lowest rates, giving way for the imperative need for future studies and approaches when it comes to studying human-machine interaction in AI settings.

Besides the VR systems, the augmented reality platforms (AR) also showed numerous applications, following the Industry 4.0 and digitalization process. Gattullo et al. (2017) described in their study the benefits of using visual technical documentation through AR in the automotive sector, such as: improving text readability and translation, providing intuitive representation of formal notations, using graphic symbols, reducing text. Moreover, in another review, Sharma, Yadav and Chopra (2020), presented the main framework where AI is applied to various domains, stating several trend features, whereas the economic and financial (e.g. fraud detection, digitalization, staffing, banking, unemployment predictions etc.) lead the way, followed by the transportation sector (e.g. traffic, routing, electric vehicles, automation etc.) and environmental sustainability (e.g. waste management, wildlife care, energy concerns, geo-science etc.). On the last positions were placed the information and communication technology (e.g. data management, networking, data security etc.), policy making and government law (e.g. public administration, legal systems, e-voting etc.) and healthcare (e.g. bio-medicine, advanced medical devices, food quality etc.). The authors state that there are gaps between the various sectors of activity, concerning the research of AI, in opposition with the implementation of these new technologies. Studies are scarce on the topics of healthcare, communication technology, education, social and cultural services, missing out probably important implications, associated risks and AI adoption in various fields (Sharma, Yadav & Chopra, 2020, p. 5).

Another sector that brings out some interesting research trends when it comes to the study and application of AI, consists of the medical field. Thus, the close collaboration between medical specialists and IT professionals contributed to developing various smart devices and technologies for medical use. For example the neurotechnical

CEEOL copyright 2023

CEEOL copyright 2023

developments in the recent years brought to attention the use of brain-computer interfaces, where, as it is expected, brain signals from a human subject are transferred and translated into commands for output devices, in order to carry out needed actions for the case of various patients in numerous disorders such as sclerosis, cerebral palsy, stroke, spinal cord injury etc. (Keskinbora, 2019). From this perspective, the applications may seem limitless, from the complex control of cursors, robotic arms, prostheses to smart wheelchairs, remote online surgery procedures and rehabilitation. Thus, Ji, Hwang and Kim (2013) showed the applied results for testing an intelligent wheelchair system for people with various disabilities, concentrating the outcome on the human subject intention recognition for movement, obstacle detection and avoidance and of course situation awareness. Haleem, Javaid and Khan (2019, p. 233) stated in their scholarly literature overview that there are five major AI technologies in relation to the medicine: machine learning, artificial neural network, natural language processing, support vector machine and heuristics analysis. The same authors conclude that AI covers at present diagnosis, personalized treatment and drug development, surgery, radiology, hospital administration and individual health records, cardiology. The same impact of AI on medical imaging market and radiology is supported and described furthermore in a research paper by Alexander et al. (2020). The use of AI assistive and integrated systems can provide multiple applications, alongside and in collaboration with the engineer domain for the future of medicine in the upcoming years.

The third major domain in which AI and the use of assistive systems create quite an impact, regards the transportation and traffic sectors. From the so-called and controversial remote guidance control, data and response analysis fully automated for air traffic or naval cargo to intelligent spacecraft module, the trend is the same: to create a human-machine optimal interaction, training and work performance. Collecting big data from different indicators such as sensory, voice, emotional, movement, biological patterns can improve the human-device interaction in terms of transportation. For example, Oungrinis et al. (2014) concluded that the AI systems should be ready to evaluate the activity of the user and exhibit a behavioral response in accordance to the activity, context and psychological condition of human users in the case of astronauts with reference to the adaptation of an environment and ambiance module in a spacecraft. Another example concerning the transportation sector could be the study of Eckhardt and Rantala (2012) which concluded that intelligent logistics centers using radio frequency identification and data analysis detain an increased efficiency in operations and customer service.

2.2. The social sciences challenges and AI

Smart technologies are built by people, to act as people, for people and the human factor is therefore of highly importance in this equation. After describing the main domains penetrated by AI, in the current section we will address the social sciences` contribution regarding technological development and implementation.

Present scholarly literature states the need for more insight and data, research and interest to study in-depth the human-machine interaction, with a more close-up on the

CEEOL copyright 2023

CEEOL copyright 2023

human part. In this view, Miller (2019, p. 33) concludes that AI research has not yet truly addressed the socially-interactive explanations and study, focusing more on symbolic models, leaving aside how people define, generate, select, present and evaluate the information, leaving so much room for integrating the social science views within this models. The same author states the imperative need for collaboration between researchers on AI topics, gathering participation from psychology, cognitive science, human-computer interaction in order to generate a better understanding, an integrative and complete framework and comprehensive approaches for future studies in building new applications and models (Miller, 2019, p. 34).

Another major topic from the social sciences perspective is the individual's personality features and patterns of emotions when working with AI technology. It is quite a challenge to build up models for research, even experimental, to track, address and evaluate the psychological patterns in association with the use of AI. Future studies will need to add more focus on the social elements and the human user in relation to smart devices, trying to integrate the socio-psychological component in the AI matrix. Professional competencies are a key element of the human-AI interaction, which can be better understood from a dynamic, adaptative perspective, despite the classical, taylorism view (Pavalache-Ilie, 2017). Shank et al. (2019), presented several results when humans work and perceive the interaction with AI devices, showing numerous cluster emotions within the interaction, indicating that the human emotional reactions are somehow tied to the perceived special outcomes of AI, its capacity to engage in human-like actions and inhabit social roles. In another work, Samsonovich (2020) presented also that human social-emotional intelligence models are yet to be analyzed and developed, creating the gateway for a virtual creative assistant for example, and studies need to focus more on how an embodiment can be made in an AI agent, creating thus a socially emotional AI, which understands human mind, emotions, moods, feelings, reactions, somatic data and biases. The proposed model included current plans and commitment, moral and ethical values and somatic comfort under moral schemas and semantic maps (Samsonovich, 2020). Beyond the present trends, where AI machines recognize images and voice input, the next level requires more in-depth data to be developed. Wang, Xie and Lu (2016) underlined the need of cognitive psychology to involve more in AI study, stressing that cognitive computing models are at the beginning, with low focus on physiological features, appearance, face, state, head posture and voice, missing the anthropomorphic thinking and behavior of the machine and its interaction with the environment. Even the need for cognition, a highly used construct (Cazan, 2016), might be of interest to be evaluated in relation to artificial human-computer interaction in the future.

The trend of managing the human element inside the AI framework will continue to add more questions on how and what to research and implement inside the human-computer interaction, with relation to individuals personal features, personality, emotions, values, also answering to moral and ethical principles while adapting the device to the human needs not the other way around. Moreover, the debate of all the research trends now cast light on the idea of integrating human values in AI systems, a red line to be resolved and adopted the years to come, targeting human centricity.

CEEOL copyright 2023

CEEOL copyright 2023

3. Discussion

Scientific research is mirrored by technological innovation, aiming to facilitate humans` life. Thus, the main task for scientific research and technological innovation is to bring the benefits of their progress in humans` daily lives and activities. Such practical implications are reached when making prototypes available for the general public.

Ethical concerns also arise when analysing the coexistence of AI and human beings. Multiple embodied intelligent applications that can replace human and social interaction used by psychologists, psychotherapists and psychiatrists already exist and their usage rises ethical concerns (Fiske, Henningsen, & Buyx, 2019). Intelligent or assistive devices are becoming part of the day-by-day interactions for many adults. Intelligent robots are used for medical assistance purposes, in the situation of ill or old people.

Future studies should also focus on ethical and legal matters when it comes to AI development, implementation and use, alongside the research on the potential outcomes of social transformation of work and skills (Gurkaynak, Yilmaz, & Haksever, 2016; Huang & Rust, 2018).

As presented in the previous sections, the overall sectors in which the AI technology reaches, are massive and expanding. Given the numerous trends for research, heterogenous results and various interests of the mentioned studies, it would be rather difficult to build a meta-analytical approach in collecting the data. There is a need for a clear problem specification, a unitary coding scheme and the use of eligible study reports (Popa, 2012, p. 175). Moreover, the social sciences` role is underrepresented so far, both in research and practice. Most existing articles, above mentioned, claim the lack of information and an inconsistent contribution of the psychological and sociological domains. Therefore, a meta-analytical perspective is still difficult to be achieved.

Given the situation, a core idea remains, stating that the human-agent interaction should be a result of a collaboration between the AI area, human-computer interaction area and the social sciences field, working together to refine AI models and applications, as recommended by other researchers also (Miller, 2019, p. 2). Even so, the trends focused primarily on the technical approaches of intelligent systems, automation performance and implementation, which developed fast, leaving the research on the social science lagging behind (Tussyadiah, 2020, p. 11). From this perspective it is vital that future research trends on AI should focus more on integrating the social key elements and approach more the associations with the psychology patterns, social interactions, the emotions dilemma, human centricity, placing actually the human factor as the core element in AI use, while introducing human principles and developing this trend into a vital guide line.

4. Conclusions

The present paper underlines the major findings regarding the development of intelligent systems and use of AI in various sectors of activity, bringing to debate the social sciences approach and need for integrated co-joint research with the technical

CEEOL copyright 2023

CEEOL copyright 2023

field. The technological domain is developing steadily, while the social aspects are left behind in this dynamic matrix, both for research and implementation. Future studies should address these challenges and concerns of how to manage the human factors inside the AI framework, considering personal features, personality dimensions, emotions, interests, values, moral and ethical principles.

Human centricity and social sciences approach should become core elements of interest and principles when researching, building, developing, implementing, training and using AI technologies, adapting such devices to the human needs not the other way around. Research trends should focus more on the life and human implications and potential within a technological process that is already on the rise.

Acknowledgements

This work is supported through the DiFiCIL project (contract no. 69/08.09.2016, ID P_37_771, web: http://dificil.grants.ulbsibiu.ro), co-funded by ERDF through the Competitiveness Operational Programme 2014-2020.

References

Alexander, A., Jiang, A., Ferreira, C., & Zurkiya, D. (2020). An intelligent future for medical imaging: A market outlook on artificial intelligence for medical imaging. *Journal of the American College of Radiology*, *17*(1), 165-170. https://doi.org/10.1016/j.jacr.2019.07.019

Bertram, P., Birtel, M., Quint, F., & Ruskowski, M. (2018). Intelligent manual working station through assistive systems. *IFAC-PapersOnLine*, *51*(11), 170-175. https://doi.org/10.1016/j.ifacol.2018.08.253

Cazan, A.-M. (2016). The factor structure of the short need for cognition scale. *Bulletin of the Transilvania University of Braşov,* Series VII, 9(58), 19-28.

Eckhardt, J., & Rantala, J. (2012). The role of intelligent logistics centres in a multimodal and cost-effective transport system. *Procedia - Social and Behavioral Sciences*, *48*, 612-621. doi: 10.1016/j.sbspro.2012.06.1039

Fiske, A., Henningsen, P., & Buyx, A. (2019). Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *Journal of Medical Internet Research,* I(5), 1-12.

Gattullo, M., Uva, A. E., Fiorentino, M., Scurati, G. W., & Ferrise, F. (2017). From paper manual to AR manual: do we still need text? *Procedia Manufacturing*, *11,* 1303-1310. doi: 10.1016/j.promfg.2017.07.258

Gurkaynak, G., Yilmaz, I., & Haksever, G. (2016). Stifling artificial intelligence: Human perils. *Computer Law & Security Review*, *32*(5), 749–758. http://dx.doi.org/10.1016/j.clsr.2016.05.003

Haleem, A., Javaid, M., & Khan, I. H. (2019). Current status and applications of Artificial Intelligence (AI) in medical field: An overview. *Current Medicine Research and Practice*, *9(*6), 231-237. https://doi.org/10.1016/j.cmrp.2019.11.005

Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research*, *21*(2), 155–172. https://doi.org/10.1177/1094670517752459

CEEOL copyright 2023

CEEOL copyright 2023

Ji, Y., Hwang, J., & Kim, E. Y. (2013). An intelligent wheelchair using situation awareness and obstacle detection. *Procedia - Social and Behavioral Sciences*, *97*(6), 620–628. DOI: 10.1016/j.sbspro.2013.10.281

Keskinbora, K. H. (2019). Medical ethics considerations on artificial intelligence. *Journal of Clinical Neuroscience*, *64*, 277-282. https://doi.org/10.1016/j.jocn.2019.03.001

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, *267,* 1-38. https://doi.org/10.1016/j.artint.2018.07.007

Oungrinis, K. A., Liapi, M., Kelesidi, A., Gargalis, L., Telo, M., Ntzoufras, S., & Paschidi, M. (2014). Intelligent spacecraft module. *Acta Astronautica*, *105(*1), 242-253. http://dx.doi.org/10.1016/j.actaastro.2014.09.010

Pavalache-Ilie, M. (2017). The professional competences – a subject to be questioned, trained and evaluated. *Bulletin of the Transilvania University of Braşov,* Series VII, 10 (59), 19-30.

Pogan, L. (2019). Emerging research milestones for social sciences in a smart society. *"C.S. Nicolăescu-Plopşor" – Anuarul* [Yearbook], *20,* 171–179. Retrieved from https://npissh.ro/wp-content/uploads/2019/12/17.-Livia-Pogan.pdf.

Popa, R. I. (2012). Data management in meta-analytical procedures: myths, realities and biases. *Bulletin of the Transilvania University of Braşov, Series VII, 5* (54), 175-180.

Samsonovich, A. V. (2020). Socially emotional brain-inspired cognitive architecture framework for artificial intelligence. *Cognitive Systems Research*, *60,* 57-76. https://doi.org/10.1016/j.cogsys.2019.12.002

Schwab, K. (2018). *The fourth industrial revolution*. Cologne: World Economic Forum.

Shank, D. B., Graves, C., Gott, A., Gamez, P., & Rodriguez, S. (2019). Feeling our way to machine minds: People's emotions when perceiving mind in artificial intelligence. *Computers in Human Behavior*, *98,* 256-266. https://doi.org/10.1016/j.chb.2019.04.001

Sharma, G. D., Yadav, A., & Chopra, R. (2020). Artificial intelligence and effective governance: A review, critique and research agenda. *Sustainable Futures*, *2,* 1-6. https://doi.org/10.1016/j.sftr.2019.100004

Tussyadiah, I. (2020). A review of research into automation in tourism: Launching the annals of tourism research curated collection on artificial intelligence and robotics in tourism. *Annals of Tourism Research, 81,* 1-13. https://doi.org/10.1016/j.annals.2020.102883

Wang, Z., Xie, L., & Lu, T. (2016). Research progress of artificial psychology and artificial emotion in China. *CAAI Transactions on Intelligence Technology*, *1*(4), 355-365. http://dx.doi.org/10.1016/j.trit.2016.11.003

Zhu, W., Fan, X., & Zhang, Y. (2019). Applications and research trends of digital human models in the manufacturing industry. *Virtual Reality & Intelligent Hardware*, *1*(6), 558-579. DOI: 10.1016/j.vrih.2019.09.005

CEEOL copyright 2023