Exploring the Relationships between Artificial Intelligence Transparency, Sources of Bias, and Types of Rationality

L. Valtonen[1], S.J. Mäkinen[2]
[1]Industrial Engineering and Management, Tampere University, Tampere, Finland
[2]Department of Mechanical and Material Engineering, University of Turku, Finland
(laura.valtonen@tuni.fi)

2022 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM) | 978-1-6654-8687-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/IEEM55944.2022.9989994

Abstract - **Artificial intelligence (AI) is permeating one human endeavor after another. However, there is increasing concern regarding the use of AI: potential biases it contains, as well as mis-judged AI use. This study continues the recent investigations into the biases and issues that are potentially introduced into human decision-making with AI. We experimentally set-up a decision-making classification task and observe human classifiers when they are guided in their decision-making either by AI or other humans. We find that over-reliance or authoritative stigmatization is present when AI is concerned and that with human guidance discursive explanatory decision-making is present. We conclude that while AI is seen as authoritative even in a low stake decision-making setting, it does not suppress choice, but combined with a lack of transparency, AI suppresses visibility into rationality creation by the decision maker. Based on the emergent explorative relationships between types of rationality, AI transparency and authoritativeness, we provide future research avenues based on our findings.**

Keywords - **Artificial intelligence, decision-making, bias, rationality**

I. INTRODUCTION

Artificial intelligence (AI) is often portrayed as a major transforming force for industry (e.g., [1,2]). A clear reason for AI utilization is its competitive promise; Organizations that exploit data quantities typical for AI, or big data analysis, have been more successful than those that do not [3-5]. However, along increasing AI importance, the need for research and dialogue between AI and management and organizational scholars is rising [6], since poor management of the idiosyncrasies associated with the use of big data can lead to failures to meet expectations [7].

Expectations fail even in cases in which the utilized AI approaches are of high accuracy and based on expert data (e.g., [8]). One suggested reason for such shortcomings is that AI is often developed and tested and optimized for accuracy in laboratory settings that do not translate into organizational realities [6]. Another possibility for failed expectations is simply that the expectations were too high to begin with [9]; The *technology effect* refers to the presence of excessive optimism regarding unfamiliar technology being equated with success. This biased thinking permeates experts and novices alike, as well as parties both pitching and investing in technology [10-12]. In reference to algorithms specifically, the term "algorithm appreciation" is used [13], and is observed in AI-augmented decision making, which can lead to failures regarding insufficient risk accounts and uncertainty considerations [14]. In addition to success, people equate novel technology with mystery, alienness, and complexity [12]. The mystery of novel technology is posed as a reason to forego critical analysis and attempts and understanding the technology in question [12,15].

The technology effect is attributed to constant exposure to technology success stories [10,11]. Such constant availability of the connection between technology and success can create biased thinking towards implicit associations between the narrated relationship [16]. In addition to overrepresented narratives of technology successes [11,12], suggested reasons for misplaced optimism and overreliance on AI include narratives of AI as an objective, unbiased, value-free [14,17-20] "supercarrier of formal rationality" [9,18].

Formal rationality is a bureaucratic means-end rationality, which relies on calculations based on universals and abstracts. It is contrasted with substantive rationality, which is based on values and allows for a plurality of rationalization of action in accordance with the actor's values. [21] However, the conceptualization of AI as a formal rationality is problematic; When talking about artificial intelligence and its contemporary successes, usually a supervised machine learning (SML) model is being applied [22,23]. SML algorithms require an originally human labelled dataset, based on which a classifier algorithm is built to map a certain input to its given label. The classifier is evaluated on its performance, or often "accuracy", which refers to how many input items it could correctly label in a certain test dataset.

The SML data can contain biased inputs or labels or both: the data can be incomplete [24] or reflect social bias in the labels (e.g., [25-27]). Moreover, outside of data, bias can creep into AI processes from the algorithms themselves [28], or the interpretations of results [29]. Hence, if AI is conceptualized as a bias-free source of objective information, it transforms what is inherently substantively rational label data created by value-laden, subjective humans into a means-end calculation optimized for a certain outcome. This suppresses substantive rationality and represents it as formal rationality. This project of rationality metamorphosis via AI could lead to the "end of choice" [18] in which we hide possible pluralities of rationalization via AI and risk substantive rationality [18] and the maintenance of unique human knowledge in organizations [30].

In the face of such obstacles and acknowledged issues with AI, the perseverance of optimism regarding this novel information technology [10-11,14] can be explained by information avoidance [9]. Information avoidance refers to active avoidance of information that people are aware that exist and would be free to access. This may be done to “bind one’s own hands while facing an inner conflict” and lead to even the abandonment of responsibility of actions and decisions. Forms of information avoidance include physical avoidance, inattention, biased interpretation of information, forgetting, and self-handicapping. Reasons for information avoidance include disappointment and regret aversion, dissonance avoidance, and optimism maintenance. [31,32]

Dissonance refers to an unpleasant mental state caused by an individual holding contradictory cognitions. The resolution of this feeling is achieved by changing cognitions to a desired logical state [33,34]. The impact of cognitive dissonance is recognized widely to impact management [35]. Dissonance emerges as new information is attained that contradicts previously held cognitions [34], thus, methods of information avoidance can help resolve or mitigate the emergent dissonance. For instance, failures of novel technology can create dissonance with cognitions primed with technology being inherently associated with success. This dissonance may prompt information avoidance regarding issues with the novel technology.

Another relevant cognitive bias is confirmation bias: Confirmation bias emerges when people apply information avoidance in an information search to find confirmatory results of their prior cognitions [29]. This bias also affects collections of people and may lead to what is called “groupthink”: Delusions of optimism due to information avoidance and willful interpretation of information are contagious in groups. In hierarchies these delusion “trickle down” from leaders [36]. Assuming that such leaders may be experts, it is interesting to note that in regarding the technology effect, experts are more prone to polarized beliefs and using their information and intelligence to enforce their prior beliefs [31,37].

We know “what”: Managers overrely on AI. We have literature suggestions as to “why”: the technology effect [10,11], algorithm appreciation [13], and dominant narratives of technology success and superior rationality [12,18]. We even have suggestions for “why” for the “whys”: information avoidance [9,31]. However, the discussion between the posed relationships between all concepts above remains hypothetical: There is scarce if any empirical research on *the reasons why* people come to overrely on AI. We begin to address the question “how does decision-making *reasoning* differ in an AI-guided setting in comparison to a human-guided setting?” We study the differences that contribute to dissimilar, possibly overly optimistic, views on AI guidance in comparison to human guidance.

In our setting people perform a simple decision-making task of categorizing news to predefined categories. After initial categorization, respondents are divided into two groups as the categories are used in decision-making; First group is advised on validating categories by AI and the other group by other people. Both groups are told the source of validation information. We find that for the group guided by their peers, people elaborated on their thinking process rationale, while for the AI guided group, people defended their decisions. Our research directs attention especially to the relationship between AI transparency and creation of decision-making rationales.

II. METHODOLOGY

To address the question “how does decision-making reasoning differ in an AI-guided setting in comparison to a human-guided setting?” we split six participants in a simple, no accountability, low risk decision-making situation into two groups of three: Those who see the recommendation for a decision given by a human agency (human group), and those who see the recommendation given by an SML-based AI (AI group). All participants perform a labelling task individually to form a subjective frame of reference on how the task should be performed, after which they complete the task again with a subset of the data along which they are provided the recommendations. The dataset and labelling were chosen as to hold inherent ambiguity to support various substantive rationalities to emerge on how the task should be done.

A. Data and Labelling

The dataset was collected using the keyword “European Organization for Nuclear Research” (CERN) from LexisNexis. The search was refined to include newswires, press releases, newspapers, and trade press news in English through the years 2016-2019. From the resulting news the sentences including the search term were extracted along with the sentence before and after. This yielded 1687 three sentence text documents.

The set was split into subsets for labelling so that every text is labelled by three separate people into one of the following categories based on CERN’s mission statement: (1) technology, (2) scientific knowledge, or (3) human capital. CERN’s mission statement “Our mission is to: (1) provide a unique range of particle accelerator facilities that enable research at the forefront of human knowledge, (2) perform world-class research in fundamental physics, (3) unite people from all over the world to push the frontiers of science and technology, for the benefit of all” [38] was given to the labelers, who were asked to label each news text according to the part of the mission they see most relevant to the text. The “human recommendation” labels were chosen as the democratic majority out of the three labels provided for each news text.

B. Artificial Intelligence Recommendations

The labelled texts were cleansed by removing extra whitespace and stopwords, punctuation, numerals, URLs, and email addresses were removed with the spaCy library for Python [39]. The corpus of text documents was vectorized for SML with the frequency inverse document frequency (TF-IDF) [40-42] as well as the bag-of-words [43] vectorizers from Python's scikit-learn library [44]. Any rows with not-a-number values were dropped, which yielded altogether 1414 documents for further analysis.

A variety of multi-class classification supervised machine learning algorithms that were suitable for a small text dataset were selected out of ready scikit-learn machine learning algorithms and tested with varied parameters to find the best performing one according to performance on the accuracy and confusion matrix with all 70/30, 85/15, and 90/10 training-test data splits [45]. The best performing algorithm over all runs was the Multilayer Perceptron (MLP) classifier [46] with the following parameters: alpha=1e-05, hidden_layer_sizes=(50, 10), solver="lbfgs". The classifications by the TF-IDF vectorization were superior to those of bag-of-words regarding the confusion matrix and thus those classifications were chosen as the "AI recommendations."

C. Interviews

The participants were surveyed for their attitudes towards AI [47] and split so that each group had a member who saw AI as in an overall positive light, an overall negative light, and a polarized light in the sense that they saw both great risks and possibilities in AI. Both groups had one participant who had experience in algorithm development and one participant who had no information technology background.

Each participant was called in for a structured interview regarding the labeling task they had performed without any knowledge as to what the interview would entail. Upon arrival to the interview the participants were informed that the interview would be recorded and given an instruction to categorize, within an hour, a set of 140 news pieces in a similar manner as previously and explain for each news piece why they selected a certain label. The difference between the groups was that the recommendation was given as either "by the best AI algorithm (from various compared ones) performing the same task" or "the best categorizations from various people performing the same task". The interviews were then analyzed in terms of common and differing themes, and the qualitative results are described below.

III. RESULTS AND DISCUSSION

Overall, all participants employed a ruling-out reasoning logic on several points of both disagreement and agreement. They weighed several label options and chose one through the negation of others, explaining why it is indeed not the other category. Some examples of this type of reasoning were answers such as: "Human capital, because does not directly concern technology facilities, and I don't really see this as science either," and "Here, there is no human mentioned, so scientific knowledge." Though not explicitly mentioned in the latter, the weighing between two categories is present. Every interviewee also employed this reasoning strategy through affirmation instead of negation: They weighed several options and chose a label by affirming some attribute of the news text. Some examples of this type of logic were answers such as: "Well, this could be human capital or technology, but they have purchased technology from elsewhere in particular, so it is human capital," and "Even though the topic is very human-centered, technology, because the facilities are referenced." Thus, no overall indication emerged to suggest that the interviewees were averse to acknowledging the variety of choice they had in the labelling process per text.

However, one interviewee in the AI group expressed that some choices were made against their will, with phrases like "Well I would want to put scientific knowledge, but". No indications were given as to why they could not have chosen according to the expressed want, but this interviewee was one of the two interviewees who assumed that the decision for the recommended label had further or better information as its basis. Out of these two interviewees, the one in the human group used the wording "is apparently" and the one in the AI group used "appears to be". Despite the similar semantic, the difference was in that the interviewee in the AI group assumed that the algorithm had been able to access the whole news text instead of the three-sentence snippet, whereas the interviewee in the human group assumed that the other labelers had more knowledge regarding the subject matter in the text. Neither knew about the functionality of the applied algorithm or the human labelers, and thus, both assumptions were spontaneous.

The lack of transparency into the process of the recommendation generation was what enabled these assumptions to take place. Had the interviewee in the AI group known what data the algorithm had access to, such unfounded assumptions of its capabilities would be less understandable. Similarly, had the interviewee in the human group known the background of the creators of the recommended labels, the assumed difference in substantive knowledge could be traced back to some concrete evidence or information, while now the pessimistic self-assessments are spontaneous and without any referable comparative framework.

All interviewees expressed uncertainty and highlighted their subjectivity with phrases like "seems like", "probably", "maybe", "I feel like", and even "well, I have to guess." Interviewees in both groups exhibited rhetoric aimed at creating rapport with phrases like "I can see that, but" In such cases, however, two interviewees in the AI group used a defensive rhetoric that was not apparent in the interviews of the human group. Both interviewees, in disagreeing with the recommendation, said they were "sticking with" the label they assigned to the text. In similar situations in the human group, the language used did not describe a defensive action, as if their opinion was being challenged in the situation. Instead, they used phrases like "this is rather" and "maybe, this is still." The interviewee in the AI group with knowledge in algorithm development was rhetorically more alike the human group in this aspect.

Both interviewees who used this rhetoric of "sticking with" spontaneously acknowledged discomfort due to seeing the AI recommendations. One said they were "annoyed to see the AI response/answer", while the other interviewee said at the end of the interview that they tried their best not to look at the recommendations, because they thought they may become biased, despite both having very polarized views of AI: The other scored a high positive AI attitude on the preliminary survey, and the other a very negative one in comparison. Despite this difference, the technology effect appeared in both seeming to regard the algorithm recommendation as something authoritative that posed a threat to their subjective opinion. Moreover, one of these interviewees complemented the view of AI as an authority figure by using rhetoric like "I'm leaning on [the AI recommended label], because there is really no sense in this" when they expressed uncertainty regarding the text and came to choose the recommended label.

Within the human group interviewees, excluding the one outstandingly conformist interviewee who chose differently from the recommended label half the time less than the two others, a type of choice justification appeared that was not present in the AI group. Despite every interviewee explaining their earlier methodology to perform the labelling task to a certain degree, when these two interviewees elaborated on their label choice and referred to their own initial conceptualization about the labeling task, they used the original mission statement as a justification. In other words, they referred to instructional documents to base their own subjective framework on, thus performing a similar type of conflation between formal and substantive rationalities as described as a function of AI previously [18]. For example, both "I put all these benefitting organizations into human capital, because no other mission statements describe", and "Patents I put into human capital, because they are not directly concerned with the other mission statements" were offered by one interviewee. The other interviewee elaborated with examples such as "Yeah, here, from the definition of the human capital, the generation of common good is probably fulfilled, so I'm putting this into human capital" and "From the definition, the creation of common good, so I'm putting it there". Thus, this dialogue between the "definition of the mission statement", formal rationality, and "I", substantive rationality, appeared throughout the interviews.

With these results, we may begin to approach empirically why people may come to overrely on AI. The differences in reasoning and rationality in our research between groups decision-making with AI and human recommendations suggest that formal rationality may be seen as inherently already present in the AI assisted decision-making situation, whereas formal rationality was both referenced and actively created within the human-assisted decision-making context. Such spontaneous assumptions of formal rationality to exist for an AI are present in common AI narratives [9, 14, 18-20].

Moreover, the lack of transparency into the creation of the AI recommendation was both an enabling factor for people to make that assumption of rationality, and a deterrent to potentially scrutinizing it. In our setting, the presence of AI suppressed a discussion and synthesis between substantive and formal rationalities: A vacuum regarding information on the decision rationale process development was created. If AI leads to less information shared regarding the decision-making rationale and its development, AI overreliance can be both a result of and a cause of information scarcity – be it serendipitous or self-induced.

IV. CONCLUSION

The emergent defensive rhetoric of "sticking with" and the dialectic creation of rationality within the human group but not the AI group reflects a situation in which the AI is perceived with a different type of authority: It is not negotiated with and there is little reward in explaining your thinking to something that is incapable of seeing your logic or point. Due to the set-up having been with no stakes for the participants, the authority yielded to the AI is notable. In the future, further research into the relationship between uncertainty and AI authority could be furthered by adding stakes into the setting. Moreover, further research requires larger sample sizes, since the small number of interviewees sets clear limitations to this study.

With this initial explorative study, we have been able to tease out details of the relationships between AI transparency, the technology effect, and types of rationality for further empirical research. Our research directs attention especially to the relationship between AI transparency and creation of decision-making rationales. Even is free choice was equally present with AI assistance as human assistance, the visibility into the process for making that choice dissolved with AI. An opaque AI does not offer its own rationale for discussion or scrutiny, and here its presence suppressed elaborations of the way people merged their substantive and given formal rationales for decision-making, furthering the lack of transparency and information in the research setting.

ACKNOWLEDGMENT

We thank the reviewers for the constructive feedback and guidance. The usual disclaimer applies.

REFERENCES

[1] K. Schwab, The fourth industrial revolution. Currency, 2017.
[2] D. Chalmers, N. G. MacKenzie, and S. Carter, "Artificial intelligence and entrepreneurship: implications for venture creation in the fourth industrial revolution," Entrepreneurship Theory and Practice, vol. 45, no. 5, pp. 1028–1053, 2021.
[3] A. McAfee, E. Brynjolfsson, T. H. Davenport, D. J. Patil, and D. Barton, "Big data: the management revolution," Harvard business review, vol. 90, no.10, pp. 60-88, 2012.
[4] G. Cao, and D. Yanqing, "How do top-and bottom-performing companies differ in using business analytics?", Journal of Enterprise Information Management, vol. 30, no. 6, pp. 874–892, 2017.
[5] O. E. Olabode, N. Boso, M. Hultman, and C. N. Leonidou, "Big data analytics capability and market performance: The roles of disruptive business models and competitive intensity", Journal of Business Research, vol. 139, pp. 1218–1230, 2022.

[6] S. Raisch, and S. Krakowski, "Artificial intelligence and management: The automation–augmentation paradox," Academy of Management Review, vol. 46, no. 1, pp. 874–892, 2021.

[7] J. Gao, A. Koronios, and S. Selle. "Towards a process view on critical success factors in big data analytics projects," in Proc. 21sr Americas Conference on Information Systems, AMCIS (2015), Puerto Rico, pp. 1-14.

[8] S. Lebovitz, N. Levina, H. Lifshitz-Assaf, "Is AI ground truth really "true"? The dangers of training and evaluating AI tools based on experts' know-what," Management Information Systems Quarterly, vol. 45, no. 3b, pp. 1501-1525, 2021.

[9] M. Zaitsava, E. Marku, and M.C. Di Guardo, "Is data-driven decision-making driven only by data? When cognition meets data." European Management Journal, in press, 2022.

[10] B. B. Clark, C. Robert, and S.A. Hampton, "The technology effect: how perceptions of technology drive excessive optimism," Journal of Business and Psychology, vol. 31, no.1, pp.87-102, 2016.

[11] T. C. Dunne, B. B. Clark, J.P. Berns, and W.C. McDowell, "The technology bias in entrepreneur-investor negotiations," Journal of Business Research, vol. 105, pp. 258-269, 2016.

[12] K.D. Elsbach, and I. Stigliani, "New information technology and implicit bias," Academy of Management Perspectives, vol. 33, no. 2, pp.185-206, 2019.

[13] J.M. Logg, J.A. Minson, and D.A. Moore, "Algorithm appreciation: People prefer algorithmic to human judgment," Organizational Behavior and Human Decision Processes, vol. 151, pp.90-103, 2019.

[14] C. Keding, and P. Meissner, "Managerial overreliance on AI-augmented decision-making processes: How the use of AI-based advisory systems shapes choice behavior in R&D investment decisions," Technological Forecasting and Social Change, vol. 171, pp. 120970, 2021.

[15] A. Vishwanath, and K.H. LaVail, "The role of attributional judgments when adopted computing technology fails: a comparison of Microsoft Windows PC user perceptions of Windows and Macs," Behaviour & Information Technology, vol. 32, no. 11, pp.1155-1167, 2013.

[16] A. Tversky, and D. Kahneman, "Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty," Science, vol. 185, no. 4157, pp.1124-1131, 1974.

[17] S. S. Sundar, "The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility," in Digital Media,Youth, and Credibility, M.J. Metzger, A.J. Flanagin, Eds., Cambridge: MIT Press, 2008, pp. 73–100.

[18] D. Lindebaum, M. Vesa, and F. Den Hond, "Insights from "The Machine Stops" to Better Understand Rational Assumptions in Algorithmic Decision Making and its Implications for Organizations," Academy of Management Review, vol. 45, no. 1, pp.247-263, 2020.

[19] K. Parry, M. Cohen, and S. Bhattacharya, "Rise of the machines: A critical consideration of automated leadership decision making in organizations," Group & Organization Management, vol. 41, no. 5, pp.571-594, 2016.

[20] K. Brunsson, and N. Brunsson, Decisions: The complexities of individual andorganizational decision-making. Cheltenham, UK: Edward Elgar, 2017.

[21] S. Kalberg, "Max Weber's types of rationality: Cornerstones for the analysis of rationalization processes in history," American journal of sociology, vol. 85, no. 5, pp.1145-1179, 1980.

[22] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. (521), vol. (7553), pp. 436-444, 2015.

[23] T. Rouleau, "What are the types of machine learning," Sama, Retrieved from https://www.sama.com/blog/types-of-machine-learning, 2020.

[24] P. Choudhury, E. Starr, and R. Agarwal, "Machine learning and human capital complementarities: Experimental evidence on bias mitigation", Strategic Management Journal, vol. 41, no. 8, pp.1381-1411, 2020.

[25] R. Binns, M. Van Kleek, M. Veale, U. Lyngs, J. Zhao, and N. Shadbolt, "'It's reducing a human being to a percentage' perceptions of justice in algorithmic decisions," Proceedings of the 2018 Chi conference on human factors in computing systems, pp. 1-14.

[26] J. Dodge, Q. V. Liao, Y. Zhang, R. K. Bellamy, and C. Dugan, "Explaining models: an empirical study of how explanations impact fairness judgment", Proceedings of the 24th international conference on intelligent user interfaces, IUI'2019, pp. 275-285.

[27] M. H.Teodorescu, L. Morse Y. Awwad, and G. C. Kane, "Failures of fairness in automation require a deeper understanding of human-ML augmentation", Management Information Systems Quarterly, vol. 45, no. 3, 2021.

[28] G. M. Johnson, "Algorithmic bias: on the implicit biases of social technology", Synthese, vol. 198, no.10, pp. 9941-9961, 2021.

[29] J. Lallement, S. Dejean, F. Euzéby, and C. Martinez, "The interaction between reputation and information search: Evidence of information avoidance and confirmation bias," Journal of Retailing and Consumer Services, vol. 53, p.101787, 2020.

[30] A. Fügener, J. Grahl, A. Gupta, and W. Ketter, "Will Humans-in-the-Loop Become Borgs? Merits and Pitfalls of Working with AI," Management Information Systems Quarterly, vol. 45, no. 3, 2021.

[31] R. Golman, D. Hagmann, and G. Loewenstein, "Information avoidance," Journal of Economic Literature, vol. 55, no. 1, pp.96-135, 2017.

[32] J. Dana, R.A. Weber, and J.X. Kuang, "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness," Economic Theory, vol. 33, no. 1, pp. 67-80, 2007.

[33] L. Festinger, A theory of cognitive dissonance (Vol. 2). Stanford, CA: Stanford university press, 1957.

[34] H. Bahnmiller, "The Intersections between Self-Deception and Inconsistency: An Examination of Bad Faith and Cognitive Dissonance Hannah Bahnmiller," Stance: An International Undergraduate Philosophy Journal, vol. 8, pp.71-80, 2015.

[35] A.S. Hinojosa, W.L. Gardner, H.J. Walker, C. Cogliser, and D. Gullifor, "A review of cognitive dissonance theory in management research: Opportunities for further development," Journal of Management, vol. 43, no. 1, pp. 170-199, 2017.

[36] R. Bénabou, "Groupthink: Collective delusions in organizations and markets," Review of Economic Studies, vol. 80, no. 2, pp. 429-462, 2013.

[37] D.M. Kahan, E. Peters, M. Wittlin, P. Slovic, L.L. Ouellette, D. Braman, and G. Mandel, "The polarizing impact of science literacy and numeracy on perceived climate change risks," Nature Climate Change, vol. 2, no. 10, pp. 732-735, 2012.

[38] CERN, "Our Mission," https://home.cern/about/who-we-are/our-mission#:~:text=Our%20mission%20is%20to%3A,for%20the%20benefit%20of%20all (accessed May 31, 2022).

[39] M. Honnibal, I. Montani, S. Van Landeghem, and A. Boyd, "spaCy: Industrial-strength Natural Language Processing in Python," Zenodo, 2020, https://doi.org/10.5281/zenodo.1212303.

[40] G. Salton, and C. Buckley, "Term-weighting approaches in automatic text retrieval", Information Processing & Management, vol. 24, no.5, pp. 513–523, 1988.

[41] H. Schütze, C. D. Manning, and P. Raghavan, Introduction to information retrieval (Vol. 39). Cambridge: Cambridge University Press, 2008.

[42] scikit-learn documentation. (n.d.-a). https://scikitlearn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html?highlight=vectorizer#sklearn.feature_extraction.text.TfidfVectorizer%7D

[43] scikit-learn documentation. (n.d.-b). https://scikit-learn.org/stable/tutorial/text_analytics/working_with_text_data.html#tokenizing-text-with-scikit-learn

[44] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[45] scikit-learn documentation. (n.d.-c). https://scikitlearn.org/stable/modules/classes.html#module-sklearn.metrics

[46] scikit-learn documentation. (n.d.-d). https://scikitlearn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html?highlight=mlpclassifier#sklearn.neural_network.MLPClassifier

[47] A. Schepman, and P. Rodway, "Initial validation of the general attitudes towards Artificial Intelligence Scale," Computers in human behavior reports, vol. 1, p.100014, 2020.