Holistic Approach to Design of Generative AI Evaluations: Insights from the Research Onion Model

Manoj Deshpande[1,*], Brian Magerko[1]

[1]Georgia Institute of Technology, Atlanta, GA 30308, USA

Abstract

In this position paper, we propose adopting a holistic approach to research design and evaluation in the rapidly evolving field of generative AI, specifically highlighting the research onion model. We emphasize the importance of an evaluation approach that encompasses not only technical efficiency but also ethical and societal implications. We present five hypothetical scenarios that apply the research onion model to evaluate a co-creative generative AI drawing application, each based on different research philosophies and strategies. While showcasing various methodologies, these scenarios also act as a reflective tool for researchers to identify potential limitations and blindspots in their research approaches. We advocate that the research onion model, or a similar holistic framework, is crucial for the responsible development and deployment of generative AI technologies, ensuring that research design and outcome are robust, relevant, and aligned with broader societal needs.

Keywords

Evaluation Design, Generative AI, Co-creative AI, Research Onion

1. Introduction

In recent years, the emergence of generative AI technologies has profoundly transformed the technological landscape. Pioneering systems like ChatGPT and DALL-E and their contemporaries have emerged as landmarks in this domain, exemplifying the profound capabilities of generative models [1, 2, 3]. These advancements have not only revolutionized technical fields but also permeated popular culture, signaling a paradigm shift [4] in how we interact with and perceive AI technology, bringing with it new possibilities and complexities.

The evolution of AI evaluation, particularly in generative AI, reflects a shift similar to that in Human-Computer Interaction (HCI). HCI moved from a focus on cognitive, positivist performance metrics to situated, contextual approaches in its third wave [5]. This shift, initially marked by quantitative methods and later evolving towards more contextual research as highlighted by Suchman [6] and Dourish [7], is mirrored in AI evaluation. The field is transitioning from purely quantitative metrics to more diverse, pluralistic methodologies that acknowledge the complex nature and broader impacts of AI technology [8, 9, 10].

The evolving landscape of AI evaluation highlights the need for AI systems designed with human needs and societal values in mind, emphasizing the creation of human-centered AI that

Joint Proceedings of the ACM IUI Workshops 2024, March 18-21, 2024, Greenville, South Carolina, USA

*Corresponding author.

manojdeshpande@gatech.edu (M. Deshpande); magerko@gatech.edu (B. Magerko)

https://www.manoj-deshpande.com/ (M. Deshpande); https://expressivemachinery.gatech.edu/ (B. Magerko)

0009-0004-1307-2553 (M. Deshpande); 0000-0003-1900-4020 (B. Magerko)

enhances human capabilities and builds trust and reliability [11, 12]. Evaluating generative AI presents challenges beyond standard performance metrics, requiring consideration of technical proficiency, ethics, and societal impact [13, 14]. This complexity calls for a nuanced, comprehensive evaluation design, enabling researchers to address the diverse aspects of AI technologies and align them with societal needs while recognizing the limitations of their research methods.

Against this backdrop, this position paper proposes adopting Saunders' research onion model [15] as an effective framework for structuring the evaluation of generative AI. Initially developed as a comprehensive guide for research methodologies in business, this model has found relevance and application in various fields. The research onion model presents a structured approach that is particularly appropriate for addressing the multi-layered complexity inherent in the evaluation of generative AI. The research onion model systematically breaks down the evaluation process into several layers: research philosophies, approaches, strategies, choices, time horizons, and techniques. This structure provides a versatile toolkit for researchers, enabling them to carefully navigate through and balance the intricacies of potential biases and blindspots inherent in evaluating generative AI systems.

The significance of our proposition is in guiding research design towards responsible AI development and deployment. Using the research onion model, we provide a comprehensive framework ensuring generative AI research is self-reflective, relevant, and in tune with both technological progress and societal needs. The paper is structured to first introduce the research onion model, then make a case for holistic AI research design, followed by demonstrating its application in five hypothetical research scenarios.

2. Research Onion Model: An Overview

Saunders' research onion model [15] presents a structured framework for designing and conducting research as shown in Figure-1.

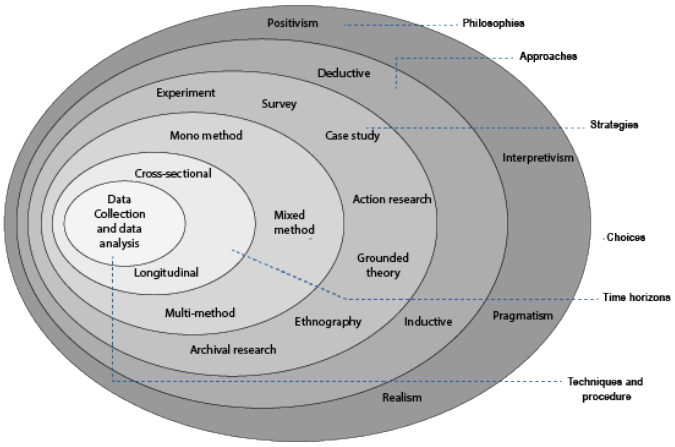

Figure 1: Saunders' research onion model [15]

The research onion model's outermost layer comprises research philosophies that shape fundamental beliefs about knowledge, including positivism with its focus on quantifiable facts,

interpretivism emphasizing subjective experiences, pragmatism blending these elements, and realism recognizing an independent reality shaped by perceptions [16]. Subsequent layers define the research approach, categorizing into deductive methods for testing theories and inductive methods for developing theories from data. The model further includes layers for research strategies (like experiments and surveys), method choices (mono-method, mixed-method, or multi-method), time horizons (cross-sectional for immediate analysis or longitudinal for long-term trends), and specific data collection and analysis procedures, each integral to structuring and guiding the research process.

The research onion model's layered structure is well-suited for the complex nature of generative AI research, guiding researchers to methodically address every aspect of their study, from philosophical foundations to data collection and analysis. This holistic approach is vital for a thorough evaluation of AI technologies, ensuring research is robust, relevant, and attuned to the complex relationship between technology and society.

3. The Need for a Holistic Approach

Evaluating generative AI requires moving beyond traditional metrics like precision and efficiency, which often miss crucial aspects such as ethical implications, user-centric design, societal impact, and the subtleties of human-AI interaction. Such a limited focus can skew the understanding and application of generative AI, overlooking its broader human and societal effects. A more nuanced approach, encompassing ethical, societal, and human-centric dimensions, is necessary, as emphasized by Bender et al. [17] and Crawford [18], who highlight the need to be aware of biases and the wider implications of AI. The research onion model offers a structured framework to help researchers identify and address these potential blindspots, ensuring a comprehensive and reflective design in AI research.

Furthermore, the Research Onion model's focus on diverse research philosophies prompts evaluators to critically examine their assumptions, which is crucial in uncovering inherent biases and limitations in their research approach. For example, a positivist approach may neglect the subjective experiences and societal impacts central to interpretivist research. This gap becomes particularly relevant in discussions of AI ethics and fairness. The satirical critique "A Mulching Proposal" [19] illustrates the inadequacies of applying highly positivist research methods when addressing the normative implications of technologies.

The model promotes diversity in research methodologies, crucial for comprehensive AI evaluation. It advocates for a mix of quantitative and qualitative methods, leading to a multi-disciplinary approach. This aligns with early AI scholars like Weizenbaum, who emphasized moral and ethical considerations [20], and Simon, who focused on the broader implications of computational systems [21]. Furthermore, from a design perspective, Friedman's concept of value-sensitive design emphasizes the importance of integrating human values into technology design and evaluation [22], supporting a balanced and holistic evaluation methodology.

4. Hypothetical Research Scenarios

In this section, we present the application of the research onion model through a series of five hypothetical scenarios designed to evaluate a co-creative generative AI drawing application. These scenarios also serve as a form of self-reflection, drawing from our previous research experiences. Four of these scenarios are grounded in different research philosophies - positivism, interpretivism, pragmatism, and realism. The fifth scenario, while maintaining the research philosophy of the first (positivism), shifts to using a design research strategy. These scenarios demonstrate how different layers of the model can inform diverse research approaches in varied contexts. At the same time, sensitize researchers to the limits and blindspots of their research approach. The hope is that by presenting a range of scenarios, each with its focus and methodology, we can emphasize the importance of critical self-reflection in research.

4.1. Scenario 1: Evaluation of AI-Assisted Drawing Efficiency

4.1.1. Approach

In this scenario, we adopt a deductive methodology, grounded in the principles of positivism, to test the hypothesis that AI assistance significantly enhances the efficiency of drawing processes. The hypothesis is that the integration of AI in artistic creation speeds up the drawing process and improves the final output's quality.

4.1.2. Strategy

To test this hypothesis, we implement a series of controlled experiments involving a diverse sample population of visual artists with varying levels of experience, backgrounds, and familiarity with AI tools. Participants are tasked with sketching an everyday object in two scenarios: one with the assistance of the generative AI drawing tool and one without it. The aim is to create a controlled environment where the direct impact of AI assistance on the drawing process can be isolated and measured.

4.1.3. Data Collection and Analysis

The data collection in this study primarily focuses on quantitative metrics: *time efficiency*, measuring the duration taken by participants to complete drawings with and without AI assistance; *precision of lines*, analyzing the accuracy and quality of visual elements in the artwork using software tools or expert evaluations; and *error rates*, counting corrections made during the drawing process to indicate AI's role in reducing revisions. These metrics, collected and analyzed statistically, aim to quantify any improvements in efficiency and artistic quality brought about by AI assistance.

4.1.4. Insight and Implications

This scenario aims to objectively evaluate the impact of AI on drawing efficiency, providing empirical data to enhance understanding of AI's role in augmenting human creativity. It offers insights into the tangible benefits of AI in artistic processes, informing both developers

and artists and encourages data-driven development of AI tools that balance technological capabilities with creative needs. Overall, it highlights the value of a positivist, quantitative approach in assessing AI's utility in creative domains, showcasing how empirical data can drive the design of AI technologies in art.

4.1.5. Potential Limitations and Blindspots

- **Quantitative Focus:** May neglect the creative process's qualitative, subjective elements, like the artist's experience and emotional connection.
- **Controlled Experiment Setting:** Could fail to mimic real-world artistic scenarios, affecting the findings' applicability.
- **Artist Diversity and Bias:** Lack of diversity in participant skills and AI familiarity might not reflect the wider artistic community.
- **Ethical and Social Implications:** Overlooks important ethical considerations such as originality and the cultural influence of AI art.
- **Human-AI Interaction Dynamics:** Potentially underestimates the changing nature of artist-AI interactions and their impact on art creation.

4.2. Scenario 2: Understanding User Experiences and Expression of AI-Assisted Drawing

4.2.1. Approach

In this scenario, we adopt an interpretivist approach to delve into the subjective experiences of artists using a co-creative generative AI drawing tool. The aim is to build theories grounded in the real-life experiences of users, emphasizing the personal and emotional aspects of the creative process. This inductive approach acknowledges the diverse and individualized nature of artistic creation and how it is influenced by the integration of AI.

4.2.2. Strategy

The core strategy involves conducting in-depth interviews and observational studies with artists who use the AI drawing tool. Artists from various backgrounds and with different levels of expertise are selected to ensure a wide range of perspectives. Observational studies are conducted in natural settings, where artists interact with the AI tool as part of their regular creative process. This allows for capturing authentic user experiences and interactions with the technology.

4.2.3. Data Collection and Analysis

In this study, artists are *interviewed* to gather insights on their experiences with the AI tool, focusing on its influence on their creative process, artistic choices, and satisfaction with their artwork. Special attention is given to *emotional responses* during the creative process, like frustration or joy, to understand AI's emotional impact on creativity. *Artists also reflect* on how the AI tool affects their artistic expression, noting any enhancements or limitations. *Observational*

notes complement this, with researchers taking detailed notes on artists' interactions with the tool and their workflow. The qualitative data collected is then subjected to thematic analysis to identify common themes and patterns in artists' experiences and perceptions.

4.2.4. Insights and Implications

This scenario offers valuable insights into the interpretive dimension of AI in art creation, exploring artists' subjective experiences to reveal how AI tools can both enhance and potentially limit creativity. These findings are crucial for AI tool developers and designers, guiding tool alignment with artists' creative needs and shaping the broader understanding of AI's role in art, not just as a tool but as a creative collaborator. Overall, it presents a comprehensive, human-centric perspective on AI's impact in art, emphasizing the significance of user experiences in developing AI tools that are technically proficient, emotionally engaging, and creatively enriching.

4.2.5. Potential Limitations and Blindspots

- **Subjective Bias:** Interpretivist methods and self-reported artist experiences could introduce biases, limiting the findings' applicability.
- **Generalizability Issues:** Individualized experiences may not be representative of all artists or user groups.
- **Emotional Response Interpretation:** Variations in interpreting emotional responses can lead to inconsistent conclusions.
- **Technical Aspect Neglect:** Emphasis on personal experiences may overlook the AI tool's technical capabilities and limitations.
- **Documentation Quality:** Observational note quality relies on researchers' skills, potentially leading to biased or incomplete process documentation.

4.3. Scenario 3: Balancing Efficiency and Creativity in AI-Assisted Art

4.3.1. Approach

We adopt a pragmatist stance in this scenario, utilizing a mixed-methods approach to evaluate a generative AI drawing tool. This approach aims to balance the objective measurements of the tool's efficiency with a nuanced understanding of its impact on creative expression. By combining quantitative and qualitative research methods, the goal is to provide a well-rounded analysis that respects both the technical aspects of the AI tool and the subjective experiences of its users.

4.3.2. Strategy

The research employs user surveys and focus groups, targeting visual artists of various experience levels and AI tool familiarity. Surveys measure metrics like time efficiency and usability, yielding quantifiable data on the AI tool's performance. Concurrently, focus groups offer in-depth qualitative insights, exploring the AI tool's impact on artists' creative processes, inspiration, and satisfaction.

4.3.3. Data Collection and Analysis

In this study, a mixed-method approach is used for analysis: *quantitative analysis* involves statistical examination of survey data to objectively measure the AI tool's efficiency and usability, while *qualitative analysis* through thematic analysis of focus group discussions reveals the tool's impact on creativity, including its effects on artistic choices and workflow. The final step *integrates these findings*, combining quantitative and qualitative insights to offer a comprehensive understanding of how the AI tool's technical capabilities influence creative processes and artistic outputs.

4.3.4. Insights and Implications

This scenario reveals key insights for optimizing generative AI tools to enhance both efficiency and creativity in art, employing a mixed-methods approach that balances technical performance with subjective experiences. These insights are beneficial for AI tool developers and the artistic community, informing the design of future tools to improve efficiency and creatively enrich the process, positioning technology as an enabler of artistic innovation. Overall, the scenario highlights the effectiveness of a mixed-methods approach in AI tool evaluation, emphasizing the need to consider both technical and creative aspects for the development of efficient, user-friendly, and creatively expressive tools.

4.3.5. Potential Limitations and Blindspots

- **Data Balance Challenge:** Difficulties in equally integrating and weighing quantitative and qualitative data may result in evaluation imbalances.
- **Limited Generalizability:** Insights from specific artist groups may not be applicable to all AI tool users.
- **Selection Bias Risk:** Participant choice for surveys and focus groups might not adequately represent the wider artistic community.
- **Inadequate Technical Analysis:** The AI tool's technical evaluation may lack the depth needed to fully understand its capabilities and limitations.
- **Subjective Creative Impact:** Assessing the AI tool's effect on artistic creativity and satisfaction is subjective and varies among artists, complicating definitive conclusions.

4.4. Scenario 4: Evaluating Ownership, Authenticity, and the Ethics of Artistic Collaboration

4.4.1. Approach

In this scenario, we adopt a realism-based approach to explore critical aspects of ownership, authenticity, and the ethics involved in artistic collaboration with an AI agent. The goal is to assess how artists perceive the role and contributions of AI in the creative process, especially in terms of ownership rights, the authenticity of the collaborative work, and the ethical implications of such partnerships.

4.4.2. Strategy

The strategy involves a comparative analysis and in-depth qualitative research. During the comparative analysis, visual artists of various experience levels and AI tool familiarity create artwork independently and in collaboration with the AI tool. This comparison aims to highlight differences in perceived ownership and authenticity between AI-assisted and traditional artworks. For the qualitative research, the idea is to conduct in-depth interviews and focus groups with artists who have used the AI tool. The discussions center around their feelings regarding the ownership of the artwork, the authenticity of the creative process when involving AI, and ethical considerations in such collaborations.

4.4.3. Data Collection and Analysis

This study explores artists' *perceptions of ownership* in collaborative artworks involving significant AI contributions, assessing their *views on authenticity* and the originality of AI-partnered creations. It also delves into *ethical considerations*, like fair recognition of the AI's role and intellectual property concerns, alongside moral implications of AI usage in art. Additionally, a *comparative evaluation of artworks* is conducted, analyzing stylistic differences, thematic depth, and overall quality between pieces created with AI assistance and those made independently.

4.4.4. Insights and Implications

This scenario delves into the complexities of AI-assisted artistic creation by exploring perceptions of ownership, authenticity, and ethics, providing insights into the subtleties of human-AI collaboration in the arts. These findings are vital for artists, AI developers, and policymakers to navigate the ethical landscape of AI in creative fields. They inform the development of AI tools that uphold artistic integrity and ethical standards, contributing to the discourse on AI's role and implications in creative practices. Overall, the scenario emphasizes the importance of considering ownership, authenticity, and ethics in AI collaborations, advocating for the responsible and ethically aligned development and use of AI tools in artistic endeavors.

4.4.5. Potential Limitations and Blindspots

- **Subjective Interpretations:** Varying artist perceptions of ownership and authenticity lead to subjective assessments.
- **Generalizability Limitations:** Results from certain artist groups may not be broadly applicable across diverse artistic contexts.
- **Ethical Complexity:** The multifaceted ethical dimensions of AI in art might not be fully explored in interviews and focus groups.
- **Limited Legal Analysis:** The scenario's ethical focus may overlook the intricacies of legal issues like intellectual property in AI art.
- **Risk of Negative Bias:** Concentrating on ethical and authenticity issues could overshadow AI collaboration's positive impacts in art.

4.5. Scenario 5: Design Research for Evaluating AI-Assisted Drawing Efficiency

4.5.1. Approach

We employ design research methods in this scenario to explore how AI assistance impacts the drawing process. Similar to scenario 1, The focus is on understanding the efficacy of AI tools in enhancing the drawing experience but from a design research perspective.

4.5.2. Strategy

The research employs a series of user studies where visual artists of various experience levels and AI tool familiarity engage with the AI drawing tool. Participants are asked to create artwork under varying conditions: using the AI tool, without it, and in a collaborative mode with the AI. This approach allows for the observation of interactions between the artist and the AI tool, providing insights into the user experience and the design effectiveness of the tool.

4.5.3. Data Collection and Analysis

Data collection is done using multiple methods: *User experience surveys* are conducted post-drawing sessions to assess usability, satisfaction, and perceived efficiency of the AI tool. *Observational studies* allow researchers to note interactions with the AI tool, workflow patterns, and any challenges or advantages during the drawing process. The quality of the artworks produced is evaluated through expert design reviews, focusing on creativity, complexity, and aesthetic appeal, and compiled into an *annotated portfolio.* Additionally, *time metrics* are recorded for each artwork, providing quantitative data to complement the qualitative insights.

4.5.4. Insight and Implications

This scenario employs a design research approach to evaluate the effectiveness of AI in art creation, blending qualitative and quantitative data for a comprehensive view of AI's enhancement of the artistic process, with a focus on user experience and creative output. The insights aim to guide the development of intuitive, user-friendly AI tools tailored to artists' preferences, highlighting the importance of user-centered design. This approach showcases the utility of design research methods in artistic AI tool evaluation, emphasizing the significance of user experience in the development of effective AI tools that resonate with artists' needs.

4.5.5. Potential Limitations and Blindspots

- **User Experience Subjectivity:** Surveys and observational studies bring subjectivity due to varying individual perceptions.
- **Limited Generalizability:** Findings from certain artist groups may not reflect the experiences of all AI tool users.
- A**rtwork Evaluation Bias:** Subjective assessments of artwork quality can be influenced by reviewer or artist biases.

- **Insufficient Quantitative Analysis:** The scenario's limited quantitative measures might not fully encompass all aspects of drawing efficiency.
- **Overlooking Technical Aspects:** A focus on user experience may neglect the AI tool's technical limitations and capabilities.

4.6. Summary

Table- 1 provides a comparative summary of five hypothetical scenarios corresponding to different layers of the research onion model while also identifying potential research blindspots.

Table 1
Overview of hypothetical scenarios

	Scenario 1	**Scenario 2**	**Scenario 3**	**Scenario 4**	**Scenario 5**
Research Philosophy	Positivism	Interpretivism	Pragmatism	Realism	Positivism
Approach	Deductive	Inductive	Mixed (Inductive & Deductive)	Mixed (Inductive & Deductive)	Deductive
Strategy	Controlled experiments	Thematic analysis	Surveys and thematic analysis	Ablation study and thematic analysis	Controlled experiments
Method Choices	Mono-method	Mono-method	Mixed-method	Multi-method	Multi-method
Time Horizon	Cross-sectional	Cross-sectional	Cross-sectional	Cross-sectional	Cross-sectional
Data Collection and Analysis	Quantitative metrics: time efficiency, precision, error rates	Artist interviews on emotional responses and artistic expression perception	Quantitative survey results, qualitative focus group themes	Ownership perceptions, authenticity assessments, ethical considerations	Surveys, observational studies, annotated portfolios
Potential Research Blindspots	May neglect qualitative aspects and real-world applicability	Subject to subjective bias and limited generalizability	Challenges in data integration and cohesiveness	Subjectivity in perceptions, ethical complexity, and generalizability issues	Potential overlook of technical applicability

5. Conclusion

In this position paper, we emphasized the importance of a holistic approach in the research and evaluation of generative AI, exemplified by the research onion model. This approach goes beyond technical efficiency to include ethical and societal considerations, reflecting a paradigm shift similar to that in Human-Computer Interaction. Through five hypothetical scenarios, we showcased the research onion model's versatility in addressing generative AI evaluation's diverse aspects and highlighted its utility in identifying potential research limitations and blindspots. Such a holistic approach is essential for understanding the complexities of rapidly evolving generative AI technologies. Future work could focus on applying this model to real case studies and developing more refined tools specifically for AI evaluation. Our advocacy for such frameworks aims to guide the responsible advancement of generative AI, aligning technological progress with human needs and societal values.

References

[1] J. Betker, G. Goh, L. Jing, T. Brooks, J. Wang, L. Li, L. Ouyang, J. Zhuang, J. Lee, Y. Guo, W. Manassra, P. Dhariwal, C. Chu, Y. Jiao, A. Ramesh, Improving Image Generation with Better Captions (2023).

[2] I. Goodfellow, Y. Bengio, A. Courville, Chapter-20, Deep Generative Models, in: Deep learning, MIT press, 2016, pp. 651–716. URL: https://www.deeplearningbook.org/.

[3] E. A. OpenAI, GPT-4 Technical Report, 2023. URL: http://arxiv.org/abs/2303.08774. doi:10.48550/arXiv.2303.08774, arXiv:2303.08774 [cs].

[4] T. S. Kuhn, The structure of scientific revolutions, University of Chicago press, 2012. URL: https://books.google.com/books?hl=en&lr=&id=3eP5Y_OOuzwC&oi=fnd&pg=PR5&dq=structure+of+scientific+revolutions&ots=xXXOzcjKoJ&sig=BaA4BOQXYYk6cOgsIv5qQVIuYgw.

[5] Y. Rogers, HCI theory: classical, modern, and contemporary, Synthesis lectures on human-centered informatics 5 (2012) 1–129. Publisher: Morgan & Claypool Publishers.

[6] L. Suchman, L. A. Suchman, Human-machine reconfigurations: Plans and situated actions, Cambridge university press, 2007.

[7] P. Dourish, Where the action is: the foundations of embodied interaction, MIT press, 2004.

[8] J. Auernhammer, Human-centered AI: The role of Human-centered Design Research in the development of AI (2020). URL: https://dl.designresearchsociety.org/drs-conference-papers/drs2020/researchpapers/89/.

[9] F. Fui-Hoon Nah, R. Zheng, J. Cai, K. Siau, L. Chen, Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration, Journal of Information Technology Case and Application Research 25 (2023) 277–304. URL: https://www.tandfonline.com/doi/full/10.1080/15228053.2023.2233814. doi:10.1080/15228053.2023.2233814.

[10] M. Muller, J. D. Weisz, W. Geyer, Mixed initiative generative AI interfaces: An analytic framework for generative AI applications, in: Proceedings of the Workshop The Future of Co-Creative Systems-A Workshop on Human-Computer Co-Creativity of the 11th International Conference on Computational Creativity (ICCC 2020), 2020. URL: https://computationalcreativity.net/workshops/cocreative-iccc20/papers/Future_of_co-creative_systems_185.pdf.

[11] M. O. Riedl, Human-centered artificial intelligence and machine learning, Human Behavior and Emerging Technologies 1 (2019) 33–36. URL: https://onlinelibrary.wiley.com/doi/10.1002/hbe2.117. doi:10.1002/hbe2.117.

[12] B. Shneiderman, Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy, International Journal of Human–Computer Interaction 36 (2020) 495–504. URL: https://www.tandfonline.com/doi/full/10.1080/10447318.2020.1741118. doi:10.1080/10447318.2020.1741118.

[13] C. O'neil, Weapons of math destruction: How big data increases inequality and threatens democracy, Crown, 2017. URL: https://books.google.com/books?hl=en&lr=&id=cbwvDwAAQBAJ&oi=fnd&pg=PA1&dq=Weapons+of+Math+Destruction:+How+Big+Data+Increases+Inequality+and+Threatens+Democracy&ots=XlvbN1LwVf&sig=V_A7hwC2e97DJmFBQnG8lLHdQNs.

[14] S. Russell, P. Norvig, Philosophy, ethics, and safety of AI, in: Artificial Intelligence: A

Modern Approach, 4th ed., 2022, pp. 1032–1062.

[15] M. Saunders, P. Lewis, A. Thornhill, Research methods for business students, Pearson education, 2009.

[16] P. Žukauskas, J. Vveinhardt, R. Andriukaitienė, Philosophy and paradigm of scientific research, Management culture and corporate social responsibility 121 (2018) 506–518. URL: https://books.google.com/books?hl=en&lr=&id=UMaPDwAAQBAJ&oi=fnd&pg=PA121&dq=Philosophy+and+Paradigm+of+Scientific+Research&ots=pHZLaLTlMB&sig=b7OhNs-ts_FYa0DYIt7khX-HEa8, publisher: IntechOpen London, UK.

[17] E. M. Bender, T. Gebru, A. McMillan-Major, S. Shmitchell, On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? ��, in: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, ACM, Virtual Event Canada, 2021, pp. 610–623. URL: https://dl.acm.org/doi/10.1145/3442188.3445922. doi:10.1145/3442188.3445922.

[18] K. Crawford, The atlas of AI: Power, politics, and the planetary costs of artificial intelligence, Yale University Press, 2021. URL: https://books.google.com/books?hl=en&lr=&id=XvEdEAAAQBAJ&oi=fnd&pg=PP1&dq=atlas+of+AI&ots=MpHzLq5UAz&sig=cRSk924Eq86np0wahxmJGvxapKQ.

[19] O. Keyes, J. Hutson, M. Durbin, A Mulching Proposal: Analysing and Improving an Algorithmic System for Turning the Elderly into High-Nutrient Slurry, in: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, ACM, Glasgow Scotland Uk, 2019, pp. 1–11. URL: https://dl.acm.org/doi/10.1145/3290607.3310433. doi:10.1145/3290607.3310433.

[20] J. Weizenbaum, Computer power and human reason: From judgment to calculation. (1976). URL: https://psycnet.apa.org/record/1976-11270-000, publisher: WH Freeman & Co.

[21] H. A. Simon, The sciences of the artificial, MIT press, 1996. URL: https://books.google.com/books?hl=en&lr=&id=k5Sr0nFw7psC&oi=fnd&pg=PR9&dq=The+Sciences+of+the+Artificial&ots=-yXKiJDJCt&sig=Q7uc_FwLq197ojwu46VA0ff6S48.

[22] B. Friedman, P. H. Kahn, A. Borning, A. Huldtgren, Value Sensitive Design and Information Systems, in: N. Doorn, D. Schuurbiers, I. Van De Poel, M. E. Gorman (Eds.), Early engagement and new technologies: Opening up the laboratory, volume 16, Springer Netherlands, Dordrecht, 2013, pp. 55–95. URL: http://link.springer.com/10.1007/978-94-007-7844-3_4. doi:10.1007/978-94-007-7844-3_4, series Title: Philosophy of Engineering and Technology.