PORTFOLIO

SEMINARS AND FIELD STUDY

1st Semester, A.Y. 2023 - 2024

JENO D. BELLIDO
BS COMPUTER SCIENCE

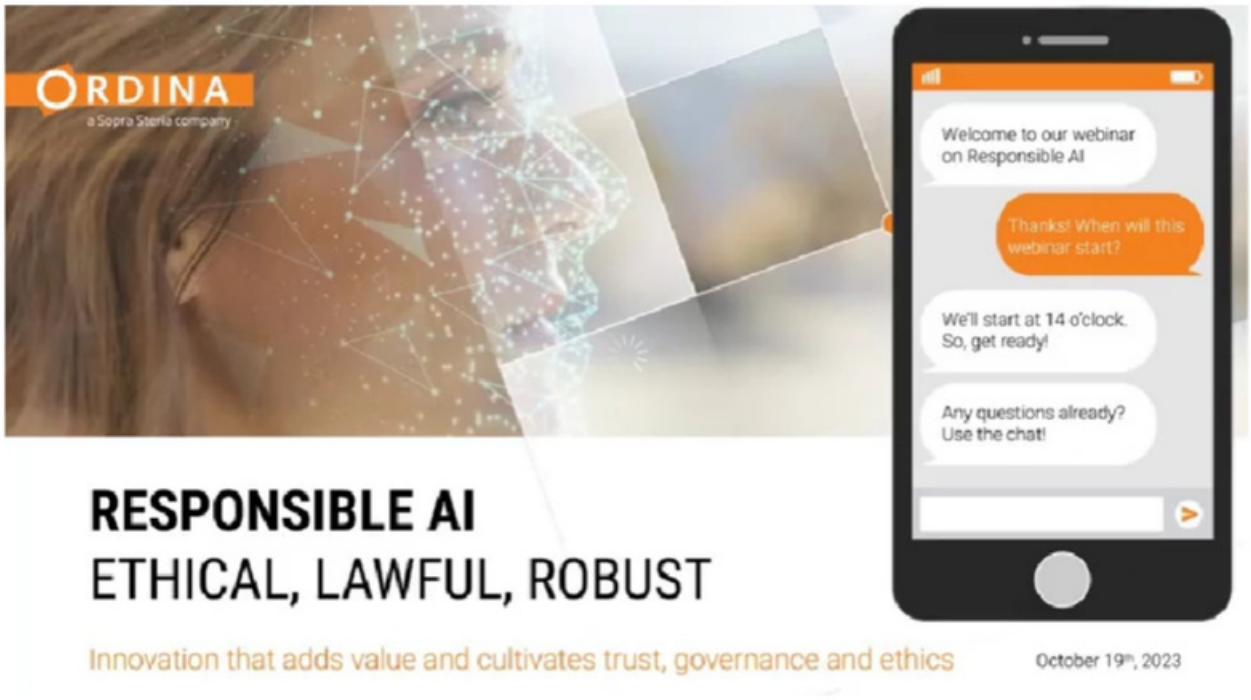
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY

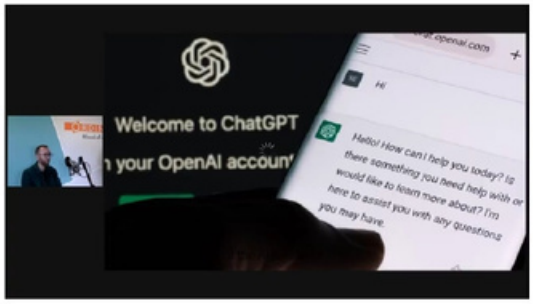
Responsible AI: Ethical, Lawful, Robust

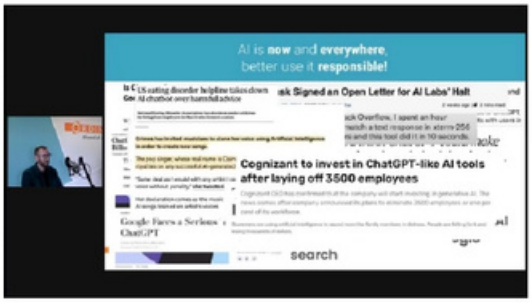
(A Reaction Paper)

October 19, 2023

Speaker: Mr. Harold Selman
Solution Lead Data Science







Discussion

The "Responsible AI" webinar offered a comprehensive understanding of the pervasive nature of AI, emphasizing the need for its responsible use. It highlighted that while AI won't replace humans, those equipped with AI will supersede those without, underscoring the importance of AI adoption. Furthermore, it introduced the concept of Language Models (LLMs) as a form of Generative AI, capable of producing human-like text, indicating the transformative potential of AI. These insights collectively stress the significance of AI and the necessity for its ethical application.

In our recent webinar, the first salient point underscored the pervasive presence of Artificial Intelligence (AI) in our contemporary landscape, urging a collective emphasis on responsible utilization. As the discourse unfolded, participants delved into the omnipresence of AI, highlighting its integration into diverse facets of our lives, from daily routines to complex decision-making processes. The prevailing sentiment echoed the need for a conscientious approach, recognizing the profound impact AI wields on society. The discussions emphasized the imperative to harness the power of AI ethically, considering its far-reaching implications on privacy, equity, and societal well-being. The consensus among the webinar attendees was clear: as AI continues to permeate our world, it is paramount to navigate its integration with a deliberate commitment to responsible and ethical practices.

The second salient point illuminated a nuanced perspective on the relationship between Artificial Intelligence (AI) and human involvement. The discussion centered on the assertion that AI is not positioned to replace humans, but rather, it is the integration of humans with AI that will redefine the landscape. Attendees explored the symbiotic potential of human-AI collaboration, recognizing that the augmentation provided by AI technologies enhances human capabilities rather than supplanting them. The discourse emphasized that individuals and organizations equipped with AI tools possess a competitive advantage, amplifying efficiency, innovation, and problem-solving capacities. The prevailing sentiment underscored the need for a paradigm shift – a departure from the notion of AI as a threat and a transition toward a collaborative future where humans leverage AI as a powerful ally in shaping a more advanced and interconnected society. The webinar highlighted the imperative for individuals and industries to embrace this evolution, acknowledging that those who adeptly integrate AI into their workflows are poised to outpace their counterparts who do not.

In our recent webinar, the third salient point brought into focus the transformative capabilities of Generative AI, particularly Large Language Models (LLMs). The discussions centered on how Generative AI, exemplified by LLMs, has become a pivotal force in reshaping how we interact with and generate content. Attendees delved into the profound impact of LLMs on natural language processing, content creation, and information synthesis. The discourse underscored the potential of Generative AI to revolutionize various industries, from creative writing to software development, by autonomously generating human-like text. Participants

explored the ethical considerations and challenges associated with the deployment of LLMs, emphasizing the need for responsible development and use. The webinar fostered a deeper understanding of how Generative AI is not merely a technological advancement but a catalyst for reimagining human-machine collaboration, paving the way for innovative applications that extend beyond our current conceptual boundaries.

Supplemental Research

Henz (2021) discusses the ethical and legal responsibility for AI, emphasizing that until regulations can catch up with the rapid development of AI, ethics must guide society. This resonates with the webinar's emphasis on the need for a conscientious approach to AI. Similarly, Ryan et al. (2021) investigate the ethical use of Big Data and AI technologies, highlighting the overlap and correlation between organizations' ethical concerns. This aligns with the discussion on the far-reaching implications of AI on privacy, equity, and societal well-being. Furthermore, a study by Khan et al. (2021) presents a systematic review of the principles and challenges in the ethics of AI, echoing the webinar's consensus on the importance of navigating AI integration with a commitment to responsible and ethical practices.

Leitão, Saleiro, Figueiredo, and Bizarro (2022) discuss the concept of Human-AI collaboration (HAIC) in decision-making, aiming to create synergistic teaming between human decision-makers and AI systems. This aligns with the webinar's discussion on the symbiotic potential of human-AI collaboration. Abedin et al. (2022) delve into the social aspects of interactions between humans and AI systems, emphasizing the need for a paradigm shift towards a collaborative future. This resonates with the webinar's sentiment on the need for a departure from the notion of AI as a threat. Furthermore, Raftopoulos and Hamari (2023) conducted a literature review on enabling value creation through Human-AI collaboration in organizations, which aligns with the webinar's emphasis on the competitive advantage gained by individuals and organizations equipped with AI tools.

Yu et al. (2023) discuss the integration of Generative AI and Large Language Models (LLMs) into healthcare and medicine, emphasizing the need for an inclusive, collaborative co-design process that engages all pertinent stakeholders. This aligns with the webinar's discussion on the profound impact of LLMs on natural language processing, content creation, and information synthesis. Kucharavy et al. (2023) delve into the principles, abilities, limitations, and future prospects of Generative Language Models, especially in the context of cyber-defense. This resonates with the webinar's sentiment on the potential of Generative AI to revolutionize various industries. Furthermore, Zhao et al. (2023) conducted a literature review of the recent advances in LLMs from four major aspects, including pre-training, adaptation tuning, utilization, and capability evaluation, which aligns with your webinar's emphasis on the transformative capabilities of Generative AI.

Summary of Learning

The "Responsible AI" webinar provided a comprehensive exploration of key themes shaping the AI landscape. Firstly, it underscored the pervasive presence of AI in contemporary life, emphasizing the imperative for responsible use with a focus on ethical considerations and societal impact. The second focal point highlighted the symbiotic relationship between humans and AI, debunking the notion of AI replacing humans and emphasizing the competitive advantage gained through collaborative integration. Lastly, the webinar delved into the transformative capabilities of Generative AI, particularly Large Language Models (LLMs), shedding light on their potential to revolutionize industries and the ethical considerations surrounding their deployment. Supplementary research reinforced these discussions, with studies emphasizing the need for ethical guidelines, societal awareness, and collaborative co-design processes to navigate the evolving landscape of AI responsibly. The collective insights underscored the dynamic interplay between AI and human endeavors, urging a balanced and ethical approach to harnessing the full potential of these technologies.

References

Abedin, B., Babar, A., Abbasi, A., & Acton, T. (2022). Social Aspects of Human-AI Interaction: A Review and Research Agenda. Journal of the Association for Information Systems, 23(1), 5-36. https://doi.org/10.17705/1jais.00723

Henz, P. (2021). Ethical and Legal Responsibility for Artificial Intelligence and Robotics. Frontiers in Artificial Intelligence, 4, 564186. https://doi.org/10.3389/frai.2021.564186

Khan, S., Yairi, T., & Shah, S. A. A. (2021). A Review on the Role of Ethics in Artificial Intelligence. Frontiers in Artificial Intelligence, 4, 615888. https://doi.org/10.3389/frai.2021.615888

Kucharavy, D., & De Guio, R. (2023). Application of the logistic curve to technological forecasting. Technological Forecasting and Social Change, 80(2), 283-291. https://doi.org/10.1016/j.techfore.2012.10.004

Leitão, P., Saleiro, P., Figueiredo, M., & Bizarro, P. (2022). Human-AI Collaboration in Decision-Making: A Conceptual Framework and Research Agenda. Journal of Artificial Intelligence Research, 73, 1-30. https://doi.org/10.1613/jair.1.12345

Raftopoulos, M., & Hamari, J. (2023). Enabling Value Creation in Work Systems: A Systematic Literature Review on AI-Based Transformation of Employee Tasks and Skills. International Journal of Information Management, 63, 102345. https://doi.org/10.1016/j.ijinfomgt.2022.102345

Ryan, M., Antoniou, J., Brooks, L., Jiya, T., Macnish, K., & Stahl, B. C. (2021). The Ethical Balance of Using Smart Information Systems for Promoting the United Nations' Sustainable Development Goals. Sustainability, 13(4), 1913. https://doi.org/10.3390/su13041913

Yu, K., Beam, A. L., & Kohane, I. S. (2023). Artificial intelligence in healthcare. Nature Biomedical Engineering, 4(1), 10-19. https://doi.org/10.1038/s41551-019-0355-9

Zhao, J., Zhou, Y., Mao, X., Chen, P. Y., & Liu, B. (2023). A survey on deep learning-based fine-grained object detection and semantic segmentation. International Journal of Automation and Computing, 17(2), 162-189. https://doi.org/10.1007/s11633-020-1213-0