AI Design to Innovation

Emergent Research Forum (ERF)

Tzu-Ling Huang
Chang Gung University
emily109emily@yahoo.com.tw

Ching-I Teng
Chang Gung University
chingit@mail.cgu.edu.tw

Abstract

Artificial intelligence (AI) is expected to create various innovations for changing human workplaces. AI is characterized by features of learning and self-growth. Efficient AI learning should depend on human inputs, particularly from human professionals (e.g., doctors and nurses). Hence, professionals' intention to facilitate AI innovation is critical. However, little is known about how to design AI to strengthen such intention, warranting our research to answer this question. We use expectancy-value theory to identify three potential AI design elements and examine how they enhance the perception that AI enhances professionals' capabilities and their intention to facilitate AI innovation. These elements are contextual-specific features of AI, extending the expectancy-value theory to the novel AI technologies. We will test our model by using two-wave data of nursing professionals' responses. The results are expected to assist AI designs that effectively motivate professionals to facilitate AI innovations.

Keywords

Artificial intelligence (AI), AI design, innovation, professional, expectancy value

Introduction

Artificial intelligence (AI) is predicted to significantly replace or reshape human jobs (Walch and Cognitive World 2019), owing to its capabilities of learning and self-growth (the perception of AI capability to grow by itself) (Creighton 2019). These capabilities enable information systems to demonstrate unprecedented potentials. Such capabilities are technology-facilitated and thus can create innovative solutions that effectively outperform humans. Such capabilities, however, still require human professionals (e.g., doctors and nurses) to engage in necessary tasks, i.e., structuring the problem, setting the goal, providing quality data, and offering initial schemes for analysis. In sum, humans and AI should work together to enhance complementary strength of each side (Wilson and Daugherty 2018), indicating the pivotal role of professionals' intention to facilitate AI innovations. These necessary tasks create burdens that are additional to human professionals' daily routines while have some uncertainty in enhancing their performance. Hence, humans encounter the decision to ride the AI technology or even questioning or resisting AI (Miller 2019). Resistance to IT has been an important issue in the IS discipline. To successfully address this issue, human professionals (e.g., doctors and nurses) in hospitals are the key persons (Lapointe and Rivard 2005). This creates the problem (i.e., human professionals resist to facilitate AI innovations) for AI adoption, AI success, and subsequently impact firms' performance that fall behind in using AI.

However, the literature on AI technology development can hardly be used to solve the aforementioned problem, warranting research aiming to understand how humans interact with AI technology. Moreover, the human-computer interaction (HCI) literature has not explained user contribution issues by using the AI-specific features, e.g., self-growth, indicating the insufficiency of HCI theories in solving this problem. In sum, both the AI and HCI literatures cannot resolve this problem.

We aim to solve the problem by using expectancy-value theory (EVT). The reason of choosing EVT is its fit with our study. EVT posits that humans would evaluate the expected payoffs and the probabilities of actual occurrences to form expectancy values, thus determining their intention to engage in a certain activity (Wigfield 1994). Hence, EVT can explain professionals' intention to facilitate AI innovations, justifying our adoption of EVT. We map EVT into the AI context. Specifically, the expected payoffs can be contextualized as what AI can do, i.e., *AI strengths*, and whether AI generates results that can be explained or rationalized by professionals (thus make AI outcome can be truly applied), i.e., *AI outcome explainability*. This concept is a major barrier to AI deployment (Arrieta et al. 2020). Moreover, the probability of actual occurrences can be translated into the perception of AI capability to grow by itself, i.e., *AI self-growth*. All these are mapped from EVT while also rooted in the AI context, justifying their inclusion as potential AI design elements for solving our target problem.

We further contextualize the expectancy value of EVT into the concept: *AI enhancement on capabilities*, which is defined as the perception that AI can help enhance professionals' capabilities. This concept is new to both the AI and the HCI literatures. The knowledge that AI will help enhance professionals' capabilities will motivate them to take actions to help AI succeed. Hence, this concept will be key to encouraging professionals to contribute their expertise to facilitate AI innovation. Given that these concepts are new to the literature, we found that the current literature cannot answer our research question, i.e., how these elements can strengthen the human professionals' intention to facilitate AI innovations.

Answers to this research question can help integrate the human professional' expertise in the AI innovation process. This is critical for AI to generate explainable outcomes, while AI software and AI designers themselves cannot have access to expertise of human professionals without their active assistance. That is, answers to our research question can enable AI designers to include human professionals' expertise, generating much better AI innovations than without such expertise. Therefore, the purpose of this study is to use expectancy-value theory and explore how to strengthen the human professionals' intention to facilitate AI innovations

Table 1 depicts how we map the EVT into our research context:

Theoretical Concept	Contextual Concept	Definition of the Contextualized Concept
Expected Payoff	Perceived AI Strength	Perceiving AI as outperforming humans
Expected Payoff	Percieved AI Outcome Explainability	Perceiving AI-generated results as rational
Probability of Occurrence	Perceived AI Self-growth	Perceiving AI as capable of learning by itself

Table 1. Mapping Theory to Our Context

Prior Research

The AI literature has disclosed various *AI strengths*, including its *self-growth* capabilities (Creighton 2019), i.e., using learning algorithms to improve AI itself (Miller 2019), thereby creating AI's outstanding performance or strength (Wilson and Daugherty 2018). Such literature depicts a promising world for AI development and innovation. However, AI business initiatives and technological development encountered obvious stumbles (Blier 2019), showing the need for research to pave a smooth avenue for AI development and innovation.

AI stumbles include a strong demand for super computational capabilities, which is forecasted to be solved by quantum computers (Miller 2019; Moret-Bonillo 2015). Nevertheless, AI stumbles should also include humans, particularly professionals those who are currently in charge of critical services, including medical, legal, and accounting ones. For example, medical doctors would challenge non-explainable AI outcomes (Schmelzer and Cognitive World 2019). This challenge is valid. If we do not know why we do better now, we cannot know whether the current solution would still work better in the future or lead us to a disaster. Hence, *AI outcome explainability* should be an important design element for securing human professionals' support, thus justifying its inclusion.

Collectively, we use these AI design elements: AI strength, AI self-growth, and AI outcome explainability, to explain why professions would evaluate AI as enhancing their capabilities and intend to facilitate AI innovation in their professions. All these can be explained by EVT, i.e., EVT could inform the development of our framework, as illustrated in Figure 1.

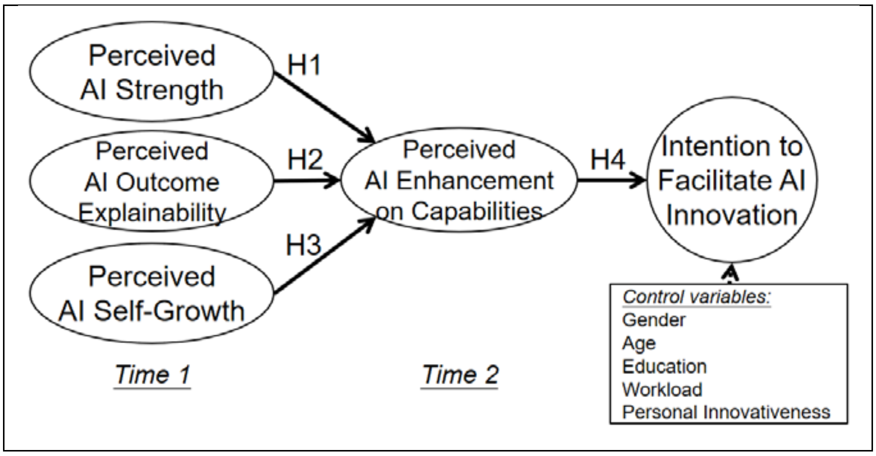

Figure 1. Research Framework

EVT perspective is used to inform the development of our hypotheses. Perceived AI strength indicates the high payoff values of using AI, thus raising the expected value of using AI (H1). Moreover, perceived AI outcome explainability is critical for obtaining workplace recognition. Without explainability, AI outcomes are frequently challenged as mysterious to human professionals, because human professionals are the ones to take the responsibility of using AI to make decisions. In this sense, AI outcome explainability is critical for making AI outcomes as any valuable, or substantial payoff of using AI to make decision. EVT posits that this payoff contributes to the expected value, i.e., the value of using AI in our research context (H2). The perception that AI can engage in self-growth, indicating that self-growth capability *likely* enables AI to outperform humans in the future, even some AI applications may not do now. This high likelihood, according to EVT, also forms the expected value, which is the expected value of using AI in our context (H3). The expected value of using AI thus should strengthen human professionals' intention to facilitate AI innovations.

Method

Participants and Data Collection

We will collect two waves of data. The first wave (time 1) will include measures of AI strength, AI self-growth, and AI outcome explainability. The second wave (time 2) will include measures of AI enhancement on capabilities and intention to facilitate AI innovation. The two waves provide temporal sequence of the data for testing the first three hypotheses. This sequence may provide preliminary evidence addressing the issue of reverse causality. The last hypothesis unlikely suffers from reverse causality, as its dependent variable is a behavioral intention, which typically comes from perceptions.

We aim to gather responses from registered nurses who work full-time for a large medical center. This medical center is aggressively introducing AI to improve hospital operations in all aspects. Hence, nurses work there have received on-the-job training of AI and recognized the hospital's determination of implementing AI. This supports the suitability of using their responses for our study.

We plan to use a proportionate random sampling method to draw a sample of 600 nurses from all nursing units. This sampling method ensures sample representativeness across nursing units. Nurses in our research context are highly aware of AI applications and their potential, because the medical center has highly dedicated to train all human professionals to understand AI. Moreover, the medical center actively embraces AI innovations. Nurses are suitable participants, as nursing jobs are known challenging for AI to replace (Wilson and Daugherty 2018). That is, AI designers urgently need nurses' active participation and strong intention to assist any AI innovations pertaining to nursing jobs. This would also be highly relevant in the current severe global nurse shortage. In this sense, the lack of current AI applications used by nurses *justifies* the relevance of our study to AI designers.

Measurement

We plan to develop our items basing on the technology adoption literature (e.g., Venkatesh et al. 2003, 2016) and articles on AI (e.g., Arrieta et al. 2020; Wilson and Daugherty 2018). These items and their constructs are new to the literature. Hence, we will ask professionals in IS, AI, and nursing to review these items and give comments for improvement. We will assess their reliability, validity, and measurement model fit. We will also evaluate common method bias, self-selection bias, and sampling representativeness.

We plan to collect information on the control variables: nurses' gender, age, education, workload, and personal innovativeness. These control variables are frequently used in the IS adoption literature. Such information will be self-reported. The impact of control variables will be assessed in the statistical significance and on the difference between including/excluding them.

Expected Results, Contributions, and Implications

We aim to solve the problem of under-realization of the full potential in AI innovations. We plan to contextualize the EVT to the AI context, build the model to explain how to enhance professionals' intention to facilitate AI innovation, and test this model. We expect that our study will find support for this model, offering solutions to the problem, e.g., verifying the impacts of the contextualized sources in driving AI innovation. Our study will push the theoretical boundary of EVT from evaluating a *static* target to a *dynamic* target, i.e., self-growing AI. Our study will offer the EVT theoretical perspective, which could amplify its usefulness and impact in explaining issues in AI development and innovation.

Although the previous IS studies has examined IT resistance issue (Lapointe and Rivard 2005; Rivard and Lapointe 2012), the approach has been a content-analytical one. Moreover, the IT resistance issue becomes more complex in the AI technology, as AI has unique features of self-growth and high likelihood of outperforming humans. These unique features should further complicate the user resistance resolutions. Hence, our findings would contribute some insights to resolve *AI resistance*, which is important for the AI designers to successfully resolve human professionals' resistance in hospitals, and therefore include their expertise.

Acknowledgements

The authors thank Chang Gung Memorial Hospital for financial support (CMRPD3J0031).

REFERENCES

Arrieta, A. B., Diaz-Rodriguez, N., Ser, J. D., Bennetot, A., Tabik, S., ... and Herrera, F. 2020. "Explainable Artificial intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI," *Information Fusion* (58), pp. 82-115.

Blier, N. 2019. "Stories of AI Failure and How to Avoid Similar AI Fails," (https://www.lexalytics.com/lexablog/stories-ai-failure-avoid-ai-fails-2019; accessed Jan 31, 2020)

Craighton, J. 2019. "The Unavoidable Problem of Self-Improvement in AI: An Interview with Ramana Kumar, Part 1" (https://futureoflife.org/2019/03/19/the-unavoidable-problem-of-self-improvement-in-ai-an-interview-with-ramana-kumar-part-1/?cn-reloaded=1; accessed Jan 31, 2020)

Lapointe, L., and Rivard, S. (2005). "A Multilevel Model of Resistance to Information Technology Implementation," *MIS Quarterly* (29:3), pp. 461-491.

Miller, A. 2019. "The Intrinsically Linked Future for Human and Artificial Intelligence interaction," (https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0202-7; accessed Jan 31, 2020)

Moret-Bonillo, V. 2015. "Can Artificial Intelligence Benefit from Quantum Computing?," *Progress in Artificial Intelligence* (3), pp .89-105.

Rivard, S., and Lapointe, L. 2012. "Information Technology Implementers' Responses to User Resistance: Nature and Effects," *MIS Quarterly* (36:3), pp. 897-920.

Schmelzer, R., and Cognitive World. 2019. "Understanding Explainable AI," (https://www.forbes.com/sites/cognitiveworld/2019/07/23/understanding-explainable-ai/#7a2ab6cd7c9e; accessed Jan 31, 2020)

Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. 2003. "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly* (27:3), pp. 425-478.

Venkatesh, V., Thong, J. Y. L., and Xu, X. 2016. "Unified Theory of Acceptance and Use of Technology: A Synthesis and the Road Ahead," *Journal of the Association for Information Systems* (17:5), pp. 328-376.

Walch, K., and Cognitive World. 2019. "Is AI A Job Killer Or Job Creator?," (https://www.forbes.com/sites/cognitiveworld/2019/11/24/is-ai-a-job-killer-or-job-creator/#75e080bf37e8; accessed Jan 31, 2020)

Wigfield, A. 1994. "Expectancy-Value Theory of Achievement Motivation: A Developmental Perspective," *Educational Psychology Review* (6:1), pp. 49-78.

Wilson, H. J., and Daugherty, P. R. 2018. "Collaborative Intelligence: Humans and AI Are Joining Forces," *Harvard Business Review* (July-August).