

Book review of Baker-Brunnbauer, J. (2023). *Trustworthy Artificial Intelligence Implementation: Introduction to the TAI Framework*. Springer. 127 pp. ISBN: 978-3-031-18275-4.

Reviewed by

Claretha Hughes ¹ ✉

Type: Book review

Citation: Hughes, C. (2023). Book review of Baker-Brunnbauer, J. (2023). *Trustworthy Artificial Intelligence Implementation: Introduction to the TAI Framework*. Springer. 127 pp. ISBN: 978-3-031-18275-4. *ROBONOMICS: The Journal of the Automated Economy*, 4, 47

¹ Professor, Human Resource and Workforce Development, College of Education and Health Professions, University of Arkansas, 255 GRAD, Fayetteville, AR 72703, USA; Email: chbanks@uark.edu

✉ Corresponding author



© 2023 The Author(s)

This work is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0).

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

This book on artificial intelligence (AI) seeks to answer the question: *What kind of awareness does the management have about the social impact of their AI products and services?* The author explores key AI issues and challenges associated with implementing trustworthy AI within organizations. He seeks to add a Trustworthy Artificial Intelligence Implementation (TAIL) Framework to the knowledge base within the management and AI communities to lessen their struggle for successfully and viably expanding usage of AI within organizations and the global economy. This book shows that with focus on the right framework, managers and AI developers can move forward together to enhance the lives of all stakeholders with ethical, trustworthy AI.

The book consists of seven chapters of an empirical research study, including the introduction. In Chapter 1 “Introduction”, the challenge associated with AI and ethics provides excellent explanations about AI. Great questions to consider are included regarding the use of AI and its influence on jobs. In Chapter 2, “Artificial Intelligence”, the author provides an in-depth review of AI and its challenges and goals as it relates to human life and behavior. In Chapter 3, “Ethics and Moral”, concerns about people’s acceptance of AI are described including evidence that people usually accept AI if it is meaningful and low risk to them. Other moral and ethical challenges of AI adoption are discussed including safety and errors, law and ethics, and social impact alongside a myriad of ways that nations and organizations are developing guidelines and 47 principles of AI ethics. In Chapter 4, “Business and Management Perspective”, the author explains the relationship between AI implementation, business strategy, and management decisions. The idea of AI as a strategic risk for organizations is explained and a direct focus on how organizations are specifically using AI for growth instead of for cost reduction is highlighted. Organizations should question whether AI is a risk, opportunity, or a pivot before use. The author provides chapter summaries of chapters two, three, and four that give the reader a very good synopsis of each chapter.

In Chapter 5, “Empirical Study”, the author describes the research methodology and definition of the material used in the research study to develop the TAIL framework. The author details all aspects of the empirical study and the findings associated with the examination of the social impact of AI products and services from a management perspective. The author used expert interviews, content analysis, a structured interview guide and recorded the responses of nine executive managers. The findings were revealed after an analysis focused without considering nonverbal communication, gestures, mimics, laughter or harrumphs.

In Chapter 6, “Findings”, the author stated findings and analysis of interview data. Some specific findings were:

- ✓ Companies were investing in AI to achieve cost savings, growth, and process optimization.
- ✓ Quality input data is required to produce quality AI outcomes and attain peoples’ trust regarding AI system usage.
- ✓ Legal, ethical frameworks should be used to govern all general software development principles.
- ✓ Robustness is achieved by keeping a human in the loop to prevent harming humans using AI systems.
- ✓ Younger employees would not work on unethical projects.
- ✓ The top two concerns with AI system usage are responsibility and protection of data privacy.
- ✓ Certification of AI products or services around a common standard that is easy to implement may be acceptable.
- ✓ Managers focus on speed of change and competition not social problems, including job losses, when choosing whether to use AI.

These and other findings led to the development of the TAIL framework.

In Chapter 7, “TAIL Framework”, the author introduces the reader to the TAIL framework and how it relates to organizations’ business models, its implementation challenges, and the TAIL canvas. The goal of the TAIL framework is to provide a holistic perspective of developing and implementing trustworthy AI systems. The 12 steps in the TAIL framework for AI system life cycle and the 12-step TAIL canvas are vital to an organization

developing and implementing AI. The framework seeks to align AI systems to specific goals of the organization to determine if AI may contribute positively or negatively to the goal. It is designed to help improve the transparency of AI through communication, decision-making processes, and complying with certification standards as applicable among the stakeholders. For the TAI framework to be successful, organizational leadership must prioritize AI ethics, allocate resources, and be committed to AI system implementation. The author provides detailed business model descriptions for the reader.

Trustworthy artificial intelligence implementation: Introduction to the TAI framework is a book about AI systems' use by managers. It asks managers to check any assumption that AI systems will be trustworthy and examine all aspects of the AI system using the TAI framework from development through implementation using the 12 step TAI canvas. Using the canvas requires in-depth discussion among all stakeholders. The book builds upon the author's expertise in exploring what has and has not worked for managers' understanding of AI and AI ethics. Many times, while reading the chapters, I thought about questions and the use value of AI. The author has provided historical data and context to support his suggestions and examples. The book provides literature that seeks to help define AI, and literature reviewing the many efforts to develop ethical AI guidelines and principles. The author recognizes the differences in cultures across countries that must be understood, respected, and valued as they seek to develop standardized ethical guidelines and principles.

Reading Chapter 1 spurred questions from this reader such as *How does a machine infer goals from human behavior when a human rarely articulates goals effectively? With almost nine billion people on the earth, how will the "right" humans with the "right" goals be selected?* The content makes one want to read additional research that supports this book. This book could have benefited from more empirical research support in Chapters 1 and 2. The topic is ever evolving so it is understood that new research is emerging quickly. Thus, the information on AI in Chapters 1 and 2 is changing. This does not diminish how this book contributes to the research knowledge base in the field. It is to alert potential readers to the expediency of the content and how new content can add to what this book provides.

The book's content on how historical data used to train AI not to be biased or discriminatory is especially important because machine learning has become so prevalent with the introduction of ChatGPT and other programs like it. The selection of historical data that is not discriminatory is going to be difficult to find so all stakeholders must be diligent. For example, how can AI predict success of Black women CEOs if no or not enough historical data exists? AI software cannot be trained to be supportive of diversity when the past shows that diversity efforts were not successful.

Explainability of AI is difficult but using a framework such as the TAI can help stakeholders begin comprehensive communications and processes that can be beneficial. The TAI framework allows an exploration of AI ethics from a non-technical perspective. Individuals using the framework can discuss opportunities and challenges associated with implementing AI ethics. AI laws and legal frameworks are designed to minimize AI risk in the marketplace. The TAI framework can be used to assess where risk exists within AI implementation. The TAI framework canvas's 12 steps provide a beginning point for organization leaders to implement trustworthy AI for social and ethical good.

This book challenges scholars and practitioners to examine how they think about human capital development as it relates to the introduction of ethical AI systems into organizations. As leaders attempt to solve complex problems related to AI system development and implementation, they along with their stakeholders must think strategically about the trustworthiness of AI from data selection, usage of historical data, machine learning, and AI system implementation from an ethical and social perspective. Teaching and training ethics is already a difficult task. Adding AI to the ethics discussion further complicates decision making for managers, but this book provides clear examples and urgency for it to be done. For practitioners and researchers who seek to help with

Book review of Baker-Brunnbauer, J. (2023). Trustworthy Artificial Intelligence Implementation: Introduction to the TAI Framework. Springer. 127 pp. ISBN: 978-3-031-18275-4.

organizational development and implementation of AI and AI ethics, this book can be a valuable asset. The scholarly studies cited, and the historical knowledge provide a rich empirical landscape from which to build a foundation for other empirical studies on TAI system implementation.

Received: 23/04/2023

Accepted: 02/05/2023