

Trusting artificial intelligence at work

Summary report

September 2021

This research summary is based on the *Trusting artificial intelligence at work* research produced by the Centre for Work Health and Safety, Charles Sturt University. Special thanks to Dr. Ben Morrison and his team for their collaboration on this project.

Background

Intelligent Decision Support Systems (IDSS) are computerised information systems that support human decision-making using artificial intelligence (AI) techniques. The use of these AI-powered systems is growing across a range of industries, as they can offer significant innovation at work, and positive economic return. The introduction of this technology in the workplace is known to greatly affect the worker's environment, e.g., role design, task allocation, time management, and yet, little is known about the potential work health and safety (WHS) harm it can generate. In particular, the extent to which working alongside an AI-powered system that a worker does not fully trust can harm its health and safety. The present research aimed to better understand workers' perceptions of IDSS and the factors that may impact their acceptance of the technology.

Method

Data were collected during three phases of investigation:

- i. A systematic review of the literature (N = 17 from 4488 citations) examined the known factors influencing the acceptance of IDSS among workers.
- ii. Current and prospective IDSS users (N = 93) responded to a questionnaire capturing their perceptions of IDSS. Respondents rated the perceived importance of several factors that may influence their acceptance of the technology.
- iii. A series of in-depth interviews (N = 10) was completed based on a Critical Decision Method procedure (CDM; Militello & Hutton, 1998) to elicit first-hand accounts of IDSS use.

Discussion

The systematic review showed a high rate of IDSS use in medical fields with workers holding largely positive views toward them. Critically, findings also suggested a lack of consideration of the WHS effects from IDSS use, and a severe lack of quality of the existing standardised assessments of workers' interactions with the technology. Data collected in this study also indicated a relatively high availability (or anticipated availability) for IDSS across a range of sectors and industries. A range of factors were found to be potentially influential regarding the worker's acceptance of advice made by IDSS technology: the quality and the importance of the decision to be made, the historical accuracy of the system; the level of accountability of the worker in making the final decision, the worker's comprehension of the system. Other factors impacting the cognitive compatibility between users and IDSS were also found of importance: the level of control and autonomy of the worker, the capacity and transparency of the technology.