

**JOURNAL OF INFORMATION
SYSTEM AND TECHNOLOGY
MANAGEMENT
(JISTM)**www.gaexcellence.com/jistm

CONCEPTUALIZING AND VALIDATING SOCIO- ALGORITHMIC AI LITERACY (SAILS) AMONG MALAYSIAN UNIVERSITY STUDENTS: A PROPOSED FRAMEWORK

Zuraidah Arif¹, Abd Latif Abdul Rahman^{2*}, Zati Atiqah Mohamad Tanuri³, Nor Azlina Azmi⁴

¹Faculty of Information Science, Universiti Teknologi MARA (UiTM) Cawangan Kedah, Kampus Sungai Petani, Malaysia

 zuraidah_arif@uitm.edu.my

 <https://orcid.org/0000-0002-3357-1649>

²Faculty of Information Science, Universiti Teknologi MARA (UiTM) Cawangan Kedah, Kampus Sungai Petani, Malaysia

 ablatif@uitm.edu.my

 <https://orcid.org/0000-0001-7970-2520>

³Faculty of Information Science, Universiti Teknologi MARA (UiTM) Cawangan Kedah, Kampus Sungai Petani, Malaysia

 zatiatiqah@uitm.edu.my

 <https://orcid.org/0000-0003-0862-5999>

⁴Faculty of Information Science, Universiti Teknologi MARA (UiTM) Cawangan Kedah, Kampus Sungai Petani, Malaysia

 nazlina@uitm.edu.my

 <https://orcid.org/0009-0000-5603-7839>

*Corresponding Author

Article Info:**Article history:**

Received date: 27.07.2026

Revised date: 30.08.2026

Accepted date: 20.09.2026

Published date: 28.09.2026

To cite this document:

Arif, Z., Rahman, A. L. A., Tanuri, Z. A. M., & Azmi, N. A. (2026). Conceptualizing And Validating Socio-Algorithmic AI Literacy (SAILS) Among Malaysian University Students: A Proposed Framework. *Journal*

Abstract:

The current measures of AI literacy tools assess technical skills, mechanics of prompt engineering, and basic ethical considerations but do not account for socio-algorithmic vulnerability or increased susceptibility to cognitive persuasion, uncritical trust, and unintended privacy breaches due to conversational interface persuasion tactics. Drawing on theoretical constructs such as CASA (Computers Are Social Actors), Dual Process Theory, and Information Literacy. Socio-Algorithmic AI Literacy (SAIL) is defined as the psychological and behavioral ability of an individual to be aware of, assess, and reject mindless social scripts triggered through the interaction of humans with human-like AI. This study provides a theoretical framework for the development and validation of Socio-Algorithmic AI Literacy Scale (SAILS) for HEI students from Malaysia. By advancing AI literacy beyond mere operational knowledge towards identification and management of socio-algorithmic risks, SAILS is able to provide a reliable instrument for assessing the readiness of students in this regard

of Information System and Technology Management, 11 (44), 604-620. at the Ministry of Higher Education Malaysia, university senates, and academic libraries.

DOI: 10.35631/JISTM.1144037 **Keyword:**

Socio-Algorithmic AI Literacy, Scale Development, CASA Paradigm, Higher Education, Information Evaluation



© The authors (2026). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC) (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact jistm@gaexcellence.com.

Introduction

The study of Information Science has always been concerned with individual information seeking, evaluation, and use (Bawden & Robinson, 2022). In the past two decades, information literacy models have evolved in response to the digital world, web search engines, and social media sites. In the contemporary period of Generative AI, there is a need for a change in the educational paradigm that is called AI literacy (Long & Magerko, 2020). The national efforts of Malaysia to ensure the development of an AI-ready workforce include the MyDIGITAL Blueprint, the National AI Roadmap, and the MOHE Guidelines on Generative AI (Yusnilzahri et al., 2025). The institutions of higher education in Malaysia are actively integrating AI into their work. Nevertheless, contemporary AI-based solutions represent something entirely different from conventional search engines. Not only does an AI-based assistant generate the links; it interacts, persuades, apologizes, and even shows empathy. As per the Computers Are Social Actors (CASA) theory, people unconsciously perceive computers as social actors whenever they detect socially humanlike cues (Gambino et al., 2020; Reeves & Nass, 1996). While the AI literacy program teaches the students that AI can hallucinate and exhibit bias, students still succumb to the "anthropomorphic halo" effect in practice (Lee, 2024; Ngo, 2026). Once an AI tutor communicates respectfully and empathically, students tend to drop their cognitive guards and accept any misinformation without double-checking it (Ngo, 2026). Consequently, there is a need for a novel metric in Information Science to evaluate not only the knowledge of students about AI code but also their ability to resist socio-algorithmic manipulation.

Socio-Algorithmic Vulnerability

Socio-Algorithmic Vulnerability is defined as the enhanced propensity of an individual to be subjected to mental manipulation, blind faith, and accidental privacy breaches that emerge due

to the interplay between the social reactivity of humans and the characteristics of persuasive algorithms (Gambino et al., 2020; Keleher et al., 2026; Lee, 2024; OECD, 2026b). In this research, Socio-Algorithmic Vulnerability is considered the initial vulnerability of a user before socio-algorithmic literacy skills become engaged. The concept is understood in three dimensions, which are interconnected:

1. **Interpersonal Social Scripting Heuristic:** The unconscious application of interpersonal social scripts, such as courtesy, empathy, and reciprocity, to conversations with social AI systems that exhibit human-like linguistic signals, which is rooted in the CASA framework (Gambino et al., 2020; Nass et al., 1994).
2. **Algorithmic Obscuration and Bias Heuristic:** The unthinking acceptance of the “machine heuristic,” which refers to a mental shortcut that makes us automatically believe that any output generated by computers is unbiased, objective, and accurate (Lee, 2024; Shin et al., 2022).
3. **Simple Ontology Heuristic:** The mistaken cognitive classification of an AI system purely as a benign social agent as opposed to an obscure corporate conduit for collecting data or a mere tool, as per the Privacy Triad framework (Keleher et al., 2026).

In terms of the SAILS-HE model, the degree of socio-algorithmic vulnerability of a person is determined by his/her initial characteristics, namely low Need for Cognition (NFC), high Interpersonal Susceptibility to Social Influence (ISSI), and the presence of a social agent-oriented conception of STIC. This increased vulnerability raises the likelihood of unintentional data leaks, cognitive offloading, and susceptibility to synthetic hallucinations, which increased SAILS can moderate and nullify (Lee, 2024; OECD, 2026b; Sotero et al., 2026).

Research Gaps

Despite the growing literature on AI literacy scales, three major research gaps exist in current Information Science and educational technology scholarship:

Gap 1: Construct Deficit in Existing AI Literacy Scales

Most validated scales, such as the Meta AI Literacy Scale (MAILS) (Koch et al., 2023), the Generative AI Literacy Scale (GAIL) (Liu et al., 2025), and general student competency tests (Do et al., 2026; Zhou et al., 2025), focus heavily on:

1. **Technical/Conceptual Knowledge** (e.g., understanding machine learning concepts),
2. **Practical Tool Usage** (e.g., prompt engineering),
3. **General Ethical Awareness** (e.g., knowing about copyright and privacy policies).

However, none of these scales measure Socio-Algorithmic Literacy, a student's psychological capacity to recognize, resist, and audit CASA-driven social cues (anthropomorphism, social presence, paralanguage, and conversational warmth).

Gap 2: The Knowledge-Behavior Disconnect in Information Evaluation

Information evaluation models assume that users critically assess source credibility using conscious, analytical thinking (System 2 processing) (Kahneman, 2011; Lee, 2024). However,

CASA research proves that social cues trigger automatic, mindless social scripts (System 1 processing) (Xu et al., 2022). Though students can achieve good scores in terms of their understanding of AI knowledge, they may still have too much faith in generated content or share their university's confidential data due to the perception of the conversation as warm and intelligent (Keleher et al., 2026; Lee, 2024).

Gap 3: Lack of Contextual Validation in Malaysian Higher Education

The majority of scales developed for AI measurement were done in Western or East Asian cultures (Almatrafi et al., 2024; Do et al., 2026). The socio-cultural environment within which higher education exists in Malaysia is distinct in that it has the following features:

- High power-distance academic relationships,
- Dual institutional structures (Public *IPTA* vs. Private *IPTS*),
- Code-switched (Bahasa Malaysia and English) communication habits.

There is no psychometrically validated scale tailored to evaluate socio-algorithmic vulnerability among university students in Malaysia.

Research Questions

What are the key construct dimensions of Socio-Algorithmic AI Literacy (SAIL) in higher education?

- What is the validity and reliability of the newly constructed Socio-Algorithmic AI Literacy Scale (SAILS) when measured among students from Malaysian HEIs?
- Does the SAILS measure exhibit measurement invariance between IPTA and IPTS university student groups in Malaysia?
- To what extent does SAILS predict information evaluation behavior and protect from AI over-trust?

Research Objectives

- To develop the multi-dimensional construct of Socio-Algorithmic AI Literacy.
- To develop and validate the SAILS measure psychometrically through robust scale development procedures (EFA, CFA, PLS-SEM).
- To test the measurement invariance of SAILS among university students in Malaysia who study in both IPTA and IPTS.
- To examine the predictive validity of SAILS towards over-trust to AI, critical verification behavior, and unintentional disclosure of personal information.

Literature Review

The Computers Are Social Actors (CASA) paradigm serves as the primary psychological foundation for the Socio-Algorithmic AI Literacy Scale (SAILS). It explains how human-like interface cues, such as conversational warmth, politeness, first-person phrasing, and perceived empathy, automatically trigger social scripts and mindless heuristic processing in human users (Gambino et al., 2020; Lee, 2024; Reeves & Nass, 1996; Xu et al., 2022). Rather than evaluating basic technical knowledge or prompting mechanics, SAILS operationalizes these

underlying CASA mechanisms into four psychometrically distinct competency dimensions designed to measure a student's capacity to audit and regulate social AI interactions.

Anthropomorphic De-biasing measures user resistance to CASA-triggered relational cues by assessing how effectively a student decouples interface warmth, politeness, and conversational rapport from factual accuracy and underlying technical capability (Lee, 2024; Ngo, 2026). This dimension evaluates a user's capacity to neutralize the "anthropomorphic halo" that often lulls individuals into suspending critical scrutiny during natural language exchanges (Carolus et al., 2018; Gambino et al., 2020). Meanwhile, Privacy Triad Calibration addresses the CASA-driven privacy paradox wherein perceiving an AI device as a "social agent", rather than a non-human conduit or tool, fundamentally relaxes user boundary management and heightens disclosure risks (Ghosh & Eastin, 2020; Keleher et al., 2026). This dimension evaluates a student's ability to maintain strict data privacy controls and resist inadvertent personal or institutional data sharing when interacting with conversational interfaces. Algorithmic Auditing evaluates the machine heuristic model of human-machine communication, which states that fluent conversation and instantaneous responses generate unwarranted epistemic trust and cognitive dependence (Kahneman, 2011; Lee, 2024). This sub-factor assesses the student's capacity to counteract their System 1 processing and perform System 2 thinking and evaluation to audit algorithmic authority created (Association of College & Research Libraries (ACRL), 2016; Kahneman, 2011). The Epistemic Vigilance responds to the CASA-induced cognitive mindlessness phenomenon, whereby students accept the plausible output of algorithms without verification (Xu et al., 2022), through the assessment of the behavioral tendency of the student to verify, triangulate, and authenticate the outputs generated by the generative AI using authoritative external sources (ACRL, 2016). Below is a summary of important literature on the relationship between the CASA approach, information science, AI literacy, and scale development.

Table 1: Comparative Analysis of Existing AI Literacy Scales, CASA Frameworks, and Identified Research Gaps

Study	CASA Construct / Mechanism	Information Science & Literacy Dimension	Sample Context	Key Findings & Identified Gap
Gambino et al. (2020)	Human-Media Social Scripts, Perceived Agency.	Human-Machine Communication, Interaction Schemas.	Theoretical Conceptualization (HCI / HMC).	Replaced mindless human-human scripts with human-media social scripts. Gap: Did not create a scale to measure individual differences in script resistance.
Koch et al. (2023)	N/A (General Competency).	Meta AI Literacy Scale (MAILS), Self-Reflection.	N = 612 German University Students.	Validated a meta-competency scale. Gap: Ignored socio-emotional interface cues and CASA mechanisms.

Xu et al. (2022)	Mindlessness vs. Mindfulness, Social Presence.	Information Perception, Trust Attribution.	$N = 834$ Adult Participants.	Proved that mindlessness explains social presence better than mindfulness. Gap: Focused on media psychology, not student information evaluation literacy.
Lee (2024)	Machine Heuristic vs. Mindlessness Account.	Source Evaluation, Credibility Attribution.	Theoretical Synthesis (HCR).	Integrated mindlessness and machine heuristics into HMC. Gap: Lacked an empirical scale to measure user socio-algorithmic auditing capacity.
Keleher et al. (2026)	Privacy Triad (<i>Social Agent</i> vs. <i>Conduit</i> vs. <i>Tool</i>).	Information Privacy, Security Behavior.	$N = 400$ Survey, $N = 12$ Grounded Theory.	Showed that treating devices as social agents alters privacy expectations. Gap: Did not operationalize the Privacy Triad into a student AI literacy scale.
Ngo (2026)	Perceived Social Presence, Transparency.	Technology Adoption, Uncertainty Reduction.	$N = 491$ Experimental Participants.	Showed that social presence builds trust, but excessive transparency triggers cognitive scrutiny. Gap: Did not measure student literacy traits that regulate this trust.
Do et al. (2026)	N/A (General Skill Test).	University Student AI Literacy Scale.	University Students in Vietnam.	Developed a 26-item student AI scale. Gap: Focused on basic technical application and general ethics; missed socio-algorithmic cues.
Liu et al. (2025)	N/A (Task Performance).	Generative AI Literacy Scale (GAIL).	$N = 500 +$ Workplace Knowledge Workers.	Created a 5-dimension GAIL scale for job performance. Gap: Designed for enterprise productivity, not information evaluation in education.

Theoretical Framework

The Socio-Algorithmic AI Literacy (SAILS) framework combines three theoretical foundations:

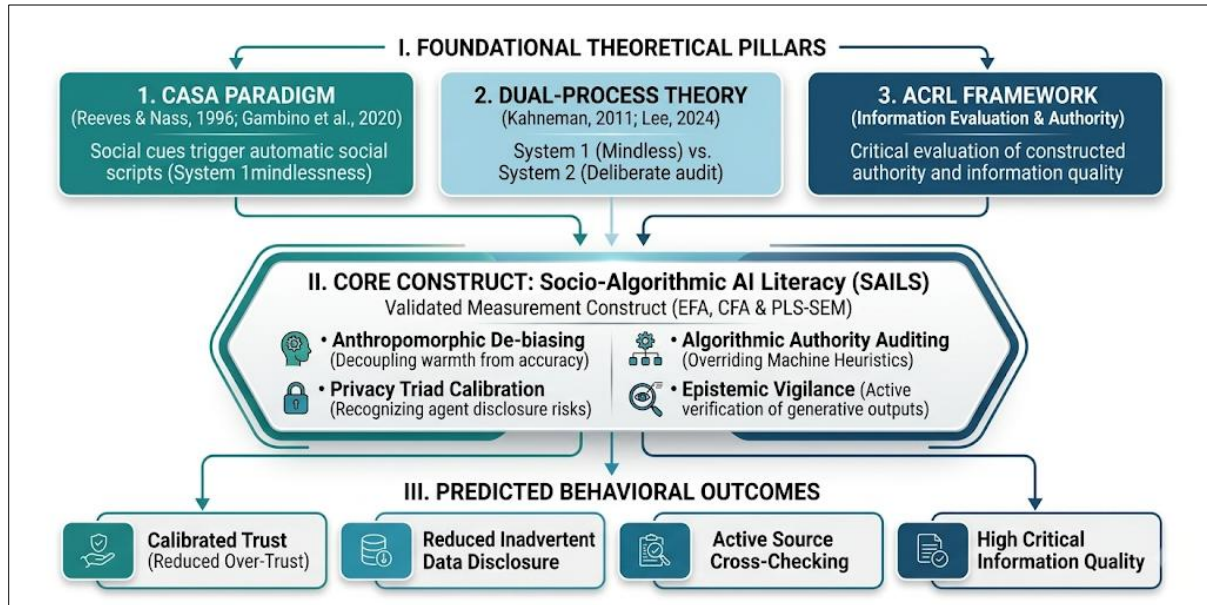


Figure 1: Theoretical Framework for the Development and Validation of the Socio-Algorithmic AI Literacy Scale (SAILS)

The SAILS framework is divided into three phases that are based on three theories, namely the CASA framework theory (Gambino et al., 2020; Reeves & Nass, 1996), Dual Process Theory (Kahneman, 2011; Lee, 2024) and ACRL Framework for Information Literacy (ACRL, 2016). These theories define the process through which social cues of an interface trigger automatic System 1 mindlessness and the cognitive shift to System 2 auditing. SAILS can be described as the main multi-dimensional measurement construct located right at the heart of the framework and consists of four separate dimensions, namely, Anthropomorphic De-biasing (disconnecting interface warmth from facts), Privacy Triad Calibration (being aware of disclosure dangers by assuming that AI is a social being), Algorithmic Authority Auditing (reversing default algorithms), and Epistemic Vigilance (verifying output of the generation process). Lastly, the framework of Information Literacy by the ACRL (2016) provides the foundation for information science in terms of standards to assess constructed authority, question the information construction process, and preserve data privacy ethics.

Methodology

Scale Development & Validation Protocol

In order to meet the need for clarity and specification of procedures, the scale development process is broken down into four operationally specific phases (Figure 2) based on psychometric scale development paradigms (DeVellis & Thorpe, 2021; MacKenzie et al., 2011).

Phase 1: Domain Specification and Item Generation

Phase 1 will commence with construct specification, whereby the four dimensions targeted by the Socio-Algorithmic AI Literacy Scale (SAILS) will be specified: Anthropomorphic De-biasing, Privacy Triad Calibration, Algorithmic Authority Auditing, and Epistemic Vigilance (Gambino et al., 2020; Keleher et al., 2026; Lee, 2024). The item pool will comprise a seven-point Likert-scale items ranging from 1=Strongly Disagree to 7=Strongly Agree, which will be derived following a thorough review of the literature and adherence to standardized scale construction procedures (DeVellis & Thorpe, 2021; MacKenzie et al., 2011). In order to determine content validity, the item pool will be subjected to review by a panel of eight experts in Malaysia who have expertise in Information Science, AI Ethics, and Psychometrics (Lynn, 1986). Content Validity Indexes will subsequently be computed to determine high performers, in line with strict psychometric guidelines, where $I''-CVI'' \geq 0.80$ and $S''-CVI/Ave'' \geq 0.90$ (Polit & Beck, 2006; Polit & Beck, 2007).

Phase 2: Pilot Testing and Exploratory Factor Analysis (EFA)

Phase 2 will involve pilot testing of the initial scale among a sample of 200 university students at two Malaysian HEIs through convenience sampling (Hair et al., 2019; Memon et al., 2020). The statistical analysis performed using IBM SPSS Statistics 29.0 will initially confirm the data factorability and sampling adequacy through the Kaiser-Meyer-Olkin measure ($"KMO">0.80$) and Bartlett's Test of Sphericity ($p<0.001$) (Field, 2018; Kaiser, 1974). The next step will include performing Exploratory Factor Analysis (EFA) by employing Principal Axis Factoring with Promax rotation to check for factor structure (Howard, 2016; Tabachnick & Fidell, 2019). Those items that meet psychometric requirements for item retention of primary factor loadings ≥ 0.60 and cross-loadings <0.30 will be included, followed by checking the internal consistency by using Cronbach's Alpha ($\alpha \geq 0.80$) (Nunnally & Bernstein, 1994).

Phase 3: Main Survey and Confirmatory Factor Analysis (CFA / PLS-SEM)

Phase 3 will focus on a more general population that includes undergraduate and postgraduate students at the Malaysian Higher Education Institutions. The use of multistage stratified random sampling techniques (Cochran, 1977; Sekaran & Bougie, 2016) will involve stratification of the sample according to the two institutional groups (IPTA and IPTS) and two academic disciplines (STEM and Social Sciences and Humanities). With SmartPLS 4.0 (Ringle et al., 2022) and R (lavaan package) (Rosseel, 2012) being the main analysis tools, the outer measurement model will be thoroughly evaluated. Construct reliability and validity will be confirmed using indicator reliability (outer loadings ≥ 0.708) (Hair et al., 2022), internal consistency (Composite Reliability $\rho_c \geq 0.70$) (Jöreskog, 1971), convergent validity (AVE ≥ 0.50) (Fornell & Larcker, 1981), and discriminant validity evaluated via the Heterotrait-Monotrait ratio (HTMT < 0.85) (Henseler et al., 2015).

Phase 4: Structural Validation and Measurement Invariance

Phase 4 completes the validation protocol by evaluating structural relationships and measurement invariance across target demographics (MacKenzie et al., 2011). The three-step Measurement Invariance of Composite Models (MICOM) procedure will be executed in SmartPLS 4 to verify that student cohorts in public (IPTA) and private (IPTS) universities interpret the scale items identically (Henseler et al., 2016). Finally, predictive validity testing

will evaluate the structural model's explanatory and predictive metrics (R^2 , f^2 , Q^2_{predict}) in predicting key downstream information behaviors (Hair et al., 2022; Shmueli et al., 2019), specifically reducing AI over-trust, increasing information verification frequency, and preventing inadvertent privacy disclosure.

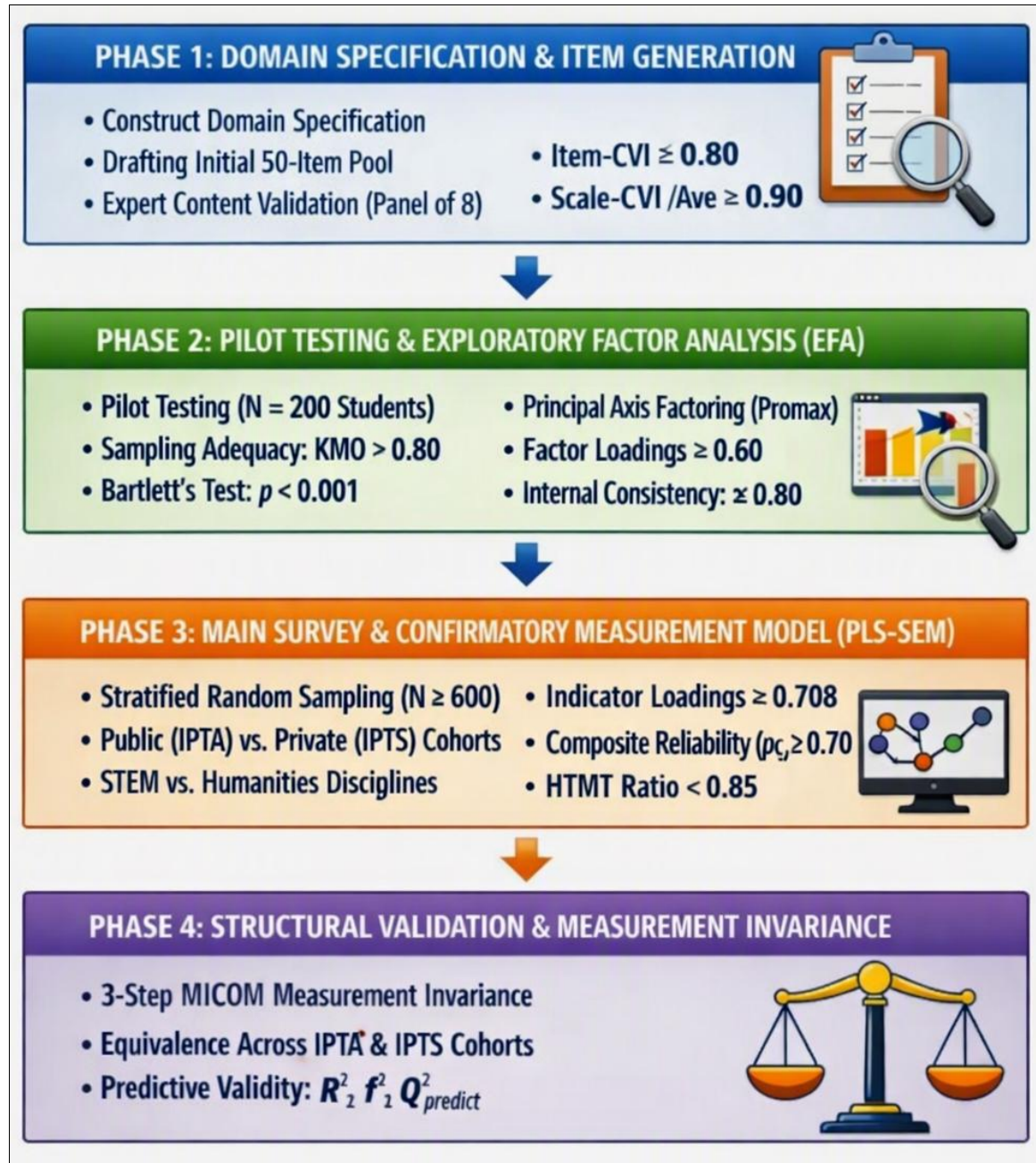


Figure 2: 4-Phase Psychometric Scale Development Protocol

Construct Definition and Measurement Scope

In scale construction of new technologies, construct validity can be established by defining skills that transcend mere mechanical operations and programming knowledge. SAILS is conceived as a multidimensional, second-order latent construct that gauges a person's ability to critically maneuver, assess, and interact with conversational AI and algorithmic

recommendation systems (Lee, 2024; Shin et al., 2022; Sotero et al., 2026). Instead of gauging technical competency, SAILS evaluates the metacognitive, evaluative, and socio-cognitive skills of the user to detect and control two separate cognitive weaknesses inherent in AI-assisted environments:

1. Automated Social Heuristics: Activation of pre-programmed social scripts in response to conversational prompts that seem human (Gambino et al., 2020; Nass & Moon, 2000; Nass et al., 1994).
2. Machine Heuristic: The tendency to always assume that the results generated through computational processes are always objective and accurate (Kahneman, 2011; Lee, 2024; Sundar, 2008).

By assessing these dual mechanisms, SAILS evaluates the psychological and behavioral competencies required to prevent uncritical over-trust, inadvertent data exposure, and cognitive over-reliance (Keleher et al., 2026; OECD, 2026; Sotero et al., 2026).

Core Competency Dimensions Measured by SAILS

The SAILS instrument operationalizes socio-algorithmic literacy across four core first-order dimensions:

Anthropomorphic de-biasing (social-cue awareness) is the user's capacity to recognize conversational social cues (e.g., politeness, empathetic phrasing, natural dialogue) and actively decouple conversational "warmth" or linguistic fluency from factual accuracy and system reliability (Gambino et al., 2020; Lee, 2024). Grounded in the Extended Computers Are Social Actors (CASA) paradigm (Gambino et al., 2020; Nass & Moon, 2000; Nass et al., 1994), this dimension measures the extent to which a user refrains from mindlessly applying human-to-human social scripts to interactive media agents. It evaluates social-cue awareness by assessing whether a user maintains analytical scrutiny despite flattering, highly polite, or anthropomorphic interface design (Heyselaar, 2023; Lee, 2024).

Privacy triad calibration (privacy awareness) is the user's metacognitive ability to classify an interactive AI along the Privacy Triad continuum, distinguishing whether the device or agent is functioning as a benevolent Social Agent, a corporate Conduit harvesting personal data, or a controllable Tool (Keleher et al., 2026). Based on the Privacy Triad framework (Keleher et al., 2026) and Communication Privacy Management theory (Ghosh & Eastin, 2020; Petronio, 2002), the present dimension gauges privacy awareness in relation to conversational AI. The measure determines whether the individual is aware of the dangers related to structural data disclosure and observes conscious boundaries relating to personal identifiers, financial data, and disclosures.

Algorithmic authority auditing (critical evaluation and mitigation of AI over-reliance) refers to the deliberate use of cognitive resources to think hard about and challenge automated outputs, decision-making, and recommendations (Kahneman, 2011; Lee, 2024; Shin et al., 2022). This dimension is based on the theories of Dual-Process Theory (Kahneman, 2011) and human-computer communication (Lee, 2024; Sundar, 2008), and it provides a direct measure of critical evaluation and AI over-reliance. It captures the extent to which users are able to overcome "machine heuristics," defined as the tendency to consider any automated output as inherently objective and neutral (Shin et al., 2022; Sotero et al., 2026).

Epistemic vigilance (epistemic trust & active verification) is a metacognitive attitude and behavior of being skeptical about generative AI claims through checking them out with external and reliable sources (ACRL, 2015; Doyle et al., 2019; OECD, 2026). Being based on the ACRL Information Literacy Framework (ACRL, 2015) and on international testing standards (OECD, 2026; OECD & European Union, 2026), this dimension tests the level of epistemic vigilance. This dimension tests whether the user sees the output of generative AI as an open question, which requires further lateral reading and truth-checking rather than the acceptance of hallucinations and biases provided by algorithms as ready facts.

Scope Summary Matrix for Scale Item Development

The table below brings together the scope, emphasis, competency, threats, and Likert indicators within each of the four first-order sub-dimensions of the SAILS test:

Table 2: Dimension Specification, Target Competencies, and Sample Measurement Items for the SAILS Instrument

Dimension	Primary Focus	Target Trait / Competency	Key Threat Mitigated	Sample Survey Indicator (5-Point Likert)	Core Reference Support
Anthropomorphic De-biasing	Social-Cue Awareness	Decoupling conversational warmth from objective factual accuracy	Mindless anthropomorphism, social flattery, and uncritical emotional bonding	<i>"Even if an AI responds in a highly polite and friendly tone, I do not assume its information is accurate."</i>	Gambino et al. (2020); Heyselaar (2023); Lee (2024); Nass et al. (1994)
Privacy Triad Calibration	Privacy Awareness	Classifying agent vs. corporate conduit vs. tool roles	Inadvertent data leakage, boundary erosion, and corporate harvesting	<i>"I avoid typing sensitive personal details into this AI because I know my data may be stored on corporate servers."</i>	Ghosh & Eastin (2020); Keleher et al. (2026); Petronio (2002)
Algorithmic Authority Auditing	Critical Evaluation	Overriding the "machine heuristic" via System 2 analytical processing	AI over-reliance, automated authority bias, and passive acceptance	<i>"I actively question the reasoning behind an AI's answer rather than assuming its output is unbiased."</i>	Kahneman (2011); Lee (2024); Shin et al. (2022); Sotero et al. (2026)

Epistemic Vigilance	Epistemic Trust Calibration	Active verification, lateral reading, and source cross-checking	Uncritical acceptance of generative AI hallucinations and filter bubbles	<i>"I regularly open a separate browser tab to double-check AI-generated claims against external sources."</i>	ACRL (2015); Doyle et al. (2019); OECD (2026); OECD & European Union (2026)
----------------------------	-----------------------------	---	--	--	---

Expected Contributions & Impact

The present study contributes to information science literature in four ways: theoretical, methodological, in terms of national policy, and in terms of educational practice. Theoretically, this work builds on information evaluation frameworks traditionally used (ACRL, 2016) to introduce the idea of Socio-Algorithmic AI Literacy through the adoption of the paradigm of Computers Are Social Actors (CASA) (Gambino et al., 2020). SAILS expands AI literacy from mere technical literacy and ethics to the capacity to detect and deal with social and algorithmic vulnerability. Methodologically, it contributes to educational technology, human-computer interaction, and library science scholarship with a psychometrically validated standardized diagnostic instrument (SAILS) developed using the most rigorous scale development procedures (MacKenzie et al., 2011). From the perspective of national policy, it is an important contribution since it offers MOHE Malaysia, university senates, and academic administrators a diagnostic measuring device for students' AI literacy in line with national strategic plans (e.g., National Guidelines on Artificial Intelligence Governance and Ethics (MOSTI), 2024). Finally, from a pedagogical standpoint, the findings empower university librarians and instructional designers to move beyond teaching generic software commands, enabling the creation of targeted workshops that address students' socio-algorithmic vulnerabilities and institutional AI integration needs in Malaysian academic libraries (Masrek et al., 2025).

Conclusion

With the pervasive penetration of Generative AI systems into education at the post-secondary level, assessment of AI literacy should go beyond assessing the skill of students in prompting and their general knowledge of ethics. Contemporary conversational AI tools often prompt social heuristics automatically, making humans vulnerable to over-trust, over-reliance, and accidental revelation of their private data. This article develops the theoretical basis and methodological blueprint of the Socio-Algorithmic AI Literacy Scale (SAILS), which is grounded in the CASA paradigm, Dual-Process Theory, and the Information Literacy model. The scale is based on four dimensions of socio-algorithmic literacy: Anthropomorphic De-biasing, Privacy Triad Calibration, Algorithmic Authority Auditing, and Epistemic Vigilance. In the development of SAILS, an intensive four-stage psychometric process of expert content validation, exploratory factor analysis, confirmatory PLS-SEM, and MICOM of the instrument for measuring the construct invariance of students at Malaysian public (IPTA) and private (IPTS) universities, this study will produce a validated diagnostic tool. Eventually, SAILS will be able to offer educators, librarians, and MOHE policymakers a mechanism that can assess students' preparedness and help plan educational interventions that create epistemically astute information users in an era of increasing AI.

-
- Acknowledgements:** The authors would like to express their sincere gratitude to Universiti Teknologi MARA for providing the necessary resources and support throughout the course of this research.
- Funding Statement:** No Funding
- Conflict of Interest Statement:** The authors declare no conflict of interest regarding the publication of this paper. All authors have contributed to this work and approved the final version of the manuscript for submission to the Journal of Information System and Technology Management (JISTM).
- Ethics Statement:** The authors would like to acknowledge the use of artificial intelligence tools in the preparation of this manuscript. Specifically, Google Gemini, Scopus AI, and NotebookLM were utilized as collaborative tools to assist in analyzing secondary research data, synthesizing qualitative insights, and structuring the drafting of the findings, discussion, and expected contribution and impact. This study was conducted in accordance with accepted academic integrity and ethical publishing standards.
- Author Contribution Statement:** All authors contributed significantly to this manuscript. Zuraidah Arif and Abd Latif Abdul Rahman were responsible for the conceptualization, methodology, data collection, analysis, and literature review. Zati Atiqah Mohamad Tanuri and Nor Azlina Azmi contributed to the drafting, and critical revision of the manuscript. All authors read and approved the final version of the manuscript before submission.
-

References

- Almatrafi, O., Johri, A., & Lee, H. (2024). A systematic review of AI literacy conceptualization, constructs, and implementation and assessment efforts (2019–2023). *Computers and Education Open*, 6, 100177. <https://doi.org/10.1016/j.caeo.2024.100177>
- Association of College and Research Libraries (ACRL). (2016). *Framework for information literacy for higher education*. Association of College and Research Libraries.
- Association of College and Research Libraries. (2016). *Framework for information literacy for higher education*. Association of College and Research Libraries.
- Bawden, D., & Robinson, L. (2022). *Introduction to information science* (2nd ed.). Facet Publishing.
- Carolus, A., Schmidt, C., Schneider, F., Mayr, J., & Muench, R. (2018). Are people polite to smartphones?: How evaluations of smartphones depend on who is asking. In *Lecture Notes in Computer Science* (Vol. 10902, pp. 500–511). Springer. https://doi.org/10.1007/978-3-319-91244-8_39
- Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). John Wiley & Sons.
- DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications* (5th ed.). SAGE Publications.
- Do, H. V., Bui, T. T., Hoa Tran, D., & Dinh, L. D. (2026). Developing and validating an AI literacy scale for university students in Vietnam. *IFLA Journal*, 52(2), 222–233. <https://doi.org/10.1177/03400352251310123>
- Doyle, M., Foster, B., & Yukhymenko-Lescroart, M. A. (2019). Development and testing of the Perception of Information Literacy Scale (PILS). *The Journal of Academic Librarianship*, 45(6), Article 102064. <https://doi.org/10.1016/j.acalib.2019.102064>
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>
- Gambino, A., Fox, J., & Ratan, R. A. (2020). Building a stronger CASA: Extending the Computers Are Social Actors paradigm. *Human-Machine Communication*, 1(1), 71–85. <https://doi.org/10.30658/hmc.1.5>
- Ghosh, C., & Eastin, M. S. (2020). Understanding users' relationship with voice assistants and how it affects privacy concerns and information disclosure behavior. *Lecture Notes in Computer Science*, 12210, 381–392. https://doi.org/10.1007/978-3-030-50309-3_25
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). SAGE Publications.
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing measurement invariance of composites using PLS. *International Marketing Review*, 33(3), 405–431. <https://doi.org/10.1108/IMR-09-2014-0304>
- Heyselaar, E. (2023). The CASA theory no longer applies to desktop computers. *Scientific Reports*, 13(1), Article 19693. <https://doi.org/10.1038/s41598-023-46527-9>
- Howard, M. C. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *International Journal*

- of *Human-Computer Interaction*, 32(1), 51–62.
<https://doi.org/10.1080/10447318.2015.1087664>
- Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36(2), 109–133. <https://doi.org/10.1007/BF02291393>
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36.
<https://doi.org/10.1007/BF02291575>
- Keleher, M., Baig, K., & Chiasson, S. (2026). The privacy triad: Understanding the influence of perceived social agency on privacy attitudes. *International Journal of Human Computer Studies*, 210. <https://doi.org/10.1016/j.ijhcs.2026.103752>
- Koch, M. J., Carolus, A., Wienrich, C., & Latoschik, M. E. (2023). MAILS-Meta AI literacy scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change-and meta-competencies. *Computers in Human Behavior: Artificial Humans*, 1(2).
<https://doi.org/10.1016/j.chbah.2023.100014>
- Lee, E.-J. (2024). Minding the source: Toward an integrative theory of human-machine communication. *Human Communication Research*, 50(2), 184–193.
<https://doi.org/10.1093/hcr/hqad034>
- Liu, X., Zhang, W., & Zhang, Y. (2025). Generative Artificial Intelligence Literacy: Scale development and its effect on job performance. *Digital Government: Research and Practice*, 6(1), 1–21. <https://doi.org/10.1145/3708123>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery.
<https://doi.org/10.1145/3313831.3376727>
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–386. <https://doi.org/10.1097/00006199-198611000-00017>
- MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in theory testing and development for leadership research. *The Leadership Quarterly*, 22(2), 293–334. <https://doi.org/10.1016/j.leaqua.2011.02.006>
- Masrek, M. N., Mohamad Rosman, M. R., Mohamed Shuhidan, S., Baharuddin, M. F., Mutia, F., Soesantari, T., Yuwinanto, H. P., & Atmi, R. T. (2025). Artificial intelligence readiness in Malaysian libraries: An assessment framework. *Journal of Applied Science, Engineering, Technology, and Education*, 7(2), 181–194.
<https://doi.org/10.35877/454RI.asci3875>
- Memon, M. A., Ting, H., Cheah, J.-H., Thurasamy, R., Chuah, F., & Cham, T. H. (2020). Sample size for survey research: Review and recommendations. *Journal of Applied Structural Equation Modeling*, 4(2), i–xx. [https://doi.org/10.47263/JASEM.4\(2\)01](https://doi.org/10.47263/JASEM.4(2)01)
- Ministry of Science, Technology and Innovation. (2024). *National guidelines on artificial intelligence governance and ethics (AIGE)*. Government of Malaysia.
<https://www.malaysia.gov.my/en/my-initiative/perkhidmatan-digital-keseluruhan-kerajaan/tadbir-urus/national-guidelines-on-artificial-intelligence-governance-and-ethics-aige>
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103. <https://doi.org/10.1111/0022-4537.00153>
- Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 72–78). Association for Computing Machinery. <https://doi.org/10.1145/259963.260288>

- Ngo, V. (2026). Humanizing AI for trust: The critical role of social presence in adoption. *AI and Society*, 41(3), 1803–1819. <https://doi.org/10.1007/s00146-025-02506-4>
- Ngo, V. (2026). Humanizing AI for trust: The critical role of social presence in adoption. *AI & SOCIETY*, 41(3), 1803–1819. <https://doi.org/10.1007/s00146-025-02506-4>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
- OECD/European Union. (2026). *Empowering learners for the age of AI: An AI literacy framework for primary and secondary education*. OECD Publishing. <https://doi.org/10.1787/65cd27d4-en>
- Organisation for Economic Co-operation and Development. (2026). *PISA 2029 Media and Artificial Intelligence Literacy (MAIL) Assessment Framework* (First Draft). OECD Publishing.
- Petronio, S. (2002). *Boundaries of privacy: Dialectics of disclosure*. State University of New York Press.
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, 30(4), 459–467. <https://doi.org/10.1002/nur.20199>
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.
- Ringle, C. M., Wende, S., & Becker, J.-M. (2024). *SmartPLS 4* [Computer software]. SmartPLS. <https://www.smartpls.com>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill-building approach* (7th ed.). Wiley.
- Shin, D., Rasul, A., & Fotiadis, A. (2022). Why am I seeing this? Deconstructing algorithm literacy through the lens of users. *Internet Research*, 32(4), 1214–1234. <https://doi.org/10.1108/INTR-02-2021-0087>
- Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J.-H., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing*, 53(11), 2322–2347. <https://doi.org/10.1108/EJM-02-2019-0189>
- Sotero, J. M., Vicente, P. N., & Granado, A. (2026). Development and validation of a scale to assess algorithmic literacy in the context of recommender systems (ALRS). *Technology in Society*, 85, Article 103237. <https://doi.org/10.1016/j.techsoc.2026.103237>
- Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. Metzger & A. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 73–100). The MIT Press.
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.
- Xu, K., Chen, X., & Huang, L. (2022). Deep mind in social responses to technologies: A new approach to explaining the Computers are Social Actors phenomena. *Computers in Human Behavior*, 134, 107321. <https://doi.org/10.1016/j.chb.2022.107321>
- Yusnilzahri, N. A. S., Rosman, M. R. M., & Bahry, F. D. S. (2025). AI literacy among undergraduates: Determinants and productivity outcomes in a Malaysian comprehensive university. *Malaysian Journal of Library and Information Science*, 30(1), 45–68. <https://doi.org/10.22452/mjlis.vol30no1.3>

Zhou, X., Lee, H., & Huang, Y. (2025). Defining, enhancing, and assessing artificial intelligence literacy and competency in K-12 and higher education: A systematic review. *Interactive Learning Environments*, 33(4), 1102–1125. <https://doi.org/10.1080/10494820.2025.2487538>