Systematic Review of AI Literacy Frameworks and Assessments

Sador Wondeson

Submitted under the supervision of Dr. Panayiota Kendeou to the University Honors Program at the University of Minnesota - Twin Cities in partial fulfillment of the requirements for the degree of Bachelor of Arts, magna cum laude in Psychology.

Acknowledgments

I would like to first and foremost thank my awesome advisor for offering this topic to me when I was very behind in my thesis planning. She also checked on my progress throughout this process, which helped me build courage when I was behind. She is the reason I was invited to join the R+L lab during my late planning in *PSY 4994V Honors Research Practicum*. Her name is Panayiota Kendeou. I also show gratitude to Victoria Johnson and Ali Fuescher for providing me with systematic review resources and offering to meet with me if I needed assistance. My fellow lab member, Luan Chau, helped enrich my introduction. Finally, I thank Amy Riegelman for guiding me through the source extraction process.

Abstract

As AI technologies increasingly permeate various aspects of society, there is a growing need to understand and evaluate AI literacy across diverse populations. Optimizing how we instill AI literacy in individuals defines the thin line between responsible/effective and irresponsible/ineffective usage of AI. To inform this work, in this paper, I conduct a comprehensive systematic review of theoretical frameworks, models, and assessment tools related to AI literacy. This review synthesized existing literature on AI literacy conceptualizations and measurements, addressing a critical gap in our understanding of how AI literacy is defined, taught, and assessed. Analysis of sixteen studies that reached the final stage of screening left prints of common themes relating to elements of frameworks adhered to in assessments, the popularity of assessment modalities used, and how frameworks and assessments built on other frameworks and assessments, respectively. Frameworks utilizing cognitive, evaluative, and sociocultural components were predominantly recognized. In comparison, most assessments included questionnaire items with expert evaluation. Studies geared toward academic applications implemented AI-assisted programs that teach students where AI belongs in their educational pursuits.

Systematic Review of AI Literacy Frameworks and Assessments

The predominant angle by which AI is interpreted is one of apparatus, efficiency, and power. The roughly five-decade-old field has been developed by various disciplines in the arts and sciences (Casal-Otero et al., 2023). Fortunately, humans remain the dictators of the usage of AI. Learning the vast abilities of this technology is void insofar as its pragmatic implementation stays locked. Rather than a mere understanding of coding, AI literacy posits cognitive, critical, and communicative competencies for students to attract to optimally integrate AI into their work (Allen & Kendeou, 2023). AI literacy is the very concept that handles the methods of understanding, evaluating, and incorporating AI into the broader sense of societal functioning. Nearly 50% of American jobs are at serious risk of automation as 2030 approaches (Ng et al., 2021). Preparing the workforce for the anticipated transition is of top priority. We are effectively boosting ourselves forward with higher efficacy if we prioritize an ethical modality that keeps us at full liberty to control AI and distribute techniques for handling it.

Optimizing our knowledge of AI literacy requires a foundation built on a consolidated definition of AI and literacy. Positive and negative effects come from GenAI systems because they support students with engaging activities but also create space for carrying them past genuine comprehension (Chiu & Sanusi, 2024). Classifying AI literacy as comprehending a specific AI ability is incomplete and unjust. AI has immense, complex applications such as household appliances, recommendation systems, and aiding medical diagnoses (Ding et al., 2024). Especially considering its vast involvement, defining AI literacy is no easy task. For the purposes of this paper, to reflect on the literature on the subject, AI is how technological machinery performs work constructed (or programmed) by humans, whereas digital literacy is the comprehension and subsequent utilization of digital tools (Chiu et. al., 2024). Synthesizing these definitions creates space for the growing need to evaluate AI literacy. A near engineering definition is far from useful in light of the advent of emerging technologies such as ChatGPT (Chiu et. al., 2024). Thereby emphasizing the earlier stages in which AI currently resides. As a hoped greater encompassing conceptualization of AI literacy arises, more aspects of its function can be included in the definition. Areas such as education see AI literacy highlighting teaching efficacy, while computer science emboldens coding efficiency. In both regards, we are left with different ethical concerns.

Educating professionals and the general community is the imperative input for generating societal flourishing. Fields often correlated with AI progression happen to be those with higher salaries. Computer science, robotics engineering, data science, etc., increase the likelihood of complacency regarding social responsibility and ethical adherence in light of lucrative compensation (Ng et al., 2021). Safe advancements in AI are on track to ensue insofar as those working with it have rigid ethical foundations behind their work. Again, whether it be computer science or education, the interpretation of "literacy" depends on the skills required in those jobs (Ng et al., 2021). Teachers also carry the potential to misuse technology. One stark example is modeling disingenuous work to their students. Students rely on their teachers to learn the basis of legitimate and illegitimate resources for their assignments. Teachers ought not propagate unreliable AI sources or leave ambiguity surrounding the extent to which AI usage remains beneficial to the students' learning. Future interpretations of AI are shaped by its representation in media, language, and school, and previous beliefs of AI also shaped by certain presentations (Bewersdorff et al., 2023). All of which urge educators to adopt effective AI-implemented teaching strategies.

Generationally speaking, the current day presents unprecedented rates of academic dishonesty, escalated by AI assets (Ndungu, 2024). Refining the role of AI in the classroom is consequently all the more essential. While much research has touched on augmenting teachers' AI literacy, there is a gap concerning maintaining positive learning environments (Chiu & Sanusi, 2024). There may be more use in distinguishing between attitudes toward AI use in classrooms and AI literacy. Extorting a methodology by which both can be measured might be relatively unproductive. Understanding self-reported surveys and questionnaires as the currently dominating tools for measuring AI literacy is an important guide for replacing them with tests (Chiu & Sanusi, 2024; Chiu et al., 2024). While self-reported surveys and questionnaires provide useful insight into teachers' and students' attitudes toward AI, they lack rigorous candor regarding the recall of AI literacy. Questionnaires are notoriously unreliable when addressing memory, as research in post-traumatic and growth beliefs would suggest. Questionnaires also need to compensate for developmental differences, which include memory appraisal improving with age (Ng et al., 2023). There are understanding-assessing links with questionnaires, however, another mechanism is needed to test if participants possess the skills necessary to determine instruction success (Ng et al., 2023). Overreliance on truthful self-reporting is another shortcoming that supports the point of pairing questionnaires with post-tests. In an effort to avoid variables such as faulty memory, misinterpretation of terms, and false appraisal of competence, some research aims at constructing courses and subsequent exams to measure knowledge of AI literacy and competence.

In pursuit of explicating the field in question, a systematic review was conducted highlighting the frameworks in place to describe how AI literacy can be understood and taught, as well as assessments performed to test AI literacy frameworks. I hope to synthesize gaps concerning understandings of AI literacy, frameworks and their corresponding efficiencies, themes and deviances, and the applicability of assessments in measuring AI literacy. The results of this review will demonstrate the main ideas of how we explain AI literacy, plus why we prioritize certain elements of such a complex network of abilities.

Method

Sources extraction

The article search for this study transpired on February 14, 2025. To ensure a reliable literature search, several keyword queries were inputted using the ERIC database, a vast database that yielded sufficient results for this review. Keyword query operators are contained, such as title (TI), abstract (AB), and descriptor (DE). The following searches and corresponding results are as follows: TI (Framework or Assessment) OR AB (Framework OR Assessment) (257,412), TI “Literacy” AND AB “Literacy” (26,433), DE “Literacy” OR DE “Adult Literacy” OR DE “Workspace Literacy” (25,907), DE “Performance Based Assessment” OR DE “Curriculum Based Assessment” (8,123), DE “Artificial Intelligence” (5,855), TI “Artificial Intelligence” OR AB “Artificial Intelligence” (3,319). Finally, a filter was added to contrive sources that met all the criteria above, leaving n=57.

Eligibility criteria

A total of 57 initial studies were pooled for the screening process. 0 articles were classified as duplicates according to the automatic data summary given by Rayyan. Three and four inclusion and exclusion criteria were established during pre-screening, respectively. For inclusion in the review: 1. Specify the target population related to AI usage; 2. Full text available; 3. Involve performance/curriculum-based assessment or AI literacy framework. For exclusion from the review: 1. Language other than English; 2. AI literacy not being the primary focus; 3. Unclear methodology; 4. Incomplete/insufficient data.

Title and abstract screening placed 37 studies into exclusion given an absence of or implication of discussing the term “AI literacy”. After analyzing the eligibility criteria, 4 studies were excluded due to not being AI literacy-focused, by either not examining AI literacy frameworks or discussing the results of AI literacy assessments. In pursuit of original ideas of the state of research, two excluded studies were systematic reviews themselves. The final pool produced 16 studies eligible for inclusion in the systematic review. The PRISMA flow chart diagram (Fig. 1) below pictures the entire screening process.

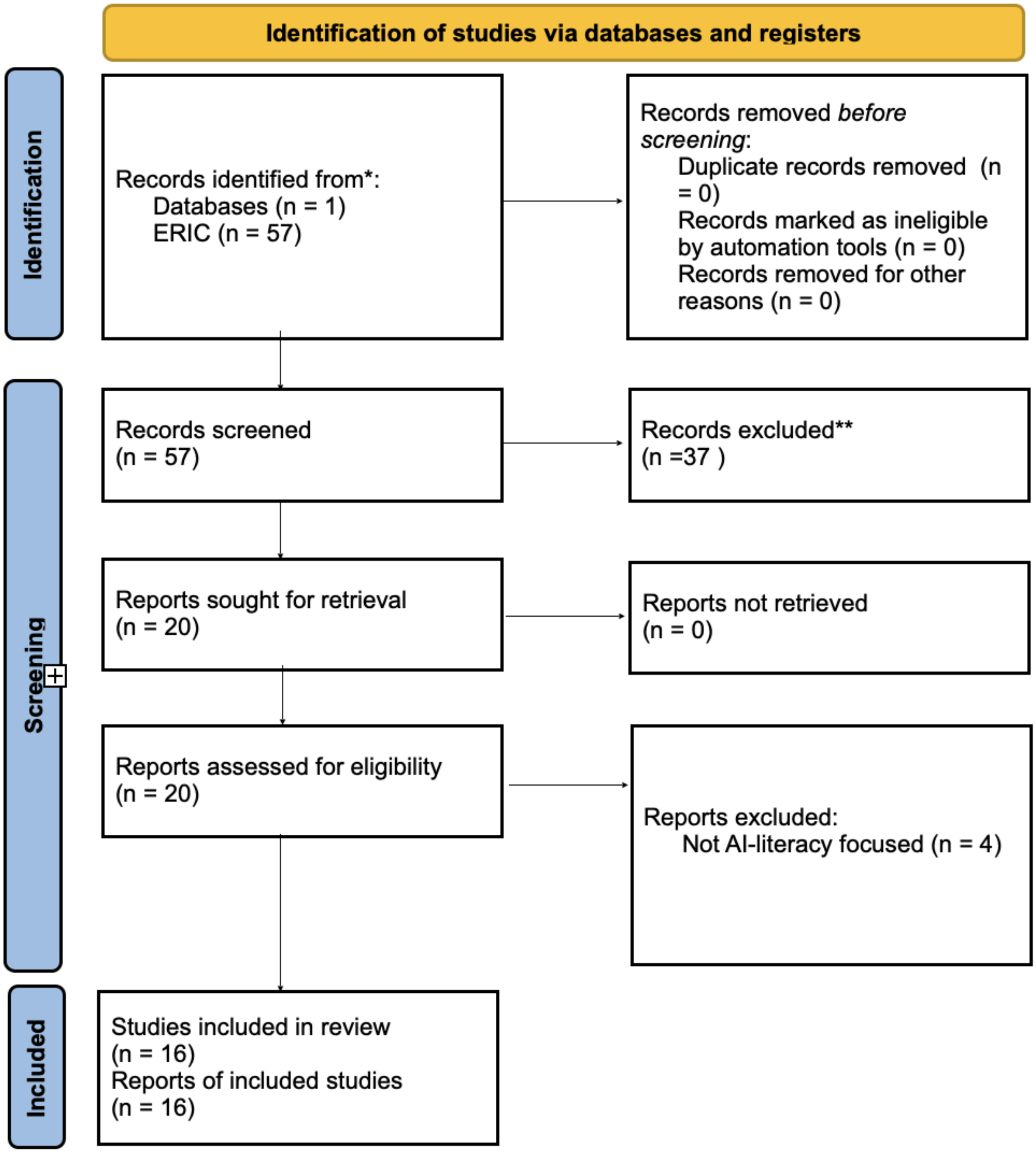

Fig. 1. Step-by-step diagram listing the process of screening studies.

Categories and Themes

First thought to note is the tendency to tailor AI literacy frameworks and assessments to specific targeted populations. Seven of the studies framed around higher education populations, including high schoolers and college students, while two studies targeted kindergarteners. The remaining seven studies in this review focused on non-student-related members of society, such as teachers, librarians, employees, and military members. Therefore, more than 50% of the studies targeted students, and roughly 75% surrounded the conversation of education. In terms of publication years, the demographic scene of AI literacy is most prevalent in recent years. Fourteen of the studies were published in 2024, with two studies published in 2023. Merging demographic and periodic themes leaves the conclusion that AI literacy is of the utmost interest, especially in recent years, among educational spheres. This isn't to say, however, that AI literacy frameworks and assessments are exclusively useful in education. The remainder of this paper partly suggests the individual and societal benefits of AI literacy frameworks and assessments, which overlap with education-focused frameworks and assessments.

Framework Outlines for AI Literacy

Zhou and Schofield (2024) support a four-step framework for implementing AI-tooled learning activities into teaching. Firstly, knowing and understanding AI allows students to build a foundational labyrinth of AI's latitude and capabilities. Students are encouraged to deepen their exploratory comprehension to make their AI-assisted note-taking and subsequent reflection of materials conducive. Improved brainstorming and organized mapping ought to yield an enhanced version of the writing process with better grammar, clarity, and an optimal final form for assignments. Secondly, using and applying AI is suitable for achieving all stages of the research process. Whether it be source gathering, data visualization/analysis, or relevant content design. Thirdly, evaluating and creating with AI elicits critical thinking skills for tackling perplexing information. Lastly, students must be taught the limitations of AI, suggesting a responsible justification for its ethical use.

In addition to stages that organize phases of AI use in academic practice, AI literacy can be split into two categories: knowledge practices and dispositions. Ndungu (2024) attempted to map AI literacy onto the ACRL Framework for Information Literacy for Higher Education by adhering to said categories and relating them to basic framework principles of information literacy. Strategic searching, valuing information, and viewing information creation as a process are all sentiments of the ACRL that Ndungu (2024) believes can be accelerated by AI knowledge practice and dispositions. McBride et al. (2024) also view AI literacy as sets of practices that can be implemented, further supporting the notion that AI literacy is not merely a comprehension of AI abilities.

Further pairing knowledge of AI with integrated practice, the Technological Pedagogical Content Knowledge (TPACK) framework transitions to understanding AI content and pedagogy as connected, as opposed to AI as an additional mechanism. The framework emphasizes applying its three tenets to practical contexts in education. However, Lo (2024) expanded the scope of the framework to academic library professionals, who are on the journey to cycle AI literacy back to information seekers. Of the three knowledge principles, technical knowledge concerns understanding AI's abilities as a technology, while not necessarily viewing AI as a distinct application from library settings. Pedagogical knowledge is translational knowledge of AI's implementation in and resulting enhancement of library services and learning facilitation. Finally, content knowledge is the second part of pedagogical knowledge because it regards fathoming the mission of library services and exactly how AI serves those services' delivery.

The multifaceted nature of AI literacy can also be understood through a dimensional approach. After reviewing the literature, Ng et al. (2023) gathered a mnemonic categorization called the ABCE framework, which consists of four learning domains. The affective domain deals with mood-related impetuses, such as intrinsic motivations like self-efficacy and extrinsic motivations like career interest. Behavioral learning refers to performance that suggests intention, management, and collaboration. The cognitive domain consists of a pyramid of level-ordered cognitive skills, akin to what is seen in Bloom's taxonomy (1956). An AI ethics domain consolidates the other three learning domains into a sagacious usage cognizant of AI's rapid development and risks of misuse.

Kong et al. (2023) also used a dimensional-conceptual framework in their assessment of AI literacy program development. There are shared affective and cognitive domains among them; however, there is no behavioral dimension, and the AI ethics dimension is called the sociocultural domain. Although similar, the domain gets its name in part because there is a focus on upholding human autonomy. Additionally, the benefits that spring forth from AI are to be dispersed equally among people so that a positive benefits-to-risks ratio of AI is accessible to all. On the other hand, Yuan et al. (2024) tuned a conceptual framework without an affective component but included a behavioral component in their study aimed at measuring competence/proficiency in AI literacy.

Key Elements

As is custom with frameworks, the studies in this systematic review can be divided into core competencies or domains. Ubiquitous presentations of knowledge principles were seen across the studies. As previously mentioned, the AI literacy frameworks discussed that used a dimensional approach had variance. Variability was seen in the inclusion and measurement of affective and behavioral components. All dimensional approaches in this review involved a cognitive domain and an AI ethics domain. This is to say that conceptions of AI literacy deem cognitive elements, or understandings of AI's abilities, as well as socioculturally responsible uses, as worthy of high regard for measurement.

Different layouts of the frameworks did not necessitate different content. Some of the frameworks list core competencies that encompass AI literacy, while others categorize them by function or intentional phase. For example, some frameworks that targeted educational implementation of AI literacy organize competencies, such as critically evaluating how AI could assist with a particular assignment, into a "thought process" group, while another framework organizes that competency as a learning outcome, say after a phase about learning objectives for an AI literacy assessment. Frameworks that contain more categories or phrases appear more exhaustive, all the while sharing many of the same ideas as others.

In conclusion, prevalence percentages can be determined for core constructs found in the studies. Approximately eighty-one percent of the studies featured knowing and understanding AI, using and applying AI, and evaluating and creating AI. Twenty-five percent discussed cognitive, AI ethics, and affective/behavioral dimensions. Another twenty-five percent discussed contextualizing (not in the sense of AI ethics) as a core construct and competency of AI literacy. This was seen in building comprehension of the integrative power of AI, unveiling different AI-powered elements to discover how to contextualize, developing non-technical and critical deciphering competencies, and engaging in practices and dispositions that address certain AI sentiments.

Evaluating Quality of Assessments

Ascertaining the appropriateness of items in assessments is necessary for determining the veracity of results. Therefore, a useful methodology for appraising meaningfulness is through measuring specific validities. The rest of this section is dedicated to defining the types of validity and reliability calculated in the studies. Construct validity measures the consistency between the scores of the assessment and the underlying hypothesis of the assessment. Content validity measures how well the assessment covers the entire construct. Face validity measures how well scores address the intended independent variable. Discriminant validity measures how discriminant one measure is from another measure, keeping unrelated constructs distinct. Convergent validity measures how much different instruments converge with their results when targeting similar constructs. Criterion validity measures how well an instrument correlates with an outcome. Composite reliability measures the internal consistency among a set of similar measures. Item-total reliability measures the extent to which an item correlates with the total score of an assessment. Split-half reliability measures how closely two split halves of a set of items measure the same construct.

Assessments Overview

One assessment with the objective of evaluating the development level of AI literacy in kindergarteners is the AI4KG children's AI literacy assessment designed by Su (2023). It takes the format of 16 multiple-choice questions, each with one answer. Each question was one of the items of the assessment that was read aloud to avoid literacy level variables. Prior to evaluating the participants, experts who researched AI literacy concerns or education professionals focusing on AI literacy were procured to undergo expert evaluations intended to measure construct validity, content validity, and face validity. Experts were surveyed through questions related to the evaluation of overall questions of the assessment, evaluation of the overall assessment, and an overall evaluation of the assessment. For AI4KG, construct validity took the form of measuring the kindergarteners' basic AI knowledge. Content validity measured whether the assessment was suitable for kindergarteners. Face validity determined whether the assessment's ability to clearly convey the questions and options to participants was appropriate. The curriculum spanned eight 35-minute sessions. Su (2023) used Cronbach's α to determine reliable construct validity, content validity, and face validity, although the list of 16 questions is not explicitly listed.

Yuan et al. (2024) contrarily used a similar AI literacy scale consisting of an expert-revised questionnaire but with a more diverse population. Participants came from different age groups and different educational backgrounds. Six constructs were organized and measured under cognitive, behavioral, and normative dimensions. For validity: construct, discriminant, convergent, and criterion are scored for AI features, AI processing, algorithm influences (cognitive), user efficacy (behavioral), ethical consideration and threat appraisal (sociocultural/normative). For reliability: composite, item-total, and split-half are scored for the same constructs.

Lo (2024) conducted an assessment in accordance with TPACK concepts such as technological and pedagogical convergence. Posting on professional listservs recruited a wide variety of librarians who would receive a survey targeting four research questions (RQ). RQ1 aimed to record familiarity with AI concepts and terminology by averaging scores on the 5-point Likert scale. RQ2 intended to highlight gaps in confidence in applying AI to their respective professional contexts. The percentages of respondents were calculated at each of five confidence levels in various administrative tasks, including evaluating the ethical implications of AI, participating in AI discussions, collaborating on AI projects, troubleshooting AI tools, and providing guidance on AI resources. RQ3 investigated perceptions surrounding integrating generative AI tools in library services and operations. Participants were asked statements about their perceptions on a 5-point Likert scale, with percent agreement recorded.

Younis (2024) also delivered a questionnaire that adheres to TPACK but additionally deployed a pretest-posttest quasi-experimental design aimed at explicating the efficacy of the AI Literacy Professional Development Program (PDP). The program lasted twenty-four training hours across three sessions. PDP contains serialized training modules that improve how well participants, who are pre-service teachers, incorporate AI mechanisms. The first stage instilled rudimentary knowledge about AI, followed by practicing such learning outcomes with hands-on ChatGPT usage. Training eventually led to the pre-service teachers creating lesson plans involving ChatGPT responsibly.

A dense questionnaire known as AILQ consists of sixty items. Learning outcomes are studied after a curriculum comprising AI learning activities terminates. In alignment with the ABCE framework previously analyzed, the questionnaire's items are divided into affective, behavioral, cognitive, and ethical dimensions. Ng et al. (2023) applied AILQ in tandem with an AI literacy program for students, which serialized twelve lessons of AI literacy workshops spanning over two months, culminating in a 5-person group project. The first ten lessons covered a topic each, while the last two were reserved for the presentation project. Many of the questionnaires used in AI literacy curricula for K-12 students lack validation (Ng et al., 2023). Therefore, the researchers implemented a four-step validation process beginning with expert revisions of the questionnaire. Then, they ensured the interviewer of the questionnaire would respond to clarity requests from the students. After, the program was piloted in a secondary school setting as designed. Finally, the results were calculated preliminarily and officially to account for extreme and missing data.

Core themes

It was clear that the assessment styles in the studies presented several similarities. Assessments that employed questionnaires to participants commonly processed expert evaluations in order to refine questionnaire items. Such a methodology is suitable for eliminating certain miscomprehensions from participants, which experts may be more keen on identifying. In practice, Ng et al. (2024) had experts discern that students would not correctly discriminate certain statements within the cognitive dimension items in their questionnaire. As is customary with the majority of assessments in this review, the presence of an expert overview ought to be a defining criterion in the quality of an assessment (Ng et al., 2024). Additionally, experts typically took control over roles dealing with calculating validities and reliabilities. All in all, expert evaluations are a decisive tool to accurately impugn the validity of assessment items in related research (Su, 2024).

Discussion

This systematic review aimed to synthesize the frameworks, models, and assessment tools concerning AI literacy in a theme-oriented fashion. Such a lens productively combated the need to understand and evaluate AI literacy across diverse populations and domains. Similarities among the frameworks and assessments expounded in this review suggest the deduction that we are approaching a stable grasp of evaluating AI literacy among various groups and are simultaneously nearing an optimized method to teach AI literacy. As much research has tested conceptions of AI literacy frameworks through assessments, curricula, etc., this review attempted to highlight key approaches from different interventions and populations, subsequently gathering implications for implementing them. In the process, informing how research thus far has addressed our crucial gap in understanding AI literacy at a commensurate pace with AI's imbuing in various aspects of society.

It was found evident that, as for AI literacy frameworks, cognitive dimensions consisting of understanding, evaluating, and applying AI components were of high favoritism. Whereas for AI literacy assessments, questionnaires were frequently prioritized as the main instrument for gathering rigorous data. However, the used instruments occasionally enforced combinations of interventions. Most commonly seen was a questionnaire-curriculum dual intervention. Other combinations were presented in framework development when new frameworks were built on previously established components from other frameworks. Younis (2024) referenced the new framework developed by Ng et al. (2023), the AI Literacy Professional Development Program (PDP), which combined digital competency premises from DigCompEdu and comprehensive understandings of non-technical, critical skills from TPACK. This supports the notion that overlap and combinations from well-researched ideas in the literature of AI literacy yield respectable results. Ndungu (2024) modifies traces of established AI literacy framework sentiments by applying practices and dispositions, and Ng et al. (2023) add learning objectives to frameworks, serving as evidence that established frameworks and assessments carry gaps in their pursuit to address gaps.

Solving said consequent gaps may require instilling learning objectives in not only researchers of AI literacy but also teachers implementing AI literacy so as to keep their influence on AI effectiveness in check (Casal-Otero et al., 2023). Learning objectives are a fundamental component of lesson plans, which are themselves a learning objective of the AI Literacy Professional Development Program (PDP). Training geared toward teachers implementing lesson plans with the assistance of ChatGPT responsibly embodies learning objectives (this time for teachers) as a crucial add-on to merely understanding AI concepts.

Limitations and Future Direction

This systematic review derived sources from one comprehensive database (ERIC), though it could have benefited from other literature databases and reference lists from contrived sources to collect more information concerning AI literacy. There was also an enhanced focus on gathering commonalities across sources through age groups, education levels, and occupations. It is possible that the search query was limited in obtaining sources from different populations due to a focus on numerous types of literacy and assessment formats. Furthermore, this review contained all but one, academic reports published in official academic journals, leaving out the chance of valuable research in other non-academic reports.

Future systematic reviews can benefit from discovering connections between populations rather than solely within. Replication or similarly applicable studies can greatly reap positive outcomes from pooling a larger sample size. Although this study had a pre-determined focus on specific assessment modalities, a broadened scope paired with a more exhaustive database collection will likely accumulate more information to synthesize. Additionally, with the help of a larger sample size yielding more elements of frameworks and assessments and their respective results, researchers' interests should be captured by formulating an optimal framework and assessment connection. Outside of the focus of this paper, yet still amusingly feasible, is the effort of comparing results from studies and producing the most encompassing conception and subsequent teaching of AI literacy.

Another point of concern for future direction is comparing studies and their populations to derive the most suitable assessment for different age groups. In line with the theory that frontal lobe development and consequent judgment vary across individuals paired in a population, be it students or workers, age-group assigned assessments may be profound data to collect. This way, specific AI literacy diversity is addressed in the avoidance of overgeneralizing which assessments specific populations are most receptive to. For example, studies indicated that kindergarteners required precise clarification of questions from questionnaires, not due to their student identity (population) but because of comprehension deficits linked to their age.

Conclusion

Overall, this systematic review analyzed and synthesized sixteen studies concerning the frameworks and assessments of AI literacy. Dating between 2023 and 2024, and targeting various populations, the studies rayed numerous approaches to conceptualizing AI and gathering data on the effectiveness of AI literacy measuring tools. Out of which the predominant dimensional, competency, and learning objective-based modalities, as well as questionnaire and program styles, were highlighted. Despite discussing the field's suffusion of several AI literacy techniques, the grand conclusion of this review supports a unifying future for the recipients of AI literacy training. As the field continues to test fits for different populations and educational systems, individuals are each at the mercy of a sturdy foundation of effective AI usage. Not only was the aim, but the results of this study were a significant step toward insights into AI literacy proposals in education and the workplace. Our heightened focus need not be on creating new frameworks and assessments but on building on the well-attested methods in this review to reach said practical proposals.

References

Allen, L. K., & Kendeou, P. (2023). ED-AI Lit: An Interdisciplinary Framework for AI Literacy in Education. *Policy Insights From the Behavioral and Brain Sciences*, *11*(1), 3–10. https://doi.org/10.1177/23727322231220339

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*, 100173. https://doi.org/10.1016/j.caeo.2024.100173

Bewersdorff, A., Zhai, X., Roberts, J., & Nerdel, C. (2023). Myths, mis- and preconceptions of artificial intelligence: A review of the literature. *Computers and Education Artificial Intelligence*, *4*, 100143. https://doi.org/10.1016/j.caeai.2023.100143

Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: a systematic literature review. *International Journal of STEM Education*, *10*(1). https://doi.org/10.1186/s40594-023-00418-7

Chiu, T. K., Ahmad, Z., Ismailov, M., & Sanusi, I. T. (2024). What are artificial intelligence literacy and competency? A comprehensive framework to support them. *Computers and Education Open*, *6*, 100171. https://doi.org/10.1016/j.caeo.2024.100171

Chiu, T. K., & Sanusi, I. T. (2024). Define, foster, and assess student and teacher AI literacy and competency for all: Current status and future Research direction. *Computers and Education Open*, 100182. https://doi.org/10.1016/j.caeo.2024.100182

*Ding, L., Kim, S., & Allday, R. A. (2024). Development of an AI literacy assessment for non-technical individuals: What do teachers know? *Contemporary Educational Technology*, *16*(3), ep512. https://doi.org/10.30935/cedtech/14619

*Griffin, J. F. (2023). Investigating Curriculum Design for Implementing Artificial Intelligence Literacy Education for Senior Leaders at the Military Education Institutes (Order No. 30317163). Available from ProQuest Dissertations & Theses Global; ProQuest Dissertations & Theses Global Closed Collection. (2798874849).

http://login.ezproxy.lib.umn.edu/login?url=https://www.proquest.com/dissertations-theses/investigating-curriculum-design-implementing/docview/2798874849/se-2

*Hazari, S. (2024). Justification and Roadmap for Artificial intelligence (AI) literacy courses in Higher education. *Journal of Educational Research and Practice*, *14*(1). https://doi.org/10.5590/jerap.2024.14.1.07

*Kong, S.-C., Cheung, W. M.-Y., & Zhang, G. (2023). Evaluating an artificial intelligence literacy programme for developing university students' conceptual understanding, literacy, empowerment and ethical awareness. *Educational Technology and Society*, 26(1), 16-30. https://doi.org/10.30191/ETS.202301_26(1).0002

Lintner, T. (2024). A systematic review of AI literacy scales. *Npj Science of Learning*, *9*(1). https://doi.org/10.1038/s41539-024-00264-4

*Lo, L. (2024). Evaluating AI Literacy in Academic Libraries: A Survey Study with a Focus on U.S. Employees. *College & Research Libraries*, *85*(5). https://doi.org/10.5860/crl.85.5.635

*McBride, C., Lee, C. H., & Soep, E. (2024). "Gotta love some human connection": Humanizing data expression in an age of AI. *Reading Research Quarterly*. https://doi.org/10.1002/rrq.550

*Ndungu, M. W. (2024). Integrating basic artificial intelligence literacy with media and information literacy in higher education. *Journal of Information Literacy*, *18*(2), 122–139. https://doi.org/10.11645/18.2.641

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education Artificial Intelligence*, *2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041

*Ng, D. T. K., Su, J., & Chu, S. K. W. (2023). Fostering Secondary School Students' AI Literacy through Making AI-Driven Recycling Bins. *Education and Information Technologies*, *29*(8), 9715–9746. https://doi.org/10.1007/s10639-023-12183-9

*Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2023). Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology*, *55*(3), 1082–1104. https://doi.org/10.1111/bjet.13411

*Picasso, F., Atenas, J., Havemann, L., & Serbati, A. (2024b). Advancing critical data and AI literacies through authentic and Real-World assessment design using a data justice approach. *Open Praxis*, *16*(3), 291–310. https://doi.org/10.55982/openpraxis.16.3.667

*Su, J. (2024). Development and validation of an artificial intelligence literacy assessment for kindergarten children. *Education and Information Technologies*, *29*(16), 21811–21831. https://doi.org/10.1007/s10639-024-12611-4

*Su, J. (2024a). Kindergarten Teachers' perceptions of AI Literacy Education for Young children. *International Journal of Technology and Design Education*, *34*(5), 1665–1685. https://doi.org/10.1007/s10798-024-09876-8

*Vallejo, M. C. G. (2024). Cómo entrenar tu dragon: developing a module to build up AI literacy of HE lecturers. *Journal of Learning Development in Higher Education*, *31*. https://doi.org/10.47408/jldhe.vi31.1278

*Younis, B. (2024). Effectiveness of a professional development program based on the instructional design framework for AI literacy in developing AI literacy skills among pre-service teachers. *Journal of Digital Learning in Teacher Education*, 1–17. https://doi.org/10.1080/21532974.2024.2365663

*Yuan, C. W., Tsai, H. S., & Chen, Y. (2024). Charting Competence: A Holistic Scale for Measuring Proficiency in Artificial Intelligence Literacy. *Journal of Educational Computing Research*, *62*(7), 1675–1704. https://doi.org/10.1177/07356331241261206

*Zhou, X., & Schofield, L. (2024). Developing a conceptual framework for Artificial Intelligence (AI) literacy in higher education. *Journal of Learning Development in Higher Education*, *31*. https://doi.org/10.47408/jldhe.vi31.1354