

EDUCATION

AI-Based Community Education Model to Improve Digital Literacy, Social Ethics, and Independent Learning

Jufri¹, Abdul Rahim², La Ode Sahrin Djalila³, Basri⁴, La Manila⁵, Gita Ramadani⁶

^{1,2,3,4,5} Pendidikan Agama Islam, Universitas Muhammadiyah Buton

ABSTRACT

Artificial intelligence (AI) is increasingly used outside formal educational institutions, but community members often encounter it without adequate critical literacy, ethical guidance, or self-directed learning skills. This study develops a human-centered AI community education model integrating digital literacy, social ethics, and self-directed learning. A structured integrative review was conducted using 25 peer-reviewed, theoretical, methodological, and international policy sources. Sources were selected for relevance to AI literacy, digital competence, responsible AI, adult and community education, and learner autonomy. A concept matrix and deductive-inductive thematic synthesis produced five findings: AI affordances depend on pedagogical purpose; digital literacy must include critical evaluation of AI-generated content and data awareness; social ethics functions as both an outcome and a governance safeguard; AI supports self-directed learning when scaffolding preserves learner agency; and implementation depends on co-design, equitable access, facilitation, and iterative evaluation. The proposed model comprises contextual diagnosis, co-design and enabling conditions, an AI-supported learning cycle, community application, and evaluation. Critical digital and AI literacy, together with social ethics, rights, and human agency, operate across all stages. The model provides theoretical propositions and operational indicators for community learning centers, local governments, libraries, educational institutions, and civil-society organizations. Empirical validation through design-based, quasi-experimental, and longitudinal research is recommended.

Keywords: artificial intelligence, community education, digital literacy, social ethics, self-directed learning, human-centered AI

1. INTRODUCTION

Artificial intelligence has moved from specialized technical environments into everyday social life. Generative AI systems can explain concepts, create text and images, translate languages, recommend resources, and provide immediate feedback. Their low entry barriers make them relevant not only to schools and universities but also to adults, community groups, local organizations, informal learners, and citizens who use AI for work, public services, entrepreneurship, communication, and personal learning. Current competency frameworks therefore place AI competence within broader digital participation rather than treating it as a narrow technical specialization (Cosgrove & Cachia, 2025; Miao et al., 2024).

This expansion creates a dual educational problem. AI can widen access to explanations, adaptive support, and learning resources; however, successful completion of a task does not automatically produce learning. When users delegate reasoning, verification, and decision making to a general-purpose AI system, output quality may improve while understanding,

metacognition, and independent judgment remain weak. The OECD (2026) distinguishes between AI that merely increases performance and AI use designed with explicit pedagogical intent. This distinction is crucial in community education because participants vary in age, schooling, digital access, language, occupation, and confidence.

Digital literacy can therefore no longer be limited to operating devices or applications. DigComp 3.0 defines digital competence as confident, critical, and responsible use of digital technologies for learning, work, and participation, and integrates AI-related knowledge, skills, and attitudes throughout its competence areas (Cosgrove & Cachia, 2025). AI literacy similarly includes understanding, using, evaluating, and communicating about AI systems and their social implications (Long & Magerko, 2020; Ng et al., 2021). Community members require practical competence in prompting, source verification, recognition of hallucination and bias, data protection, copyright, disclosure, and responsible content creation.

A second concern is social ethics, defined here as the values and norms guiding responsible conduct toward other people and communities in digitally mediated environments. It includes privacy, fairness, non-discrimination, intellectual honesty, respectful communication, accountability, and protection of human dignity. UNESCO (2022, 2023) places human rights, transparency, inclusion, data protection, and human oversight at the center of responsible AI. Ethics must therefore shape tool selection, learning activities, and evaluation from the start rather than appearing as an optional lecture after technical training.

A third concern is learner autonomy. Community education is closely related to adult learning, lifelong learning, empowerment, and practical problem solving. Self-directed learning involves diagnosing needs, setting goals, choosing resources and strategies, monitoring progress, and evaluating outcomes (Garrison, 1997; Knowles, 1975). AI can assist these processes through questioning, examples, translation, feedback, and planning. Yet poorly designed use can reduce productive struggle and create dependence. The educational objective is not maximum AI use, but calibrated support that progressively transfers responsibility to the learner.

Research on AI in education has concentrated mainly on formal institutions, students, teachers, assessment, and academic integrity. Reviews report opportunities in personalization, feedback, accessibility, and teaching support alongside risks of inaccuracy, bias, privacy violations, unequal access, and reduced critical engagement (Chiu et al., 2023; Lo, 2023; Tlili et al., 2023). Less attention has been given to an integrated community education framework connecting critical digital literacy, social ethics, and self-directed learning. This study addresses that gap through a human-centered model grounded in critical pedagogy and empowerment (Freire, 1970), communities of practice (Wenger, 1998), adult learning (Merriam & Bierema, 2013), self-directed and self-regulated learning (Garrison, 1997; Zimmerman, 2002), digital and AI literacy, and human-AI augmentation (Molenaar, 2022).

The study asks: (1) What AI affordances, risks, and enabling conditions are most relevant to community education for digital literacy, social ethics, and self-directed learning? (2) What components and mechanisms should form a human-centered AI community education model? and (3) What theoretical propositions, practical implications, and measurable indicators follow from the model? Because this is a conceptual review, it does not test causal hypotheses; it develops propositions for subsequent empirical validation. Its novelty lies in treating AI-based community education as a socio-technical and pedagogical system and in integrating literacy, ethics, and learner autonomy within one operational framework.

2. METHODS

2.1. Review Design

The study used a structured integrative literature review combined with conceptual model development. An integrative review is appropriate when a topic draws on empirical, theoretical, methodological, and policy literature and when the aim is to construct a framework rather than calculate a pooled effect (Snyder, 2019; Torraco, 2016). The reporting logic was informed by PRISMA 2020 principles of transparency, although this study was not a clinical systematic review and did not conduct meta-analysis (Page et al., 2021).

2.2. Source Identification and Eligibility

The review was updated in July 2026. Sources were identified through publisher platforms, ERIC records, backward reference searching, and the official repositories of UNESCO, OECD, and the European Commission. Search combinations paired “artificial intelligence” or “generative AI” with “community education,” “adult learning,” “digital literacy,” “AI literacy,” “digital ethics,” “responsible AI,” “self-directed learning,” “self-regulated learning,” “lifelong learning,” and “human-centered AI.” Foundational works published before 2020 were retained when they supplied established theories or analytical procedures. The eligibility decisions are summarized in Table 1.

Table 1. Review protocol and eligibility criteria

Criterion	Operational decision
Inclusion	Peer-reviewed research, foundational theory, or authoritative policy/framework documents directly relevant to at least two focal areas: AI in education, digital/AI literacy, social ethics, community/adult education, and self-directed learning.
Time frame	Priority to 2020–2026 AI literature; earlier sources retained only for foundational theories or established review methods.

Criterion	Operational decision
Exclusion	Purely technical AI papers without educational or social relevance; product descriptions; duplicates; unsupported opinion pieces; and incomplete bibliographic records.
Final corpus	25 documents: peer-reviewed articles or proceedings, foundational books, a reporting guideline, and international competency or policy frameworks.

2.3. Data Extraction and Thematic Synthesis

Each source was entered into a concept matrix containing context, theoretical constructs, AI affordances, risks, enabling conditions, and implications for program design. Analysis followed four steps. First, deductive codes were derived from the three focal outcomes: digital literacy, social ethics, and self-directed learning. Second, inductive coding identified recurring mechanisms and contextual conditions. Third, themes were compared across empirical studies, theory, and policy documents. Fourth, the themes were translated into inputs, learning processes, safeguards, outputs, and feedback loops. Thematic synthesis followed the familiarization, coding, theme development, review, definition, and reporting logic described by Braun and Clarke (2006). Credibility was strengthened through source triangulation, attention to counter-evidence on dependency and cognitive offloading, and explicit separation between literature-derived findings and the authors' proposed model.

2.4. Scope and Validity

The review was designed for conceptual integration rather than exhaustive database saturation. It did not calculate effect sizes or evaluate one intervention. Consequently, the final product is an evidence-informed theoretical model rather than a validated causal model. Its validity depends on future testing in communities differing in age, language, infrastructure, culture, and educational background.

3. RESULTS AND DISCUSSIONS

The synthesis generated five interrelated findings. AI was neither uniformly beneficial nor inherently harmful; its educational effects depended on pedagogical purpose, learner agency, ethical safeguards, facilitator support, and local conditions. Table 2 links each finding to its principal risk and required model response.

3.1. Conditional Educational Affordances of AI

The most consistent affordances were immediate explanations, adaptive examples, feedback, translation, idea generation, and personalized learning pathways (Chiu et al., 2023; Lo, 2023; Tlili et al., 2023). These functions can reduce barriers associated with time, location, language, and limited access to specialist support. Nevertheless, general-purpose AI may improve task completion without producing durable understanding. OECD (2026) identifies

cognitive offloading and metacognitive disengagement as central risks when AI is not tied to learning principles. The effective mechanism is therefore not access alone, but guided interaction requiring explanation, comparison, verification, application, and reflection.

3.2. Digital Literacy as Critical AI Competence

Operational ability is insufficient in an environment where generated content may appear coherent while being inaccurate, biased, or fabricated. The reviewed frameworks converge on multidimensional competence: locating information, evaluating evidence, understanding data and AI limitations, producing content, protecting privacy, and participating responsibly (Cosgrove & Cachia, 2025; Long & Magerko, 2020; Ng et al., 2021). Community learning should therefore use authentic tasks, such as comparing an AI response with reliable sources, identifying missing perspectives, checking claims or images, and revising prompts to reduce ambiguity. These tasks connect functional skills with epistemic judgment.

3.3. Social Ethics as Outcome and Safeguard

Ethical competence is more than knowledge of rules. It involves identifying affected stakeholders, anticipating harm, justifying choices, and accepting responsibility. Recurrent concerns include privacy, surveillance, bias, misinformation, intellectual property, plagiarism, opacity, and unequal access (Floridi et al., 2018; UNESCO, 2022, 2023). Social ethics therefore occupies two positions in the model. It is an outcome because participants should demonstrate responsible behavior, and it is a safeguard because ethical criteria must shape tool selection, data practices, learning activities, and evaluation before implementation.

3.4. Self-Directed Learning and Calibrated Scaffolding

AI can support learners in defining goals, planning activities, accessing explanations, practicing, receiving feedback, and reflecting. These functions correspond to the phases of self-directed and self-regulated learning (Garrison, 1997; Zimmerman, 2002). Support becomes counterproductive when the system performs essential cognitive work for the learner. The synthesis therefore supports calibrated scaffolding: facilitators first model productive AI use; participants practice prompting and verification; and assistance is gradually reduced while learners explain decisions and evaluate progress. This arrangement reflects human-AI augmentation rather than substitution (Molenaar, 2022).

3.5. Context, Co-Design, and Equity

Community education cannot assume uniform devices, connectivity, literacy, language, motivation, or trust. DigComp 3.0 recognizes access, basic literacy, guidance, and support as prerequisites for digital competence (Cosgrove & Cachia, 2025). Critical pedagogy and communities-of-practice theory also indicate that learning should begin with locally experienced problems and develop through dialogue and participation (Freire, 1970; Wenger, 1998). Effective implementation thus requires needs assessment, community co-design,

culturally and linguistically appropriate materials, low-bandwidth alternatives, and facilitators capable of mediating both technology and discussion.

Table 2. Thematic synthesis and model responses

Theme	Evidence-based interpretation	Principal risk	Required model response
Conditional AI affordances	AI can widen access, feedback, explanation, translation, and personalization when linked to explicit learning tasks.	Task completion without durable understanding.	Use guided inquiry, comparison, practice, and reflection rather than answer substitution.
Critical digital/AI literacy	Competence includes information evaluation, data awareness, AI limits, content creation, safety, and participation.	Misinformation, hallucination, bias, and uncritical trust.	Teach verification, source comparison, prompt critique, and responsible content production.
Social ethics	Responsible use requires dignity, fairness, privacy, attribution, transparency, and human oversight.	Privacy breaches, discrimination, plagiarism, and diffusion of responsibility.	Embed ethical scenarios, disclosure rules, data minimization, and accountability protocols.
Self-directed learning	AI can scaffold goal setting, resource selection, feedback, monitoring, and reflection.	Dependency, cognitive offloading, and reduced metacognition.	Transfer control gradually and assess reasoning, not only outputs.
Community context	Access, language, culture, trust, and facilitation shape outcomes.	Exclusion and mismatch with lived realities.	Co-design programs, provide flexible access, and anchor learning in local problems.

3.6. Proposed Human-Centered AI Community Education Model

The five findings were integrated into the cyclical model shown in Figure 1. The model contains five implementation stages and two safeguards that operate across all stages. The

stages organize program delivery; the safeguards ensure that efficiency does not override critical judgment, rights, or human agency.

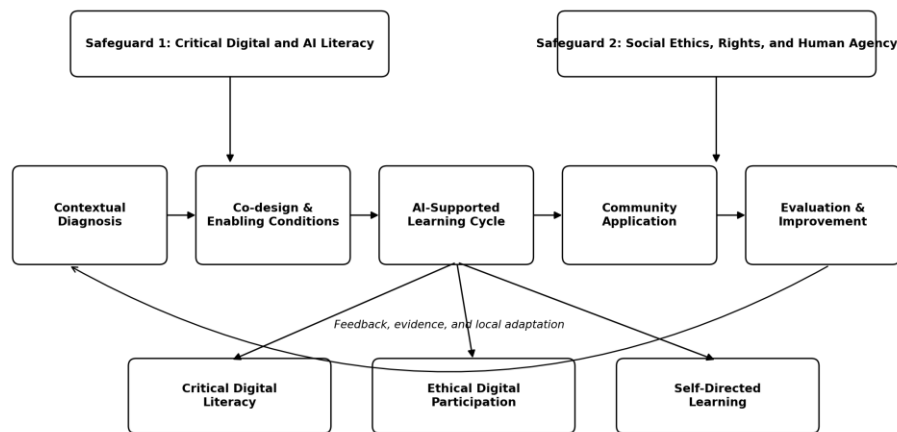


Figure 1. Human-centered AI community education model

Stage 1, contextual diagnosis, maps participants, devices, connectivity, language, digital practices, learning goals, trusted actors, and prevalent risks. Stage 2, co-design and enabling conditions, engages facilitators and community representatives in selecting problems, learning formats, tools, access arrangements, and ethical protocols. Stage 3 uses an AI-supported learning cycle: define a real problem, formulate a learning goal, interact with AI, verify the output, apply the knowledge, and reflect on the result and learning process. Stage 4 transfers learning to a socially meaningful product or action, such as a verified public-information resource, a digital micro-business plan, an anti-misinformation campaign, a community archive, or a service-access guide. Stage 5 evaluates competence, behavior, learner autonomy, participation, and implementation quality and feeds the evidence into the next cycle.

Critical digital and AI literacy operates across the model through source evaluation, awareness of model limitations, verification, and responsible content creation. Social ethics, rights, and human agency operate through privacy, fairness, attribution, respectful communication, transparency, and retention of human responsibility. As detailed in Table 3, AI plays supporting roles, while facilitators and community members retain contextual and ethical judgment.

Table 3. Operational components and evaluation indicators

Component	Core activities and AI role	Human role	Illustrative indicators
Contextual diagnosis	Map access, goals, practices, risks, language, and baseline competence; AI may assist survey adaptation and data organization.	Interpret findings, protect sensitive data, and identify priorities.	Baseline digital/AI literacy, access profile, priority risks, and learning needs.
Co-design and enabling conditions	Select problems, tools, formats, and access strategies; AI may generate alternatives.	Approve tools, establish safeguards, and ensure inclusion.	Participation in design, privacy protocol, accessibility, and cultural fit.
AI-supported learning cycle	Goal setting, prompting, verification, application, and reflection; AI serves as tutor, translator, simulator, and feedback partner.	Model critical use, challenge unsupported claims, and gradually reduce scaffolding.	Verification accuracy, reasoning quality, learning goals, monitoring, and reflection.
Community application	Create locally relevant products or actions; AI assists drafting, visualization, rehearsal, and revision.	Validate relevance, ethics, cultural appropriateness, and social impact.	Product quality, stakeholder feedback, collaboration, and community uptake.
Evaluation and improvement	Organize formative feedback and portfolios.	Triangulate outcomes, examine incidents, and revise the program.	Ethical behavior, autonomy, participation, implementation fidelity, and sustainability.

3.7. Linkage to Theory and Conceptual Propositions

The model integrates complementary explanatory perspectives. Critical pedagogy explains why lived problems, dialogue, and agency must determine the learning process rather than technology-led content transmission (Freire, 1970). Communities of practice explain

competence development through participation, peer exchange, and socially meaningful products (Wenger, 1998). Digital and AI literacy frameworks specify the critical knowledge, skills, and attitudes needed for responsible use. Self-directed and self-regulated learning theories explain goal setting, monitoring, strategy use, motivation, and reflection. Human-AI augmentation clarifies that control and responsibility can be deliberately distributed and gradually transferred (Molenaar, 2022).

This theoretical linkage produces five propositions for empirical testing. First, AI-supported community education should improve critical digital literacy more strongly when interaction is coupled with structured verification and source comparison than when training focuses only on tool operation. Second, explicit ethical scenarios, disclosure rules, privacy practices, and facilitated reflection should mediate the relationship between AI use and ethical participation. Third, AI should strengthen self-directed learning when it scaffolds goal setting, monitoring, and reflection while progressively returning control to the learner; unrestricted answer generation may weaken these outcomes. Fourth, community co-design, equitable access, local-language support, and facilitator competence should moderate effectiveness and sustainability. Fifth, community application projects should strengthen transfer by converting individual competence into collaborative and publicly accountable practice.

3.8. Theoretical and Practical Implications

Theoretically, the model extends AI-in-education discourse beyond formal classrooms by treating communities as sites of AI learning, governance, and participation. It links literacy, ethics, and autonomy through a common mechanism: preservation of informed human agency. It also positions social ethics as both an antecedent and an outcome, explaining why ethical principles must influence design before behavior is evaluated. Finally, the model specifies gradual transfer of control as a pedagogical mechanism within human-AI augmentation and identifies community application as a mechanism connecting individual competence to civic participation.

Practically, community learning centers, public libraries, village programs, religious organizations, civil-society groups, and local governments can use the model to move beyond demonstrations of AI tools. A program can begin with a brief needs and risk assessment, select one locally relevant problem, teach verification and ethical-use routines, and culminate in a community product. Facilitators require enough AI literacy to explain limitations, protect data, challenge unsupported claims, and decide when AI should not be used. Universities can provide facilitators, instruments, and research support; community organizations can secure relevance and trust; and government can support access, safeguards, and continuity. Funding should cover mentoring and evaluation rather than devices alone.

4. CONCLUSION

AI-based community education should not be judged by the number of tools introduced or the speed with which participants produce outputs. Its educational value lies in whether community members become more capable of evaluating information, acting ethically, directing their learning, and participating responsibly in digital life. The review indicates that AI can expand access, feedback, translation, and personalization, but these benefits depend on pedagogical purpose, verification, ethical safeguards, human oversight, calibrated scaffolding, co-design, and equitable access.

The proposed model consists of contextual diagnosis, co-design and enabling conditions, an AI-supported learning cycle, community application, and iterative evaluation. Critical digital and AI literacy and social ethics, rights, and human agency operate across every stage. The model offers an integrated theoretical framework, practical program sequence, and measurable indicators for future implementation.

The main limitation is methodological: the model is based on a selective integrative review and has not yet been tested as a complete intervention. Global frameworks and formal-education studies may not fully represent the infrastructural, linguistic, religious, cultural, and socioeconomic diversity of Indonesian communities. Rapid changes in AI systems may also make particular tool features obsolete. Future studies should use design-based research to co-design modules, quasi-experimental designs to compare tool-use training with the complete model, qualitative methods to examine trust and dependency, and longitudinal evaluation to assess whether critical, ethical, and independent learning practices persist and transfer to work, family, and civic contexts.

5. REFERENCES

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). <https://doi.org/10.1145/3442188.3445922>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Cosgrove, J., & Cachia, R. (2025). *DigComp 3.0: European Digital Competence Framework* (5th ed.). Publications Office of the European Union. <https://doi.org/10.2760/0001149>
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018).

- AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28, 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Freire, P. (1970). *Pedagogy of the oppressed*. Continuum.
- Garrison, D. R. (1997). Self-directed learning: Toward a comprehensive model. *Adult Education Quarterly*, 48(1), 18–33. <https://doi.org/10.1177/074171369704800103>
- Knowles, M. S. (1975). *Self-directed learning: A guide for learners and teachers*. Association Press.
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>
- 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). <https://doi.org/10.1145/3313831.3376727>
- Merriam, S. B., & Bierema, L. L. (2013). *Adult learning: Linking theory and practice*. Jossey-Bass.
- Miao, F., Shiohira, K., & Lao, N. (2024). AI competency framework for students. UNESCO.
- Molenaar, I. (2022). Towards hybrid human–AI learning technologies. *European Journal of Education*, 57(4), 632–645. <https://doi.org/10.1111/ejed.12527>
- 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>
- OECD. (2023). *OECD Digital Education Outlook 2023: Towards an effective digital education ecosystem*. OECD Publishing. <https://doi.org/10.1787/c74f03de-en>
- OECD. (2026). *OECD Digital Education Outlook 2026: Exploring effective uses of generative AI in education*. OECD Publishing. <https://doi.org/10.1787/062a7394-en>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Roe, J., & Perkins, M. (2025). Generative AI in self-directed learning: A thematic scoping review. *Interactive Learning Environments*. Advance online publication. <https://doi.org/10.1080/10494820.2025.2603431>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using

- chatbots in education. *Smart Learning Environments*, 10, Article 15.
<https://doi.org/10.1186/s40561-023-00237-x>
- Torraco, R. J. (2016). Writing integrative literature reviews: Using the past and present to explore the future. *Human Resource Development Review*, 15(4), 404–428.
<https://doi.org/10.1177/1534484316671606>
- UNESCO. (2022). Recommendation on the ethics of artificial intelligence. UNESCO.
- UNESCO. (2023). Guidance for generative AI in education and research. UNESCO.
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2