

Agentic Artificial Intelligence for Subnational SDG Localisation: a Human-in-the-Loop Framework for Voluntary Local Reviews in Sarawak, Malaysia

Michael King^{1*}, Poline Bala², Nwanesi Peter Karubi³

^{1,2,3}Institute of Borneo Studie, Universiti Malaysia Sarawak (UNIMAS), 94300 Kota Samarahan, Sarawak, Malaysia

¹bpoline@unimas.my, ²pknwanesi@unimas.my

* Corresponding author

Michael King

Institute of Borneo Studie, Universiti Malaysia Sarawak (UNIMAS), 94300 Kota Samarahan, Sarawak, Malaysia

Email ID:-22010285@siswa.unimas.my | ORCID: 0009-0002-7313-1477

ABSTRACT

Voluntary Local Reviews (VLRs) have become the main instrument through which subnational governments report their contribution to the Sustainable Development Goals (SDGs), yet most remain qualitative because of local data gaps, uneven indicator coverage and limited staff capacity. At the same time, artificial intelligence has advanced from single-task models to agentic systems that plan, retrieve evidence and use tools across multi-step workflows. This article asks how agentic AI can support the full VLR cycle in a data-constrained, multilingual region of the Global South without displacing local ownership and accountability. Using a design science research approach, it derives six design requirements from the literature and develops the Agentic Voluntary Local Review (A-VLR) framework: a four-layer architecture in which an orchestrator coordinates specialised agents for policy–SDG mapping, indicator localisation, data harvesting and estimation, and evidence-based drafting, under mandatory human-in-the-loop checkpoints and a cross-cutting assurance agent. The framework embeds evidence traceability, multilingual community input, Indigenous data governance and sovereign deployment in its design. Its application is illustrated through Sarawak, Malaysia, where an SDG-aligned state strategy, a growing number of councils preparing VLRs and state investment in AI coincide. The article introduces the concept of computational localisation, proposes an evaluation protocol and offers four testable propositions. It argues that frontier AI can help local governments move towards traceable, inclusive and evidence-based SDG reporting, provided that humans and communities remain in charge of what is reported.

Keywords: agentic AI; large language models; SDG localisation; Voluntary Local Reviews; design science; data sovereignty; Sarawak; Borneo

INTRODUCTION

A decade after the adoption of the 2030 Agenda (United Nations, 2015), it is widely accepted that most Sustainable Development Goal (SDG) targets depend on decisions taken by subnational governments (Fenton & Gustafsson, 2017). Voluntary Local Reviews (VLRs) have become the main instrument through which cities, municipalities and regions report that contribution. They are encouraged by United Nations guidance (UN DESA, 2020; UN-Habitat & UCLG, 2020), and their number continues to grow. Yet the evidence on how VLRs are actually produced points to a persistent capacity problem. In its sixth annual review, the Institute for Global Environmental Strategies (IGES) found that local governments still favour qualitative reviews of policies and actions over quantitative assessments of progress, largely because of local data limitations and the complexity of the SDG indicator framework (Ortiz-Moya et al., 2025). An analysis of 120 English-language VLRs found that local monitoring frameworks average 87 indicators, are unevenly spread across goals, and leave goals such as SDG 15 and SDG 17 frequently uncovered (Stamos et al., 2024).

In parallel, artificial intelligence (AI) has moved from narrow predictive models to large language models (LLMs) and, more recently, to agentic systems: LLM-based agents that plan, call tools, retrieve documents and act in multi-step workflows (Wang et al., 2024; Yao et al., 2023). Early studies show that LLMs can map texts to SDGs with useful accuracy (Yin et al., 2024) and can align SDG indicators with heterogeneous open data (Benjira et al., 2025). Much of the AI-for-SDGs literature, however, remains framed at the global or national scale (Vinuesa et al., 2020), focuses on single tasks such as classification, and pays limited attention to the political, multilingual and data-sovereignty conditions under which local governments in the Global South would use such systems. The latest IGES review of VLRs makes no reference to AI or digital tools at all (Ortiz-Moya et al., 2025).

This article addresses that gap. It asks two research questions:

RQ1. What design requirements must an AI system meet to support the full VLR cycle of a subnational government in a developing, culturally diverse and data-constrained region?

RQ2. How can agentic AI be configured to meet those requirements while keeping local ownership, accountability and inclusion at the centre of SDG localisation?

The study adopts a design science research approach (Hevner et al., 2004; Peffers et al., 2007) and develops the **Agentic Voluntary Local Review (A-VLR) framework**: a four-layer, human-in-the-loop architecture of specialised AI agents. The framework is demonstrated through an illustrative application to Sarawak, Malaysia. Sarawak is an instructive case. Its Post COVID-19 Development Strategy 2030 (PCDS 2030) is officially presented as aligned with the SDGs (Economic Planning Unit Sarawak, 2021; Premier of Sarawak, 2023). Its local authorities have joined the SDG Cities movement, with six councils reported to have submitted VLRs and a further 20 committed to begin preparation (Premier of Sarawak, 2024). The state is also building sovereign AI capacity through the Sarawak Artificial Intelligence Centre and a state AI blueprint (UKAS, n.d.).

The article makes three contributions. First, it proposes the concept of **computational localisation**: treating the translation of global goals into local targets, indicators and narratives as a socio-technical workflow that AI agents can partly carry out, under human authority. Second, it specifies six design requirements and a reference architecture that move beyond single-task SDG classification towards an end-to-end, auditable VLR process. Third, it sets out an evaluation protocol and four testable propositions to guide empirical work in Sarawak and comparable regions.

The remainder of the article is structured as follows. Section 2 reviews the literature on SDG localisation, AI for sustainable development and agentic AI. Section 3 explains the methodology. Section 4 presents the design requirements and the A-VLR framework. Section 5 illustrates its application in Sarawak and proposes an evaluation protocol. Section 6 discusses theoretical and practical implications, risks and propositions. Section 7 concludes.

2. LITERATURE REVIEW

2.1 SDG localisation and Voluntary Local Reviews

The SDGs were designed as a form of “governance by goal-setting”: they set aspirational targets without binding implementation mechanisms, leaving delivery to national and subnational actors (Biermann et al., 2017). Localisation is the process by which these global goals are interpreted, adapted and implemented in specific places (Fenton & Gustafsson, 2017). VLRs emerged as a voluntary, bottom-up counterpart to the Voluntary National Reviews (VNRs) presented at the UN High-Level Political Forum, and UN DESA (2020) and UN-Habitat and UCLG (2020) have issued guidance on their structure and content.

Three recurring weaknesses appear in the VLR literature. The first is **data scarcity at the local scale**: many SDG indicators are defined for national statistical systems, and local governments often lack disaggregated data (Ortiz-Moya et al., 2025; Stamos et al., 2024). The second is **uneven goal coverage**: goals tied to municipal mandates, especially SDG 11, dominate, while environmental and partnership goals are under-reported (Stamos et al., 2024). The third is **resource-intensive production**: VLRs require document review, indicator selection, data collection, stakeholder consultation and drafting, which strain small councils. These weaknesses are most acute in middle-income regions, which now produce most VLRs (Ortiz-Moya et al., 2025).

2.2 Artificial intelligence for the SDGs

Vinuesa et al. (2020) estimated that AI could act as an enabler for a large majority of SDG targets, while also inhibiting some, and called for regulatory oversight. Satellite imagery combined with machine learning has been used to estimate poverty and other outcomes where survey data are sparse (Burke et al., 2021). Text-based approaches have been developed to classify documents by SDG, including the open-source OSDG tool (Pukelis et al., 2020). More recently, Yin et al. (2024) compared seven LLMs on mapping publications to SDGs and found that several open and proprietary models achieved strong F1 scores, with trade-offs between precision and recall. Benjira et al. (2025) combined LLMs with a knowledge graph to map open data to SDG indicator 11.7.1 and reported performance well above traditional schema-matching methods.

Critical scholarship cautions against treating AI as inherently sustainable. Sætra (2021) argues that AI must be assessed within the sociotechnical system it reinforces, including its environmental and social costs. The energy

and carbon costs of training large models are well documented (Strubell et al., 2019), and ethical frameworks emphasise beneficence, non-maleficence, autonomy, justice and explicability (Floridi et al., 2018). At the global level, the UN High-level Advisory Body on AI has highlighted capacity gaps and the risk that developing regions become consumers rather than co-creators of AI (United Nations AI Advisory Body, 2024).

2.3 From large language models to agentic AI

Foundation models are general-purpose models trained on broad data and adapted to many tasks (Bommasani et al., 2021). Two techniques have extended their usefulness for evidence-based work. Retrieval-augmented generation (RAG) grounds a model's output in retrieved documents, improving factuality and enabling citation (Lewis et al., 2020). Reasoning-and-acting approaches such as ReAct allow a model to interleave reasoning with tool use, such as searching, querying a database or running a calculation (Yao et al., 2023). Building on these, LLM-based autonomous agents combine planning, memory and tool use, and can be organised into multi-agent systems in which specialised agents collaborate on complex tasks (Wang et al., 2024). The main limitation is hallucination: fluent but unsupported content (Ji et al., 2023). This is a serious risk for official reporting, where every claim must be traceable to evidence.

2.4 Research gap

Table 1 summarises the gap. Existing AI applications address individual tasks in the SDG evidence chain. None, to the author's knowledge, integrates them into a governed, end-to-end workflow for producing VLRs at the subnational level, and few address multilingual communities, Indigenous data rights or sovereign deployment, all of which are central in regions such as Sarawak.

Table 1. Positioning of this study against representative prior work

Stream	Representative work	Focus contribution and	Gap addressed here
VLR practice and monitoring	Ortiz-Moya et al. (2025); Stamos et al. (2024)	Trends, indicator use and data limitations in VLRs	No digital or AI-based response to the documented capacity gap
AI as SDG enabler	Vinuesa et al. (2020); Burke et al. (2021)	Macro-level potential of AI and remote sensing for SDGs	Global/national framing; little on local governance processes
Text-to-SDG mapping	Pukelis et al. (2020); Yin et al. (2024)	Classifying documents by SDG with ML and LLMs	Single task; no link to local indicators, data or reporting
Indicator–data alignment	Benjira et al. (2025)	LLM plus knowledge graph to map open data to an SDG indicator	One indicator, European data; not an end-to-end VLR workflow
Agentic AI	Yao et al. (2023); Wang et al. (2024)	Planning, tool use and multi-agent collaboration	Not yet applied to SDG localisation or public reporting
This study	A-VLR framework	Governed, multi-agent, human-in-the-loop workflow for the full VLR cycle in a multilingual Global South region	Integrates the above; adds sovereignty, inclusion and assurance

Source: Author.

3. METHODOLOGY

The study follows design science research (DSR), which aims to produce and evaluate artefacts that solve identified organisational problems while contributing to knowledge (Hevner et al., 2004). The process model of Peffers et al. (2007) was applied in five steps: (1) problem identification, (2) definition of solution objectives, (3) design and development, (4) demonstration and (5) evaluation design. Communication is achieved through this article.

Problem identification drew on the VLR literature reviewed in Section 2 and on public documents describing Sarawak's development strategy and SDG Cities activity. **Solution objectives** were expressed as design requirements, each grounded in a documented problem or principle. **Design and development** produced a reference architecture and a specification of agent roles and human checkpoints. **Demonstration** takes the form of an illustrative application to a hypothetical Sarawak local authority, built from publicly reported conditions; it is explicitly not a field implementation. **Evaluation** follows the Framework for Evaluation in Design Science (FEDS) of Venable et al. (2016). At this stage the evaluation is ex ante and artificial: the framework is assessed analytically against the requirements, and a protocol is specified for subsequent naturalistic evaluation with councils.

This approach is suited to an emerging technology where empirical deployments are not yet available. Its limitation is that the framework's performance claims remain to be tested; Section 5.3 and Section 6.4 therefore state measurable criteria and propositions.

4. THE AGENTIC VOLUNTARY LOCAL REVIEW (A-VLR) FRAMEWORK

4.1 Design requirements

Six design requirements (DRs) were derived from the literature. They are summarised in Table 2.

Table 2. Design requirements for AI-supported subnational SDG reporting

DR	Requirement	Rationale (problem or principle)	Framework response
DR1	End-to-end coverage of the VLR cycle	VLR production spans mapping, indicators, data, consultation and drafting (UN DESA, 2020)	Specialised agents A1–A4 under an orchestrator
DR2	Local indicator adaptation	Local frameworks differ from UN indicators and are unevenly distributed (Stamos et al., 2024)	Indicator Localisation Agent and local knowledge graph
DR3	Evidence traceability	LLMs can hallucinate (Ji et al., 2023); official reporting must be verifiable	RAG with sentence-level citation; Assurance Agent
DR4	Human authority and accountability	Localisation is political, not only technical (Fenton & Gustafsson, 2017)	Mandatory human-in-the-loop checkpoints H1–H4
DR5	Inclusion and data rights	Multilingual and Indigenous communities; CARE principles (Carroll et al., 2020)	Multilingual intake; CARE-aligned data rules; inclusion audit
DR6	Sovereign and sustainable deployment	Capacity gaps and dependence risks (United Nations AI Advisory Body, 2024); energy costs (Strubell et al., 2019)	State-hosted deployment; right-sized and open models where feasible

Source: Author.

4.2 Architecture

Figure 1 presents the A-VLR framework. It has four layers.

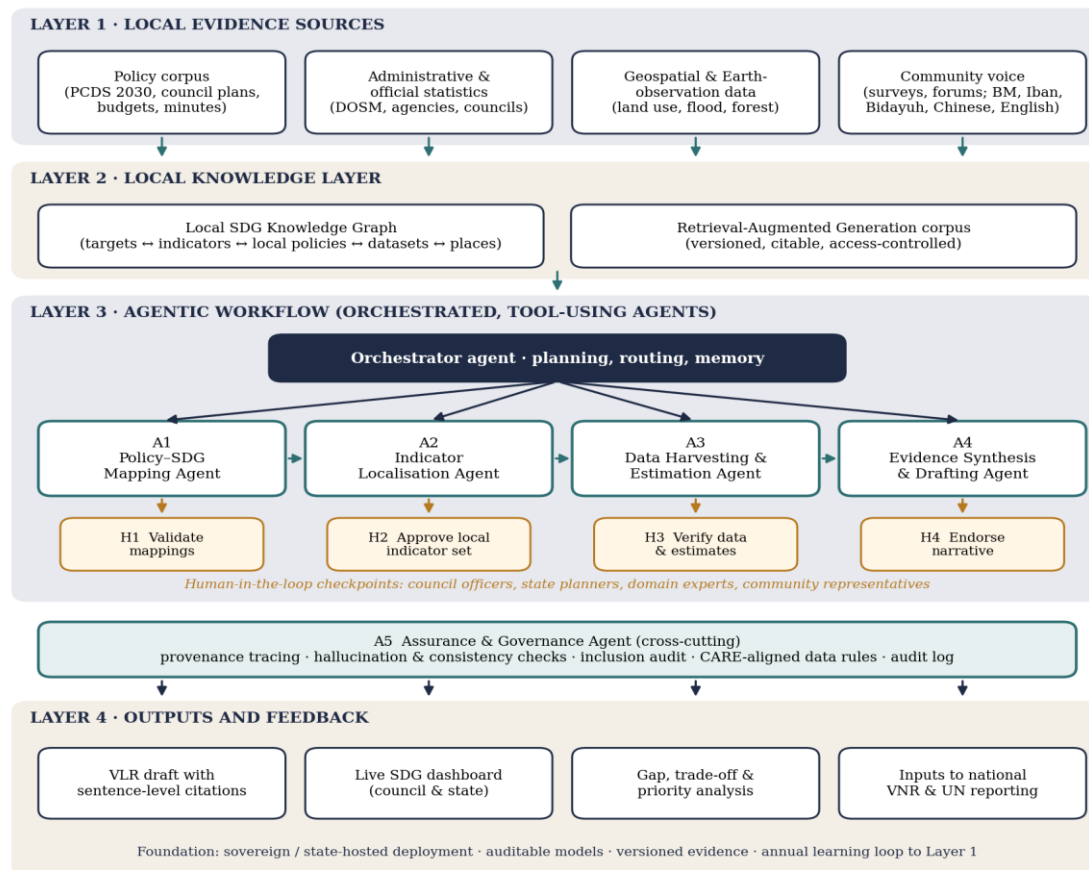


Figure 1. The Agentic Voluntary Local Review (A-VLR) framework. Source: Author.

Layer 1: Local evidence sources. The framework draws on four kinds of evidence: the policy corpus (state and council plans, budgets, minutes and project reports); administrative and official statistics; geospatial and Earth-observation data; and community voice gathered through surveys and forums. In Sarawak, community input must be accepted in Bahasa Malaysia, English, Chinese and Indigenous languages such as Iban and Bidayuh.

Layer 2: Local knowledge layer. Evidence is organised into a local SDG knowledge graph linking SDG targets and indicators to local policies, datasets and places, building on the approach of Benjira et al. (2025). A versioned RAG corpus stores source documents so that every generated statement can cite its origin (Lewis et al., 2020). Access controls apply to sensitive and community-owned data.

Layer 3: Agentic workflow. An orchestrator agent plans tasks and routes them to four specialised agents:

A1 Policy-SDG Mapping Agent maps local policies, programmes and budget lines to SDG targets, producing a coverage matrix that exposes under-addressed goals.

A2 Indicator Localisation Agent proposes a local indicator set by matching UN indicators with locally available data and suggesting proxies where official indicators cannot be computed, with a justification for each.

A3 Data Harvesting and Estimation Agent retrieves data through approved tools and interfaces, computes indicators, and, where data are missing, applies documented estimation methods (for example, Earth-observation proxies), clearly labelled as estimates.

A4 Evidence Synthesis and Drafting Agent drafts VLR chapters, case narratives and summaries, with every factual sentence linked to evidence in the RAG corpus.

Each agent's output passes through a human-in-the-loop checkpoint (H1–H4), where council officers, state planners, domain experts and community representatives validate, amend or reject it. A fifth agent, **A5 Assurance and Governance Agent**, operates across the workflow: it traces provenance, checks for unsupported claims and

internal inconsistencies, audits whether rural and Indigenous communities and all relevant goals are represented, enforces data-use rules, and records an audit log.

Layer 4: Outputs and feedback. Outputs include a VLR draft with sentence-level citations, a live SDG dashboard for council and state, an analysis of gaps, trade-offs and priorities, and structured inputs for national VNR reporting. The foundation of the framework is sovereign or state-hosted deployment, auditable models and versioned evidence, with an annual learning loop that feeds lessons back into Layer 1.

4.3 Distinctive features

The A-VLR framework differs from existing tools in three respects. First, it treats SDG reporting as a **workflow of linked decisions**, not a single classification task. Second, it **embeds human authority structurally**: the agents prepare, but humans decide, at four defined points. Third, it **builds inclusion and sovereignty into the architecture** rather than leaving them to policy statements, which responds directly to concerns about AI widening capacity gaps (United Nations AI Advisory Body, 2024).

5. ILLUSTRATIVE APPLICATION: SARAWAK, MALAYSIA

5.1 Context

Sarawak is Malaysia's largest state, with a dispersed population, extensive rural and riverine settlements and a diverse population that includes many Indigenous communities. PCDS 2030 is the state's main development framework, emphasising economic transformation, digitalisation and environmental sustainability (Economic Planning Unit Sarawak, 2021); the state government has described it as fulfilling 11 of the 17 SDGs (Premier of Sarawak, 2023). At the local level, 24 local councils were reported in 2023 to be implementing SDG Cities roadmaps, six had submitted VLRs, and 20 more had committed to begin preparing them (Premier of Sarawak, 2024). Sarawak has also established the Sarawak Artificial Intelligence Centre and initiated a state AI blueprint aligned with PCDS 2030 (UKAS, n.d.). Nationally, Malaysia has reported SDG progress through its VNR (Economic Planning Unit, 2021), and Subang Jaya's VLR is an established Malaysian reference point (Subang Jaya City Council, 2021).

This combination of an SDG-aligned state strategy, an expanding number of councils preparing VLRs, and state investment in AI makes Sarawak an appropriate setting to illustrate the A-VLR framework.

5.2 Illustrative walk-through

The following walk-through describes how the framework could support a hypothetical Sarawak municipal council preparing its first VLR. It is illustrative and does not report results from an actual deployment.

Mapping (A1, H1). The council uploads its strategic plan, annual budget and relevant PCDS 2030 documents. A1 maps programmes and budget lines to SDG targets. The coverage matrix shows, for example, strong coverage of SDG 6, 9 and 11 and weak coverage of SDG 15 and 17, mirroring patterns in the wider VLR literature (Stamos et al., 2024). Council planners review and correct the mapping at H1.

Indicator localisation (A2, H2). A2 proposes a local indicator set, flagging which UN indicators can be computed directly and which need proxies, such as household access to treated water from utility records, or forest cover change from satellite imagery. Planners and domain experts approve a final set at H2.

Data and estimation (A3, H3). A3 retrieves administrative data through approved interfaces, computes indicators and produces clearly labelled estimates where data are missing. Statistical officers verify methods and values at H3.

Community voice and drafting (A4, H4). Community submissions in Iban, Bidayuh, Bahasa Malaysia, Chinese and English are summarised, with the original text retained for verification. A4 drafts the VLR, citing every claim. A5 flags unsupported sentences and checks that rural longhouse communities are represented. The council's SDG committee and community representatives endorse the narrative at H4.

5.3 Evaluation protocol

Following FEDS (Venable et al., 2016), the next stage is a naturalistic evaluation with two or three Sarawak councils, comparing A-VLR-supported preparation with conventional preparation. Table 3 sets out the proposed criteria and metrics.

Table 3. Proposed evaluation criteria and metrics

Criterion	Metric	Method
Mapping accuracy	Precision, recall and F1 of policy–SDG target mapping	Comparison with an expert-annotated gold standard
Indicator coverage	Share of SDGs and targets with at least one valid local indicator	Before–after comparison with existing VLRs
Traceability	Share of factual sentences linked to a verifiable source	Automated check plus manual audit sample
Factual reliability	Rate of unsupported or incorrect claims in drafts	Blind expert review
Efficiency	Staff hours and calendar time to complete a VLR draft	Time logs in pilot councils
Inclusion	Representation of rural, Indigenous and language groups in evidence and narrative	Content analysis; community feedback
Legitimacy and trust	Perceived usefulness, accountability and ownership among officers and communities	Surveys and focus groups
Sustainability of AI use	Energy use and cost per VLR cycle	Infrastructure monitoring

Source: Author.

6. DISCUSSION

6.1 Theoretical implications

The framework contributes to localisation theory by conceptualising **computational localisation**. Localisation has been studied mainly as a governance and political process (Fenton & Gustafsson, 2017). The A-VLR framework shows that several of its steps, namely interpreting goals against local policy, translating indicators and assembling evidence, can be represented as explicit, inspectable computational operations. This makes previously tacit choices visible: which policy was mapped to which target, which proxy replaced which indicator, and on what evidence. In this sense, agentic AI can make localisation more transparent, provided that the human checkpoints remain the locus of decision.

The framework also extends the AI-for-SDGs literature (Vinuesa et al., 2020) from asking whether AI can enable SDG targets to asking how AI should be governed when it participates in the public accounting of SDG progress.

6.2 Practical implications

For state and local governments, the framework offers a route to scale VLR production across many councils without proportional increases in staff, which is relevant to Sarawak's plan to extend VLRs to more local authorities (Premier of Sarawak, 2024). For state AI programmes, it identifies a concrete public-sector use case that aligns AI investment with PCDS 2030 and the SDGs. For national agencies, standardised and traceable local outputs could strengthen the subnational evidence base of future VNRs.

6.3 Risks and safeguards

Four risks require attention. **Hallucination and false precision**: drafts could present unsupported or overstated claims; DR3, A5 and H3–H4 are the safeguards (Ji et al., 2023). **Automation of a political process**: VLRs are also exercises in local accountability and participation, and over-automation could hollow them out; mandatory human checkpoints and community endorsement address this. **Data rights and sovereignty**: community and

Indigenous data should be governed according to collective benefit, authority to control, responsibility and ethics (Carroll et al., 2020), and hosted under state control. **Environmental cost:** the energy footprint of AI should be measured and minimised through right-sized models (Strubell et al., 2019; Sætra, 2021).

6.4 Propositions for future research

P1. Councils using A-VLR will report a larger share of SDGs and targets with quantitative local indicators than comparable councils using conventional methods.

P2. Sentence-level evidence linking and an assurance agent will reduce unsupported claims in VLR drafts compared with unconstrained LLM drafting.

P3. Mandatory human-in-the-loop checkpoints will be associated with higher perceived ownership and legitimacy of the VLR among council officers and community representatives.

P4. Multilingual community intake will increase the representation of rural and Indigenous perspectives in VLR narratives.

7. LIMITATIONS AND FUTURE RESEARCH

This study presents a design artefact evaluated ex ante. Its performance has not yet been tested in operation, and the Sarawak application is illustrative. Future research should implement a prototype, conduct the naturalistic evaluation in Table 3 with pilot councils, test the propositions in Section 6.4, and examine transferability to other subnational settings in Southeast Asia and beyond. Research is also needed on the performance of LLMs in Indigenous Bornean languages, on cost and energy benchmarks for public-sector agentic systems, and on the institutional arrangements that would allow a state to host and govern such systems.

8. CONCLUSION

VLRs have made subnational governments visible in the global SDG process, but their quality and spread are limited by data, capacity and resources. This article proposed the Agentic Voluntary Local Review framework as a design response. Its six design requirements and four-layer architecture combine specialised AI agents with mandatory human checkpoints, a local knowledge layer, and built-in assurance, inclusion and sovereignty safeguards. Illustrated through Sarawak, the framework suggests that frontier AI can help local governments move from largely qualitative reporting towards evidence-based, traceable and inclusive SDG reviews, provided that humans and communities remain in charge of what is reported and why. The proposed evaluation protocol and propositions provide a basis for testing this claim in practice.

DECLARATIONS

Author contributions. Michael King: conceptualisation, methodology, framework design, writing – original draft, writing – review and editing. [Add co-authors / supervisors and CRediT roles if applicable.]

Funding. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflict of interest. The author declares no conflict of interest.

Data availability. No new data were generated or analysed in this study. All sources cited are publicly available.

Ethics statement. This conceptual study did not involve human participants. Future empirical evaluation will obtain ethical approval from the UNIMAS research ethics committee.

Declaration of generative AI use. During the preparation of this manuscript, the author used generative AI tools to assist with drafting and structuring the text and preparing the figure. The author reviewed and edited all content and takes full responsibility for the publication

REFERENCES

1. Benjira, W., Atigui, F., Bucher, B., Grim-Yefsah, M., & Travers, N. (2025). Automated mapping between SDG indicators and open data: An LLM-augmented knowledge graph approach. *Data & Knowledge Engineering*, 156, 102405. <https://doi.org/10.1016/j.datak.2024.102405>
2. Biermann, F., Kanie, N., & Kim, R. E. (2017). Global governance by goal-setting: The novel approach of the UN Sustainable Development Goals. *Current Opinion in Environmental Sustainability*, 26–27, 26–31. <https://doi.org/10.1016/j.cosust.2017.01.010>

3. Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., et al. (2021). On the opportunities and risks of foundation models (arXiv:2108.07258). arXiv. <https://doi.org/10.48550/arXiv.2108.07258>
4. Burke, M., Driscoll, A., Lobell, D. B., & Ermon, S. (2021). Using satellite imagery to understand and promote sustainable development. *Science*, 371(6535), eabe8628. <https://doi.org/10.1126/science.abe8628>
5. Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., et al. (2020). The CARE principles for Indigenous data governance. *Data Science Journal*, 19(1), 43. <https://doi.org/10.5334/dsj-2020-043>
6. Economic Planning Unit Sarawak. (2021). Post COVID-19 Development Strategy 2030. Sarawak Government.
7. Economic Planning Unit. (2021). Malaysia Voluntary National Review (VNR) 2021. Prime Minister's Department, Malaysia.
8. Fenton, P., & Gustafsson, S. (2017). Moving from high-level words to local action—Governance and implementation of the UN 2030 Agenda in municipalities. *Current Opinion in Environmental Sustainability*, 26–27, 129–133. <https://doi.org/10.1016/j.cosust.2017.07.009>
9. Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., et al. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
10. Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75–105. <https://doi.org/10.2307/25148625>
11. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., et al. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), Article 248. <https://doi.org/10.1145/3571730>
12. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459–9474.
13. Ortiz-Moya, F., Yoshida, T., & Kataoka, Y. (2025). State of the Voluntary Local Reviews 2025: At a crossroads for local SDG action. Institute for Global Environmental Strategies. <https://doi.org/10.57405/iges-14390>
14. Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45–77. <https://doi.org/10.2753/MIS0742-1222240302>
15. Premier of Sarawak. (2023, May 29). PCDS 2030 fulfils 11 out of 17 UN Sustainable Development Goals [News release]. https://premier.sarawak.gov.my/web/subpage/news_view/5474
16. Premier of Sarawak. (2024, January 4). Local councils in S'wak aggressively embark on SDG Cities Roadmap [News release, reproduced from The Borneo Post]. https://premier.sarawak.gov.my/web/subpage/news_view/5694
17. Pukelis, L., Bautista-Puig, N., Skrynik, M., & Stanciasukas, V. (2020). OSDG – Open-source approach to classify text data by UN Sustainable Development Goals (SDGs) (arXiv:2005.14569). arXiv. <https://doi.org/10.48550/arXiv.2005.14569>
18. Sætra, H. S. (2021). AI in context and the Sustainable Development Goals: Factoring in the unsustainability of the sociotechnical system. *Sustainability*, 13(4), 1738. <https://doi.org/10.3390/su13041738>
19. Stamos, I., Vivas, L., Enrique Regueira, I., & Bertozzi, C. (2024). What does SDG monitoring practice tell us? An analysis of 120 Voluntary Local Reviews. *Sustainability*, 16(23), 10649. <https://doi.org/10.3390/su162310649>
20. Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 3645–3650). Association for Computational Linguistics. <https://doi.org/10.18653/v1/P19-1355>
21. Subang Jaya City Council. (2021). Voluntary Local Review 2021: Subang Jaya. Majlis Bandaraya Subang Jaya.
22. UKAS (Unit Komunikasi Awam Sarawak). (n.d.). Sarawak AI Blueprint strengthens state AI development direction. Retrieved October 1, 2026, from https://ukas.sarawak.gov.my/web/subpage/news_view/45270
23. UN DESA. (2020). Global guiding elements for Voluntary Local Reviews (VLRs) of SDG implementation. United Nations Department of Economic and Social Affairs.
24. UN-Habitat & UCLG. (2020). Guidelines for Voluntary Local Reviews, Volume 1: A comparative analysis of existing VLRs. United Nations Human Settlements Programme and United Cities and Local Governments.
25. United Nations AI Advisory Body. (2024). Governing AI for humanity: Final report. United Nations.
26. United Nations. (2015). Transforming our world: The 2030 Agenda for Sustainable Development (A/RES/70/1). United Nations General Assembly.

27. Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: A Framework for Evaluation in Design Science research. *European Journal of Information Systems*, 25(1), 77–89. <https://doi.org/10.1057/ejis.2014.36>
28. Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., et al. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications*, 11, 233. <https://doi.org/10.1038/s41467-019-14108-y>
29. Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., et al. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6), 186345. <https://doi.org/10.1007/s11704-024-40231-1>
30. Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. In *Proceedings of the Eleventh International Conference on Learning Representations (ICLR 2023)*.
31. Yin, H., Aryani, A., & Nambiar, N. (2024). Evaluating the performance of large language models for SDG mapping (arXiv:2408.02201). *arXiv*. <https://doi.org/10.48550/arXiv.2408.02201>