Dear Reviewer,

Thank you for your comments. Responses (in blue) are written below the respective comments (in black).

MAJOR COMMENTS

COMMENT*: 1. Section 4.1 is bound to a specific technology (LLMs) rather than to the function it serves - and is too narrowly text-centric*

This is my central concern. Section 4.1 is framed as "LLM-driven enhancements..." and the entire argument is built on large language models as THE mechanism for co-design and cross-domain translation. This is a problem for a Perspective paper, which should age well, on two counts:

(a) Technology-binding. LLMs are the dominant technology today, but the field is moving fast and the specific technology may change within the lifetime of this paper. The underlying capability the authors actually need is an interface: technical or domain-specific content goes in, and content optimized for a different audience comes out (e.g., soil scientist <-> modeler <-> farmer <-> policy-maker). That capability is defined by its function - content in, audience-appropriate content out - not by whichever model happens to implement it. It could be an LLM now; it could be a different architecture, a hybrid symbolic/neural system, or something not yet named later. I strongly recommend the authors reframe Section 4.1 (and its heading) around the function/interface rather than the named technology, and treat "LLMs" as the current-best instantiation cited as an example, not as the organizing concept. This would future-proof the section and make its argument more robust.

(b) Text-centrism / missing modalities. The section treats co-design and science-to-action almost entirely as a text problem ("translate plans, findings, and implications," "convert highly technical content into content... to convert to policy"). But effective communication and co-design across disparate audiences - farmers, land managers, rights holders, the public, policy-makers - very often depends on non-text modalities: infographics, data visualizations, maps, diagrams, decision-support dashboards, and audio/visual material. These require different underlying technologies (generative imagery, automated data visualization, cartographic tools), not just language models. Given that the paper elsewhere rightly emphasizes multi-modal engagement (Sect. 3, AmeriFlux example) and reducing barriers for non-technical participants, it is inconsistent for the AI-outreach section to collapse back to a purely text-in/text-out view. I recommend the authors explicitly broaden Section 4.1 to acknowledge that communication is multi-modal and that realizing this vision will require a portfolio of technologies serving the same co-design function.

(c) Internal inconsistency of "altitude" across Section 4. Reinforcing the point above: the three subsection headings mix levels of abstraction. Section 4.2 is framed neutrally as "AI-guided..." (a function/capability), whereas 4.1 and 4.3 lead with the specific technology ("LLM-driven...",

and LLM-centric text in 4.3). I suggest the authors adopt a consistent framing throughout Section 4: name each subsection by the challenge or function being addressed (translation/co-design, adaptive sampling, harmonization), and cite specific technologies (LLMs, agents, etc.) as current examples within. This single change would resolve both the technology-binding concern and the inconsistency.

RESPONSE: We thank the reviewer for this helpful comment. We agree that the original Section 4.1 foregrounded LLMs and described communication largely in terms of translating technical text. We will revise the subsection heading and reframe the discussion around the broader function of AI-enabled communication across domains and modalities, presenting LLMs as one current approach rather than the organizing technology. We will also change Figure 1 from saying "LLM-driven co-design" to "AI-driven co-design".

The revised text will recognize that communication may involve text, audio, video, and visual or spatial materials, including captions, maps, diagrams, and data visualizations. It will describe how different AI-enabled tools might help prepare or adapt these materials for different audiences, while distinguishing that function from co-design itself.

We will also broaden the discussion of limitations. In addition to noting the potential for technical errors and bias, the revised text will highlight that transformed visual and audio-visual materials may omit uncertainty, distort emphasis, or be inaccessible or inappropriate for their intended audiences. We therefore emphasize review and evaluation with intended users.

We do want to note that Section 4 as a whole already discusses AI approaches beyond LLMs, including AI-guided sampling, optimization, hybrid AI–process models, agents, and AI methods for quality control and data harmonization. However, we agree that Section 4.1 did not make its broader, multimodal framing sufficiently clear, and plan to revise it accordingly.

> **New text excerpts for section 4.1:**
>
> "We note that these are current examples and that the most useful tools in the future may diverge from contemporary tools.
>
> …
>
> "The opportunity here is to leverage existing and emerging AI technologies to help participants access, interpret, and discuss scientific information across domains and communication formats.
>
> "Different AI approaches may support this communication function. Speech and language AI could assist with transcription, captions, and translation. Generative AI, including LLMs, could help draft or adapt explanations for different audiences. Furthermore, multimodal generative AI could help develop draft visual or audio-visual materials. AI-assisted visualization and mapping tools could help participants explore spatial patterns or present findings in forms suited to their needs. Lastly, agentic workflows could iterate (with human-in-the-loop check points) between different

interested parties and come up with a range of possibilities for co-design, ranking each one in their ability to meet different priorities.

…

"For example, an LLM could draft a plain-language summary, while a map or data visualization tool could help participants discuss the spatial pattern represented by the same findings. Interactive scenario tools that combine AI with simulations or process-based models could also help participants compare possible interventions and discuss their assumptions and trade-offs.

…

"AI-generated or transformed visual, spatial, and audio-visual materials also require review. Materials may distort emphasis, or be inaccessible or inappropriate for their intended audiences. Summaries of participant input may obscure dissent or minority perspectives."

***COMMENT**: 2. Justify the collapse of standardization and centralization into a single participant axis*

Section 2.1 builds on SanClements et al. (2022), collapsing "degree of standardization" and "degree of centralization" into one axis while reframing it around participants (formalized <-> flexible). The manuscript notes SanClements et al. proposed a 1:1 relationship between the two, but this 1:1 assumption is not obviously robust: one can imagine efforts that are highly centralized in funding/management yet loosely standardized in protocols, or vice versa. Because this collapse is foundational to the framework, please add a sentence or two justifying it and acknowledging cases where standardization and centralization decouple (and where a project might therefore sit ambiguously on the axis). Relatedly, "formalized/flexible" (engagement structure) and "standardized/unstandardized" (protocols) are related but not identical concepts; a brief clarification of exactly what the axis measures would help readers place their own efforts.

RESPONSE: We plan to edit the manuscript (including Section 2.1) to align with these suggestions. This will entail further justifying the axis collapse and adding caveats that acknowledge situations in which standardization and centralization decouple from each other. We will also add text to clarify what the collapsed axis represents to help readers place their own efforts along it. This will include the following next text (new text in bold): "**In the SanClements gradients, long-term research networks occupy low centralization and low standardization, coordinated distributed experiments and observations occupy the midpoint, and centralized observing networks occupy high centralization and high standardization. SanClements points to one example that was intentionally designed outside of this relationship: NEON's Assignable Assets program, which allows**

researchers to fulfill independent activities within NEON's infrastructure, therefore decoupling the standardization and centralization.

Our conceptual framework (Fig. 2) builds upon these prior frameworks, collapsing the degree of standardization and centralization into a single axis. However, unlike the SanClements et al. (2022) framework, our focus in the first axis is on the participants generating the distributed data, rather than on the management approach. **We describe this axis as a participant-based continuum that ranges from highly formalized to highly flexible. This is similar in spirit to the SanClements axes - highly formalized participants often align with greater centralization of funding, management, education and cyberinfrastructure and greater standardization of sensors, protocols, and algorithms, and the opposite is often true of highly flexible participants; however it is not identical, and there is variation how formalization/flexibility of participants connects to the concepts of centralization and standardization. Since SanClements already describes standardization/centralization well, here we explore a participant-based continuum.“**

***COMMENT**: 3. Scope statement vs. examples (human-on-the-ground coordination)*

Section 1 explicitly scopes the paper to "efforts with human-to-human coordination requiring on-the-ground efforts." However, several featured examples are largely autonomous/sensor-based - most notably Argo (Table 1), a global array of autonomous profiling floats, and to a degree the flux-tower and streamgaging networks. Please reconcile the scoping statement with these examples: either broaden the scope statement, or clarify that the human-coordination criterion applies to the network's design and governance rather than to each measurement. As written there is a tension a careful reader will notice.

RESPONSE: We plan to clarify in the introduction that the scope includes on-the-ground efforts that use sensor deployments requiring human-to-human coordination to deploy but not necessarily to collect singular measurements. New text: “On-the-ground efforts include settings where data collection may be autonomous (e.g., sensors) but still require human coordination, or may require sample/observation collection by a human.”

***COMMENT**: 4. Evidence base and its own geographic/institutional bias*

The paper rightly flags that existing networks are heavily biased toward western Europe and North America (Sect. 3). It is worth noting explicitly that the manuscript's own evidence base - a 2025 workshop and the 35+ exemplar projects, many of them author-affiliated and US-DOE/PNNL-centric (WHONDRS, EXCHANGE, MONet, NGEE Arctic, CROCUS, GROWdb) - is subject to the same selection bias. This does not undermine the synthesis, but a sentence acknowledging the provenance and potential representativeness limits of the exemplar set would improve transparency and match the self-critical tone the authors apply elsewhere.

RESPONSE: We plan to revise the introduction to Section 3 to note the geographical bias of the workshop attendees and the assessed distributed data generation efforts. The new text will read: "It is important to note that both the workshop attendees and the assessed data generation efforts are geographically biased, with many from the United States, which informs the perspectives represented."

***COMMENT**: 5. Missing: CARE principles alongside FAIR*

The output pillar and conclusions lean heavily on FAIR (Wilkinson et al., 2016), which is appropriate. But the paper repeatedly invokes co-design with "rights holders," communities, and the public, and the equitable-participation literature. In that context, the CARE Principles for Indigenous Data Governance (Collective benefit, Authority to control, Responsibility, Ethics) are now a standard complement to FAIR for community-generated and Indigenous data and would strengthen both Section 3 (output) and the equity framing. Their absence is a notable gap given the paper's emphasis on broad, equitable engagement.

RESPONSE: We plan to revise the text to incorporate the CARE principles in several places including in the introduction third paragraph and in the Section 3 Output sub-section. We plan to revise the text to the following (**new text in bold**): "For example, ensuring broad participation and codesign with interested parties (i.e., stakeholders, organizations, agencies, landowners, land managers, rights holders, and the public), **including considerations for Indigenous Data Sovereignty (Carroll et al, 2022),** remains a key limitation for distributed data generation (Molla et al., 2025; Muraina and Jimoh, 2022)."

We will also add, in Section 3, the following: "Decisions around sharing of data and outcomes must be made with intentionality, particularly with regard to Indigenous Data Sovereignty where application of the CARE principles (Collective Benefit, Authority to Control, Responsibility, and Ethics) may be appropriate (Carroll et al. 2020; Jennings et al., 2023)."

MINOR / SPECIFIC COMMENTS

*- **COMMENT**: Citation mismatch (Sect. 4.1, p. 13, ~l. 291): The sentence "expediting cross-domain translation, similar to LLM use in cross-language translation (Gui et al., 2019)" is supported by Gui et al. (2019), Geoforum - "Globalization of science and international scientific collaboration: A network perspective." That paper is about collaboration-network geography and predates modern LLMs; it does not support a claim about LLM-based (machine) cross-language translation. Or do I oversee the point? Please replace with an appropriate reference on neural/LLM machine translation, or reword.*

RESPONSE: We plan to revise the sentence and include appropriate references. This sentence will read as follows (**new text in bold): "**LLMs could improve collaborations across similar or connected distributed data generation projects by expediting cross-domain translation, similar to

LLM use in cross-language translation **in text and audio/visual media (Ke and Ribeiro 2026).”**

Ke and Ribeiro 2026: https://doi.org/10.1007/s44163-026-01589-2

---

*- **COMMENT**: Citation mismatch (Sect. 2.2.1, ~l. 186): "DroughtNet (Poyatos et al., 2021)." The cited reference (Poyatos, Werner & Martinez-Vilalta, 2021) is the SAPFLUXNET sap-flow database, not DroughtNet (the International Drought Experiment network). Please correct the citation or the network name.*

RESPONSE: Thank you for flagging this. We will include the appropriate reference for Drought-Net (Knapp et al. 2016; https://doi.org/10.1111/gcb.13504).

---

*- **COMMENT**: Consistency of the core term. The abstract defines distributed data generation as "data collected from multiple sources and locations," while the Introduction uses "geographically dispersed locations." Minor, but please make the definition verbatim consistent between abstract and main text.*

RESPONSE: The definition will be revised so that it is identical in both places. Revised abstract text (new text in bold): “Distributed data generation, or **data collected from geographically dispersed locations** using standardized approaches, and involving coordination among **numerous** investigators, has emerged as a powerful approach to meet contemporary demands for scalable environmental knowledge.”

---

*- **COMMENT**: Section 3 reads partly as an annotated bibliography. The best-practices section is valuable but at times enumerates examples more than it distills transferable principles. Consider foregrounding the principle for each best practice (one crisp sentence) and relegating the project instances to Table 1, so the takeaways are easier to extract.*

RESPONSE: We plan to revise the structure of Section 3 to begin each sub-section with the best practices and to split Table 1 into three tables (one for each-subsection) to make it easier for readers to navigate.

---

*- **COMMENT**: The authors already note "concrete examples of AI-enabled distributed efforts remain limited." It may help readers to state explicitly that Sections 2-3 (framework and best practices) are the durable contribution while Section 4 is deliberately speculative/horizon-scanning - this manages expectations and protects the paper's longevity.*

RESPONSE: We plan to revise the text in the introduction to Section 4 to include a sentence to convey this message to readers. New text now reads as follows (new text in bold): “Below we propose three growth areas for distributed data generation facilitated by AI advancements, **which differ from Sect. 2 and 3 in that they are deliberately exploratory and forward-looking**.”

TECHNICAL CORRECTIONS

*- **COMMENT**: Fig. 1 caption: "using multi-model engagement" appears to be a typo for "multi-modal engagement" (the term used consistently in Sect. 3).*

RESPONSE: Thank you. This will be corrected.

---

*- **COMMENT**: Sect. 4.2 (~l. 347): "(U.S. DOE 2026b." - missing closing parenthesis.*

RESPONSE: Thank you. This will be corrected.