Summary of the CLeAR Documentation Framework for AI Transparency:

Recommendations for Practitioners and Context for Policymakers

With growing attention to AI regulation, rights-based principles, data equity, and risk mitigation, this is a pivotal moment to think about the social impact of AI, including its risks and harms, and the implementation of accountability and governance more broadly. Most proposed AI regulation mandates some level of transparency, as transparency is crucial for addressing the ways in which AI systems impact people. Transparency can be realized, in part, by providing information about how the **data** used to develop and evaluate the AI system was collected and processed, how AI **models** were built, trained, and fine-tuned, and how models and **AI systems** were evaluated and deployed. Towards this end, documentation has emerged as an essential component of AI transparency and a foundation for responsible AI development.

This report introduces the **CLeAR Documentation Framework,** designed to help practitioners and policymakers understand what principles should guide the process and content of AI documentation and how to create such documentation. The report introduces four principles for documentation and offers definitions, recommends approaches, explains tradeoffs, highlights open questions, and helps guide the implementation of documentation. It builds on and is aligned with previous principles-based frameworks for documentation. The CLeAR Principles state that documentation should be:

- **Comparable:** Able to be compared; having similar components to documentation of other datasets, models, or systems to permit or suggest comparison; enabling comparison by following a discrete, well-defined format in process, content, and presentation.
- **Legible:** Able to be read and understood; clear and accessible for the intended audience.
- **Actionable:** Able to be acted on; having practical value, useful for the intended audience.
- **Robust:** Able to be sustained over time; up to date.

Recommendations for AI Documentation

A healthy documentation ecosystem requires the participation of both practitioners and policymakers. Both audiences benefit from understanding the context of documentation, current approaches, and tradeoffs when it comes to the implementation of AI documentation. Thus, we offer additional context and recommendations:

1. **Importance of documenting datasets, models, and AI systems.**
   Documentation is worthwhile for various stakeholders. It enhances systems reliability by creating opportunities to reflect on development decisions, enables knowledge transfer across organizational silos, and encourages responsible use. Further, it provides information that can be used to determine the appropriateness of an AI system or its underlying data or models, thus helping inform consumer choice, advocacy work, regulation development, and regulation enforcement. This information can also enable recourse in the event of harms caused by or inquiries into the AI system and drive accountability.

2. **Documentation should occur throughout the lifecycle.**
   Policies about documentation should address the whole lifecycle of algorithmic systems, which begins at the ideation phase, rather than at the time of product launch. Documentation should be developed alongside AI models or systems and include CLeAR (**C**omparable, **L**egible, **A**ctionable, **R**obust) information to facilitate meaningful transparency into datasets and models. Practitioners should also document the development process itself, and expand the focus of documentation to be context-aware.

3. **Consider risk and impact assessments as complementary to other documentation.**
   Risk and Impact Assessments are increasingly common mechanisms for evaluating threats and vulnerabilities of systems, as well as considering and measuring, implications for individuals, communities, and their environments. This is in contrast to other documentation efforts that cover other parts of the AI system lifecycle, for a diversity of audiences, and focused on a wide variety of system components (such as datasets and models). That said, these efforts are complementary. Dataset, model, and AI systems documentation can also be informed by and should be updated following findings from impact assessments.

4. **The opportunity to drive behavior through documentation requirements.**
   The intention and value of documentation can shift dramatically depending on what is shared, with which audiences, and when. For example, while documentation is a mechanism for transparency even if not shared with external audiences, policymakers can drive behaviors by requiring documentation be disclosed at certain times and for certain audiences.

5. **Documentation requires robust organizational support.**
   Documentation takes time and requires appropriate skills. Therefore, documentation requirements are unlikely to be successful without robust organizational support, including capacity building and additional resources for technical and workflow processes. Having buy-in at the executive level and aligning the goals of documentation with the organization's goals is a first critical step to building momentum around real change in an organization's approach to documentation.

Dataset, model, and AI system documentation are straightforward mechanisms for transparency into AI systems. The CLeAR Framework enables a foundation for designing and implementing documentation, while considering tradeoffs and encouraging holistic thinking about documentation needs. Our hope is that this framework will help practitioners to create and use documentation, and support policymakers to better understand the importance of documentation and tradeoffs that should be considered for area-specific documentation regulation. Only through collective efforts can we ensure that AI is created and deployed responsibly.

Authors

Kasia Chmielinski* (Data Nutrition Project & Harvard University)
Sarah Newman* (Data Nutrition Project & Harvard University)
Chris N. Kranzinger* (Data Nutrition Project)
Michael Hind (IBM Research)
Jennifer Wortman Vaughan (Microsoft Research)
Margaret Mitchell (Hugging Face)
Julia Stoyanovich (New York University)
Angelina McMillan-Major (independent researcher)
Emily McReynolds (independent researcher)
Kathleen Esfahany (Data Nutrition Project & Harvard University)
Mary L. Gray (Microsoft Research, Indiana University & Harvard University)
Maui Hudson (TKRI, University of Waikato)
Audrey Chang (Data Nutrition Project & Harvard University)

**Joint first authors*

Acknowledgements

The authors would like to thank the following colleagues for their thoughtful feedback on and contributions to this report: Emily Bender (University of Washington), Ashley Casovan (IAPP, Carleton University), Mahima Pushkarna (Google Research), Kush R. Varshney (IBM Research).

We would additionally like to thank Lindsay Underwood, Alicia Pritt, and Laura Manley at the Shorenstein Center for their support publishing this paper.

Contact: Kasia Chmielinski and Sarah Newman