

Annotation and Benchmarking on Understanding and Transparency of Machine learning Lifecycles (ABOUT ML)

Chapter 1: Project Overview

1.1 Statement of Importance for ABOUT ML project

As machine learning (ML) becomes central to many decision-making processes - including high stakes decisions in criminal justice and banking - the organizations deploying such automated decision-making systems face increased pressure for transparency on how these decisions are made. Annotation and Benchmarking on Understanding and Transparency of Machine learning Lifecycles (ABOUT ML) is a project of the Partnership on AI (PAI) working towards establishing new norms on transparency via documentation by evaluating best practices throughout the ML system lifecycle from design to deployment.

Presently, there is neither consensus on which practices work best nor on what information needs to be disclosed and for which goals. Moreover, the definition of transparency itself is highly contextual. Because there is currently no standardized process across the industry, each team that wants to improve transparency in their ML systems must address the entire suite of questions about what transparency means for their team, product, and organization, given their specific goals and constraints. Our goal is to provide a head start to that process of exploration. We will offer a summary of recommendations and practices that is mindful of the variance in transparency expectations and outcomes. We hope to provide an adaptive resource to highlight common themes about transparency, rather than a rigid list of requirements. This should serve to guide teams to identify and address context-specific challenges.

While substantial decentralized experimentation is currently taking place,¹ the ABOUT ML project aims to accelerate progress by pooling insights more quickly, sharing resources, and reducing redundancy of efforts. PAI aims to provide a gathering place for researchers, AI practitioners, civil society organizations, and especially those affected by AI products to discuss, debate, and ultimately decide on broadly applicable recommendations. ABOUT ML thus seeks to bring together representatives from a wide range of relevant stakeholders to improve public discussion, and promulgate best practices into new industry norms that will reflect diverse interests and chart a path forward for greater transparency in ML. We encourage any organization undertaking transparency initiatives to share their practices and lessons learned to PAI for incorporation into future versions of this document.

This is an ongoing project with regular evaluation points to keep up with the rapidly evolving field of AI. PAI has a broad partnership of over a hundred organizations, including corporate developers of AI, civil society organizations, and academic institutions, that will be

¹Holstein, K., Vaughan, J.W., Daumé, H., Dudík, M., & Wallach, H.M. (2018). Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need? CHI.

involved in the drafting and vetting of documentation themes recommended in this document. In addition, PAI will be engaging with the Tech Policy Lab at the University of Washington to run a Diverse Voices² panel to gather opinions from stakeholders whose perspectives might not otherwise be captured. Through this process, PAI will also seek to gain deeper insights into the Diverse Voices process in order to inform the ABOUT ML recommendations on how to incorporate the perspectives of diverse stakeholders.

We will begin by highlighting recurrent themes in ML research about documentation, but our ambitious aim is to identify all practices that have sufficient positive data of efficacy to be deemed best practices in ML transparency. PAI will then aim to disseminate these best practices broadly into new norms to improve transparency in the AI industry.

1.1.1 About this document

Given the growing influence of AI transparency research at a global scale, leveraging PAI's position as a multi-stakeholder organization to amalgamate key contributions across separate partner initiatives will serve to organize and advocate for the underlying themes of current ML system documentation proposals. Initiatives like ABOUT ML that offer guidance on ML documentation will give companies a head start on a path of overseeing, auditing, and monitoring ML technologies, contributing to the key business goal of earning and keeping trust from consumers and policymakers.

The focus in this version 0 draft is on extracting major themes from recent research on recommendations for transparency documentation. Future versions will merge existing practices with insights from research, formalize best practices through an investigation into attempts to implement recommendations, and set new industry norms for documentation in ML lifecycles. Future drafts may also include commentary on other enablers of ML transparency, such mechanisms to adjust team and institutional settings, model interpretability tools, test suites and modified evaluation procedures, and more detail on necessary feedback loops for transparency. See below section "ABOUT ML project process overview" for further detail.

²Young, M., Magassa, L. and Friedman, B. (2019) Toward inclusive tech policy design: a method for underrepresented voices to strengthen tech policy documents. *Ethics and Information Technology* 21(2), 89-103.

1.1.2 ABOUT ML project process overview

ABOUT ML Process

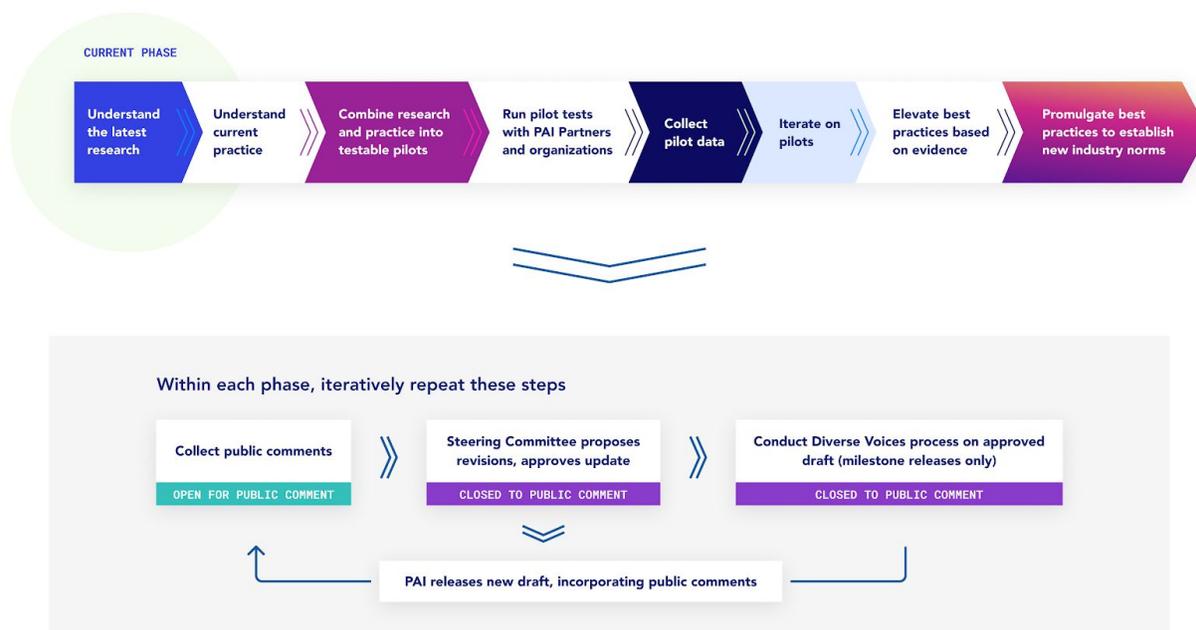


Figure 1.1 : ABOUT ML project process details

PAI is launching the ABOUT ML iterative multistakeholder process with this initial draft v0 in order to initiate a broader community discussion. The goal is to update this draft into future releases and move towards best practices by going through the following phases:

- Phase 1 : Understand the latest research (project is currently here)
- Phase 2 : Understand current practice
- Phase 3 : Combine research theory and results of current practice into testable pilots
- Phase 4 : Run pilot tests with PAI Partners and organizations
- Phase 5 : Collect data from pilot tests for transparency practices
- Phase 6 : Iterate on pilots with the latest research and practice
- Phase 7 : When there is sufficient body of evidence for a certain practice, elevate it to a best practice
- Phase 8 : Promulgate effective practices to establish new industry norms for transparency

PAI recognizes that this effort can only succeed with input from as broad a set of stakeholders as possible, and will be seeking input not only from our 90+ partners, but also from stakeholders from academia, civil society organizations, companies designing and deploying ML technology, and the general public. We welcome your [participation](#).

The process is modeled after iterative ongoing processes to design internet standards (such as [W3C](#)³, [IETF](#)⁴, and [WHATWG](#)⁵) and includes a public [forum for discussion](#) and a place to submit any proposed changes. We welcome you to join in the public discussion and to submit proposed changes as many times as desired.

Public comments will be collected and batch evaluated by the [ABOUT ML Steering Committee](#), a group of ~30 experts, researchers and practitioners recruited from a diverse set of PAI Partner organizations. The Steering Committee will guide the process of updating ABOUT ML drafts based on the public comments submitted and new developments in research and practice. They will vote to approve new releases by “rough consensus”⁶ commonly used by other multi-stakeholder working groups. They will convene 1-3 times a year, depending on the volume of proposed changes and the velocity of change of research and practice.

To prevent contributors from commenting on an out-of-date draft, the document will be closed to proposed changes for ~2-4 weeks prior to each Steering Committee meeting. This closure will be announced here with at least 2 weeks of notice. The first meeting of the Steering Committee will be on September 28th, 2019, so the v0 draft will be closed for comment from September 14th until the end of the update process.

To ensure that diverse perspectives— especially those from communities historically excluded from technology decision-making—contribute to any ABOUT ML recommendations, PAI is engaging with our Partner, the Tech Policy Lab at the University of Washington to conduct [Diverse Voices](#)⁷ panels. This method was designed to gather feedback from stakeholders whose perspectives might not otherwise be consulted and to ensure that those perspectives are reflected in the released text. Thus, for any ABOUT ML releases that go through the Diverse Voices process, the panel feedback will be the last edits incorporated before a new release. This also means that each round of Diverse Voices panels will cause public comment on the document to be closed for several months, although the [public forum](#) will remain open for discussion during that time. Public comment on the document itself will re-open with the new release of the draft. The first round of Diverse Voices panels for ABOUT ML will be held between October and December 2019. Thus, draft v0 will be closed from September 14th, 2019 until early 2020, and will reopen upon the release of draft v1 in early 2020. Additional panels will be convened approximately once a year, especially for milestone releases when the ABOUT ML project progresses through the phases outlined above.

³ World Wide Web Consortium Process Document (W3C) process outlined here: <https://www.w3.org/2019/Process-20190301/>

⁴ Internet Engineering Task Force (IETF) process outlined here: <https://www.ietf.org/standards/process/>

⁵ The Web Hypertext Application Technology Working Group (WHATWG) process outlined here: <https://whatwg.org/faq#process>

⁶ Oever, N., Moriarty, K. The Tao of IETF: A novice's guide to the Internet Engineering Task Force. <https://www.ietf.org/about/participate/tao/>.

⁷ Young, M., Magassa, L. and Friedman, B. (2019) Toward inclusive tech policy design: a method for underrepresented voices to strengthen tech policy documents. *Ethics and Information Technology* 21(2), 89-103.

ABOUT ML Timeline

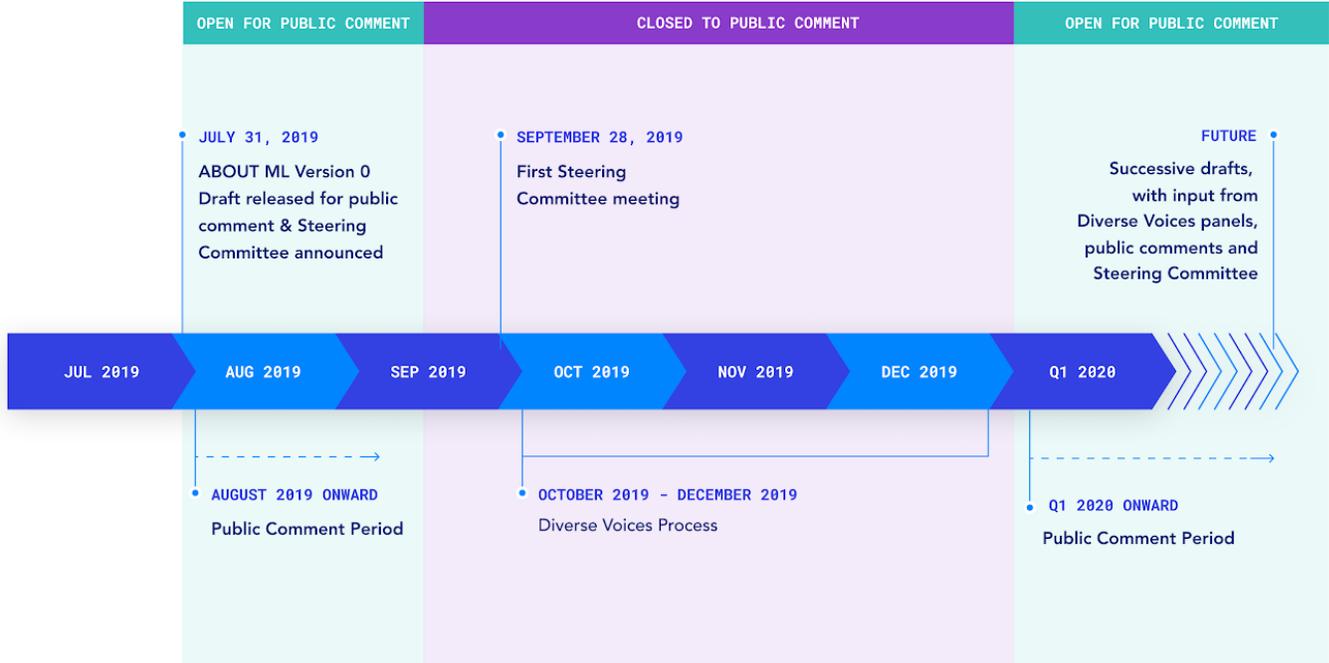


Figure 1.2: ABOUT ML project timeline details