JusTech
Standards, Open Source, and IP Transactions

Unlocking AI Transparency:
The Access Methodology Framework for Classifying AI System Openness

Abstract:

AI regulations, standards, and guidelines are rapidly evolving to address the risks and ethical concerns associated with AI systems. However, the fragmented and inconsistent nature of these rules presents significant challenges for businesses trying to comply, especially in the absence of a global AI governance framework or standardized conformity assessment tests. A key focus of many regulations is transparency—ensuring that AI systems' operation and decision-making processes can be understood. Yet even developers often struggle to explain their own AI systems' behavior. Although standards bodies are working on solutions such as conformity testing, progress lags behind the pace of AI innovation. To tackle these challenges, some stakeholders, including policymakers, are advocating for open licensing schemes to enable the evaluation of AI system components and ensure they cannot be misused for harmful, unlawful or immoral purposes.

Such open licensing schemes are inspired by the Open Source Initiative (OSI) definition for open source software, the Open Source Definition (OSD). Current open licensing schemes largely align with the OSD, which mandates that the licensed software be royalty-free, modifiable, and reusable for any purpose, with source code made publicly available. OSI has relied on the same principles in adopting an Open Source AI definition. While these principles may enhance transparency, they would also prevent AI system producers from keeping their components as trade secrets and would require them to license these components for free to competitors, potentially undermining the producers' business interests.

Concerns have been raised that AI developers could engage in "open-washing," branding their AI systems as open source although some of the AI system components are not licensed under an OSD-compliant license. Those raising such concerns believe "open-washing" undermines transparency obligations and allows larger entrenched companies to exploit innovations from smaller startups without contributing back. In response to these concerns, researchers have proposed frameworks to classify and evaluate the openness of AI systems.

The Model Open Framework, for example, categorizes AI components into three levels of completeness, requiring components to be available under OSD-compliant licenses for transparency. Another proposed approach uses a multidimensional gradient framework, assessing openness across several factors while still emphasizing adherence to OSD principles. These proposals, like OSI's Open Source AI definition, equate transparency with licensing terms that prohibit monetary compensation, expressly authorize the use of AI system components in competitive systems for free, and prevent AI system producers from maintaining their trade secrets.

While these proposals aim to ensure compliance with AI regulations, they may stifle innovation, discourage investment, and disadvantage AI startups that rely on proprietary licensing to protect their competitive edge. Moreover, fully open AI systems can pose risks by enabling malicious actors to exploit AI technologies for harmful, illicit, or immoral purposes.

This paper introduces an alternative approach: the **Access Methodology Framework (AMF)**. The AMF offers a flexible, use-case-specific approach to balancing transparency with innovation and risk mitigation. It emphasizes black-box conformity tests where feasible, avoiding

the need for OSD-style licenses. When black-box conformity tests are not available and transparency is warranted, the AMF merely requires that the applicable license permit testing, evaluation, and analysis solely for determining how the AI system operates and makes decisions. The AMF proposes two access options to the AI system's source code, data, model weights, or documentation: the Private Access (PA) License where access is conditioned on an obligation of confidentiality; and the Open Access (OA) License where access is granted without confidentiality restrictions.

The choice between PA and OA licenses depends on the size and definition of the applicable stakeholder groups given a particular use case. PA licenses suit smaller, well-defined groups, while OA licenses work better for larger, less defined groups. The AMF accommodates a much wider range of licenses beyond the OSD open source approach, enabling business-friendly terms that encourage innovation and investment. For example, both PA and OA license may prohibit commercial use or redistribution since neither right is needed to assess how the AI system operates or makes decisions. By providing this flexibility, the AMF strikes a balance between fostering AI's benefits and addressing its risks, supporting sustainable growth in the AI industry while maintaining accountability and transparency.