

AI-CONTROL Workshop

IABOPEN Summary

Problem Statement

Large Language Models and other machine learning techniques require voluminous input data, and one common source of such data is the Internet -- usually, "crawling" Web sites for publicly available content, much in the same way that search engines crawl the Web.

This similarity has led to an emerging practice of allowing the Robots Exclusion Protocol (RFC 9309) to control the behavior of AI-oriented crawlers.

This emerging practice raises many design and operational questions. It is not yet clear whether robots.txt (the mechanism specified by RFC 9309) is well-suited to controlling AI crawlers. A content creator or host may not be able to distinguish a crawler used for search indexing from a crawler used for LLM ingest – and indeed some crawlers may be used for both purposes. Potential use cases may extend across many different units of content, policies to be signaled, and types of content creators. Before robots.txt becomes a de facto solution to AI crawling opt-out, it is necessary to examine whether it is an appropriate mechanism: in particular, whether the creator of a particular unit of content can realistically and fully exercise their right to opt-out, and the scope of data ingest to which that opt-out applies.

This workshop aims to explore practical opt-out mechanisms for AI, and build an understanding of use cases, requirements, and other considerations in this space.

Questions we set out to answer

- What are the success criteria for AI Control?
- What needs to be expressed?
- What technical capabilities are necessary to achieve specific objectives?
- What challenges/advantages does each approach have (opt-out, opt-in, technical blocking, etc.)?
- What are the boundaries, overlaps, and limitations in technical mechanisms and legal regulation?

Workshop summary

- Broad participation
 - AI vendors, Large publishers, Government representatives, civil society, legal experts
- Focused discussion
 - Did not try to resolve conflicts between rights and innovation
- Core insight
 - Improving communication of preferences benefits everyone

Outcomes

- Focus on:
 - Overall framework for preferences
 - Attachment mechanisms (where IETF-controlled)
 - Simple common vocabulary
- Out of initial scope:
 - Authentication/authorization
 - Audit/transparency mechanisms
 - Registries

Observations

- Feedback was very positive – good combination of technical and policy focus
- Some said they would not have participated if it were not in-person + Chatham House Rule – perceived safety
- Positive model for starting work that has policy implications

After the workshop

- First impressions blog
 - <https://www.ietf.org/blog/impressions-ai-control-workshop/>
- Draft workshop report published Monday
 - <https://datatracker.ietf.org/doc/draft-iab-ai-control-report/>
- Mailing list
 - <https://mailman3.ietf.org/mailman3/lists/ai-control@ietf.org/>
- **Side meeting to discuss potential IETF work**
 - **Thursday 11:30-13:00, Wicklow 2A**