Network Working Group
Internet-Draft
Intended status: Informational
Expires: 10 March 2026

M. Nottingham

S. Krishnan

6 September 2025

IAB AI-CONTROL Workshop Report
draft-iab-ai-control-report-02

Abstract

The AI-CONTROL Workshop was convened by the Internet Architecture Board (IAB) in September 2024. This report summarizes its significant points of discussion and identifies topics that may warrant further consideration and work.

Note that this document is a report on the proceedings of the workshop. The views and positions documented in this report are those of the workshop participants and do not necessarily reflect IAB views and positions.

Discussion Venues

This note is to be removed before publishing as an RFC.

Source for this draft and an issue tracker can be found at https://github.com/intarchboard/draft-iab-ai-control-report.

Status of This Memo

Table of Contents

1. Introduction 2
1.1. Chatham House Rule 3
1.2. Views Expressed in this Report 3
2. Workshop Scope and Discussion 4
2.1. Crawl Time vs. Inference Time 5
2.1.1. Multiple Uses for Crawl Data 5
2.1.2. Application of Preferences 5
2.2. Trust 6
2.3. Attachment 6
2.3.1. robots.txt (and similar) 6
2.3.2. Embedding 7
2.3.3. Registries 8
2.4. Vocabulary 8
3. Conclusions 8
3.1. Potential Standards Work 9
3.1.1. Out of Initial Scope 9
4. Security Considerations 9
5. Informative References 9
Appendix A. About the Workshop 10
A.1. Agenda 10
A.1.1. Thursday 2024-09-19 11
A.1.2. Friday 2024-09-20 11
A.2. Attendees 11
IAB Members at the Time of Approval 12
Acknowledgements 12
Authors' Addresses 13

1. Introduction

The Internet Architecture Board (IAB) holds occasional workshops designed to consider long-term issues and strategies for the Internet, and to suggest future directions for the Internet architecture. This long-term planning function of the IAB is complementary to the ongoing engineering efforts performed by working groups of the Internet Engineering Task Force (IETF).

The Internet is one of the major sources of data used to train large language models (Large Language Models (LLMs), or more generally, "Artificial Intelligence (AI)"). Because this use was not envisioned by most publishers of information on the Internet, a means of expressing the owners' preferences regarding AI crawling has emerged, sometimes backed by law (e.g., in the European Union's AI Act [AI-ACT]).

The IAB convened the AI-CONTROL Workshop on 19-20 September 2024 to "explore practical opt-out mechanisms for AI and build an understanding of use cases, requirements, and other considerations in this space" [CFP]. In particular, the emerging practice of using the Robots Exclusion Protocol [RFC9309] -- also known as "robots.txt" -- has not been coordinated between AI crawlers, resulting in considerable differences in how they treat it. Furthermore, robots.txt may or may not be a suitable way to control AI crawlers. However, discussion was not limited to consideration of robots.txt, and approaches other than opt-out were considered.

To ensure many viewpoints were represented, the program committee invited a broad selection of technical experts, AI vendors, content publishers, civil society advocates, and policymakers.

1.1. Chatham House Rule

Participants agreed to conduct the workshop under the Chatham House Rule [CHATHAM-HOUSE], so this report does not attribute statements to individuals or organizations without express permission. Most submissions to the workshop were public and thus attributable; they are used here to provide substance and context.

Appendix A.2 lists the workshop participants, unless they requested that this information be withheld.

1.2. Views Expressed in this Report

This document is a report on the proceedings of the workshop. The views and positions documented in this report are expressed during the workshop by participants and do not necessarily reflect IAB's views and positions.

Furthermore, the content of the report comes from presentations given by workshop participants and notes taken during the discussions, without interpretation or validation. Thus, the content of this report follows the flow and dialogue of the workshop but does not attempt to capture a consensus.

2. Workshop Scope and Discussion

The workshop began by surveying the state of AI control.

Currently, Internet publishers express their preferences for how their content is treated for purposes of AI training using a variety of mechanisms, including declarative ones, such as terms of service, embedded metadata, and robots.txt [RFC9309], and active ones, such as use of paywalls and selective blocking of crawlers (e.g., by IP address, User-Agent).

There was disagreement about the implications of AI opt-out overall. Research presented at the workshop [DECLINE] indicates that the use of such controls is becoming more prevalent, reducing the availability of data to AI (for purposes including training and inference-time usage). Some of the participants expressed concern about the implications of this -- although at least one AI vendor seemed less concerned by this, indicating that "there are plenty of tokens available" for training, even if many opt out. Others expressed a need to opt out of AI training because of how they perceive its effects on their control over content, seeing AI as usurping their relationships with customers and a potential threat to whole industries.

However, there was quick agreement that both viewpoints were harmed by the current state of AI opt-out -- a situation where "no one is better off" (in the words of one participant).

Much of that dysfunction was attributed to the lack of coordination and standards for AI opt-out. Currently, content publishers need to consult with each AI vendor to understand how to opt out of training their products, as there is significant variance in each vendor’s behaviour. Furthermore, publishers need to continually monitor both for new vendors, and for changes to the policies of the vendors they are aware of.

Underlying those immediate issues, however, are significant constraints that could be attributed to uncertainties in the legal context, the nature of AI, and the implications of needing to opt out of crawling for it.

2.1. Crawl Time vs. Inference Time

Perhaps most significant is the "crawl time vs. inference time" problem. Statements of preference are apparent at crawl time, bound to content either by location (e.g., robots.txt) or embedded inside the content itself as metadata. However, the target of those directives is often disassociated from the crawler, either because the crawl data is not only used for training AI models, or because the preferences could be applicable at inference time.

2.1.1. Multiple Uses for Crawl Data

A crawl’s data might have multiple uses because the vendor also has another product that uses it (e.g., a search engine), or because the crawl is performed by a party other than the AI vendor. Both are very common patterns: operators of many Internet search engines also train AI models, and many AI models use third-party crawl data. In either case, conflating different uses can change the incentives for publishers to cooperate with the crawler.

Well-established uses of crawling, such as Internet search, were seen by participants as at least partially aligned with the interests of publishers: they allow their sites to be crawled, and in return, they receive higher traffic and attention due to being in the search index. However, several participants pointed out that this symbiotic relationship does not exist for AI training uses -- with some viewing AI as hostile to publishers, because it has the capacity to take traffic away from their sites.

Therefore, when a crawler has multiple uses that include AI, participants observed that "collateral damage" was likely for non-AI uses, especially when publishers take more active control measures, such as blocking or paywalls, to protect their interests.

Several participants expressed concerns about this phenomenon’s effects on the ecosystem, effectively "locking down the Web" with one opining that there were implications for freedom of expression overall.

2.1.2. Application of Preferences

When data is used to train an LLM, the resulting model does not have the ability to only selectively use a portion of it when performing a task, because inference uses the whole model, and it is not possible to identify specific input data for its use in doing so.

This means that while publishers' preferences may be available when content is crawled, they generally are not when inference takes place. Those preferences that are stated in reference to use by AI -- for example, "no military uses" or "non-commercial only" cannot be applied by a general-purpose "foundation" model.

This leaves a few unappealing choices to AI vendors that wish to comply with those preferences. They can simply omit such data from foundation models, thereby reducing their viability. Or, they can create a separate model for each permutation of preferences -- with a likely proliferation of models as the set of permutations expands.

Compounding this issue was the observation that preferences change over time, whereas LLMs are created over long time frames and cannot easily be updated to reflect those changes. Of particular concern to some was how this makes an opt-out regime "stickier" because content that has no associated preference (such as that which predates the authors' knowledge of LLMs) is allowed to be used for these unforeseen purposes.

2.2. Trust

This disconnection between the statement of preferences and its application was felt by participants to contribute to a lack of trust in the ecosystem, along with the typical lack of attribution for data sources in LLMs, lack of an incentive for publishers to contribute data, and finally (and most noted) a lack of any means of monitoring compliance with preferences.

This lack of trust led some participants to question whether communicating preferences is sufficient in all cases without an accompanying way to enforce them, or even to audit adherence to them. Some participants also indicated that a lack of trust was the primary cause of the increasingly prevalent blocking of AI crawler IP addresses, among other measures.

2.3. Attachment

One of the primary focuses of the workshop was on _attachment_ -- how preferences are associated with content on the Internet. A range of mechanisms was discussed.

2.3.1. robots.txt (and similar)

The Robots Exclusion Protocol [RFC9309] is widely recognised by AI vendors as an attachment mechanism for preferences. Several deficiencies were discussed.

First, it does not scale to offer granular control over large sites where authors might want to express different policies for a range of content (for example, YouTube).

Robots.txt is also typically under the control of the site administrator. If a site has content from many creators (as is often the case for social media and similar platforms), the administrator may not allow them to express their preferences fully, or at all.

If content is copied or moved to a different site, the preferences at the new site need to be explicitly transferred, because robots.txt is a separate resource.

These deficiencies led many participants to feel that robots.txt cannot be the only solution to opt-out: rather, it should be part of a larger system that addresses its shortcomings.

Participants noted that other, similar attachment mechanisms have been proposed. However, none appear to have gained as much attention or implementation (both by AI vendors and content owners) as robots.txt.

2.3.2. Embedding

Another mechanism for associating preferences with content is to embed them into the content itself. Many formats used on the Internet allow this; for example, HTML has the <meta> tag, images have XMP and similar metadata sections, and XML and JSON have rich potential for extensions to carry such data.

Embedded preferences were seen to have the advantage of granularity, and of "travelling with" content as it is produced, when it is moved from site to site, or when it is stored offline.

However, several participants pointed out that embedded preferences are easily stripped from most formats. This is a common practice for reducing the size of a file (thereby improving performance when downloading it), and for assuring privacy (since metadata often leaks information unintentionally).

Furthermore, some types of content are not suitable for embedding. For example, it is not possible to embed preferences into purely textual content, and Web pages with content from several producers (such as a social media or comments feed) cannot easily reflect preferences for each one.

Participants noted that the means of embedding preferences in many formats would need to be determined by or coordinated with organisations outside the IETF. For example, HTML and many image formats are maintained by external bodies.

2.3.3. Registries

In some existing copyright management regimes, it is already common to have a registry of works that is consulted upon use. For example, this approach is often used for photographs, music, and video.

Typically, registries use hashing mechanisms to create a "fingerprint" for the content that is robust to changes.

Using a registry decouples the content in question from its location, so that it can be found even if moved. It is also claimed to be robust against stripping of embedded metadata, which is a common practice to improve performance and/or privacy.

However, several participants pointed out issues with deploying registries at Internet scale. While they may be effective for (relatively) closed and well-known ecosystems such as commercial music publishing, applying them to a diverse and very large ecosystem like the Internet has proven problematic.

2.4. Vocabulary

Another major focus area for the workshop was on _vocabulary_ -- the specific semantics of the opt-out signal. Several participants noted that there are already many proposals for vocabularies, as well as many conflicting vocabularies already in use. Several examples were discussed, including where existing terms were ambiguous, did not address common use cases, or were used in conflicting ways by different actors.

Although no conclusions regarding exact vocabulary were reached, it was generally agreed that a complex vocabulary is unlikely to succeed.

3. Conclusions

Participants generally agreed that on its current path, the ecosystem is not sustainable. As one remarked, "robots.txt is broken and we broke it."

Legal uncertainty, along with fundamental limitations of opt-out regimes pointed out above, limit the effectiveness of any technical solution, which will be operating in a system unlike either

robots.txt (where there is a symbiotic relationship between content owners and the crawlers) or copyright (where the default is effectively opt-in, not opt-out).

However, the workshop ended with general agreement that positive steps could be taken to improve the communication of preferences from content owners for AI use cases. In discussion, it was evident that the discovery of preferences from multiple attachment mechanisms is necessary to meet the diverse needs of content authors, and that therefore defining how they are combined is important.

We outline a proposed standard program below.

3.1. Potential Standards Work

The following items were felt to be good starting points for IETF work:

* Attachment to Web sites by location (in robots.txt or a similar mechanism)
* Attachment via embedding in IETF-controlled formats (e.g., HTTP headers)
* Definition of a common core vocabulary
* Definition of the overall regime; e.g., how to combine preferences discovered from multiple attachment mechanisms

It would be expected that the IETF would coordinate with other SDOs to define embedding in other formats (e.g., HTML).

3.1.1. Out of Initial Scope

It was broadly agreed that it would not be useful to work on the following items, at least to begin with:

* Enforcement mechanisms for preferences
* Registry-based solutions
* Identifying or authenticating crawlers and/or content owners
* Audit or transparency mechanisms

4. Security Considerations

This document is a workshop report and does not impact the security of the Internet.

5. Informative References

[CHATHAM-HOUSE]
Chatham House, "Chatham House Rule", n.d., <https://www.chathamhouse.org/about-us/chatham-house-rule>.

[CFP] Internet Architecture Board, "IAB Workshop on AI-CONTROL", n.d., <https://datatracker.ietf.org/group/aicontrolws/about/>.

[PAPERS] Internet Architecture Board, "IAB Workshop on AI-CONTROL Materials", n.d., <https://datatracker.ietf.org/group/aicontrolws/materials/>.

[AI-ACT] European Parliament, "Regulation (eu) 2024/1689 of the European Parliament and of the Council", 13 June 2024, <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.

[DECLINE] Longpre, S., Mahari, R., Lee, A., and C. Lund, "Consent in Crisis: The Rapid Decline of the AI Data Commons", 2025, <https://www.ietf.org/slides/slides-aicontrolws-consent-in-crisis-the-rapid-decline-of-the-ai-data-commons-00.pdf>.

[RFC9309] Koster, M., Illyes, G., Zeller, H., and L. Sassman, "Robots Exclusion Protocol", RFC 9309, DOI 10.17487/RFC9309, September 2022, <https://www.rfc-editor.org/rfc/rfc9309>.

Appendix A. About the Workshop

The AI-CONTROL Workshop was held on 2024-09-19 and 2024-09-20 at Wilkinson Barker Knauer in Washington DC, USA.

Workshop attendees were asked to submit position papers. These papers are published on the IAB website [PAPERS], unless the submitter requested it be withheld.

The workshop was conducted under the Chatham House Rule [CHATHAM-HOUSE], meaning that statements cannot be attributed to individuals or organizations without explicit authorization.

A.1. Agenda

This section outlines the broad areas of discussion on each day.

A.1.1. Thursday 2024-09-19

Setting the stage An overview of the current state of AI opt-out, its impact, and existing work in this space

Lightning talks A variety of perspectives from participants

A.1.2. Friday 2024-09-20

Opt-Out Attachment: robots.txt and beyond Considerations in how preferences are attached to content on the Internet

Vocabulary: what opt-out means What information the opt-out signal needs to convey

Discussion and wrap-up Synthesis of the workshop’s topics and how future work might unfold

A.2. Attendees

Attendees of the workshop are listed with their primary affiliation. Attendees from the program committee (PC) and the Internet Architecture Board (IAB) are also marked.

* Jari Arkko, Ericsson
* Hirochika Asai, Preferred Networks
* Farzaneh Badiei, Digital Medusa (PC)
* Fabrice Canel, Microsoft (PC)
* Lena Cohen, EFF
* Alissa Cooper, Knight-Georgetown Institute (PC, IAB)
* Marwan Fayed, Cloudflare
* Christopher Flammang, Elsevier
* Carl Gahnberg
* Max Gendler, The News Corporation
* Ted Hardie
* Dominique Haza谷l-Massieux, W3C
* Gary Ilyes, Google (PC)
* Sarah Jennings, UK Department for Science, Innovation and Technology
* Paul Keller, Open Future
* Elizabeth Kendall, Meta
* Suresh Krishnan, Cisco (PC, IAB)
* Mirja K端hlewind, Ericsson (PC, IAB)
* Greg Leppert, Berkman Klein Center
* Greg Lindahl, Common Crawl Foundation
* Mike Linksvayer, GitHub
* Fred von Lohmann, OpenAI
* Shayne Longpre, Data Provenance Initiative

* Don Marti, Raptive
* Sarah McKenna, Alliance for Responsible Data Collection; Sequentum
* Eric Null, Center for Democracy and Technology
* Chris Needham, BBC
* Mark Nottingham, Cloudflare (PC)
* Paul Ohm, Georgetown Law (PC)
* Braxton Perkins, NBC Universal
* Chris Petrillo, Wikimedia
* Sebastian Posth, Liccium
* Michael Prorock
* Matt Rogerson, Financial Times
* Peter Santhanam, IBM
* Jeffrey Sedlik, IPTC/PLUS
* Rony Shalit, Alliance For Responsible Data Collection; Bright Data
* Ian Sohl, OpenAI
* Martin Thomson, Mozilla
* Thom Vaughan, Common Crawl Foundation (PC)
* Kat Walsh, Creative Commons
* James Whymark, Meta

The following participants requested that their identity and/or affiliation not be revealed:

* A government official

IAB Members at the Time of Approval

Internet Architecture Board members at the time this document was approved for publication were:

* Matthew Bocci
* Roman Danyliw
* Dhruv Dhody
* Jana Iyengar
* Cullen Jennings
* Suresh Krishnan
* Mirja K端hlewind
* Warren Kumari
* Jason Livingood
* Mark Nottingham
* Tommy Pauly
* Alvaro Retana
* Qin Wu

Acknowledgements

The Program Committee and the IAB would like to thank Wilkinson Barker Knauer for their generosity in hosting the workshop.

We also thank our scribes for capturing notes that assisted in the production of this report:

* Zander Arnao
* Andrea Dean
* Patrick Yurky

Authors' Addresses

Mark Nottingham
Melbourne
Australia
Email: mnot@mnot.net
URI: https://www.mnot.net/

Suresh Krishnan
Email: suresh.krishnan@gmail.com