

Internet Engineering Task Force
Internet-Draft
Updates: 9309 (if approved)
Intended status: Informational
Expires: 6 October 2025

F. Canel, Ed.
K. Madhavan
Microsoft Corporation
4 April 2025

Robots Exclusion Protocol Extension to communicate AI preferences
vocabulary
draft-canel-robots-ai-control-01

Abstract

This document extends the RFC9309 by facilitating the communication of content usage specifically within the area of Artificial Intelligence (AI).

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 6 October 2025.

Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	2
2. Requirements Language	2
3. Specification	2
3.1. Robots Control Rules	2
3.2. Application Layer Response Header	3
3.3. HTML Meta Element	3
4. IANA Considerations	3

1. Introduction

Although the Robots Exclusion Protocol allows service owners to manage the access of automated clients, known as crawlers, to the URIs on their services as specified by [RFC8288], it does not offer any mechanisms to regulate how the data retrieved by these services can be used in the field of Artificial Intelligence (AI).

Application developers are requested to honor these tags. The tags are not a form of access authorization however.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Specification

3.1. Robots Control Rules

This RFC defines solely the communication method for the AI preference vocabulary, without detailing the vocabulary itself. The term `AIPreferenceVocabularyTerm` is a non-null string that serves as a placeholder for the AI preference vocabulary terms.

`DisallowAIPreferenceVocabularyTerm` - These lines specify whether a URI and its associated data are prohibited to be used for the scenario outlined by this AI preference vocabulary term.

`AllowAIPreferenceVocabularyTerm` - These lines specify whether a URI and its associated data are allowed to be used for the scenario outlined by this AI preference vocabulary term.

The values are case insensitive and honor the same matching and prioritization logic as existing Allow and Disallow robots.txt rules.

3.2. Application Layer Response Header

The same rules can also be set in the Application Layer Response Header X-Robots-Tag:

noAIPreferenceVocabularyTerm - instructs the parser that the data is prohibited to be used for the scenario outlined by this AI preference vocabulary term.

AIPreferenceVocabularyTerm - instructs the parser that the data can be used for the scenario outlined by this AI preference vocabulary term.

The values are case insensitive and honor the same matching and prioritization logic defined for X-Robots-Tag.

3.3. HTML Meta Element

Same rules can also be set via an HTML meta tag (Meta Robots):

```
<meta name="robots" content="noAIPreferenceVocabularyTerm">
```

```
<meta name="examplebot" content="AIPreferenceVocabularyTerm">
```

4. IANA Considerations

TODO: <https://www.rfc-editor.org/rfc/rfc9110.html#name-field-name-registry>