Internet-Draft
Intended status: Experimental
Expires: June 13, 2026

A. Schulze-Hueneke
JamOne-DE

December 10, 2025

The Ethical Crawler Agreement Protocol (ECAP)
draft-schulze-ecap-00

Abstract

This document specifies the Ethical Crawler Agreement Protocol (ECAP), an application-layer protocol for managing consent and ethical verification between web hosts and automated agents (crawlers, AI scrapers). Unlike the voluntary robots.txt standard, ECAP utilizes cryptographic signatures and a declarative policy handshake to ensure verifiable, consent-based access to web resources.

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 June 13, 2026.

Table of Contents

1. Introduction

The proliferation of automated agents, particularly for AI training

and data harvesting, has rendered traditional access control mechanisms like robots.txt insufficient. ECAP proposes a new standard where access is negotiated via a cryptographic handshake, ensuring that agents explicitly declare their intent and hosts explicitly grant consent.

2. Terminology

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.

Agent: The automated client initiating the request.
Host: The server hosting the resource.
Intent: The declared purpose of the data access (e.g., "research").

3. Protocol Overview

ECAP operates over HTTP/1.1 or HTTP/2. It introduces a set of headers and a well-known policy file located at "/.well-known/ecap-policy".

4. The ECAP Handshake

The handshake consists of the following steps:

1. The Agent requests "/.well-known/ecap-policy".
2. The Host returns the JSON policy containing allowed intents and a public key.
3. The Agent sends a signed request to the protected resource using ECAP headers.
4. The Host verifies the signature and intent against its policy.
5. The Host responds with "210 Accepted" or an appropriate error code (e.g., "403E Ethical Denied").

5. Data Formats

5.1. HTTP Headers

This document defines the following HTTP headers:

ECAP-Agent-ID: A unique identifier for the crawler (e.g., email).
ECAP-Intent: The purpose of the crawl (e.g., "indexing").
ECAP-Signature: Ed25519 signature of the request metadata.
ECAP-Timestamp: ISO 8601 timestamp to prevent replay attacks.

5.2. Policy File

The policy file MUST be a JSON document containing version, allowed intents, and cryptographic material.

6. Security Considerations

Implementations MUST use secure signature algorithms (Ed25519 recommended). Hosts MUST validate timestamps to prevent replay attacks (max skew +/- 60 seconds).

7. IANA Considerations

This document requests the registration of the HTTP headers listed in Section 5.1 in the "Permanent Message Header Field Names" registry.

8. References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, May 2017.

Author’s Address

Adnan Schulze-Hueneke
JamOne-DE
Schwuelper
Germany

Email: hallo@jamone.de
URI: https://ecap-protocol.com