Workgroup: Network Working Group
Internet-Draft: draft-colwell-privacy-txt-00
Published: 15 April 2024
Intended Status: Informational
Expires: 17 October 2024
Authors: N. Sullivan L. V. D. Peet G. Smaragdakis
TU Delft TU Delft
B. Colwell
BringYour, Inc.

# A File Format to Aid in Consumer Privacy Enforcement, Research, and Tools

## Abstract

This proposal outlines a new file format called privacy.txt. It follows similar placement on a web server as robots.txthttps://datatracker.ietf.org/doc/html/rfc9309, security.txthttps://datatracker.ietf.org/doc/html/rfc9116, or ads.txthttps://iabtechlab.com/ads-txt/, in the / directory or /.well-known directory.

The file format adds structured data for three areas: 1. A machine parsable and complete privacy policy 2. Consumer actions under their privacy rights 3. Cookie disclosures

## About This Document

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

The latest revision of this draft can be found at https://example.com/LATEST. Status information for this document may be found at https://datatracker.ietf.org/doc/draft-colwell-privacy-txt/.

Discussion of this document takes place on the WG Working Group mailing list (mailto:WG@example.com), which is archived at https://example.com/WG.

Source for this draft and an issue tracker can be found at https://github.com/USER/REPO.

## 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 17 October 2024.

## Copyright Notice



## Table of Contents



## 1. Introduction

Consumers in many parts of the world have extensive privacy rights under laws such as the GDPR and the CPRA. However, without some formalization of a service's privacy policy, it is difficult or often intractable for consumers to exercise those rights; enforcement to verify compliance with laws and develop effective monitoring; and researchers and technologists to develop tools to allow greater adoption and success of privacy practices.

Consumer data originally gets into the cloud by connections from consumer devices to web servers in the cloud. To be able to audit

and technically enforce privacy it must be possible to track the privacy policies applied to every byte of consumer data entering the cloud. However, currently the association between a web request and the privacy policy is tenuous, leading to the possibility of incorrect or unverifiable consumer data usage at the very source. This proposal fills that hole by associating structured privacy data with every web server. Just like HTTPS security can be technically enforced, this proposal makes it possible to technically enforce privacy by verifying that the structured privacy information exists and is in good standing before sending data to the server.

This proposal outlines a new file format called privacy.txt. It follows similar placement on a web server as robots.txthttps://datatracker.ietf.org/doc/html/rfc9309, security.txthttps://datatracker.ietf.org/doc/html/rfc9116, or ads.txthttps://iabtechlab.com/ads-txt/, in the / directory or /.well-known directory [1,2,3].

The file format adds structured data for three areas: 1. A machine parsable and complete privacy policy 2. Consumer actions under their privacy rights 3. Cookie disclosures

The file format is UTF8 text and lists Field:Value, one per line. Whitespace and lines that start with # are ignored.

### 1.1. 1. A machine parsable and complete privacy policy

It is currently difficult to associate a complete privacy policy text with a service for a number of reasons. First, even though it must be linked from the company webpage, there is not a canonical URL. Second, it is common for services to use client-side rendering, interactive elements, break out links for addendums, and server rules to prevent machine parsing/scraping.

This file format proposes two fields for the privacy policy. One or both can be used, depending on the policy format.

Entity: NAME,COUNTRY_CODE

The entity issuing the privacy policy. A name that contains a comma should escape the comma as \,. The country code should follow 2-letter ISO 3166-1.

Privacy-policy-text: URL

A complete privacy policy in a single UTF8 text file that can be downloaded by any user agent or machine tool. This must include all addendums in the text file. It must not include links. Information about contact and consumer actions are covered in this file format and do not need to be linked to in the policy text.

Privacy-policy: URL

If Privacy-policy-text is present, this can simply point to the existing privacy policy, in whatever form it currently exists. Otherwise, it must point to a machine parsable/scrapable static HTML file that contains the complete policy on a single page.

### 1.2. 2. Consumer actions under their privacy rights

This file format proposed fields to structure the consumer actions described in the privacy policy and commonly required by law. Currently it is difficult to get even an email that can service privacy requests from many top-100 site privacy policies. There is currently no law about how easy it should be to take privacy actions, similar to the US CAN-SPAM Acthttps://www.fcc.gov/general/can-spam, which led to an industry standard one-click link for marketing emails. The spirit of these fields is similar, to make it as easy as possible for a consumer to exercise their privacy rights.

Below a one-click URL refers to a URL that can process a request without requiring a customer password or login. The URL should take customer identification such as email and verify as necessary to complete the request.

Contact: mailto:EMAIL

An email contact for the privacy office must be given. This email must be able to handle consumer requests via email where there is not an applicable Action-* field for the request. Responses can ask for additional verification but should not require customer password or login. If Action-* fields are defined for all applicable consumer requests, this email does not need to handle any requests. This proposal imagines companies would build self-service one-click URLs for all consumer actions as the most scalable outcome.

Action-delete-account-and-data: mailto:EMAIL|URL

Email or one-click URL to process an account and data deletion request.

Action-delete-personal-data: mailto:EMAIL|URL

Email or one-click URL to process a personal data deletion request.

Action-opt-out-sharing:mailto: EMAIL|URL

Email or one-click URL to opt out of personal data sharing with third parties.

Action-shared-list:mailto: EMAIL|URL

Email or one-click URL to get a list of all third parties where personal data has been shared.

Action-opt-out-marketing: mailto:EMAIL|URL

Email or one-click URL to opt out of marketing.

### 1.3. 3. Cookie disclosures

Common privacy laws call for transparency in cookie storage. In order to audit and enforce transparency, this file format proposes fields that describe the cookies used by a web site, following a previously published formatprivacy.txt Implementation Guide Louise van der Peet November 2022. A web browser could technically enforce this declaration by refusing access to undeclared cookies.

Banner: CONSENT_PRESENT,CONSENT_PLATFORM

A boolean attribute whether a consent banner is present, and the consent management platform name, which can be set to non-specific-custom or any identifying name if it is a custom banner, or set to none detected when there is not banner.

Cookie: FIELD#1,FIELD#2,...FIELD#7

The field values are given as a complete septuple with each field defined by the following table, taken fromprivacy.txt Implementation Guide Louise van der Peet November 2022. From these fields, the most important cookie attributes related to privacy and compliance can be derived.

| Field | Name | Description |
|---|---|---|
| FIELD#1 | Cookie name | The name of the cookie. This identifies which cookie is set. The website uses this together with the value to identify the cookie. |
| FIELD#2 | Domain name of the cookie | The domain attribute of a cookie specifies which domain may receive the cookie. If this is the same as the host domain, that means it is a first party cookie. |
| FIELD#3 | Duration of the cookie | The duration attribute contains the storage limit of the cookie. This is in the form of the amount of seconds the cookies will remain on the user’s device before it is expired and deleted. |
| FIELD#4 | First or Third party cookie | This is a boolean attribute that indicates whether the cookie is a third party cookie. Thus means that the target do- main is different from the host domain. It is placed on the website by someone other than the owner and collects data for that third party. |

| Field | Name | Description |
|---|---|---|
| FIELD#5 | Optional cookie | This is a boolean attribute which indicates whether this is an options cookie or not. Optional cookies can be refused by the user, using the consent banner. When cookies are not optional they will always be placed on the user’s device when they access the website, with or without consent. |
| FIELD#6 | Httponly | This is a boolean attribute which indicates whether the httpOnly flag is set. This means that the cookie can only be transferred via HTTP, and therefor the cookie can only be accessed by the cur- rent server. This helps mitigate client- side scripts accessing the cookie data. |
| FIELD#7 | Secure status | This is a boolean attribute which indicates whether the secure flag is set on the cookie. The secure flag causes the browser to only send the cookie over encrypted channels, therefor securing the communication between the user’s de- vice and the server. |

Table 1

## 2. Conventions and Definitions

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. Security Considerations

Following this file format makes it easier for consumers to take privacy actions, similar to one-click unsubscribe. Removing the barrier to actions makes it easier to make mistakes. It would be reasonable to allow some grace period of undo in case of a security incident.

## 4. IANA Considerations

This document has no IANA actions.

## 5. Normative References

**[RFC2119]** Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/rfc/rfc2119>.

**[RFC8174]**
Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/rfc/rfc8174>.

## Acknowledgments

## Authors' Addresses

Nick Sullivan

Email: nicholas.sullivan+ietf@gmail.com

Louise Van der Peet
TU Delft

Email: L.VanderPeet@tudelft.nl

Georgios Smaragdakis
TU Delft

Email: g.smaragdakis@tudelft.nl

Brien Colwell
BringYour, Inc.

Email: brien@bringyour.com