Unified User-Agent String (UUAS)
draft-karcz-uuas-00

Abstract

User-Agent is a header used by certain protocols, e.g. HTTP. Unified User-Agent String is intended to unification of that complicated strings.

Status of this Memo

This memo defines an Experimental Protocol for the Internet community. This memo does not specify an Internet standard of any kind. Discussion and suggestions for improvement are requested. Distribution of this memo is unlimited.

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 http://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."
Copyright Notice

Copyright (c) 2012 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 (http://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.

This document may not be modified, and derivative works of it may not be created, except to format it for publication as an RFC or to translate it into languages other than English.
1. Introduction

So far User-Agent strings are long, complicated and often ambiguous. (e.g. "Mozilla/4.0 (compatible; MSIE 6.0; X11; Linux i686; en) Opera 8.01" - it is Opera Browser, but it can be read as Internet Explorer or Netscape Navigator.) This document specifies a new, easy and clear format of Unified User-Agent String (UUAS).

2. Structure

Unified User-Agent String is separated by space into two parts:

- Agent signature (REQUIRED) - overview Section 2.1
- Engine signature (OPTIONAL) - overview Section 2.2

General ABNF notation of UUAS is:

\[ \text{uuas} = \text{agent} \ast 1(\text{SP engine}) \]

2.1. Agent signature

In Unified User-Agent String signature contains two main parts separated by space:
Agent identifier, containing name and version of Agent - Section 2.1.1

Agent features, containing information about Agent platform. - Section 2.1.2

General ABNF notation of Agent signature is:

agent = agent_id SP agent_ft

2.1.1. Agent identifier

Agent identifier contains two parts separated by a slash sign:

- Agent name, e.g. "FooBrowser"
- Agent version, e.g. "1.0"

Characters allowed to use in Agent name are: small and capital letters and a minus sign. Characters allowed to use in Agent version are: digits, dot, a minus sign and small and capital letters.

2.1.2. Agent Features

Data contained in Agent features are (in turn):

1. Security level flag - Section 2.1.2.1
2. Operating system - Section 2.1.2.2
3. Hardware platform - Section 2.1.2.3
4. Client language - Section 2.1.2.4
5. OPTIONAL Contact - Section 2.1.2.5
6. OPTIONAL Addition data - Section 2.1.2.6

All of Agent features is contained in brackets. Data are separated by semicolon and space.

2.1.2.1. Security level flag

Security level flag indicates client.s security level. This field may contain these values:

- "N" - no encryption
2.1.2.2. Operating system

Operating system field indicates client's operating system, e.g. "Windows NT 6.1", "FreeBSD 8.7" or "Mac OS X 10.7".

Characters allowed to be used in this field are: small and capital letters, digits, dot, a minus sign and space.

2.1.2.3. Hardware platform

Hardware platform field indicates client's processor's Instruction Set Architecture, e.g. "i386", "ppc" or "m68k".

Character allowed to be used in this field are: small and capital letters, digits, dot and a minus sign. Maximum length of this field is 10 characters.

2.1.2.4. Client language

This field indicates client application's language. It may contain language tags defined in [RFC5646].

2.1.2.5. Contact

This field contains one HTTP URL and zero or more client application-related e-mail addresses.

URL and e-mail addresses are preceded by a plus sign. All subfields are separated by a semicolon and a space.

ABNF notation of this field is here:

```
contact = curl *(break cemail)
curl = plus <rfc1738-httpurl>
cemail = plus <rfc5322-address-spec>
```

2.1.2.6. Additional data

Additional data field can contain any data. All subfields are separated by semicolon and space.

Characters allowed to be used in this field are: characters x20-x24, x26-x27, x2A-x2E, x30-x3A, x3C-x7E and hexadecimal-defined characters
Hexadecimal-defined character ABNF notation is here:

```
hexchar = %x25 2HEXDIG
```

This is ABNF notation of Additional data (one):

```
addit_data = break 1*adchar
```

2.2. Engine signature

Engine signature indicates information about basis of the client application, e.g. HTML rendering engine. This signature contains two parts separated by space:

- Engine identifier - structure identical as in Agent identifier
- OPTIONAL Engine information - Section 2.2.1

ABNF notation of Engine signature is:

```
engine = engine_id *1(SP engine_if)
```

2.2.1. Engine information

This field contains contact information and/or other data about engine separated by semicolon and space.

This is ABNF notation of Engine information:

```
engine_if = obracket (1* adchar / contact) *addit_data cbracket
```
2.3. ABNF definition of UUAS

; Unified User-Agent String general definition
uuas = agent *1(SP engine)

; Agent signature definition
agent = agent_id SP agent_ft
agent_id = agent_nm slash agent_vr
agent_nm = 1*nmchar
agent_vr = 1*vrchar

agent_ft = obracket security_flg break os break hwplatform break
    language *1(break contact) *addit_data cbraket
security_flg = %x4E / %x49 / %x55
os = 1*(vrchar / SP)
hwplatform = 1*10vrchar
language = <rfc5646-Language-Tag>

; Non-required data definitions
contact = curl *(break cemail)
curl = plus <rfc1738-httpurl>
cemail = plus <rfc5322-addr-spec>
addit_data = break 1*adchar

; Engine signature definition
engine = engine_id *1(SP engine_if)
engine_id = agent_id
engine_if = obracket (1*adchar / contact) *addit_data cbraket

; Additional definitions
slash = %xF2
obracket = %x28
cbracket = %x29
plus = %x2B
break = %x3B SP
hexchar = %x25 2HEXDIG

nmchar = ALPHA / "-"
vrchar = nmchar / DIGIT / "."
adchar = %x20-24 / %x26-27 / %x2A-2E / %x30-3A / %x3C-7E / hexchar

3. Security Considerations

Security is truly believed to be irrelevant to this document.
4. References

4.1. Normative References


Appendix A. Acknowledgments

I would like to thank my English teacher, who devoted their time to conduct a linguistic revision of this Memo.

This document was prepared using 2-Word-v2.0.template.dot.

Authors. Addresses

Mateusz Karcz
Independent
18/6 Sambora Street
83-110 Tczew
Poland

Email: sami.wiec.e.kto@interia.eu