

World Wide Web

Jeszenszky Péter

Debreceni Egyetem, Informatikai Kar

jeszenszky.peter@inf.unideb.hu

Utolsó módosítás: 2022. szeptember 5.

Alapötlet

- A Web alapötlete eredetileg az alábbi sarokköveken alapul:
 - Erőforrások azonosítása globális azonosítókkal (URI-kkal)
 - Kliens-szerver modell
 - Hiperszöveg jelölőnyelv (HTML)

Web architektúra

- A Web architektúrája kortárs nézőpontból:
 - *Architecture of the World Wide Web, Volume One*
(W3C ajánlás, 2004. december 15.)
<https://www.w3.org/TR/webarch/>
 - A kliens-szerver modell egyáltalán nem is kerül említésre a szövegében!

Web architektúra: fogalmak (1)

- **Világháló, Web (*World Wide Web*):** információs tér, melynek (erőforrásoknak nevezett) lényeges elemeit URI-k azonosítják.
- **Erőforrás (*resource*):** bármi, ami azonosítható egy URI-val.
 - **Információ erőforrás:** azzal a tulajdonsággal rendelkező erőforrás, hogy minden lényeges jellemzője továbbítható egy üzenetben.
- **Egységes erőforrás-azonosító (URI):** a Weben használt globális azonosító.
- **Reprezentáció (*representation*):** erőforrás állapotról információkat kódoló adatok.

Web architektúra: fogalmak (2)

- **Tartalomegyeztetés (*content negotiation*):** egy erőforráshoz több reprezentáció kínálása és ezek közül a legmegfelelőbb kiválasztása, amikor egy reprezentációt kell szolgáltatni.
- **Hivatkozás-feloldás (*dereferencing*):** egy URI használata a hivatkozott erőforrás eléréséhez:
 - Az elérésnek számos formája van, köztük az erőforrás egy reprezentációjának letöltése, létrehozása, módosítása, vagy az erőforrás néhány vagy összes reprezentációjának törlése.

Web architektúra: fogalmak (3)

- **Web ágens (*web agent*):** a Weben egy személy, entitás vagy folyamat nevében cselekvő személy vagy szoftver.
 - Például egy keresőrobot (*web crawler*).
- **Felhasználói ágens (*user agent*):** a web ágensek egy fajtája, egy személy nevében cselekvő szoftver.
 - Például egy webböngésző.

A Web architektúráis alapjai

- **Azonosítás:**

- Az erőforrások azonosítása URI-knak nevezett globális azonosítókkal történik.

- **Kölcsönhatás:**

- A web ágensek szabványos protokollok segítségével kommunikálnak, melyek üzenetcsere alapján.
 - A web protokollok közé tartozik például a HTTP, HTTPS és WebDAV.
- Egy üzenet tartalmazhat adatokat valamint metaadatokat az erőforrásról, az üzenet adatokról és magáról az üzenetről.

- **Adatformátumok:**

- A protokoll korlátozza az átvihető reprezentáció adatok és metaadatok formátumát.
- A Web maga nem korlátozza a tartalomszolgáltatók által használható adatformátumokat.
 - Ahhoz, hogy egy adatformátum egyaránt használható legyen két fél között, a két félnek (egy ésszerű mértékig) meg kell állapodni annak szintaxisában és szemantikájában.

Web architektúra: példa szcenárió

URI

`http://weather.example.com/debrecen`

Azonosít

Erőforrás

Debrecen időjárása

Reprezentáció

Metaadatok:

Content-Type:
application/xhtml+xml; charset=utf-8

Adatok:

```
<!DOCTYPE html>
<html xmlns="//www.w3.org/1999/xhtml">
  <head>
    <title>10 Day Weather Forecast for
      Debrecen</title>
    ...
```

Ábrázol

Szabványok (1)

- Egy szabvány egy olyan dokumentum, mely követelményeket, előírásokat, irányelveket vagy jellemzőket fogalmaznak meg, melyek következetesen használhatók annak biztosítására, hogy anyagok, termékek, folyamatok és szolgáltatások megfeleljenek a rendeltetésüknek.
 - Lásd:
<https://web.archive.org/web/20200101101550/https://www.iso.org/standards.html>

Szabványok (2)

- Eredetünk szerint a szabványoknak három fajtája van:
 - **De facto szabványok:** a gyakori használatból vagy a piaci elfogadottságból származnak.
 - Példa: QWERTY billentyűzetkiosztás, TeX, PDF (2008 előtt).
 - **De jure szabványok:** helyi, állami és/vagy nemzetközi szintű szabályozók által kötelezőként előírt szabványok.
 - Példa: Nemzetközi Mértékegységrendszer (SI), PDF (2008-tól).
 - **Önkéntes közmegegyezéssel szabványok:** különböző magánintézmények (például szakmai egyesületek és szervezetek, akkreditált szabványügyi szervezetek és ipari konzorciumok) által meghatározott szabványok.
 - Példa: az Internet protokollkészletet (közismert nevén TCP/IP), HTML, CSS.
- Lásd:
 - Andrew L. Russell. *Open Standards and the Digital Age*. Cambridge University Press, 2014. <https://arussell.org/open/>

Nyílt szabvány

- Nincs egyetlen általánosan elfogadott meghatározás:
 - *OpenStand: The Modern Paradigm for Standards* (IEEE, ISOC, IETF, IAB, W3C, ...) <https://open-stand.org/>
 - *Open Standards Requirement for Software* (Open Source Initiative) <https://opensource.org/osr/>
 - ...
- További információ: *Open standard* https://en.wikipedia.org/wiki/Open_standard

Web szabványok

- Az alábbi szervezetek felelnek web szabványokért:
 - *Ecma International* <https://www.ecma-international.org/>
 - *International Organization for Standardization (ISO)* <https://www.iso.org/>
 - *Internet Engineering Task Force (IETF)* <https://www.ietf.org/>
 - *Unicode Consortium* <https://unicode.org/consortium/consort.html>
 - *Web Hypertext Application Technology Working Group (WHATWG)* <https://whatwg.org/>
 - *World Wide Web Consortium (W3C)* <https://www.w3.org/>
 - ...

Internet Assigned Numbers Authority (IANA)

- A név jelentése kb. „Internet számkiosztó hatóság”.
- Az Internet működésének alapjául szolgáló kódok és számok kiosztását koordinálja. <https://www.iana.org/>
 - A DNS-gyökérzóna felügyelete, az .int és .arpa felső szintű tartományok üzemeltetése.
 - Az IP-címek kiosztásának globális koordinálása.
 - Nyilvántartja a különféle Internet protokollokhoz használt kódokat és számokat.
 - Lásd: *Protocol Registries* <https://www.iana.org/protocols>
- Az IANA egy funkció, melyet jelenleg egy nonprofit cég, az *Internet Corporation for Assigned Names and Numbers* (ICANN) működtet.

Internet Engineering Task Force (IETF)

- A név jelentése kb. Internet mérnöki munkacsoport.
- Internet szabványokat fejlesztő nemzetközi szabványügyi szervezet.
 - Az IETF fejleszti például az Internet protokollkészletet (közismert nevén TCP/IP).
 - Nincs formális tagság, sem tagsági díj, bárki számára nyílt a részvétel lehetősége.
 - Levelezési listák: <https://www.ietf.org/list/>
 - A szakmai munka munkacsoportokban történik.
- Alapítás: 1986. január 16.
 - Lásd: *IETF Turns 25 on 16 January 2011*
<https://www.ietf.org/mail-archive/web/ietf-announce/current/msg08366.html>
- Az Internet szabványokhoz kötődő specifikációkat az RFC dokumentum-sorozatban publikálja.

Request for Comments (RFC) (1)

- Az RFC sorozat az Internetről szóló műszaki és szervezeti dokumentumokat tartalmaz.
 - Jelentés: „kéretik megkritizálni”.
- Az RFC dokumentumsorozat 1969-ben indult az ARPANET projekt részeként.
 - Az első RFC:
 - Steve Crocker, *Host Software*, RFC 1, 7 April 1969.
<https://www.rfc-editor.org/rfc/rfc1>

Request for Comments (RFC) (2)

- Az RFC Editor szerkeszti, adja ki és katalogizálja az RFC-ket.
<https://www.rfc-editor.org/>
- Eredet szerint az RFC sorozatot négy folyamra osztják:
 - Az Internet Engineering Task Force (IETF) folyam
 - Az Internet Architecture Board (IAB) folyam
 - Az Internet Research Task Force (IRTF) folyam
 - A független beadványok folyama
- További információk az RFC sorozatról:
 - Russ Housley (ed.), Leslie L. Daigle (ed.), *The RFC Series and RFC Editor*, RFC 8729, February 2020.
<https://www.rfc-editor.org/rfc/rfc8729>

Request for Comments (RFC) (3)

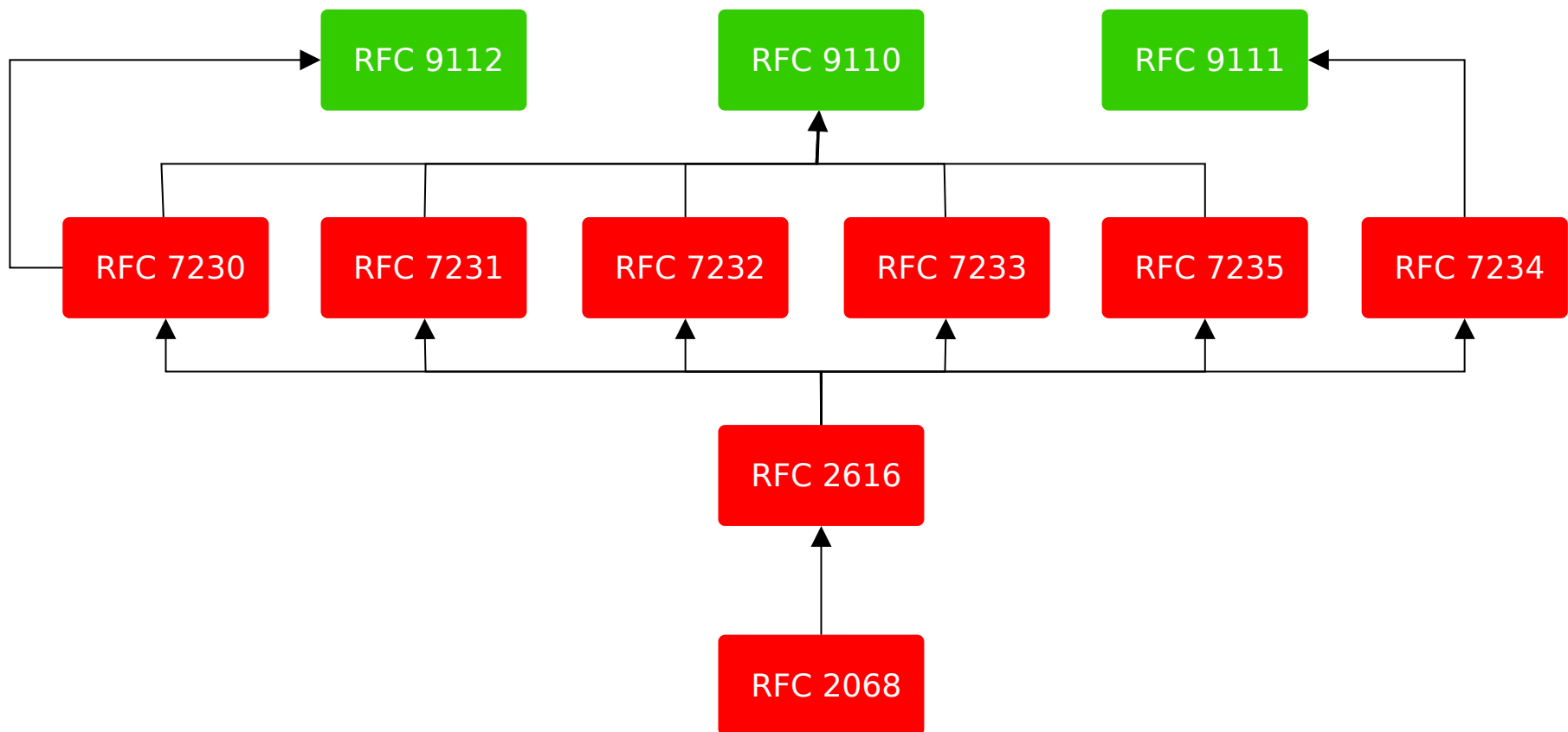
- Minden RFC-t egy szám azonosít, mint például RFC 9110.
- Minden RFC elérhető ASCII szöveggként, mint például: <https://www.rfc-editor.org/rfc/rfc9110.txt>
 - Ugyanez az RFC HTML-ben:
<https://www.rfc-editor.org/rfc/rfc9110.html>
- Az összes RFC listája:
<https://www.rfc-editor.org/rfc-index.html>

Request for Comments (RFC) (4)

- A kiadott RFC-k soha nem módosulnak.
- A különféle hibákat hibajegyzékek javítják.
- Változtatások egy javított RFC írásával és kiadásával is eszközölhetők.
 - Egy RFC elavulttá tehet (*obsolete*) vagy frissíthet (*update*) korábbi RFC-ket.

Request for Comments (RFC) (5)

- Példa: Hypertext Transfer Protocol – HTTP/1.1



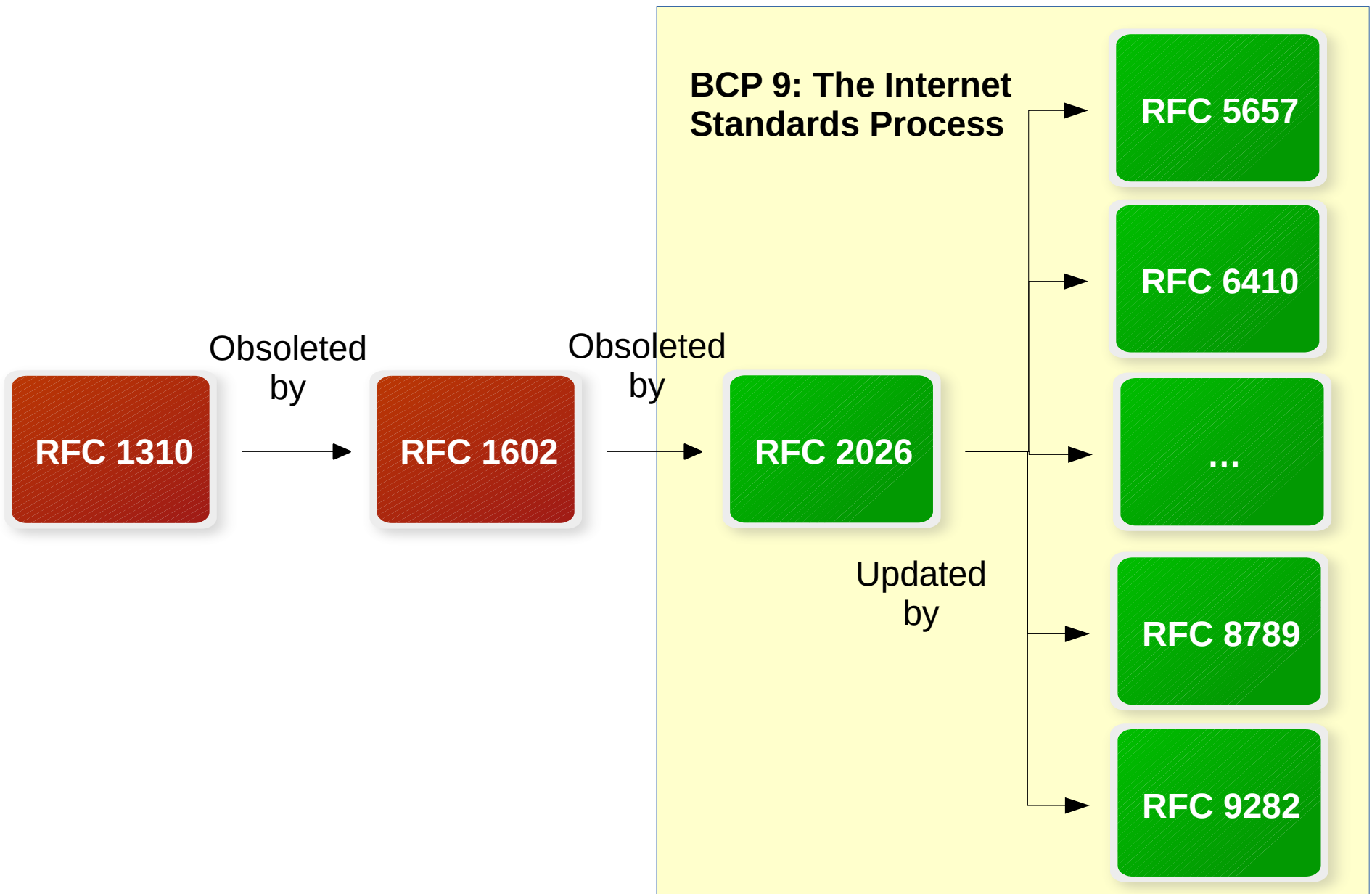
Request for Comments (RFC) (6)

- Az IETF RFC-k sorozata az alábbi két fontos alsorozatot tartalmazza:
 - ***Best Current Practice (BCP)***:
 - Jelentés: „jelenlegi legjobb gyakorlat”.
 - A BCP-k irányelveket és folyamatokat dokumentálnak, vagy az IETF működését írják le.
 - *BCP Index*: <https://www.rfc-editor.org/rfc/bcp/>
 - ***Internet Standard (STD)***:
 - *STD Index* : <https://www.rfc-editor.org/rfc/std/>

Request for Comments (RFC) (7)

- A BCP-k és STD-k egy számot kapnak az alsorozatukban, bár megtartják az RFC számukat is.
 - Példa:
 - Scott O. Bradner, *The Internet Standards Process – Revision 3*, **BCP 9**, RFC 2026, October 1996. <https://www.rfc-editor.org/rfc/rfc2026>
 - Tim Berners-Lee, Roy T. Fielding, Larry Masinter, *Uniform Resource Identifier (URI): Generic Syntax*, **STD 66**, RFC 3986, January 2005. <https://www.rfc-editor.org/rfc/rfc3986>
- Több RFC-hez is tartozhat ugyanaz a BCP vagy STD szám.
 - Például egy STD szám egy szabványt azonosít, nem pedig egy dokumentumot.

Request for Comments (RFC) (8)



Request for Comments (RFC) (9)

- **Standards Track:** Az Internet szabványnak szánt RFC-k érettségi szintjeinek összefoglaló neve.
 - Eredetileg 3 érettségi szintet (*maturity level*) használtak:
 - *Proposed Standard* (javasolt szabvány)
 - *Draft Standard* (szabványtervezet)
 - *Internet Standard* (Internet szabvány)
 - Jelenleg a *Proposed Standard* és az *Internet Standard* érettségi szinteket használják.
- Lásd:
 - Scott O. Bradner, *The Internet Standards Process – Revision 3*, BCP 9, RFC 2026, October 1996. <https://www.rfc-editor.org/rfc/rfc2026>
 - Russell Housley, Dave Crocker, Eric W. Burger, *Reducing the Standards Track to Two Maturity Levels*, BCP 9, RFC 6410, October 2011. <https://www.rfc-editor.org/rfc/rfc6410>

Request for Comments (RFC) (10)

- **Internet-Draft:** egy specifikáció a fejlesztése közben átnézésre és észrevételezésre elérhetővé tett előzetes verziója.
 - Végül vagy RFC-ként kerülnek publikálásra, vagy nem.
 - Bármikor változhatnak vagy törlésre kerülhetnek.
 - Legfeljebb 6 hónapig érvényesek.
 - Nem ajánlott hivatkozni vagy idézni őket, kivéve mint folyamatban lévő munkát.
 - Például:
 - Austin Wright (ed.), Henry Andrews (ed.), Ben Hutton (ed.), Greg Dennis, *JSON Schema: A Media Type for Describing JSON Documents*, 10 June 2022. <https://datatracker.ietf.org/doc/id/draft-bhutton-json-schema-01.html>

Request for Comments (RFC) (11)

- 1989 óta csaknem minden április 1-jén egy vagy több vicces RFC került kiadásra.
 - Példa:
 - Jogi Hofmueller (ed.), Aaron Bachmann (ed.), IOhannes Zmoelnig (ed.), *The Transmission of IP Datagrams over the Semaphore Flag Signaling System (SFSS)*, RFC 4824, April 1 2007. <https://www.rfc-editor.org/info/rfc4824>
- Lásd:
 - *April Fools' Day Request for Comments*
https://en.wikipedia.org/wiki/April_Fools%27_Day_Request_for_Comments

World Wide Web Consortium (W3C)

- A W3C egy nemzetközi közösség, ahol tagszervezetek, főállású alkalmazottak és a nyilvánosság munkálkodnak együtt webszabványok fejlesztésén. [...] A W3C küldetése a web lehetőségeinek maximális kiaknázása.
 - Lásd: *About W3C* <https://www.w3.org/Consortium/>
- A W3C ajánlásoknak nevezett, webtechnológiákat meghatározó és webszabványoknak számító dokumentumokat publikál.
 - A szabványok fejlesztése a Nyílt Szabványok Alapelveinek megfelelően történik.
 - Lásd: *W3C Mission* <https://www.w3.org/Consortium/mission>
 - Lásd: *Standards FAQ – What does Web standard mean? What is a Recommendation?* <https://www.w3.org/standards/faq#std>

A W3C tervezési alapelvei

- **Web mindenkinek:** a web elérhető kell, hogy legyen mindenki számára, a hardverektől, szoftverektől, hálózati infrastruktúrától, anyanyelvtől, kultúrától, földrajzi elhelyezkedéstől, vagy a fizikai vagy szellemi képességektől függetlenül.
 - Kapcsolódó fogalmak: **akadálymentesítés** (*web accessibility*), **nemzetköziesítés** (*internationalization*)
- **Web mindenhol:** a web elérhető kell, hogy legyen a legkülönbözőbb eszközökről.
 - Például mobiltelefonok, okostelefonok, PDA eszközök, háztartási készülékek, ...
- Lásd: <https://www.w3.org/Consortium/mission.html#principles>

A W3C története

- A MIT-n alapították 1994 októberében.
- Igazgatója Tim Berners-Lee, a web feltalálója és megalkotója.
<https://www.w3.org/People/Berners-Lee/>
- 1996 óta 300-nál több ajánlást jelentetett meg.
 - Lásd: <https://www.w3.org/TR/?status=REC>

W3C: néhány mérföldkő (1)

- **1996 október:** *PNG (Portable Network Graphics) Specification Version 1.0*
<https://www.w3.org/TR/REC-png-961001>
- **1996 december:** *Cascading Style Sheets, level 1*
<https://www.w3.org/TR/REC-CSS1-961217>
- **1998 február:** *Extensible Markup Language (XML) 1.0*
<https://www.w3.org/TR/1998/REC-xml-19980210>
- **1998 április:** *Mathematical Markup Language (MathML) 1.0 Specification*
<https://www.w3.org/TR/1998/REC-MathML-19980407/>

W3C: néhány mérföldkő (2)

- **1998 október:** *Document Object Model (DOM) Level 1 Specification* <https://www.w3.org/TR/REC-DOM-Level-1/>
- **1999 november:** *XSL Transformations (XSLT) Version 1.0* <https://www.w3.org/TR/1999/REC-xslt-19991116>
- **1999 december:** *HTML 4.01 Specification* <https://www.w3.org/TR/html401/>
- **2000 január:** *XHTML 1.0: The Extensible HyperText Markup Language* <https://www.w3.org/TR/2000/REC-xhtml1-20000126/>
- **2001 május:** *XHTML 1.1 – Module-based XHTML* <https://www.w3.org/TR/2001/REC-xhtml11-20010531/>

W3C: néhány mérföldkő (3)

- **2004 október:** *XML Schema* <https://www.w3.org/TR/xmlschema-0/>
<https://www.w3.org/TR/xmlschema-1/> <https://www.w3.org/TR/xmlschema-2/>
- **2011 június:** *Cascading Style Sheets Level 2 Revision 1 (CSS 2.1) Specification*
<https://www.w3.org/TR/CSS2/>
- **2012 június:** *Media Queries*
<https://www.w3.org/TR/2012/REC-css3-mediaqueries-20120619/>
- **2011 szeptember:** *Selectors Level 3*
<https://www.w3.org/TR/2011/REC-css3-selectors-20110929/>
- **2014 október:** *HTML5 – A vocabulary and associated APIs for HTML and XHTML*
<https://www.w3.org/TR/2014/REC-html5-20141028/>
- **2019 december:** *WebAssembly Core Specification* <https://www.w3.org/TR/wasm-core-1/>
- **2020 április:** *Web of Things (WoT) Architecture* <https://www.w3.org/TR/wot-architecture/>
- **2021 január:** *WebRTC 1.0: Real-Time Communication Between Browsers*
<https://www.w3.org/TR/webrtc/>
- **2022 április:** *Media Queries Level 3* <https://www.w3.org/TR/mediaqueries-3/>

A W3C működése (1)

- Jelenleg 472 tagszervezete van a világ minden részéről (2022. szeptember 5.)
 - A tagok listája: <https://www.w3.org/Consortium/Member/List>
 - Adobe, Amazon, Apple, CERN, Google, IBM, Intel, Meta, Microsoft, SZTAKI, ...
- A munkát az alábbi anyaintézmények koordinálják:
 - MIT/CSAIL (USA), ERCIM (Franciaország), Keio University (Japán), Beihang University (Kína)
- A W3C tevékenységéhez kapcsolódó területileg szerveződő közösségek: *W3C Chapters* <https://chapters.w3.org/>
 - *Hungary Chapter: W3C Magyar Iroda* <https://chapters.w3.org/hungary/>

A W3C működése (2)

- A fejlesztést munkacsoportok végzik.
- A munkacsoportok által előállított eredmények többek között szakmai jelentések, tesztkészletek és nyílt forrású szoftverek.
- A munkacsoportok a szóban forgó terület szakértőiből állnak, akik az alábbiak lehetnek:
 - W3C alkalmazott,
 - egy tagszervezetet képviselő egyén (általában egy tagszervezet alkalmazottja),
 - egy meghívott szakértőként közreműködő egyén.

A W3C működése (3)

- Jelenleg 44 munkacsoportja van (2022. szeptember 5.). <https://www.w3.org/groups/wg/>
 - *Cascading Style Sheets (CSS) Working Group*
<https://www.w3.org/Style/CSS/members>
 - *HTML Working Group* <https://www.w3.org/groups/wg/htmlwg>
 - *Web Applications Working Group*
<https://www.w3.org/groups/wg/webapps>
 - *Web Machine Learning Working Group*
<https://www.w3.org/groups/wg/webmachinelearning>
 - ...

W3C részvétel (4)

- Nyitott a részvétel lehetősége a nyilvánosság számára:
 - Csatlakozás levelezési listákhoz <https://www.w3.org/Mail/>
 - Csatlakozás közösségi és üzleti csoportokhoz <https://www.w3.org/community/>
 - Közreműködés nyílt forrású szoftverek fejlesztésében <https://www.w3.org/Status>
 - Specifikációk és egyéb erőforrások fordítása <https://www.w3.org/Consortium/Translation/>
 - Részvétel a W3C által rendezett eseményeken (például konferenciákon, *workshop*okon) <https://www.w3.org/participate/eventscal>
- Lásd: <https://www.w3.org/participate/>

W3C szakmai jelentések

- A W3C által publikált különféle dokumentumokról lásd:
 - *W3C Process Document* (2021. november 2.)
<https://www.w3.org/Consortium/Process/>
 - *Documents published at W3C* <https://www.w3.org/standards/types>
- Az összes szakmai jelentés: <https://www.w3.org/TR/>
- A W3C dokumentumok felhasználását a *W3C Document License* szabályozza.
 - Lásd: *W3C Document License*
<https://www.w3.org/Consortium/Legal/copyright-documents>
 - További információk: *Intellectual Rights FAQ*
<https://www.w3.org/Consortium/Legal/IPR-FAQ-20000620.html>

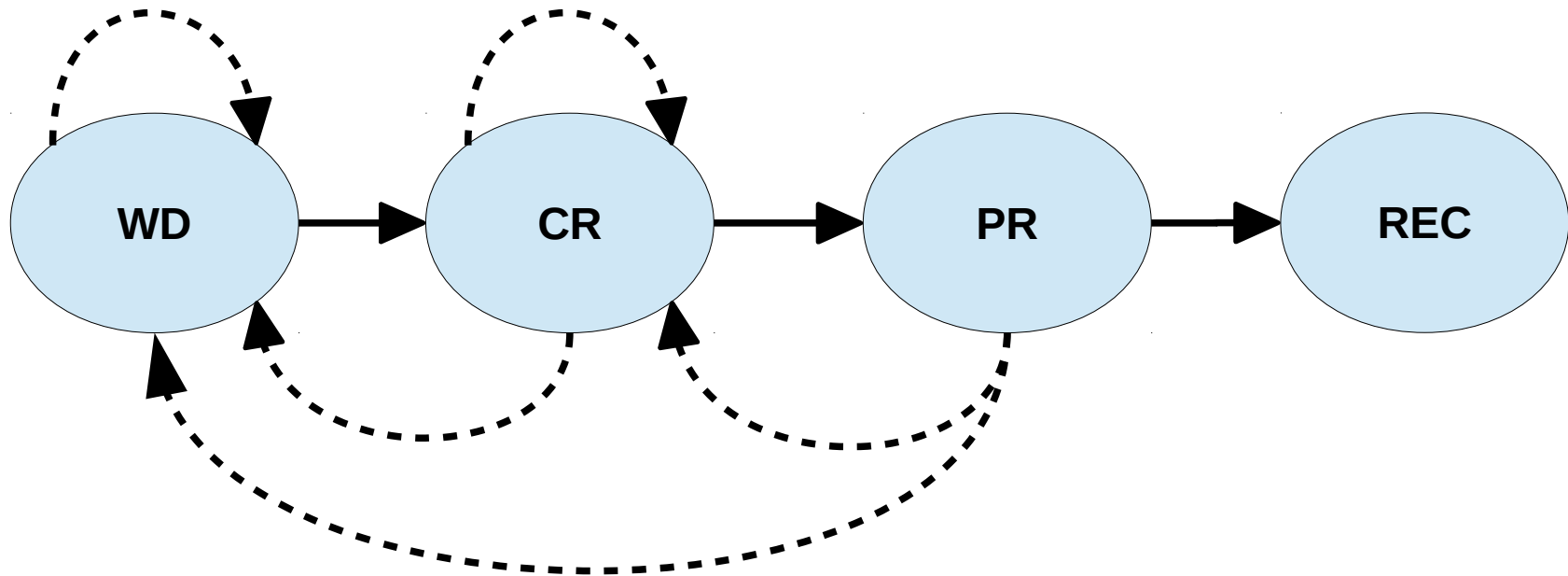
W3C szakmai jelentések érettségi szintjei (1)

- **Munkaterv (*Working Draft* – **WD**)**: a közösség általi áttekintésre közzétett dokumentum, beleértve a W3C tagokat, a nyilvánosságot és más műszaki szervezeteket.
 - Néhány, de nem minden munkaterv célja, hogy ajánlássá lépjen elő.
- **Előzetes javaslattev (Candidate Recommendation – **CR**)**: már széles körben áttekintett dokumentum, melyet implementációs tapasztalatok szerzése céljából tesznek közzé.
- **Javaslattev (Proposed Recommendation – **PR**)**: egy olyan dokumentum, mely megfelelő minőségű ahhoz, hogy ajánlássá váljon.
- **Ajánlás (Recommendation – **REC**)**: széles körben alkalmazható webszabvány.
- **Csoport feljegyzés (Group Note – **NOTE**)**: olyan dokumentum, melyet nem szánnak hivatalos szabványnak.
 - A műszaki specifikációktól eltérő információk dokumentálására teszik közzé őket, mint például egy specifikációt motiváló felhasználói esetek és a használatra vonatkozó legjobb gyakorlatok.

W3C szakmai jelentések érettségi szintjei (2)

- Egy ajánlás túlhaladottá vagy elavulttá válhat:
 - **Túlhaladott ajánlás (*Superseded Recommendation*)**: olyan specifikáció, melyet egy újabb verzió helyettesít.
 - Példa:
 - *XHTML 1.0 The Extensible HyperText Markup Language (Second Edition)*
<https://www.w3.org/TR/xhtml1/>
 - **Elavult ajánlás (*Obsolete Recommendation*)**: egy olyan specifikáció, melyről a W3C úgy döntött, hogy nincs elegendő piaci jelentősége ahhoz, hogy a továbbiakban is implementálásra ajánlja.
 - Példa:
 - *The 'view-mode' Media Feature* <https://www.w3.org/TR/view-mode/>

W3C szakmai jelentések érettségi szintjei (3)



WHATWG (1)

- *Web Hypertext Application Technology Working Group* (WHATWG) <https://whatwg.org/>
 - A Web fejlődése iránt elkötelezett közösség, mely böngészőkben implementálható szabványokat fejleszt.
 - Kiejtés: *what-wee-gee*, *what-wig*, *what-double-you-gee*
 - Lásd: *How do you spell and pronounce WHATWG?*
<https://whatwg.org/faq#spell-and-pronounce>
 - További információk: *WHATWG – FAQ*
<https://whatwg.org/faq>

WHATWG (2)

- Szabványok:
 - DOM <https://dom.spec.whatwg.org/>
 - Fullscreen API <https://fullscreen.spec.whatwg.org/>
 - HTML <https://html.spec.whatwg.org/>
 - URL <https://url.spec.whatwg.org/>
 - XMLHttpRequest <https://xhr.spec.whatwg.org/>
 - ...
- Lásd: *WHATWG – Standards*
<https://spec.whatwg.org/>

WHATWG (3)

- Történet:
 - 2004-ben alapították az *Apple*, a *Mozilla Foundation* és az *Opera Software* programozói, akik elégedetlenek voltak a W3C a HTML fejlesztésére irányuló tevékenységével.
- Működés:
 - Működését az irányítócsoporthoz (*Steering Group*) koordinálja, melynek tagjai az Apple, Google, Microsoft és a Mozilla.
- Részvétel:
 - Nyitott a részvétel lehetősége a nyilvánosság számára.
 - Lásd: *WHATWG – Participation* <https://participate.whatwg.org/>

WHATWG (4)

- Fejlesztési modell:
 - A WHATWG „élő szabványoknak” nevezett specifikációkat fejleszt, melyek folyamatosan frissülnek.
- Az élő szabványok felhasználását a CC BY 4.0 licenc szabályozza.
<https://creativecommons.org/licenses/by/4.0/>
 - Lásd: *WHATWG – Intellectual Property Rights Policy* <https://whatwg.org/ipr-policy>