

PUBLISHING@W3C

Ivan Herman
W3C/CWI, ivan@w3.org

[Workshop on E-Publishing at W3C](#)

15 September, 2017
Frankfurt, Germany

About W3C: “Leading the Web to its Full Potential”



Tim Berners-Lee, Web Inventor and Director of W3C

- ▶ Founded 1994
- ▶ \approx 480 Members
- ▶ \approx 65 staff in US (MIT), China (Beihang), Europe (ERCIM), and Japan (Keio)
- ▶ Focused on Web ecosystem: users, developers, browsers, etc.
- ▶ Developing new technologies for the Open Web Platform
- ▶ W3C focuses both on the Open Web, as well as specific industry requirements brought by industry segments

Role of W3C

Bring diverse ecosystem players together to develop Web technologies



PUBLISHING@W3C

Lead Publishing to its Full Potential with the Web

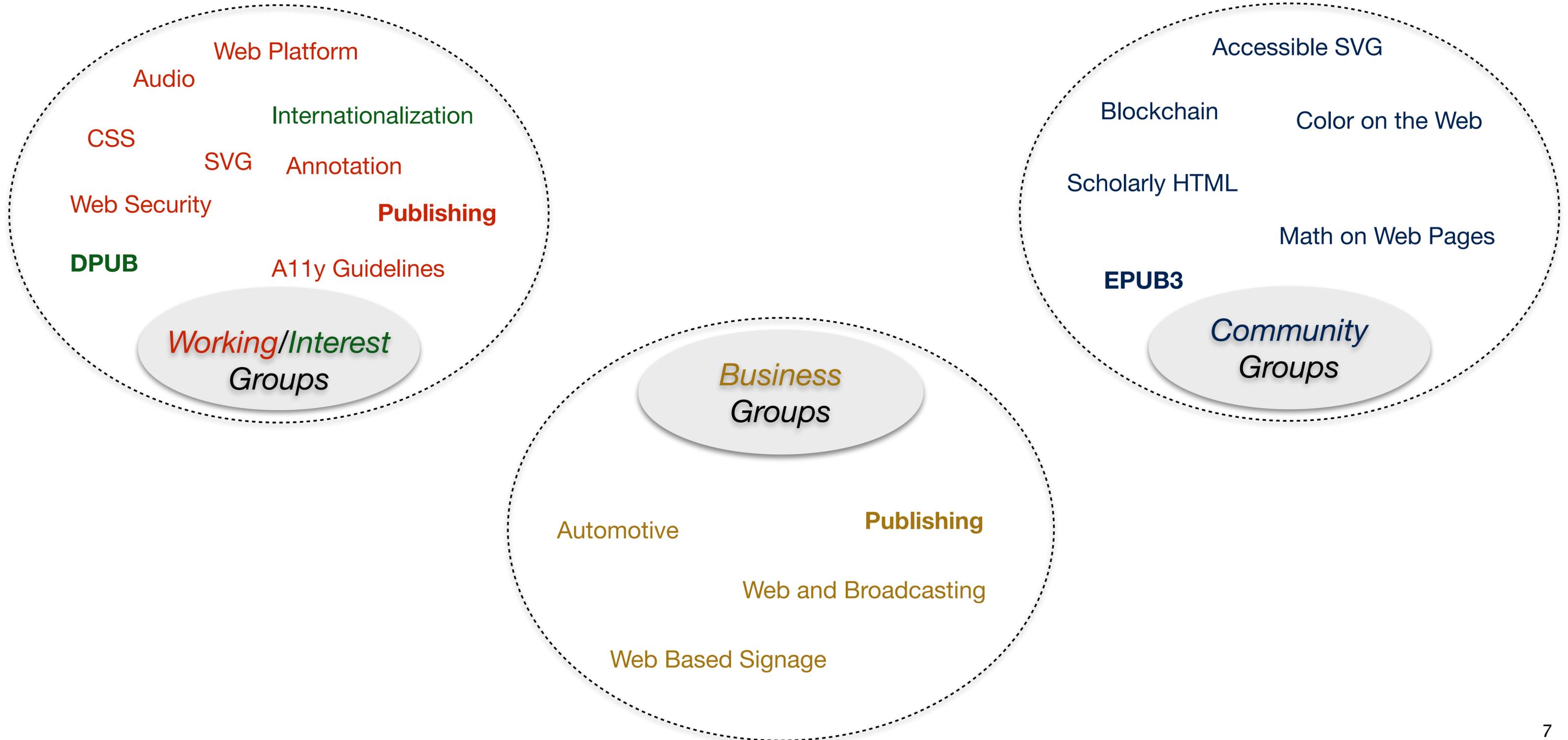
Publishing@W3C: the WWW

- ▶ Four aspects that we will touch upon:
 - ▶ **Where**, i.e., what are the groups within W3C that make up Publishing@W3C?
 - ▶ **What**, i.e., what will the initiative work on?
 - ▶ **When**, i.e., what are the timings of the activities?
 - ▶ **Who**, i.e., who can participate and how?

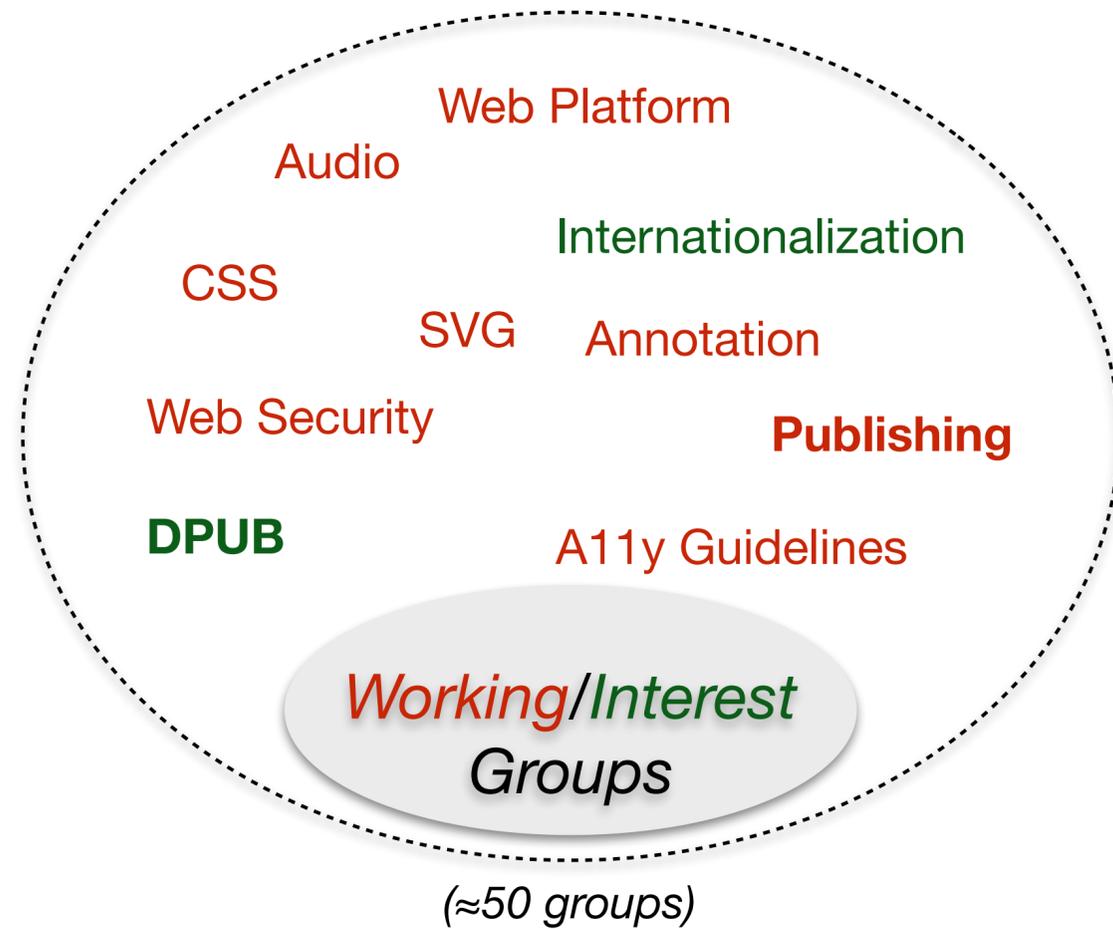


Where?

Crash Course on W3C groups...



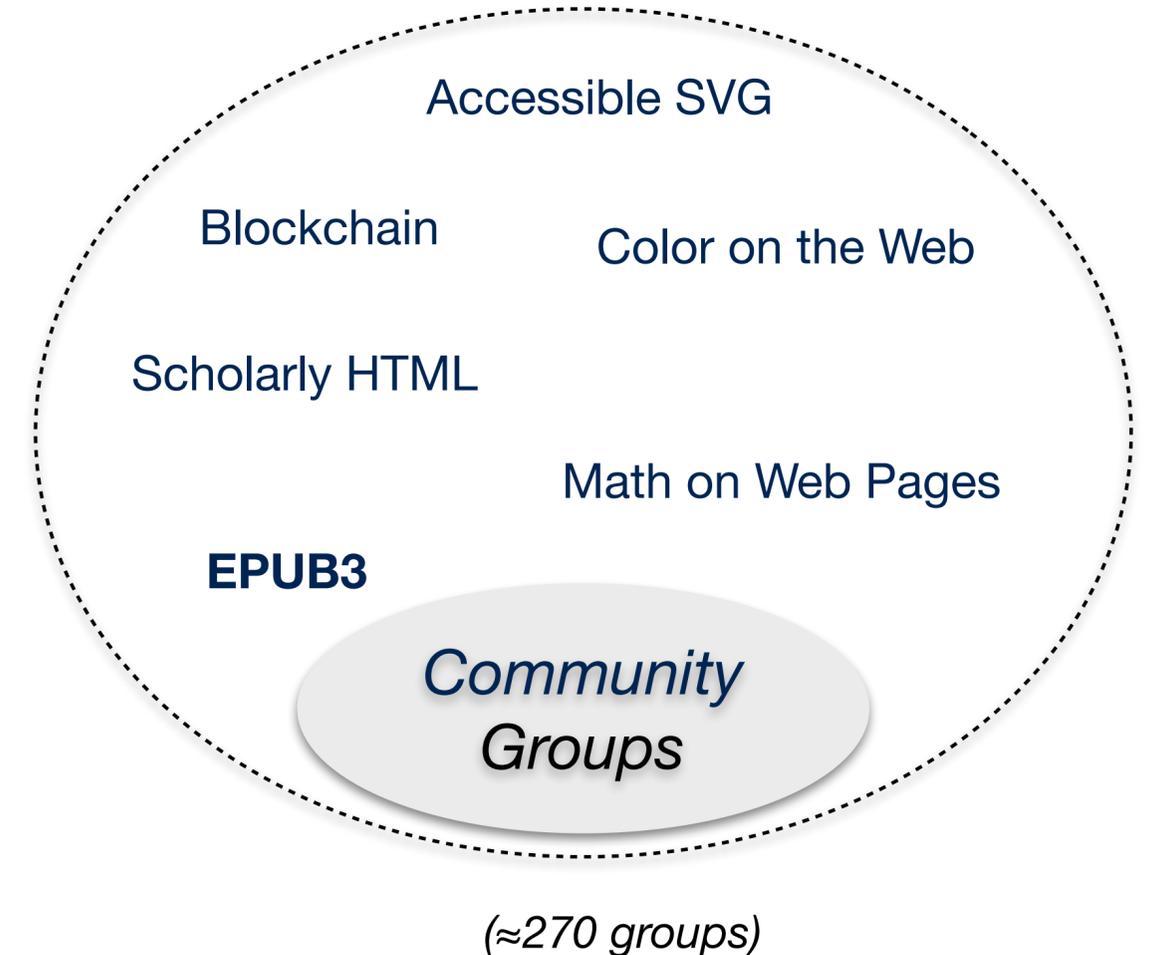
Working and Interest Groups



- ▶ Strict charters
- ▶ Voted on by the W3C membership
- ▶ Populated by W3C members
- ▶ W3C staff participate
- ▶ ***Only Working Groups can publish formal W3C Recommendations***
- ▶ ***There is an IPR protection on the results***

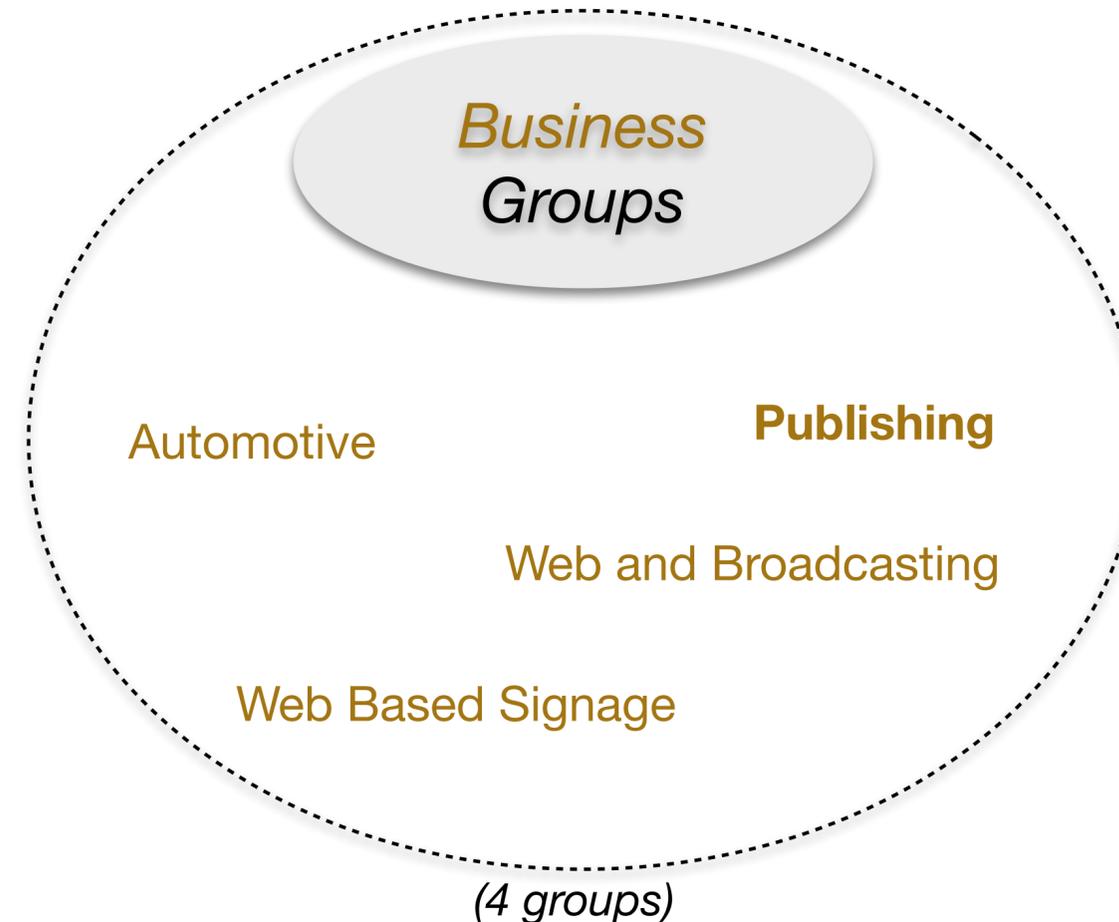
Community Groups

- ▶ Loose charters
- ▶ Anybody can start a CG, provided at least 5 other persons support it
- ▶ Anybody can join it for free
- ▶ Typically no (official) W3C staff
- ▶ May publish specifications but ***these are not formal W3C Recommendations***

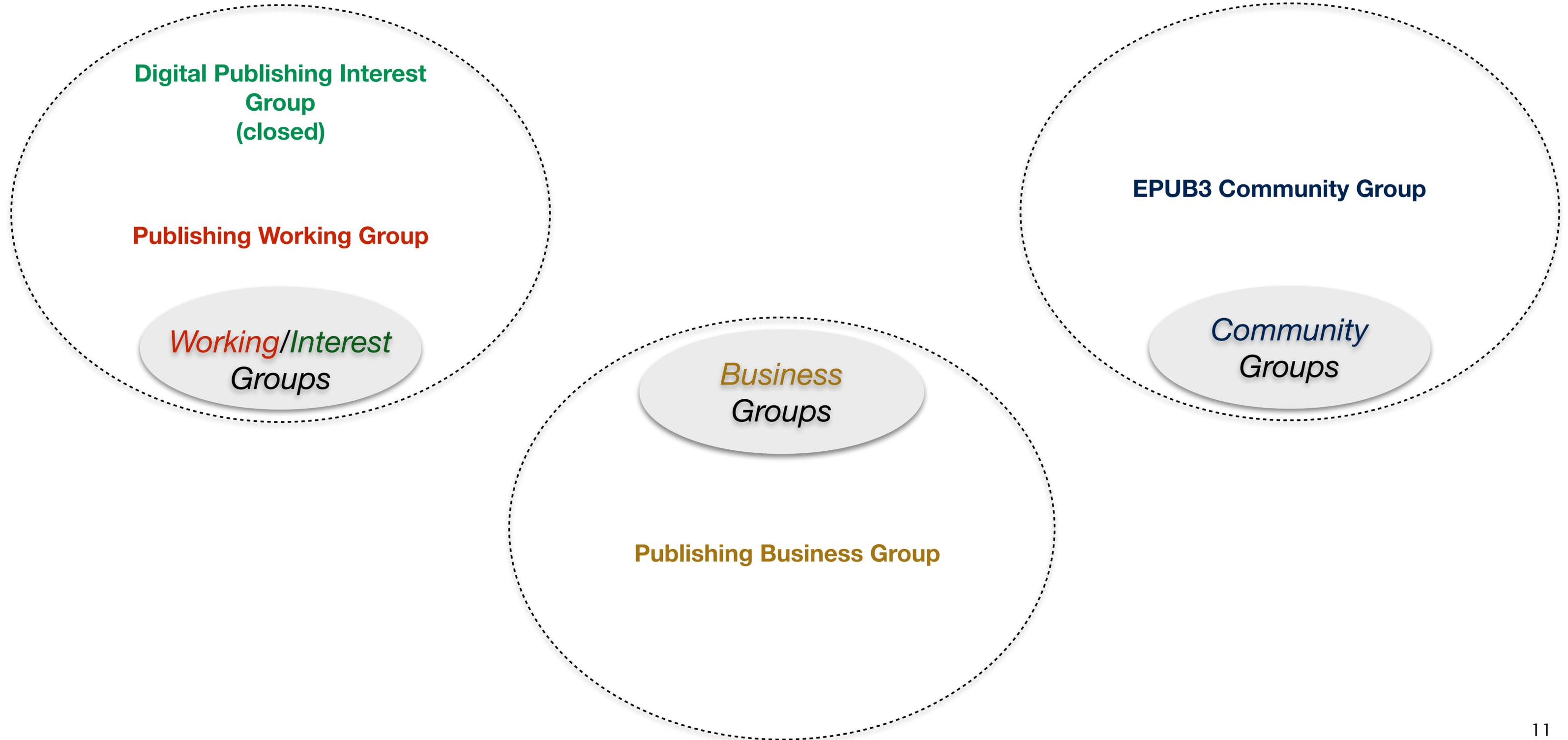


Business groups

- ▶ Much like Community Groups but concentrate on a specific business area
- ▶ Populated by either W3C members or “Business Group” members (who join only this group but for a smaller fee)
- ▶ W3C staff participate
- ▶ May publish specifications but ***these are not formal W3C Recommendations***



“Where”: Publishing@W3C Groups



“Where”: Publishing@W3C groups



“Where”: there is actually more

- ▶ The groups above are the strictly speaking “publishing” groups
- ▶ However, there are number of issues that
 - ▶ are very relevant to publishing, *but not only*
 - ▶ whose technical solution is to be worked on by other Working Groups, e.g.:
 - ▶ pagination and general page control — CSS WG
 - ▶ accessibility issues — ARIA WG and Accessibility Guidelines WG
 - ▶ offline access — Web Platform WG and Service Workers WG
- ▶ All Publishing@W3C groups will have active relationships with such W3C groups



What?

Publishing Business Group



Publishing Business Group

- ▶ *Provides strategic directions for Publishing@W3C:*
 - ▶ concentrates on the business aspects of publishing
 - ▶ serves as a conduit for feedback between the publishing ecosystem and W3C
 - ▶ maintains contacts with other organizations (EDItEUR, BISG, ISO, etc)
- ▶ Oversight and approval of the EPUB 3 Community Group's work
- ▶ Community outreach in general (conferences and symposia, possible online courses, etc...)

Some (possible) BG discussion topics

- ▶ Future evolution of epubtest and epubcheck tools
- ▶ Oversight of the evolution of “EPUB for Education” (a.k.a. EDUPUB)
- ▶ Rendering preferences in EPUB, page transition requirements
- ▶ Issues around possible ISO Standard version of EPUB 3.X and/or EPUB Accessibility guidelines
- ▶ STEM publishing issues
- ▶ etc.

W3C Publishing Summit 2017

- ▶ Some speakers:
 - ▶ Abhor Parasnis, Adobe
 - ▶ Ben Dugas, Kobo
 - ▶ Tzviya Siegman, Wiley
 - ▶ Rick Johnson, VitalSource
 - ▶ Jake Archibald, Google
 - ▶ Marcos Caceres, Mozilla
 - ▶ Liisa McCloy-Kelley, Pinguin
 - ▶ Ken Brooks, Macmillan
 - ▶ Dave Cramer, Hachette
 - ▶ James English, NYPL
 - ▶ Tim O'Reilly, O'Reilly Media
 - ▶ Micah Bowers (Bluefire)
 - ▶ and others...



The screenshot shows a web browser window with the URL <https://www.w3.org/publishing/events/summit2017.html>. The page features a dark navigation bar with the text "PUBLISHING@W3C" and menu items: "HOME", "SPEAKERS", "PROGRAM", "LOGISTICS", and "SPONSORS". The main content area has a green background with the text "W3C PUBLISHING SUMMIT 2017" in large, bold, blue letters. Below this is the main title "How Web Technologies are Shaping Publishing Today, Tomorrow, and Beyond" in bold black text. A dark blue "REGISTER" button is centered below the title. At the bottom, a light green box contains the dates "November 9-10, 2017", the location "San Francisco CA, USA", and a partial sentence: "The web is **the** universal publishing platform. [Join us](#) for the first-ever".



EPUB3 Community Group

EPUB 3 Community Group

- ▶ EPUB 3.1 related technical activities, e.g.,
 - ▶ upgrade of the “EPUB for Education” profiles to EPUB 3.1
 - ▶ handling errata for EPUB 3.1 (with a possible republication)
 - ▶ exploring means to keep up EPUB Checker, reading systems’ test suite, or the EPUB 3 samples repository
- ▶ The membership of the CG *may* take up further work
 - ▶ re-publication of EPUB 3.1 in ISO
 - ▶ publication specific CSS transition techniques (e.g., transitions of images in a cartoon/manga)



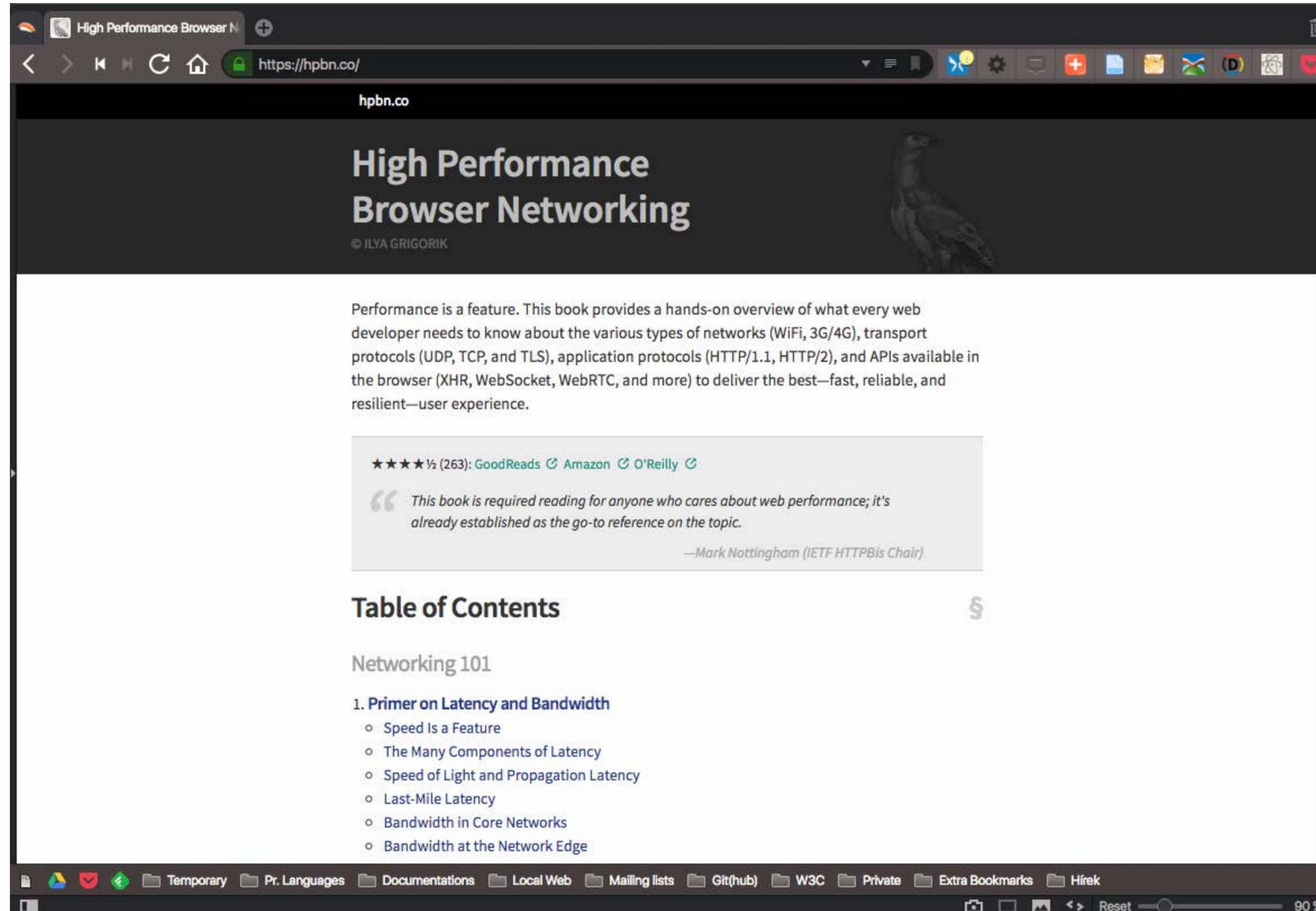
Publishing Working Group

“What”: Web Publications

Publications—with all their specificities and traditions—should become first class entities on the Web.

- ▶ This means:
 - ▶ it should be possible to load the publication content into a browser or a specialized reader, whatever the user prefers
 - ▶ it should be possible to read the book either offline or online, whatever the circumstances dictate
 - ▶ contents could be authored regardless of where they are used

Example: book through a Web browser, visible offline and online



Example: scientific article on the Web

PeerJ
CompSci

Related research

✓ PEER-REVIEWED

Achieving human and machine accessibility of cited data in scholarly publications

Human-Computer Interaction | Data Science | Digital Libraries | World Wide Web and Web Science

Joan Starr¹, Eleni Castro², Mercè Crosas², Michel Dumontier³, Robert R. Downs⁴, Ruth Duerr⁵, Laurel L. Haak⁶, Melissa Haendel⁷, Ivan Herman⁸, Simon Hodson⁹, Joe Hourclé¹⁰, John Ernest Kratz¹, Jennifer Lin¹¹, Lars Holm Nielsen¹², Amy Numberger¹³, Stefan Proell¹⁴, Andreas Rauber¹⁵, Simone Sacchi¹³, Arthur Smith¹⁶, Mike Taylor¹⁷, Tim Clark¹⁸

Published May 27, 2015

Note that a [Preprint of this article](#) also exists, first published December 14, 2014.

PubMed [26167542](#)

> Author and article information

Abstract

Reproducibility and reusability of research results is an important concern in scientific communication and science policy. A foundational element of reproducibility and reusability is the open and persistently available presentation of research data. However, many common approaches for

Enter your institution
To find colleagues at PeerJ

Enter to search

Download

Content Alert ^{NEW}

Just enter your email

Tools & info

Peer Review history

See citing articles **23**

Ask questions **4**

Add links

Visitors **5,642** click for details

Views **9,340**

Downloads **845**

Report problem with article

Outline

Introduction

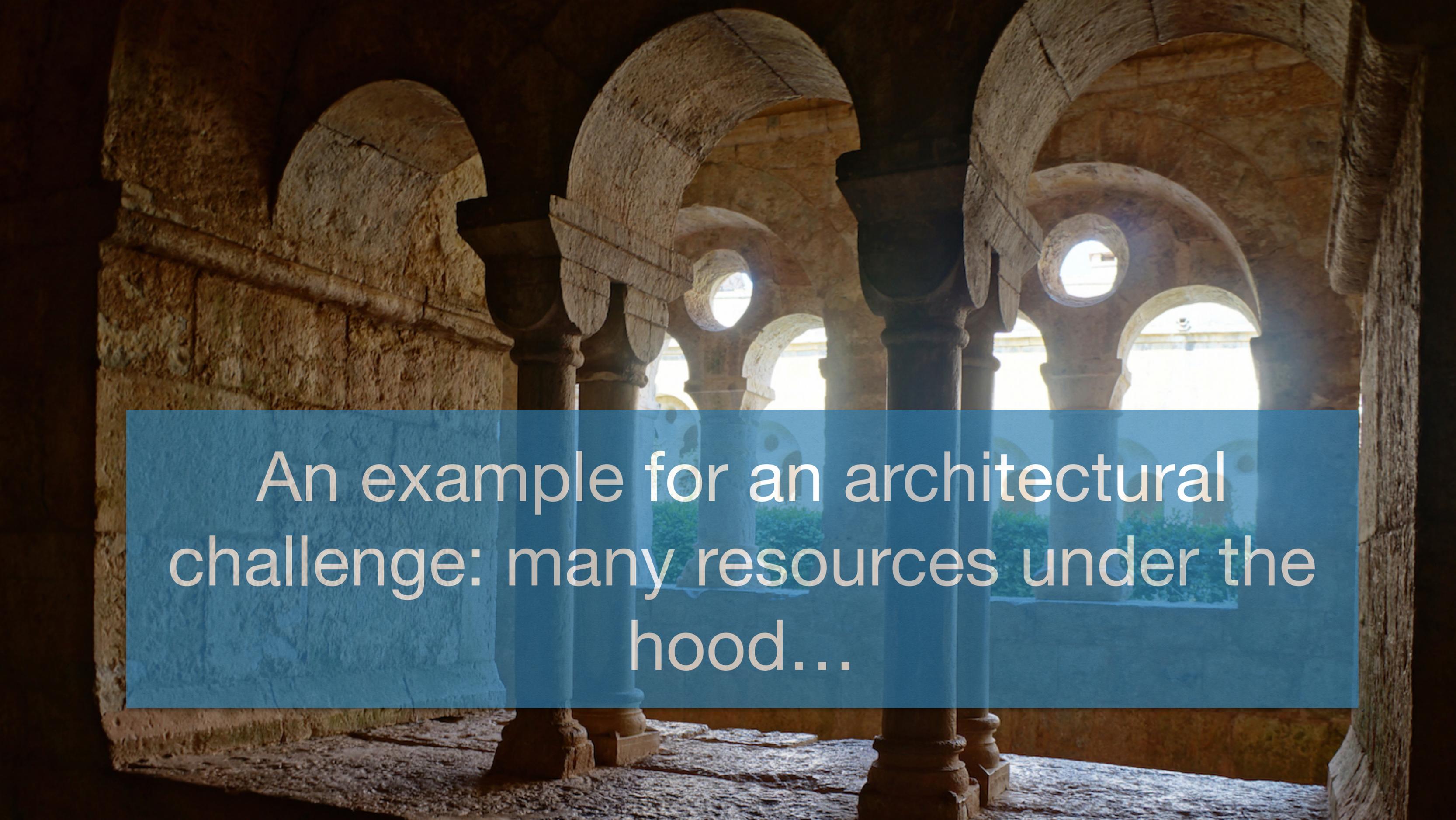
Recommendations for Achieving

These are still experimental

- ▶ Use some bleeding edge browser technologies (Web Workers, Service Workers, manifests)
- ▶ Many aspects are handcrafted in specialized scripts
 - ▶ or not yet done, e.g., downloading the scholarly article in PDF, i.e., a different format
- ▶ Lots of details remain to be clarified
- ▶ There should be an alignments with the technologies used in Web browsers, by Web developers, ...
- ▶ But: they indicate the direction!

Use Cases

- ▶ Some are listed in the “Web Publications Use Cases and Requirements”:
 - ▶ articles (scholarly journals, news agency reports, business publications)
 - ▶ magazines
 - ▶ product documentations
 - ▶ governmental and legal publications
 - ▶ and, of course, books of any kind
- ▶ All these use cases raise a number of technical requirements

A photograph of a stone building with a series of arches and columns, viewed from an interior perspective looking out through the arches. The architecture features a repeating pattern of semi-circular arches supported by columns. The stone is a warm, earthy tone. Light streams in from the openings, creating a play of light and shadow. A semi-transparent blue rectangular box is overlaid on the lower half of the image, containing white text.

An example for an architectural challenge: many resources under the hood...

Need the concept of “publication” over many resources on the Web

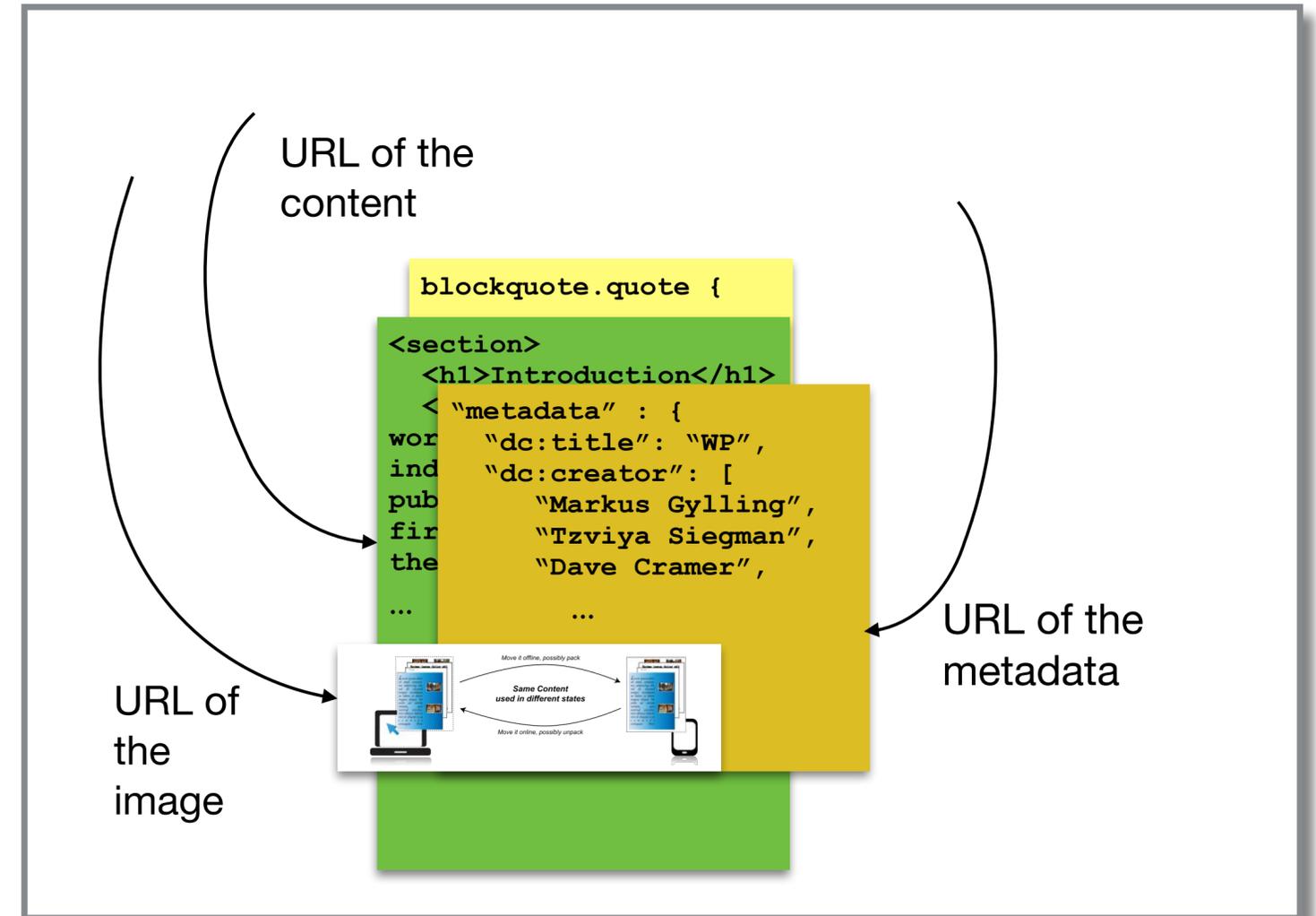
- ▶ A publication may consist of many resources (text, images, videos, etc)
- ▶ On the Web, these are linked together, but they keep their complete “independence”
- ▶ For publishing the abstract “work” should be a single, conceptual unit on the Web: **a Web Publication (WP)**

Why do we need the WP concept?

- ▶ The publication has to be checked for integrity as a whole (e.g., in a legal publication)
- ▶ Value counters (counters, section numbering, footnotes, endnotes, etc.) should be consistent across the publication
- ▶ Any search, text analysis, etc., facility for a book should have access to the full publication
- ▶ User preferences should apply to the “work” rather than to a single Web page
- ▶ Publishing and library business models are based on the concept of “works”

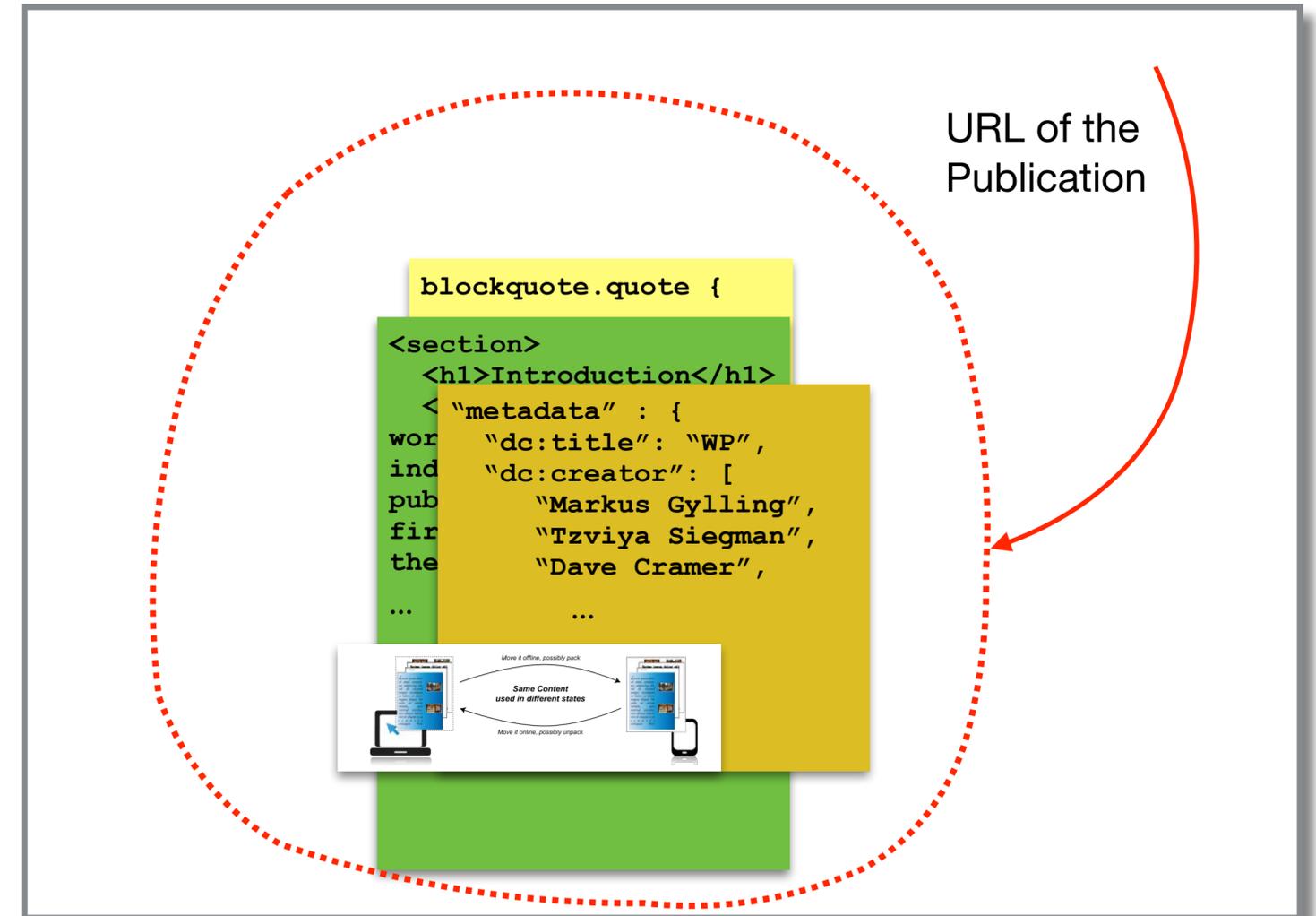
How does that translate to the Web?

- ▶ The current Web builds on the notion of a single resource:
 - ▶ HTML source, metadata, CSS style sheet, etc.
 - ▶ each has its own URL
- ▶ Rendering is based on the interoperation of many such resources



How does that translate to the Web?

- ▶ A WP is a collection of Web resources, considered as a single Web resource
- ▶ there is a unique Web address for the collection
- ▶ metadata is for the publication *as a whole*
- ▶ Web Publications are part of the Web!



A number of technical issues are being discussed

- ▶ Definition of the “WP InfoSet”, i.e., a collection of information that characterizes a WP
 - ▶ title, constituent Web resources and their reading order, identifiers, default natural language,...
- ▶ Specific security/privacy considerations
- ▶ Are there specific accessibility issues to consider
- ▶ What information are necessary for, e.g., unified search
- ▶ How to achieve offline access to publications
- ▶ ...

“What”: further concepts

- ▶ Packaged Web Publication
 - ▶ the Web Publication may have to be stored and distributed in a separate file (“package”)
 - ▶ what packaging format should be used?
- ▶ EPUB 4
 - ▶ a *profile* of PWP, with additional requirements (e.g., on accessibility)
 - ▶ trying to reduce backward incompatibilities with EPUB 3



When?

Now!

These groups are now operational

- ▶ The Publishing@W3C merger happened in February '17
- ▶ The Publishing Business Group has been established in March '17
- ▶ The EPUB 3 Community Group has been established in March '17
- ▶ The Publishing Working Group charter approval was in June '17, chartered until July 2020

Some upcoming events

- ▶ Community outreach at the Frankfurt Book Fair, 11-15 October (not confirmed yet)
- ▶ Several meetings in and around the W3C Technical Plenary in Burlingame, USA:
 - ▶ Working Group F2F Meeting on the 6-7 November
 - ▶ W3C Publishing Summit, 9-10 November
 - ▶ Business Group F2F Meeting on 10 of November
- ▶ EPUB Summit and a WG F2F meeting in Berlin, May 2018 (not confirmed yet)



Who?

This is not how development work at W3C...





...it is more like that!



W3C members move things forward!

In general...

- ▶ Different groups are staffed a bit differently
 - ▶ **Community Groups:** anybody can join
 - ▶ **Working or Interest Groups:** staffed by W3C members
 - ▶ **Business Groups:**
 - ▶ staffed by W3C members, or
 - ▶ special “Business Group” members can staff a specific BG on a lower fee (but not any other group)
- ▶ The way the work is done: ***find consensus*** among participants!

For Publishing@W3C

- ▶ IDPF members in good standing (in March '17) can join W3C gradually through the “Transitional Publishing Industry (TPI)” membership
 - ▶ similar fees to IDPF’s
 - ▶ can join the Publishing Business Group and the Publishing Working Group
 - ▶ (and, of course, the EPUB3 Community Group)
- ▶ This “transition” is for two years, then the standard W3C rules will apply

For practicalities on membership...

- ▶ Contact Georg Rehm, leader of the German Office of the W3C at DFKI, in Germany:
 - ▶ +49 30 23895-1833
 - ▶ Georg.Rehm@w3.org





Publishing@W3C: combine traditions with
modernity...

Some Useful Links

- ▶ “Home” of Publishing@W3C:
 - ▶ <http://www.w3.org/publishing/>
- ▶ Publishing Business Group:
 - ▶ <https://www.w3.org/publishing/groups/publ-bg/>
- ▶ Publishing Working Group:
 - ▶ <https://www.w3.org/publishing/groups/publ-wg/>
- ▶ EPUB 3 Community Group:
 - ▶ <https://www.w3.org/publishing/groups/epub3-cg/>
- ▶ Web Publication (WG Editors’ draft)
 - ▶ <https://w3c.github.io/wpub/>
- ▶ Web Publications Use Cases and Requirements (Interest Group Note):
 - ▶ <https://www.w3.org/TR/pwp-ucr/>
- ▶ Web Publication (Interest Group Note):
 - ▶ <https://www.w3.org/TR/pwp/>
- ▶ W3C Publishing Summit
 - ▶ <https://www.w3.org/publishing/events/summit2017.html#>

Slides are available at:

https://www.w3.org/2017/Talks/Publishers_DE-IH/

©W3C



Thank you for your attention!