Web 1,2,3?
The future of library services

Patrick Hochstenbach, Ghent University Library, Belgium
ELAG 2022, Riga, Latvia

Welcome

- Introduction
  - Working since 1996 for academic libraries
  - SFX, OpenURL, OAI-PMH Static Repositories
  - Ghent Library, Los Alamos National Laboratory, Lund University LIbrary
  - Currently PHD Computer Science on Decentralized Web at IDLab, IMEC, Ghent
  - Since 2010 part of ELAG Programme Committee
- Familiar with Web 3.0/Web3/Semantic Web?
  - My main interest is in the Semantic Web. I've been a (late) adopter, I am (still) a proponent
  - I've been following Blockchain. I'm not (yet) an adopter, I am (still) a critic

Patrick Hochstenbach

Sources

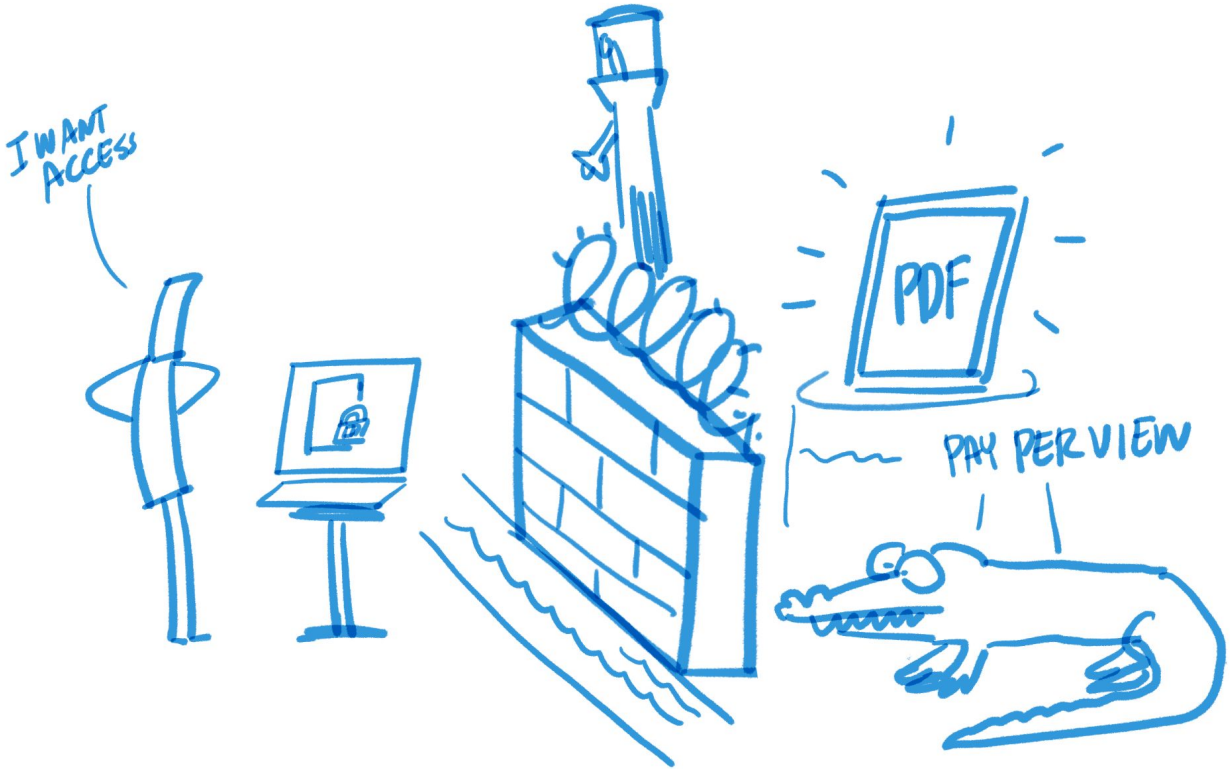
- Lots of material in this presentation is based and adapted from Ruben Verborgh's Ghent University class *Web Fundamentals* . Slides: {12-20 + 32-36}
- David Rosenthal's Standford EE380 talk. Slides: {51-52}
- Pascal Hitzler's ACM *A Review of the Semantic Web Field* article. Slides: {59-62}
- And David Deutsch's foreword in *The Science of Can and Can't* ISBN: 978-0-241-31094-6. Slide: {65}

Patrick Hochstenbach

A sort of introduction

From which, remarkably enough, nothing develops

Patrick Hochstenbach



Medicine or remedy?

- Open Access
- OpenURL
- Seamless Access (GEANT)
- GetFTR (publishers)
- CASA (Google)
- Libkey (Third Iron)
- QuickLinks (Exlibris)
- RightFind (CCC)
- Shibolleth (national)
- VPN
- IP Based
- Citrix
- EZProxy
- ….

Patrick Hochstenbach

Aaron Tay @aarontay
Library Analytics Manager
Singapore Management University

Patrick Hochstenbach

Aaron Tay @aarontay · Dec 2, 2021

So it is 2031, what is the fate of access/delivery technologies eg seamlessaccess, getftr, CASA, access broker extensions, libkey etc for library resources? Here are some scenarios... (1)

> **Aaron Tay** @aarontay · Nov 30, 2021
>
> [blogged] A relook at GetFTR , Libkey, Exlibris Quicklinks, and other linking and authentication technologies musingsaboutlibrarianship.blogspot.com/2021/11/a-relo...
>
> Show this thread

2 4 3

Aaron Tay @aarontay · Dec 2, 2021

Scenario 1 - Seamless Access, GetFTR are heavily adopted by publishers (even aggregators for the later) & most institutions move to federated access. IP auth is mostly dead. Access broker extensions mostly gone except those that have pivoted to discovery/recommender features (2)

1

Aaron Tay @aarontay · Dec 2, 2021

In this scenario i envision library link resolvers still around cos GS stubbornly refuses to support getFTR and well we need link resolvers for non-doi material that isnt supported by getFTR (3)

1

Patrick Hc

Aaron Tay @aarontay · Dec 2, 2021

scenario 2 - This is status quo. The landscape is a mess of different diverse, fragmented technologies. GetFTR never signs up aggregators so it is not a complete solution, users are confused with all the options. (4)

1 1

Aaron Tay @aarontay · Dec 2, 2021

Scenario 3 - OA triumphant! Most acad books, journals are open access/free to read. While you still need to authenicate for A&Is (which are mostly dying) and other resources at least for full-text journals and books, people expect "real seamless access" aka no login (5)

2 1

Patrick Hochstenbach @hochstenbach · Dec 2, 2021

We'll try for option 3 in our Mellon funded research project on (re)decentralisation of all aspects of scholalry publication : peer review, publishing, search and retrieve, archiving

1

Patrick Hochstenbach

Aaron Tay
@aarontay

Replying to @hochstenbach

wishing you all the luck but wild guessing I would say Scenario 3 is even less likely than Scenario 1 - Seamlessaccess+GetFTR wins completely and most likely is 2 - more or less status quo with mix of everything. We shall see...

2:14 PM · Dec 2, 2021 · Twitter for Android

How we got here?

Patrick Hochstenbach


I didn't invent the Internet
I did not invent the Web
Tim Berners-Lee
Vint Cerf
©2014 W3C

The world before the Web was highly **heterogeneous**.

- **Exchanging information was hard.**
  - different hardware
  - different software
- **Innovation was hard.**
  - For which machines do we build?
  - For which operating systems do we build?

Patrick Hochstenbach

The Web strives to be **universal** through independence of many factors.

Anyone can use the Web, regardless of:

- *Hardware:* desktop, phone, tablet, watch…
- *Software:* operating system, browser, app…

Developers are free to innovate.

- Build for the Web.
- Standards provide interoperability.

Patrick Hochstenbach

The Web’s universality helped accelerate its growth.

HTTP and HTML were OS-independent.

- Yet the server and browser were NeXT-specific.

Support for other systems followed in 1992.

- Notably Mosaic gave the Web a broader audience.

It provides compatibility at an unprecedented scale.

- The text-based browser Lynx is still developed and used.
- Your phone can access websites created 20 years ago.

Patrick Hochstenbach

The Web brings **freedom of expression** to **everyone** across the world.

- Anyone should **access** and **benefit** from the Web #ForEveryone
- Anyone can say **anything** about **anything**.
- We all have our own **spaces**,
  so we don't have to **agree**.
- We can **link** to opinions of others
  to discuss about them.

Patrick Hochstenbach

The Web brings **permissionless innovation** at a global scale

- **Anyone** can build **anything** for **any** reason.
- The technologies are **open**.
- You **don’t need** anyone’s **permission** to join the Web and launch a new idea.

Patrick Hochstenbach

Permissionless innovation has brought unprecedented **creativity** to the world.

- What permissions Google and Facebook needed to create their catalog?
  - In contrast: what permissions your library need(ed) to access our own data?

Patrick Hochstenbach

The number of websites started growing at an explosive rate.

year	number of websites
1991	1
1992	10
1993	130
1994	2,738
1995	23,500
2000	17,000,000
2010	207,000,000
2022	1,955,000,000

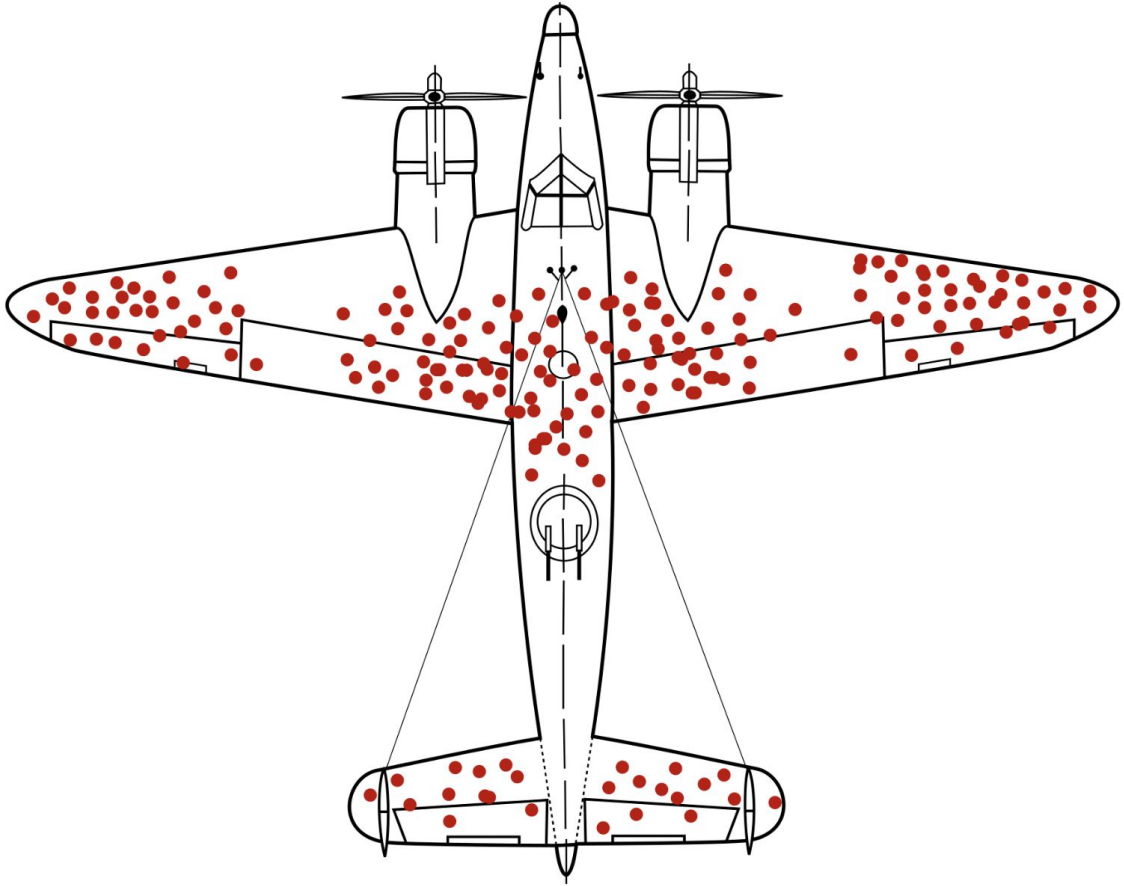
Patrick Hochstenbach

And then … oops

Patrick Hochstenbach

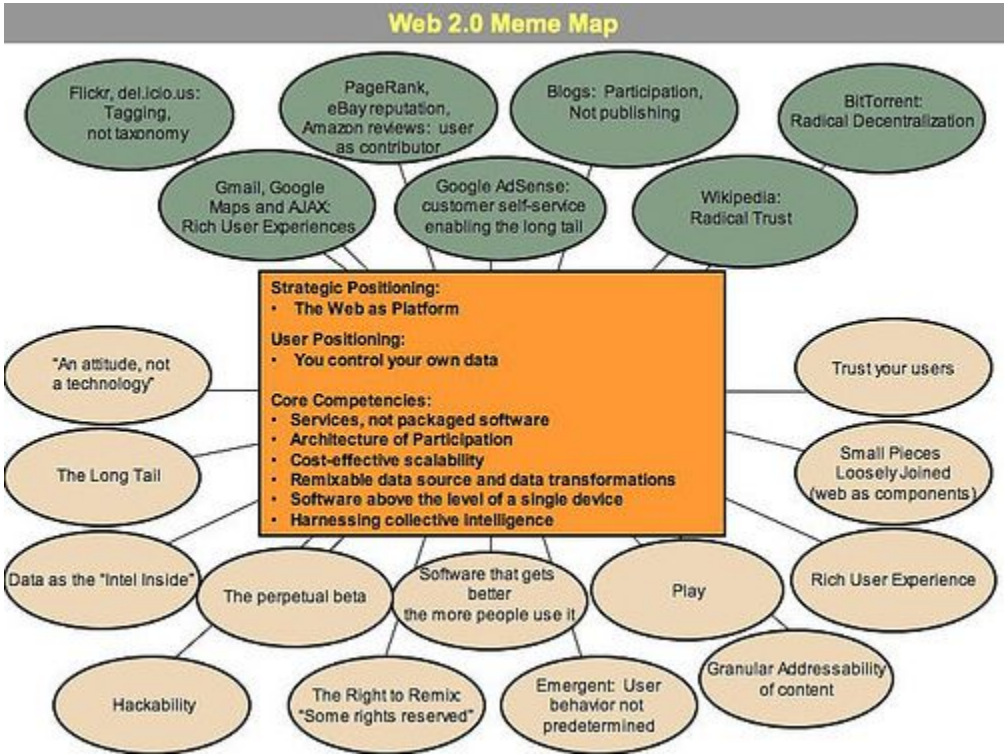
Oops for whom? Did the Web crash?

Patrick Hochstenbach



Patrick Hochstenbach

O'Reilly 2005 - What did survivors do well? => Web 2.0



Patrick Hochstenbach

https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html

Contrast it with Web 1.0 : the global **library**

- The publishing, read-only Web
- Static HTML + binary content + hyperlinks
- Give away browsers, sell hardware and software
- You run your own server
- You did content management
- Screen scraping to provide interoperability

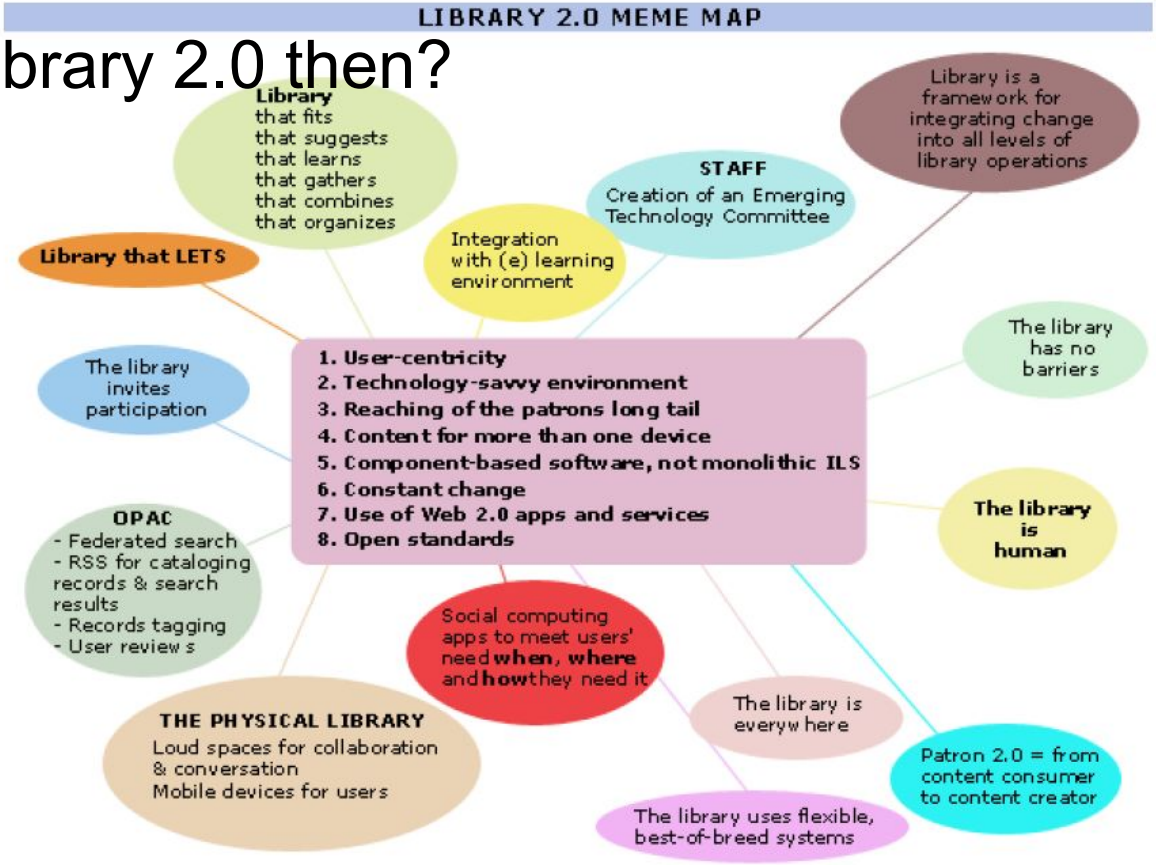
Patrick Hochstenbach

Web 2.0 should **collaborate** and **share**

- Radical decentralisation (Bit Torrent)
- Radical trust (Wikipedia)
- Participation (Blogs), Tagging (Flickr , del.icio.us)
- The web as platform (for humans)
- Services (not packaged software)
- Remixing of data
- Data!!
  - Who own it
  - "The race is on to own certain classes of core data [..] creating a critical mass via aggregation and turn that data into a service"

Patrick Hochstenbach

What is Library 2.0 then?



Pa http://www.flickr.com/photos/42538191@N00/113222147/

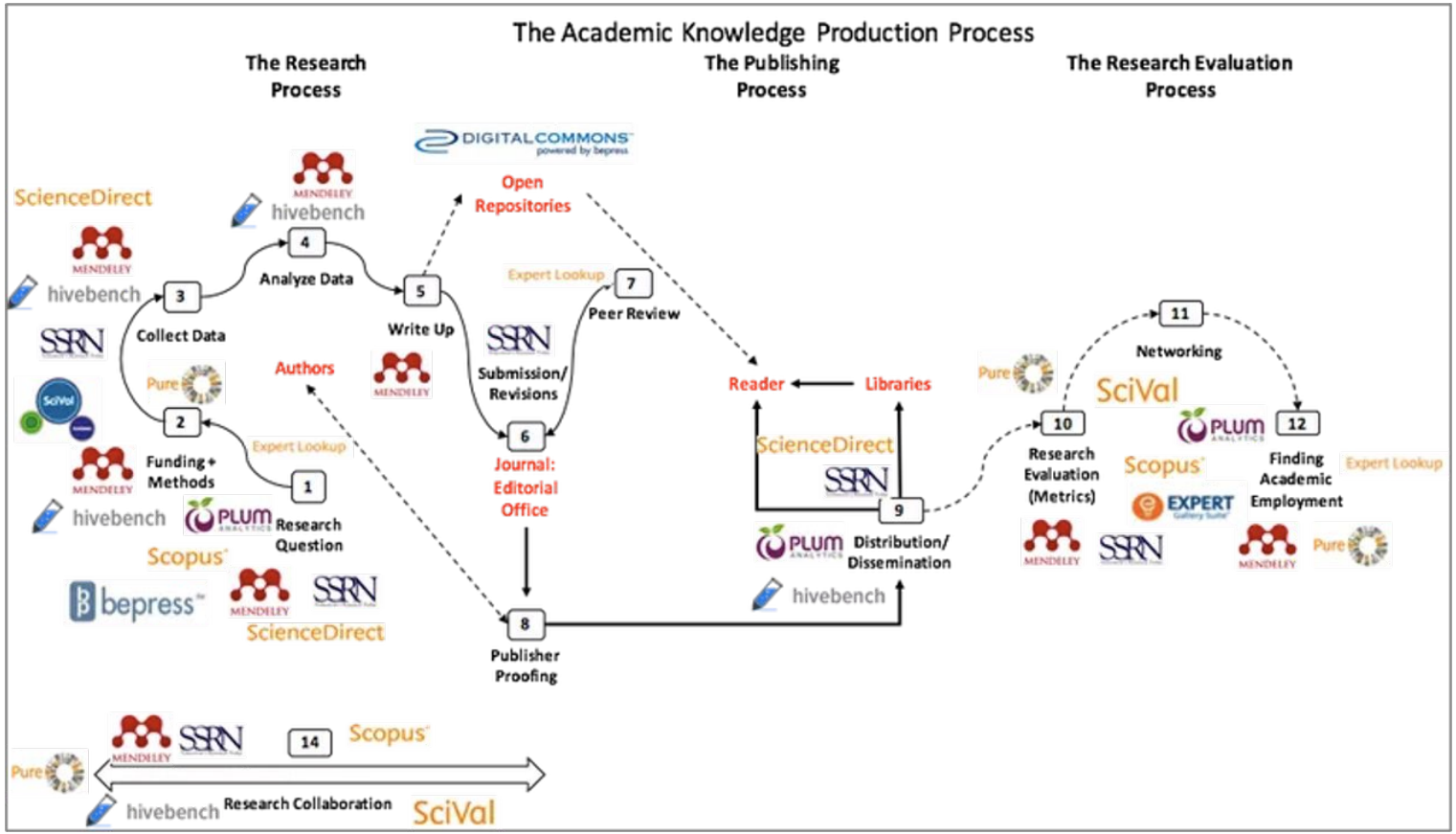
Library 2.0 meant **participation** & **fun**!

- Participation
  - Comment
  - Tagging
  - RSS feeds
  - Trust our users more than ever
- Integration & Flexibility
- Dis-Integrated Library Systems
  - Open source
  - Open standards
  - Pick and choose from a variety of sources

Patrick Hochstenbach

Did we have fun yet?

Patrick Hochstenbach



Posada, A., Chen, G. (2017) Publishers are increasingly in control of scholarly infrastructure and why we should care
http://bit.ly/2xKRnSr

Patrick Hochstenbach

Challenges

Patrick Hochstenbach

Threats to universality

- **Centralization** from multiple angles is threatening the Web.
- **Technological decentralization** can differ from practice:
  - Only certain devices and software can access parts of the Web.
  - A few companies make or break websites (e.g. search engines)
  - Platforms restrict the web (e.g. you need a specific account to access parts of the Web)
- Platforms that became big due to permissionless innovation are **preventing** this mode

Patrick Hochstenbach

The Facebook founder has no intention of allowing anyone to build anything on his platform that does not have his express approval.

Having profited mightily from the Web's openness, he has kicked away the ladder that elevated him to his current eminence.

—John Naughton, The Guardian

Patrick Hochstenbach

Do we give governments the power to censor the Web?

Patrick Hochstenbach

Patrick Hochstenbach

Patrick Hochstenbach

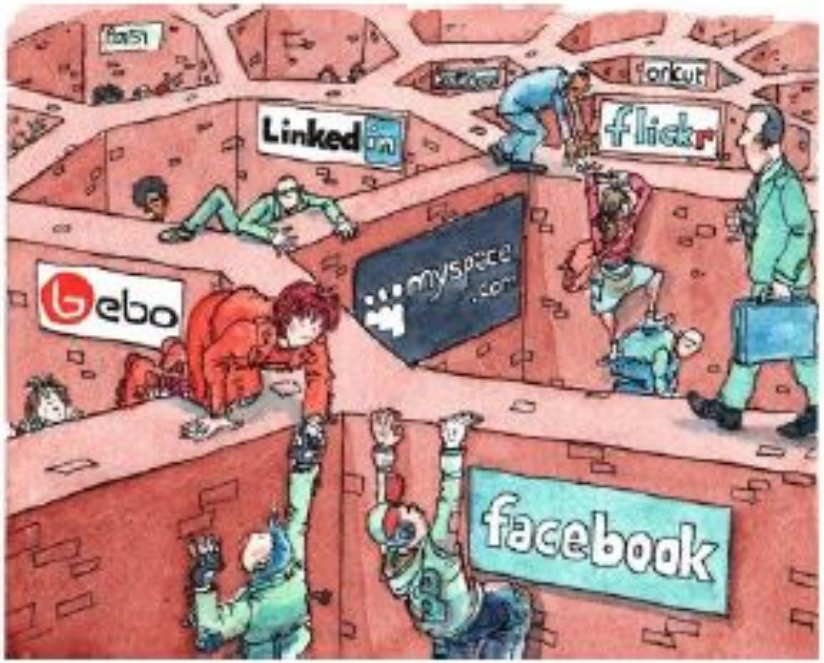
What did we get?

Patrick Hochstenbach

What did we get?

- Instead of **being** Web 2.0 we are **controlled** by Web 2.0
- We are giving away our **identity**, our **data** & our **services**
- Platforms have **tight control** over all these aspects in silos with **tight coupled** services
- Dr Darren Saunders: "In some ways the scientific publishing model resembles the economy of the social [Web 2.0] internet: labour is provided for **free** in exchange for the hope of **status**, while **huge profits** are made by a few big firms who run the market places."

Patrick Hochstenbach



Patrick Hochstenbach

If there is a sense of reality,
there must also be a sense
of possibility

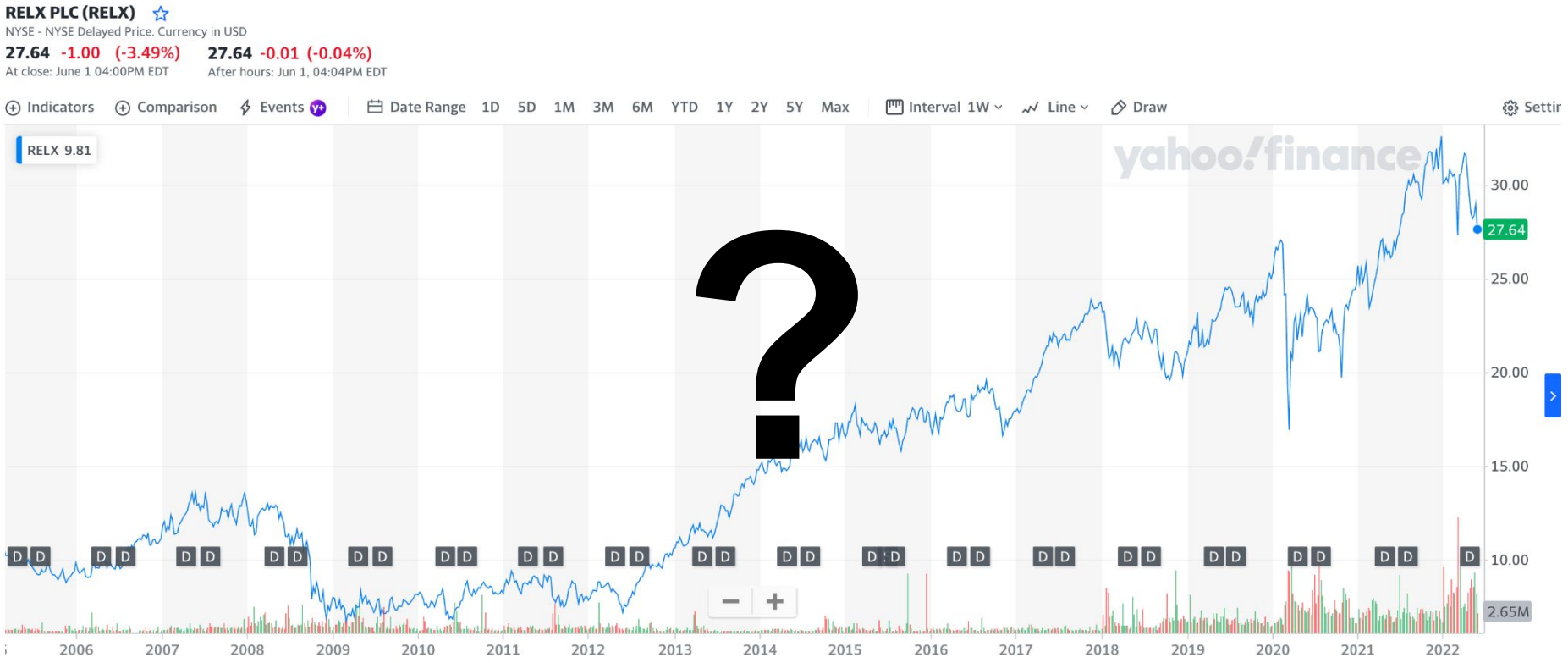
Patrick Hochstenbach

The web is dominated by a small number of large companies.
Could it be different?

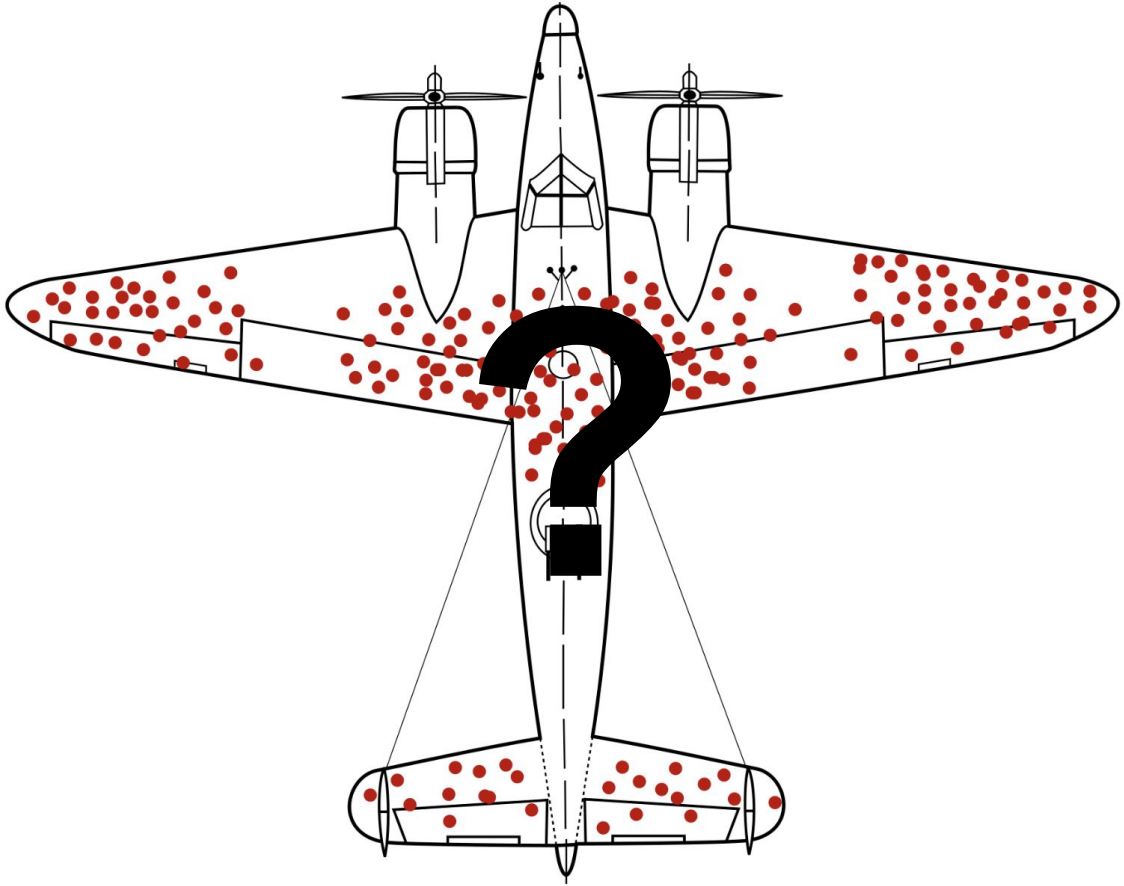
Patrick Hochstenbach

Need a Web 3.0 , Web3, something with a 3?

Patrick Hochstenbach



Patrick Hochstenbach



Patrick Hochstenbach


ETHER
DEBT
DIGITAL CASH
PEER-TO-PEER
WEALTH
NETWORK
E-COMMERCE
DATA
RENDER
CONTRACT
ANONYMOUS
EXCHANGE
MARKET
BITCOIN
DIGITAL
BLOCKCHAIN
PAY
EXCHANGE
SMART
CRYPTOCURRENCY
VIRTUAL
ETHEREUM
ALTCOIN
FINANCE
WEB
INTERNET
INTERNATIONAL
PAYMENTS
TECHNOLOGY
BUY
INVESTMENT
LEDGER

What are the options?

Patrick Hochstenbach

Web 3.0, Web3, Semantic Web

- There is a bit of confusion of terminology. The combination of Web with the number 3 is used by different technologies.
- They all talk about providing solutions for:
  - Decentralization
  - Democratization
  - Online safety
  - Digital divides
  - Digital colonialism
  - Accessibility
- But the solutions are radical different
  - Web3 technologies such as Blockchain - rebooting the Web from scratch
  - Web 3.0 technologies, Semantic Web - adding on top of existing Web

Patrick Hochstenbach

A review of the Web3

Patrick Hochstenbach

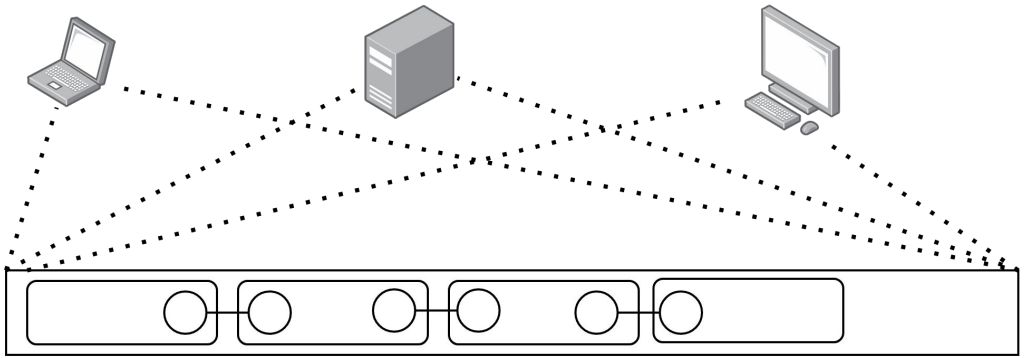
Web3 *coined* [sic] by Gavin Wood (co-founder Ethereum)

- Blockchain
- Cryptocurrency
- NFTs
- Catch all for decentralized digital infrastructure

Patrick Hochstenbach

Blockchain Review

- Origins of the technology itself is not new
  - Blockchain technologies were already available in 1991 (Stornetta & Haber)
- Global shared add-only ledger



Patrick Hochstenbach

Blockchain Review, the Good

- Comes in two forms
  - **Permissioned** blockchain - require a central authority
  - **Permissionless** blockchain - decentralized network
- Three use-cases:
  - **Data authentication** and **verification**: encryption, data that can't change , digital signatures
  - **Asset management:** cryptocurrencies
  - **Smart contracts**: allows some kind of distributed computation
- Great technology to **prevent disagreement** on transactions in decentralized environment where there is **no trust** between the participating nodes
  - Who did what on the shared ledger
- Since 2008 used in cryptocurrency by Satoshi Nakamoto (not much is known about the person, persons or group behind that name)
  - Financial exchanges without central banking

Patrick Hochstenbach

Blockchain Review, the Bad

- Almost all (crypto) blockchains are permissionless
  - No central authority required to add things to the ledger
  - But, need some way to prevent anyone messing with the chain, or prevent fraudulent changes to the chain
- There blockchains need some kind of cryptocurrency to operate and protect the network; cryptocurrency needs speculation to function (numbers should go up)
- Proof of {...resources…}
- Massive computations with linear rewards
  - The more you can compute, the more revenue
- In an economy of scale this leads to an effective centralisation
  - A handful (~50) miners control the majority of mining power

Patrick Hochstenbach

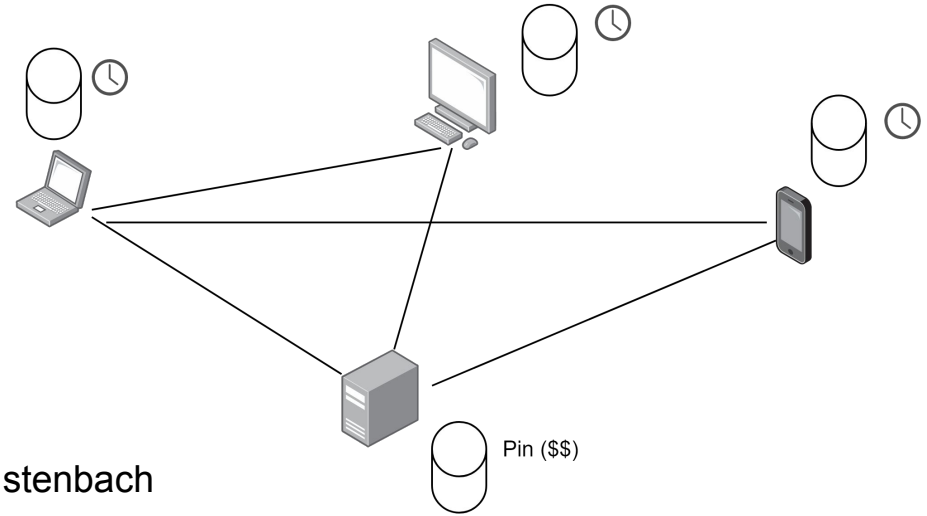
Blockchain Review, the Ugly

- Can one ignore the externalities?
- Speculation with big volatility
- Massive carbon footprint
- Disrupting hardware supply chains
- Magnet for crime
  - Money and (rogue) state actors are involved, anonymous exchanges of funds, tax evasion, fraud, hacks (involving billions of dollars)
  - Actions that can't be undo (immutability)
  - One small mistake results in many oopsies involving $12 billion, 5% of all funds, lost so far
- In effect what you get as public is not a decentralised service but dependent on the API of a few large coin markets

Patrick Hochstenbach

IPFS Review

- Juan Benet
- Origins in the P2P era of Napster in the later 1990s
- Benet combined in 2013~2015 the benefits of Git and BitTorrent into a global filesystem and founded Protocol Labs in 2014 to support it



Patrick Hochstenbach

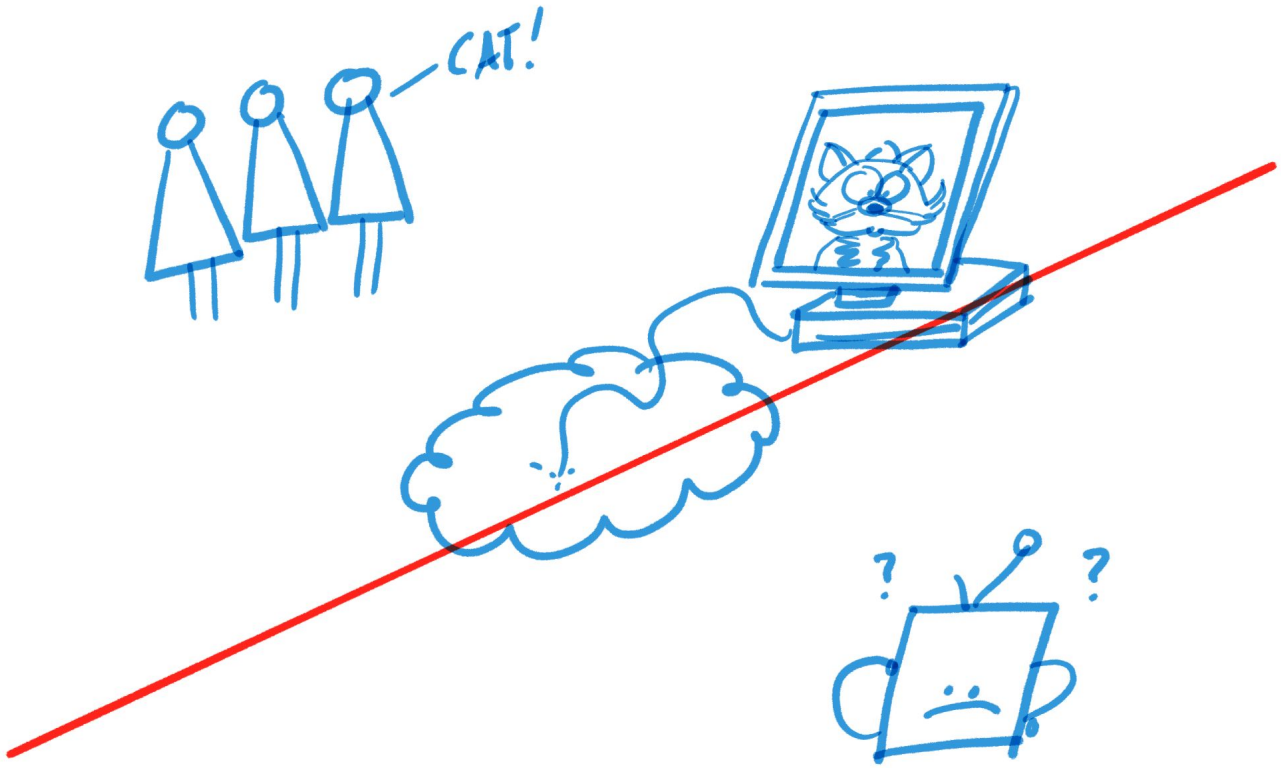
IPFS Review

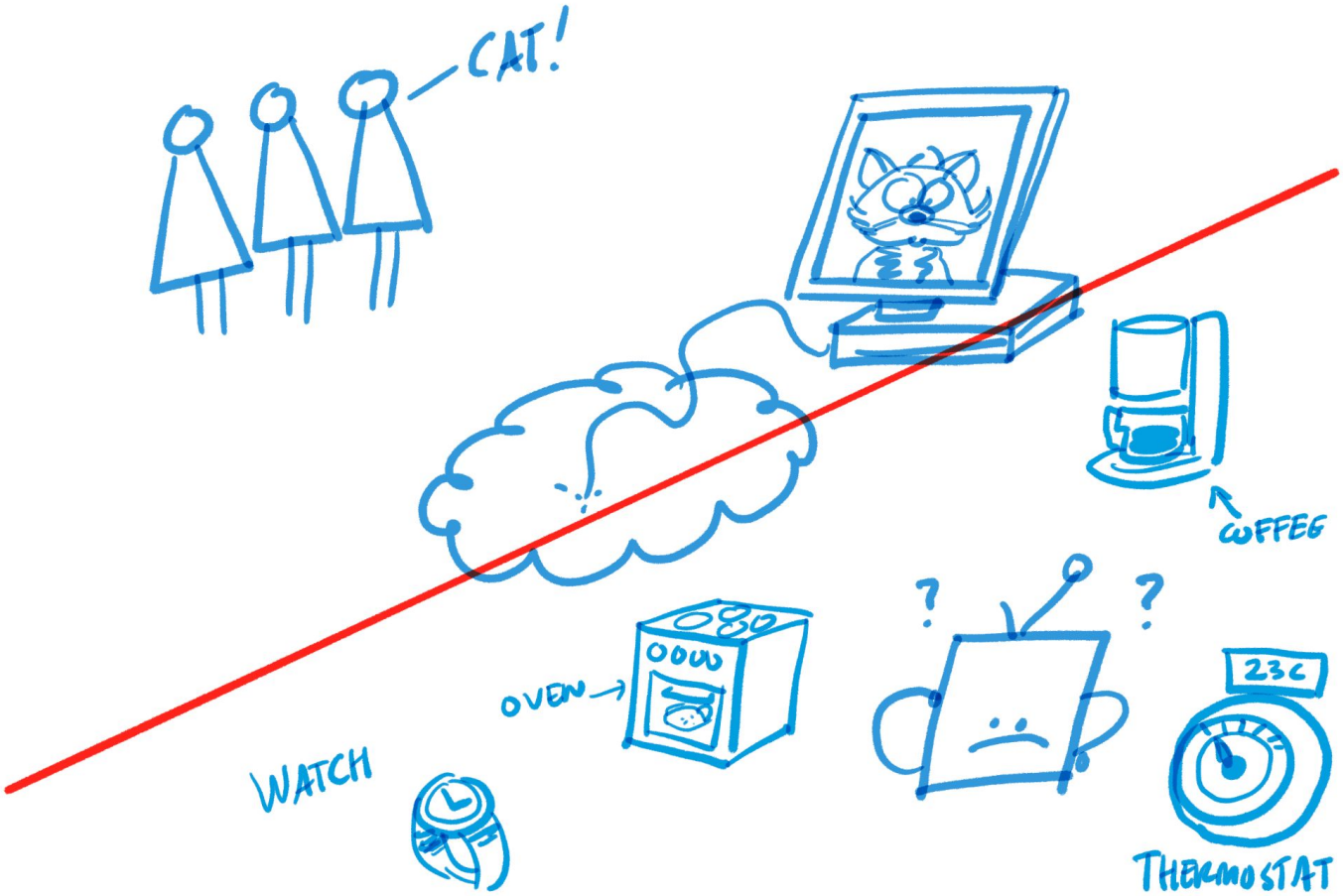
- It is not "free" storage device
- Your data is stored on your disk but is shared with the world
  - They can cache the data as long as they want
- Popular data will live as long as IPFS is available
- Unpopular data gets eventually forgotten
- …Unless you pay a Pinning server to keep the data available in the network
  - A kind of rent free payed in FileCoin
- Use cases:
  - sharing public data
  - private data only with strong encryption and a level of centralisation who can access
- IPFS says how to store, not what it is that you store

Patrick Hochstenbach

A review of Web 3.0

Patrick Hochstenbach





(Sem|Ped)antic Web 1997 -

- RDF
  - Machine understandable metadata
- Ontologies
  - Data models
    - Concepts ("cat" , "mammal" and "live birth")
    - Relationships (:"mammals give live birth")
  - Standards such as OWL
- Rule languages
  - RIF, Notation3
- Query languages
  - SPARQL
- Autonomous agents

Patrick Hochstenbach

Linked (Open) Data, Pragmatic Web 2006 -

- 4 rules deployment scheme
  - Use URIs as names for things
  - Use HTTP URIs so that people can look up those names
  - When someone looks up a URL, provide useful information, using standards (RDF,...)
  - Include links to other URIs to that they can discover more things
- Less emphasis on OWL and reasoning
- Open & crowdsourcing
- DBPedia, Wikidata
- BBC, Facebook, Google
- Schema.org (in 2015 30% of all web pages were tagged this way)

Patrick Hochstenbach

Knowledge Graphs, Corporate Web 2012 -

- The techniques of the Semantic Web can be used outside the web too within the border of a platform
- Google Knowledge Graph (the infoboxes in their search results)
  - No downloads, but with an API
- Linked ~~Open~~ Data ©
- Dangers of new centralisation

Patrick Hochstenbach

Social Web & Decentralization 2016 -

- Separate the application from storage
  - Breaking down the solis
  - Store all your data on a server you trust (your tweets, your profile, your images, your blog posts, your likes, your fitness statistics)
- Separate the service from implementation
  - No more need for a tight coupling of services to a platform
- Create a market for applications on top of this data
  - Installed on your phone, desktop, browser or remote
- It allow you to:
  - Control access to your data, whatever applications are used
  - Data from various applications be cross-linked, breaking silos
  - Permissionless innovation
  - Persistence of applications and data may be different.

Patrick Hochstenbach

Now what?

Patrick Hochstenbach

Summary

- There is no easy solution to solve the world problems
- Economic of scale leads to centralisation
  - If revenue grows faster than is linearly with level of contribution, you get centralisation
- Some level of centralisation is hard to avoid
- What we have access to might be more important than where it is stored

Patrick Hochstenbach

Take alternatives seriously

- The **actual** is maintaining the status quo
- The **counterfactual** is how it all could/can be different
- Focussing on the actual we can become to risk-averse with some pessimism and fatalism as norm
- Does all "low-hanging fruit" already have been covered and solved?
- This is a time with more contradictions, gaps and unresolved vagueness in accessing and using information with exciting prospects to explore them.
- Sometimes this will require us to adopt radically different modes of exploration.

Patrick Hochstenbach