# Nástroje a možnosti internetu

Organizace webu

8. 10. 2021

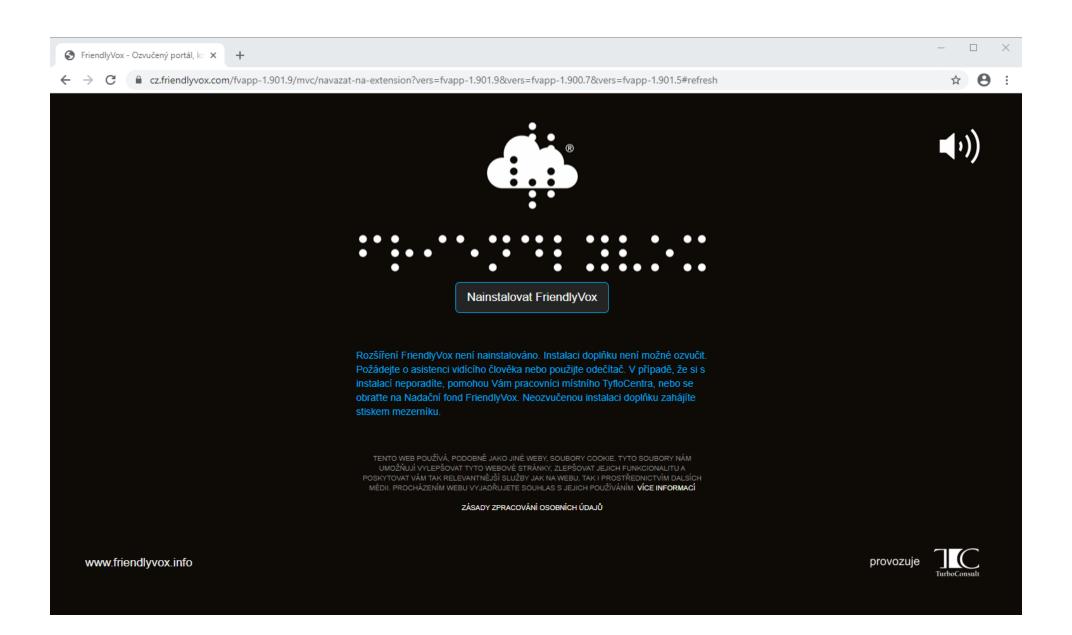
## Doplnění: přístupnost webu

dobrý web je přístupný

Google -> Tab

- mnoho bariér
- musí to řešit i knihovny!
- je to dané i zákonem
- WCAG 2.1

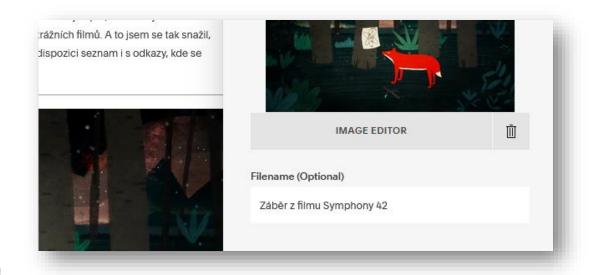
"Access by everyone regardless of disability is an essential aspect."

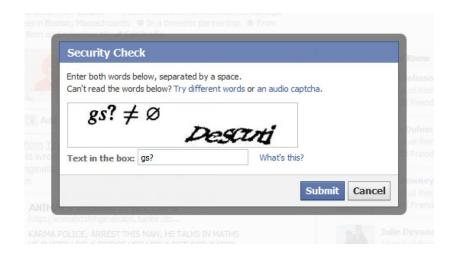


### Textová alternativa

## Alternativní text Jak byste tento objekt a jeho kontext popsali někomu, kdo je nevidomý? (Doporučujeme napsat 1 až 2 věty.) Obsah obrázku pes, exteriér, tráva, hnědá Popis byl vytvořen automaticky Označit jako dekorativní 🕕

### <img src="obrazek.jpg" alt="Pejsek">





## Přístupnost webu

- možná automatizace
- ale co třeba grafy?



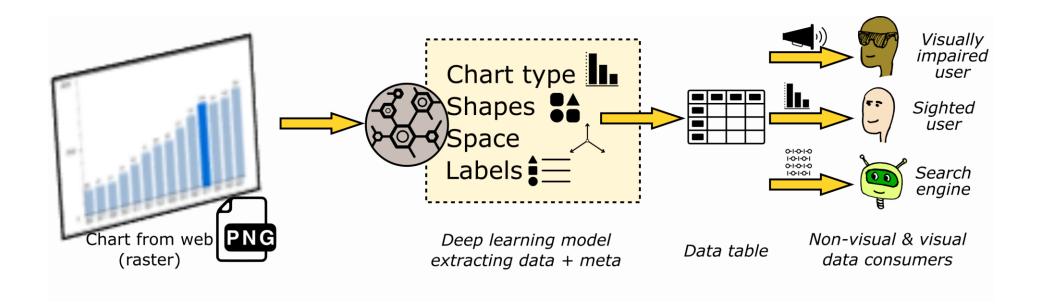
#### R-číslo: ukazatel, jakou rychlostí se epidemie šíří

R-číslo udává průměrný počet dalších osob, které přímo nakazí jeden nakažený pacient. Čím vyšší je reprodukční číslo, tím se nemoc rychleji šíří. Pokud je R-číslo menší než jedna, počet nakažených v populaci klesá.

Jak uvádí ministerstvo zdravotnictví, model pro výpočet R-čísla pracuje s počty nově nakažených osob, které prozatím nevykazují příznaky, průběhem jejich onemocnění včetně infekčního období, nástupem příznaků onemocnění, přesunem pacienta do izolace a potvrzením onemocnění laboratorním testem.

Vývoj hodnoty R-čísla v Česku od začátku epidemie





Eurographics Conference on Visualization (EuroVis) 20 (Guest Editors) Volume 38 (2019), Number 3

### Visualizing for the Non-Visual: Enabling the Visually Impaired to Use Visualization

J. Choi, D S. Jung, D D. G. Park, D J. Choo, D and N. Elmqvist

<sup>1</sup>Korea University, Seoul, South Korea, <sup>2</sup>University of Texas at Arlington, TX, USA, <sup>3</sup>University of Maryland, College Park, MD, USA

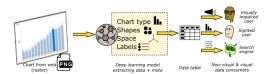


Figure 1: Overview of our pipeline enabling visualizing data for visually impaired users. A bornt stored as a ruser lindge is retrieved to from the web. Our pipeline automatically detects the chart type, extracted the shapes, recovers the substrate, and it is extracted as a data table. This data table can then be used for several purposes (from the top): with a screen reader for a visually impaired person, new standards for a subject person, or as travel data for a robot (search engine).

#### Abstract

The majority of visualizations on the web are still stored as ruster images, making them inaccessible to visually impaired users. We propose a deep-neural-network-based approach that automatically recognizes key elements in a visualization, including a visualization type, graphical elements, labels, legends, and most importantly, the original data conveyed in the visualization. We leverage such extracted information to provide the reading of the extracted information to visually impaired people. Based on interviews with visually impaired users, we built of Google Chrome extension designed to work with screen reader software to automatically decode charts on a webpage using our pipeline. We compared the performance of the back-end algorithm with existing methods and evaluated the utility using qualitative feedback from visually impaired users.

## A nyní: všichni publikují...

- informační (webové) přehlcení web nekončí
- jak v tom dělat pořádek?
- jak (si) informace organizovat?
- · využít v dalších procesech, vytvářet hodnotu?

"And many people become "infovores", force-feeding themselves endless books, articles, and courses, in the hope that something will stick.
[...] How do I make what I'm consuming right now easily discoverable for my future self?"

navigační problém & nowness

#### Five Minutes to Your Own Website

hile lum is a great option to start blogging because the limited



### Practical Advice for Data Science Writing

https://medium.com/p/cc842795ed52?source=user\_profile-----3-----

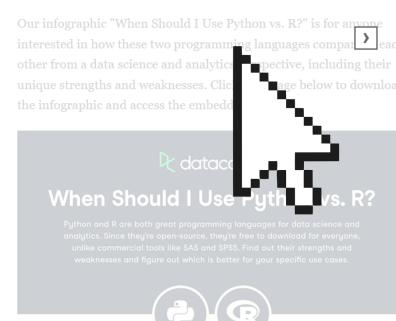
#### The Best Tool is the One that Gets the Job Done

Windows vs MacOS. Rvs Python. Sublime vs Atom vs PyCharm. Medium vs your own blog. These arguments are all unproductive. The correct response is to use whatever tool lets you solve the problem (within the confines of your environment). Moreover, the tool with more options is not always better.

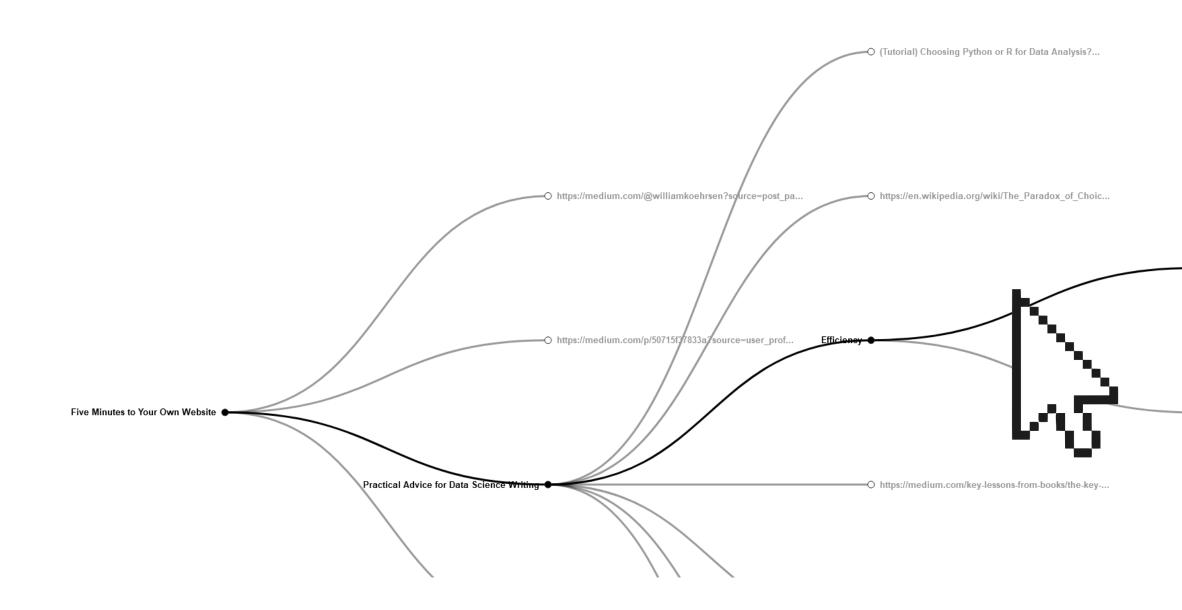
While more features can sound great, they often get in the way of you doing work. Generally, I try to keep things as simple as possible. When people ask me for recommendations for a writing platform, I say Medium because it has a limited amount of features. When I write, I want to focus on the content instead of spending time trying to format everything exactly as I want.

More customization options means more time customizing those options and less time doing what you should be doing — writing or coding.

I've gotten stuck in the tool optimization loop before: I've been persuaded to switch to a new technology and spent time to learn the features only to be told that this technology is obsolete and the next







hypertext hypermedia

## Přístupy k organizaci webu

- <u>Gopher</u> adresářová struktura
- Web propojená hypermédia
- navigation problem
- harmonie mezi propojováním a řádem

The navigation problem is becoming even more acute with the continuous growth of web sites in terms of their structure, which is becoming more complex, and their size in terms of the vast amount of information they intend to deliver. In contrast, users are not willing to invest time to learn this structure and expect transparent delivery of relevant content without delay. This problem needs to be tackled by orientation and navigation tools which improve user interaction, in analogy to the diverse set of tools ranging from road signs, the compass and map, to global positioning systems (GPS), all of which we use to orient ourselves in physical spaces.

### HYPERTEXT GARDENS

## Hypertext Gardens: Delightful Vistas

Mark Bernstein
Eastgate Systems, Inc.

**Into the Garden** 

Beyond the Navigation Problem
The Limits of Structure

**Gardens and Paths** 

The Virtue of Irregularity
Gates and Signposts

Statuary and Follies: Punctuating the Reader Experience
Planning Pathways

**Seven Lessons from Gardening** 

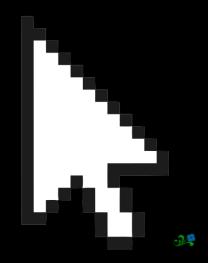
25

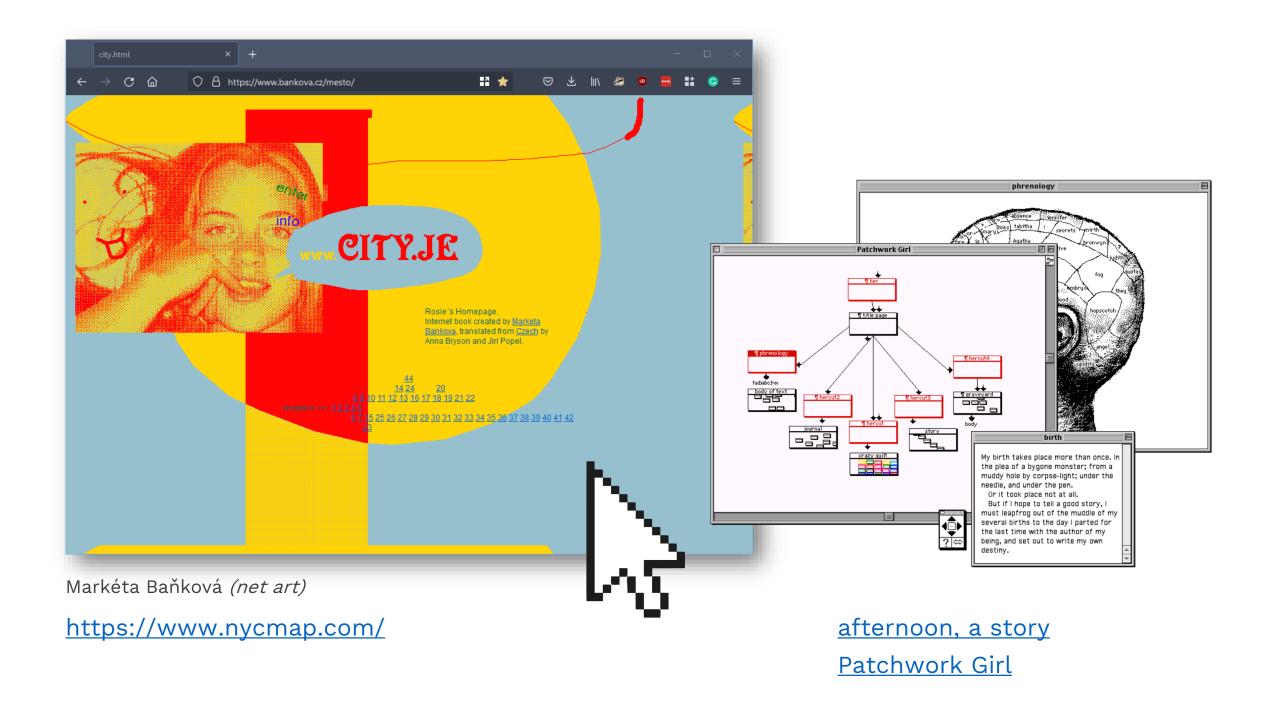
### EASTGATE

Copyright (c) 1998 by Eastgate Systems, Inc. All Rights Reserved.

Eastgate Systems, Inc. 134 Main Street, Watertown MA 02472 USA. Email: <u>info@eastgate.com</u>. Phone: (617) 924-9044 (800) 562-1638

http///www.eastgate.com/garden/Enter.html





## Co je dnes hlavním organizačním hlediskem webu?

## Je web v nepořádku?

Hossein Derakhshan: "The web was not envisioned as a form of television when it was invented. But, like it or not, it is rapidly resembling TV: linear, passive, programmed and inward-looking."

- internet jako knihovna -> internet jako televize
- uzavřené ekosystémy -> opravdu chcete odejít z FB?
- chronologie na úkor kurátorství/propojování

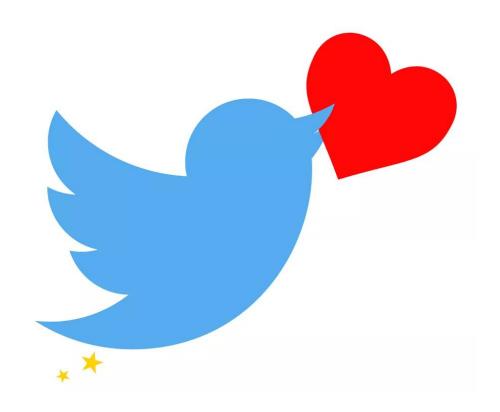
## Organizační metafory webu

řeka táborák zahrada



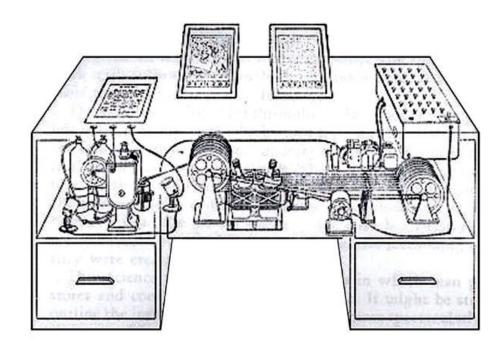






web jako druhý mozek

"Consider a future device [...] in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory."



## Organizace webu

- "není třeba si nic pamatovat, všechno je on-line"
- second brain / druhý mozek
- "recollection in the age of forgeting"
- Knowledge Management Systems

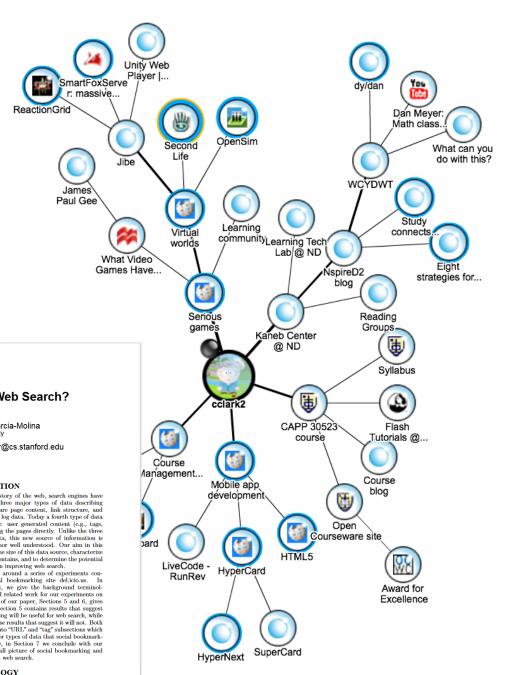
Záložkujete weby chaoticky nebo si v tom držíte nějaký systém?



### Záložkování

- social bookmarking
- social tagging
- folksonomie

NASA Webtagger Del.icio.us (2002)



#### Can Social Bookmarking Improve Web Search?

Paul Heymann, Georgia Koutrika, and Hector Garcia-Molina Dept. of Computer Science, Stanford University Stanford, CA, USA heymann@stanford.edu, koutrika@stanford.edu, hector@cs.stanford.edu

#### ABSTRACT

Social bookmarking is a recent phenomenon which has the potential to give us a great deal of data about pages on the web. One major question is whether that data can be used to augment systems like web search. To answer this question, over the past year we have gathered what we believe to be the largest dataset from a social bookmarking site yet analyzed by academic researchers. Our dataset represents about forty million bookmarks from the social bookmarking site del.icio.us. We contribute a characterization of posts to del.icio.us: how many bookmarks exist (about 115 million). how fast is it growing, and how active are the URLs being posted about (quite active). We also contribute a characterization of tags used by bookmarkers. We found that certain tags tend to gravitate towards certain domains, and vice versa. We also found that tags occur in over 50 percent of the pages that they annotate, and in only 20 percent of cases do they not occur in the page text, backlink page text, or forward link page text of the pages they annotate. We conclude that social bookmarking can provide search data not currently provided by other sources, though it may currently lack the size and distribution of tags necessary to make a significant impact.

#### Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval; H.3.5 [Information Storage and Retrieval: On-line Information Services; H.1.2 [Models and Principles]: User/Machine Systems—Human infor-

#### 1. INTRODUCTION

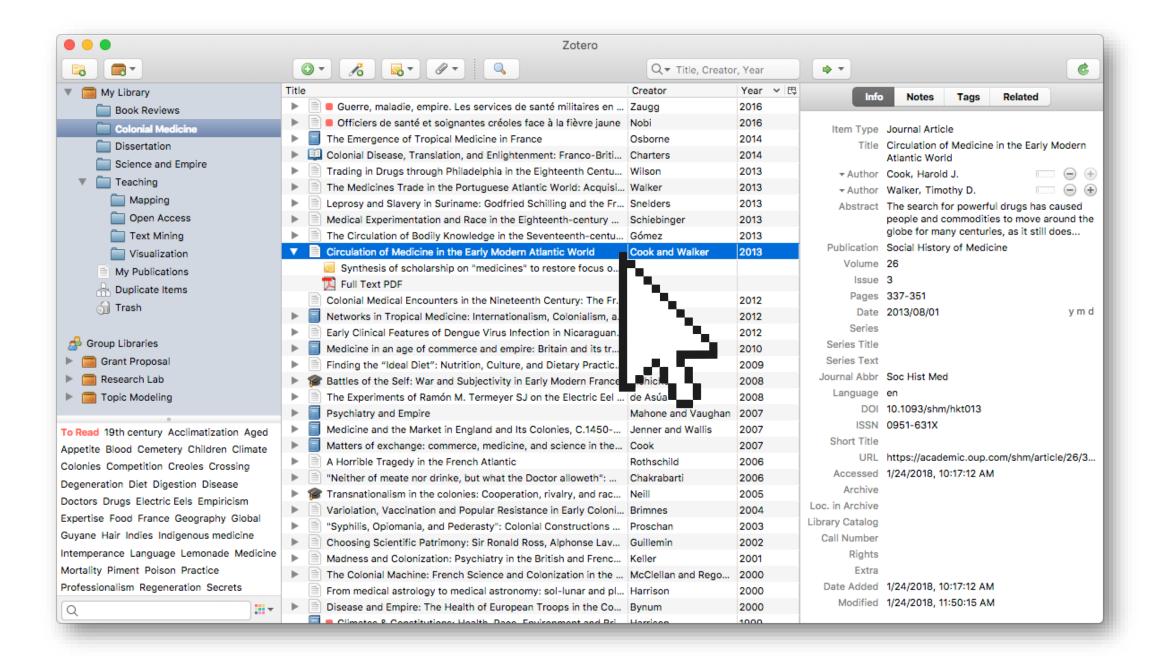
For most of the history of the web, search engines have only had access to three major types of data describing pages. These types are page content, link structure, and query or clickthrough log data. Today a fourth type of data is becoming available: user generated content (e.g., tags, bookmarks) describing the pages directly. Unlike the three previous types of data, this new source of information is neither well studied nor well understood. Our aim in this paper is to quantify the size of this data source, characterize what information it contains, and to determine the potential impact it may have on improving web search.

This paper centers around a series of experiments conducted on the social bookmarking site del.icio.us. In Sections 2, 3, and 4, we give the background terminology, methodology and related work for our experiments on del.icio.us. The core of our paper, Sections 5 and 6, gives two sets of results. Section 5 contains results that suggest that social bookmarking will be useful for web search, while Section 6 contains those results that suggest it will not. Both sections are divided into "URL" and "tag" subsections which focus on the two major types of data that social bookmarking provides. Finally, in Section 7 we conclude with our thoughts on the overall picture of social bookmarking and its ability to augment web search.

We differentiate social bookmarking from other social sites involving shared bookmarks, like social news sites. The two major social bookmarking sites are del.icio.us and Stumble

### Záložkování

- záložkovací nástroj <u>Memex</u>
- sociální referenční nástroj Hypothes.is
- historie v prohlížeči <u>Vivaldi</u>
- archivace webu s **Conifer**
- referenční nástroje <u>Zotero</u> a <u>BibSonomy</u>





Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction

Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools

What links here Related changes Upload file Article Talk Read Edit View history Search Wikipedia Q

### Memex

From Wikipedia, the free encyclopedia

For the software company, see Memex Technology Limited.

The **memex** (originally coined "at random",<sup>[1]</sup> though sometimes said to be a portmanteau of "memory" and "index"<sup>[2]</sup>) is the name of the hypothetical protohypertext system that Vannevar Bush described in his 1945 *The Atlantic Monthly* article "As We May Think". Bush envisioned the memex as a device in which individuals would compress and store all of their books, records, and communications, "me hanized so that it may be consulted with exceeding speed and flexibility". The money are memory. The money are memory and provide an "enlarged intimate supplement to one's memory". The money influenced the development of early hypertext systems (evaluated by a specific provided by and personal knowledge by the provided by the creation of the World Wide Web) and personal knowledge by the provided by the provi

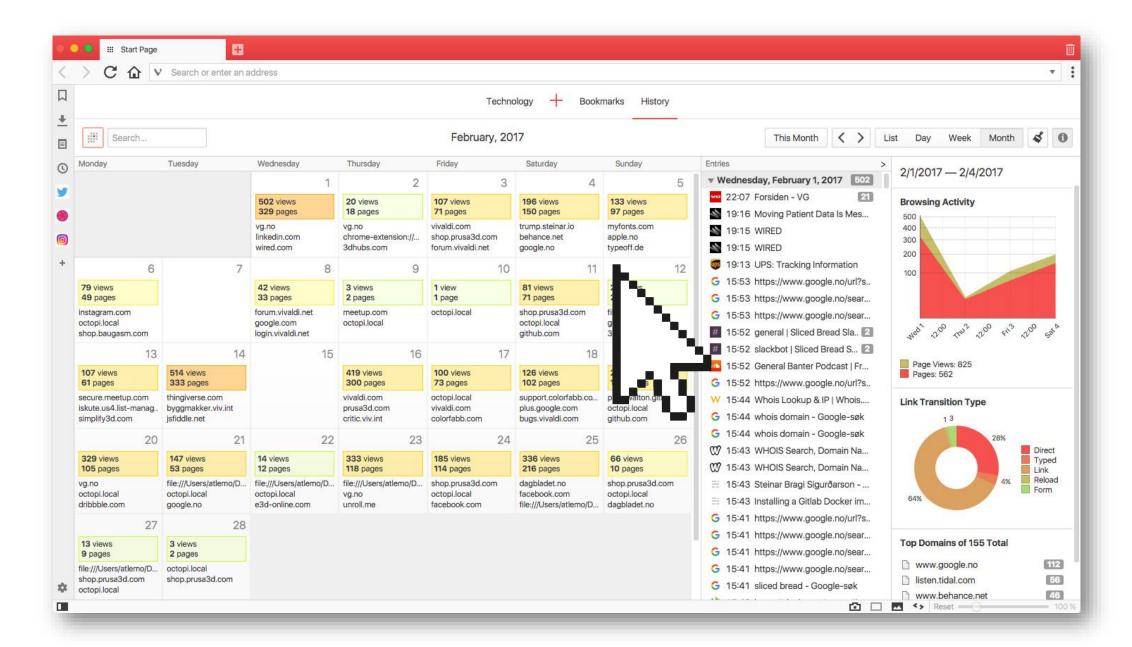


Vannevar Bush

Bush for the purpose Leone Lillustration was based upon a document

bookmark list of static microfile pages and lacked a true hypertext system, where parts of pages would have internal structure beyond the common textual format. Early electronic hypertext systems were thus inspired by memex rather than modeled directly upon it.

Contento [hide]







### Conifer

### Collect and revisit web pages.

Conifer is a web archiving service that creates an interactive copy of any web page that you browse, including content revealed by your interactions such as playing video and audio, scrolling, clicking buttons, and so forth.



# Jak si děláte poznámky?



## Hypertextové poznámky

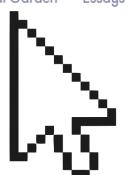
- Zettelkasten, wiki, Markdown atp.
- -> jiné přístupy k *rozpomínání* na webu

"The Garden is the web as topology. The web as space. It's the integrative web, the iterative web, the web as an arrangement and rearrangement of things to one another. Things in the Garden don't collapse to a single set of relations or canonical sequence, and that's part of what we mean when we say "the web as topology" or the "web as space". Every walk through the garden creates new paths, new meanings, and when we add things to the garden we add them in a way that allows many future, unpredicted relationships" – Mike Caulfield

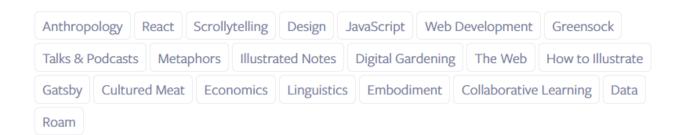
#### M

Digital Garden Essays Illustration Bookshelf About

## The Digital Garden



An open collection of notes, resources, sketches, and explorations I'm currently cultivating. Some notes are Seedlings, some are budding, and some are fully grown Evergreen.



Seedling Budding Evergreen

A Brief Introduction to Digital Anthropology



GreenSock Animations with React Hooks

ß

Drawing the Invisible: React Explained in Five Visual Metaphors The Bare Essentials of Greensock

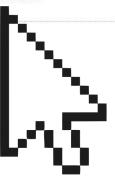
Douglas, Dirt, and

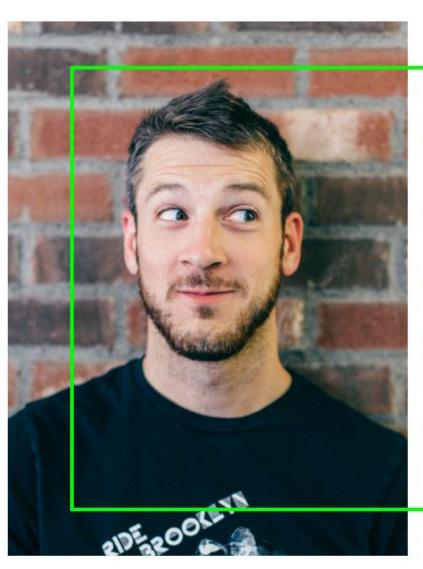
A Brief History &

Why You Own an

Illustrating Gatsby's

### Tom Critchlow





### Hello there!

I do consulting work for startups & media companies. Read more about that here.

I'm passionate about art & technology and am co-founder of Brooklyn art collective Fiercely Curious.

I also know a little kung fu (but I'm rusty)

Learn more about my background here



SITE ME CHANGES

† NEWS REDDIT

THIS IS THE WEBSITE of Gwern Branwen. I write about psychology, statistics, and technology. I am best known for work on the darknet markets & Bitcoin, blinded self-experiments, dual n-back & spaced repetition, and anime neural networks.

For information about my site's philosophy, method, traffic statistics, and implementation, see the About page; for information about myself, my use of other websites, and contact information, see the Links page; for information about new pages, see the Changelog; to receive updates, news, & reviews, subscribe = to the newsletter (archives).

#### NEWEST

- August 2020 newsletter
- Sidenotes In Web Design
- July 2020 newsletter
- ♦ GPT-3 & The Scaling Hypothesis
- June 2020 newsletter
- GPT-3 Creative Writing
- "PALM Anime Hands Model/Dataset"
- "Holy Wars are Coordination Problems"
- May 2020 newsletter
- Prototype anime GAN
- Danbooru2019 Figures dataset
- April 2020 newsletter
- ♦ Generating MIDI Music With GPT-2

#### **POPULAR**

- ♦ Silk Road 1: Theory & Practice
- DNM Archives (2013-2015)
- DNM Arrests (2011-2015)
- Modafinil
- LSD microdosing self-experiment
- Zeo sleep self-experiments
- Dual N-Back FAQ
- Spaced repetition
- Death Note: L, Anonymity & Entropy
- Commoditize Your Complement
- Google survival analysis
- Banner Ads Considered Harmful
- This Waifu Does Not Exist (details)
- Neural Net/Tank Urban Legend
- Time-lock cryptography
- Archiving URLs
- Terrorism is not about Terror
- Melancholy of Subculture Society



#### NOTABLE

- ♦ Why Correlation ≠ Causation
- When Should I Check The Mail?
- Why Tool AIs Want to Be Agent AIs
- Complexity no Bar to AI
- Making Anime Faces w/Neural Nets
- Death Note script authenticity
- The Existential Risk of Math Errors
- Iodine & Adult IQ meta-analysis
- The Melancholy of Kyon
- In Defense of Wikipedia Inclusionism
- Bitcoin is Worse is Better
- Are Sunk Costs Fallacies?
- Book reviews



#### **STATISTICS**

- ♦ Why Correlation Usually ≠ Causation
- When Does The Mail Come? Decision Analysis
- Predicting Google service closures

### **CRYPTO/BITCOIN**

- Time-lock encryption
- Bitcoin is Worse is Better
- A Silk Road 1 Guide

- Danbooru2019 Anime Dataset
  - PALM: Hand Detection/Crops
- Figures

 $\mathbf{AI}$ 

### **GARDENS** private classic notes blog posts 0 cozy web B books & groups research twitter papers 6 LEARNING IN PUBLIC

DIGITAL

CHAOS STREAMS CULTIVATED PERFORMANCE

## Digitální zahrady

- předchůdci: osobní wiki
- např. <u>Tiddlywiki</u>
- bliki blog a wiki (cca 2003)
- vlastní řešení
- specifické platformy: <a href="https://www.are.na/">https://www.are.na/</a>
- Digital Gardeners
- Awesome Knowledge Management



Welcome! You are now browsing a Digital Garden. This is my personal space for **Learning In Public**. I am a lifelong learner so everything is a Work-In-Progress like me, but I do not let **perfectionism** get in the way. That means that what you read here is not authoritative or complete, and is not representative of my best work. However, it is representative of my interests and current state of knowledge, and if you have the same interests, then this space is also yours to use as a reference. Feedback and social sharing is welcome...

I have a **right to be wrong or incomplete** in my Digital Garden, either due to paucity of time or knowledge. You will not hold this, or my readership, against me because I will keep learning, **with your help**. Everything in the Digital Garden is a **living document** and I will **retract or rephrase** things I no longer agree with.

I will report how strongly I hold my beliefs, always reserving the **right to be wrong** and change my mind.

I will report how much experience I have in the topic, by disclosing how much work I have done so far on it and linking to others who will know more.

I will link to **further resources** so that readers can discover influencing and contrasting opinions from the original source.

I will not get discouraged if I don't receive feedback. I plant ideas in my Digital Garden for my own use, not solely to get visitors.

I will reward feedback by listening and immediately correcting things I got wrong.

I do not promise to agree with or respond to all feedback.

