

Enabling Social Information Exchange via Dynamically Robust Annotations

Vishwajeet Pattanaik*

Shweta Suran

Dirk Draheim

Tallinn University of Technology

Estonia

Motivation

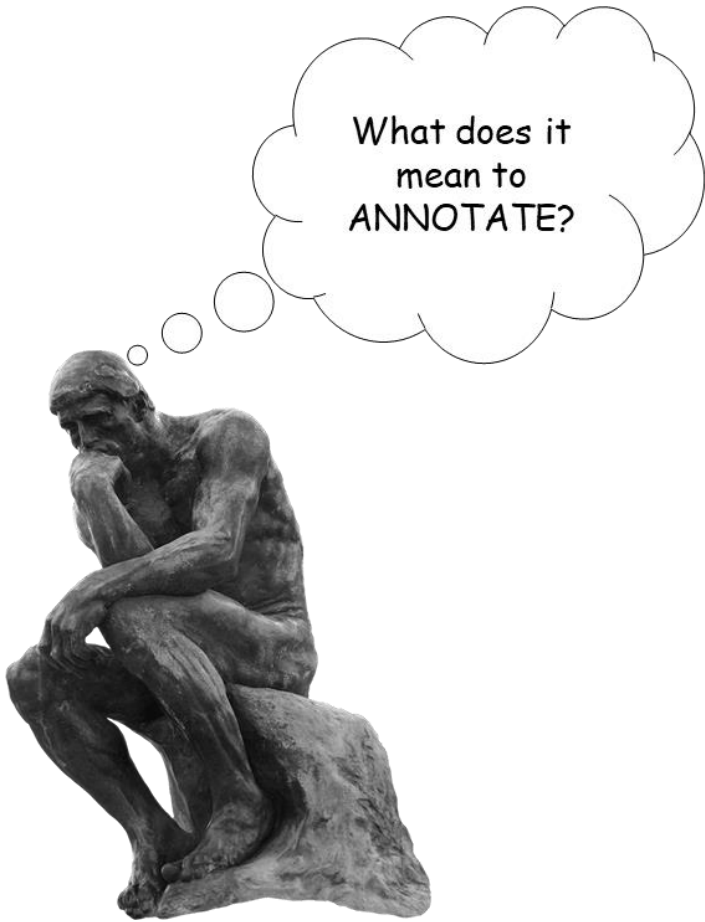
- “The Web is starting to wane in the face of a ‘nasty storm’ of issues”
– Tim Berners-Lee ([2017](#))
 - political ads
 - behavioural micro-targeting i.e., “weaponised AI propaganda” ([10.1353/jod.2017.0025](#))
 - clickbait ([10.1007/978-3-319-63751-8](#))
 - fake news ([10.1126/science.aao2998](#))
 - filter bubble (Pariser, [2011](#))
 - link rot i.e., web page decay ([10.1007/s00799-016-0171-9](#))

Dokieli
MicroBlogging
Mobile
Social Linked Data
Social App Editor Summarization
Network Integrated
MakingSense
Tagging Client
Client-Side
SOLID
Collaborative
Musubi

How can we solve the **nasty storm** of issues with Web, using the **wisdom of the crowd**?

...while not relying on developers and content providers...

Background



What does it mean to ANNOTATE?

- An annotation is a note added to a book, drawing or any other kind of text as a **comment** or explanation. ([2015](#))
- Web Annotations have emerged as a **First-Class Object** ([10.1109/MIC.2013.123](#))
- Web annotation tools like: Hypothes.is are gaining tremendous interest. ([10.1038/528153a](#), [10.1038/d41586-019-01427-9](#))

1. Before Breakfast *alliteration*

HERE'S Papa going with that ax?" said Fern to her mother as they were setting the table for breakfast. "Out to the hoghouse," replied Mrs. Arable "Some pigs were born last night." "I don't see why he needs an ax," continued Fern, who was only eight. *plowable* *nature*

"Well," said her mother, "one of the pigs is a runt. It's very small and weak, and it will never amount to anything. So your father has decided to do away with it." *death*

"Do *away* with it?" shrieked Fern. "You mean *kill* it? Just because it's smaller than the others?" *kill*

Mrs. Arable *put a pitcher of cream* on the table. "Don't yell, Fern!" she said. "Your father is right. The pig would probably die anyway." *plant*

Fern pushed a chair out of the way and ran outdoors. The grass was wet and the earth smelled of springtime. Fern's sneakers were sopping by the time she caught up with her father. *goes writing eating in a farm*

Spring pig

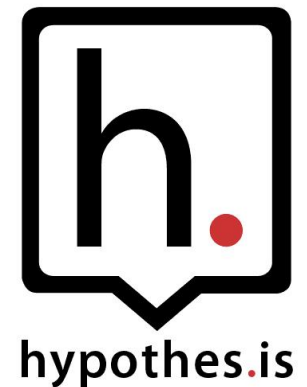
Sense

Related Work – Hypothes.is

- non-profit, started in 2011
- open-source software
- Goal: “to enable a conversation over the world’s knowledge”

It’s 215,000 users have added 5 million comments on scholarly sites

- Rina Shaikh-Lesko ([10.1038/d41586-019-01427-9](https://doi.org/10.1038/d41586-019-01427-9))



Before Hypothes.is'

- XPath (e.g. `/html/body/div[3]/div[3]/div[4]/div/p[2]/b[3]`) Matching

The **World Wide Web (WWW)**, also called **the Web**, is an information space where documents and other web resources are identified by Uniform Resource Locators (URLs), interlinked by hypertext links, and accessible via the Internet.^[1] English scientist Tim Berners-Lee invented the World Wide Web in 1989. He wrote the first web browser in 1990 while employed at CERN near Geneva, Switzerland.^{[2][3]} The browser was released outside CERN in 1991, first to other research institutions starting in January 1991 and to the general public on the Internet in

A global map of the web index for countries in

```

<b>World Wide Web</b>
(
<b>WWW</b>
), also called
<b>the Web</b>
, is an
<a href="/wiki/Information_space" title="Information space">information space</a>
where documents and other
<a href="/wiki/Web_resource" title="Web resource">web resources</a>

```

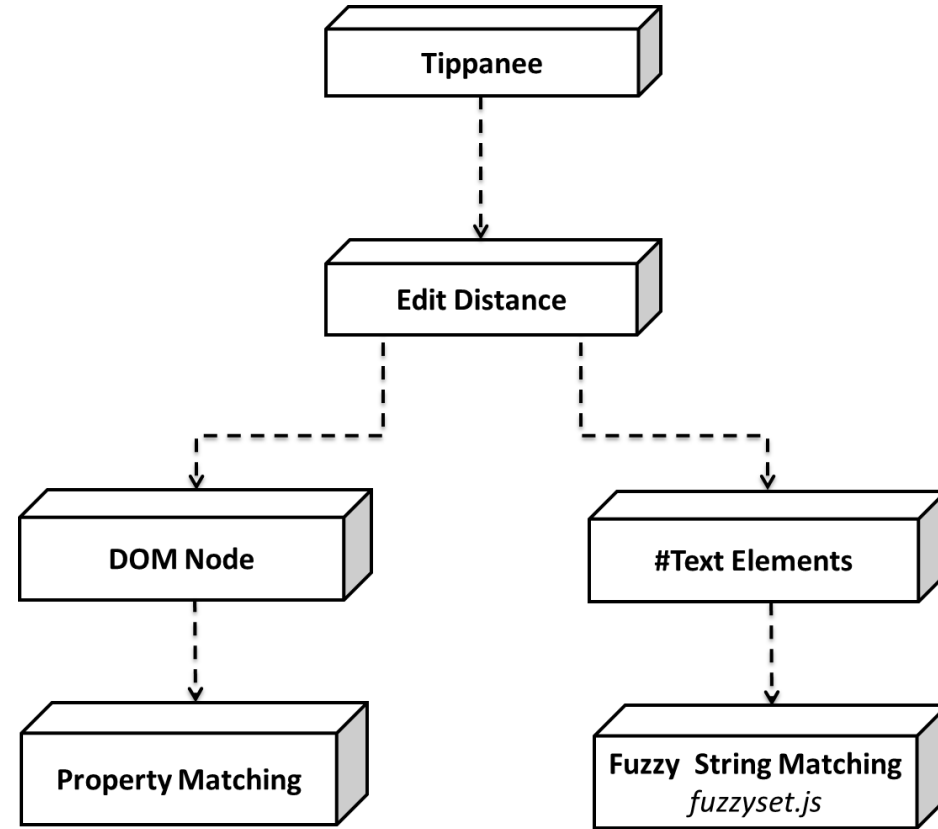
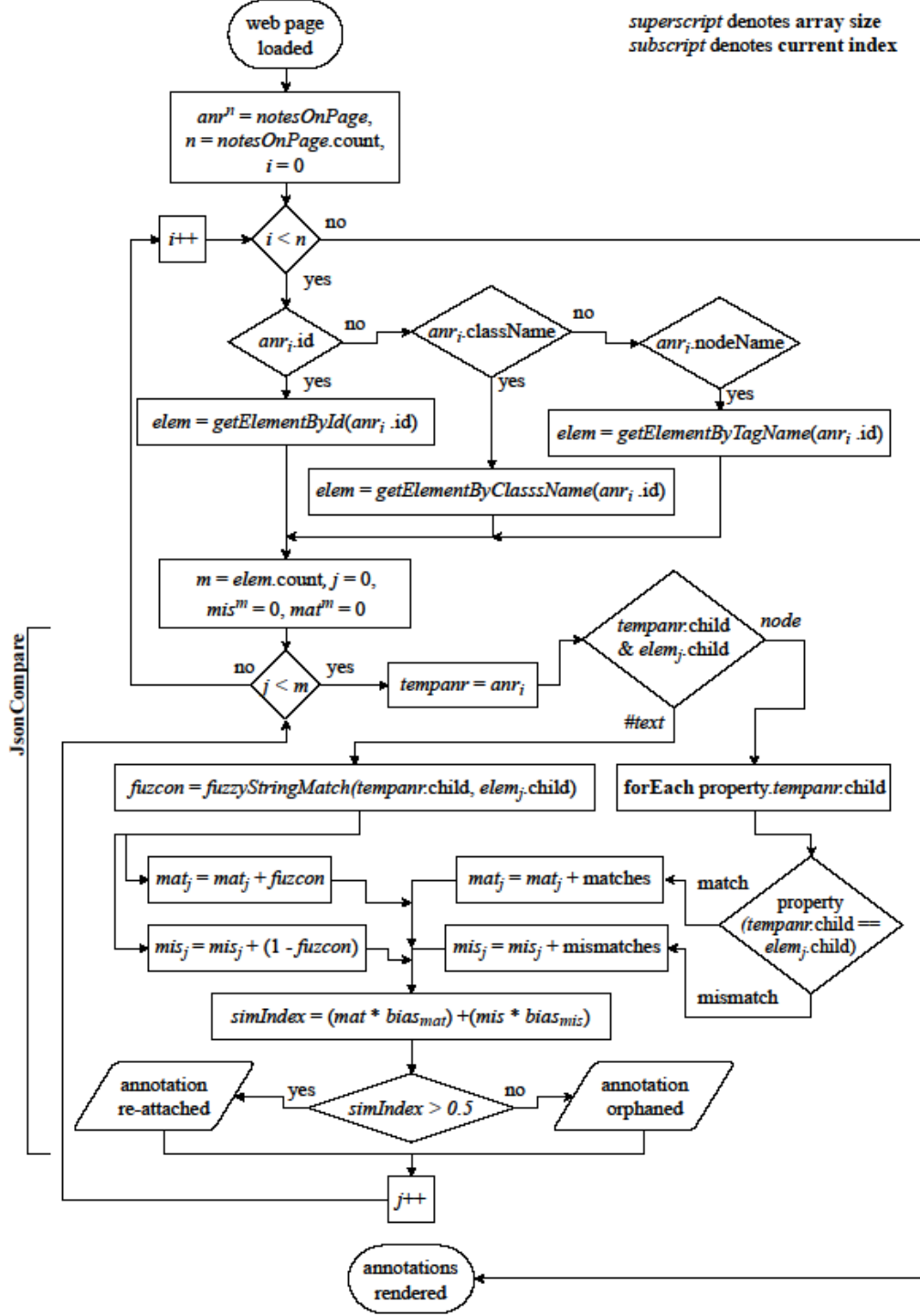
Hypothes.is' Fuzzy Anchoring

- Robustly anchoring annotations using keywords (Brush et al. [2001](#))
- Robust anchoring of annotations to content (Brush et al. [2010](#))
 - String-matching algorithms:
 - Bitap matching ([1992](#))
 - Myers diff ([1986](#))
 - Levenshtein distance ([1965](#))
- **27%** Hypothes.is annotations already orphaned
- ...and **61%** at risk of being orphaned soon ([10.1007/978-3-319-24592-8_2](#))

Objectives

- Design and evaluate an artefact
 - web-based Crowdsourcing Information System (CIS)
 - Interoperable
 - Support *activities* on-the-fly
 - Act as conversation layer over the Web
 - Provide a stable and robust approach for tracking textual contextual
 - Provide a social environment that promotes *co-creation (contribution & selection)*
 - Based on the principles for *Collective Intelligence*

Tippane to the rescue!



Tippane Chrome Extension

... as well as demonstration of innovative syst...
... sessions at the CAiSE conference will f...
discussion, and exchange of ideas among presenters and participar...
CAiSE'18 Forum are welcome to address any of the CAiSE'18 confer...
particular the theme of this year's conference: *Information Systems i*

| Important Dates | |
|-----------------------------|---------------|
| Paper submission deadline: | 4 March 2018 |
| Notification of acceptance: | 6 April 2018 |
| Camera-ready deadline: | 13 April 2018 |

Notes

April 13, 2018 3:30 PM
Paper submission deadline: 4 March 2018
Strict Deadline!!!
Apr 13, 2018 3:35 PM

Reconstruct Anchor

Link Annotations

Add semantic description









Hypothes.is vs Tippanee

Welcome to Wikipedia
the free encyclopedia that anyone can
5,745,710 articles in English
from today's featured article


victor_013 Sep 13
Public
5,710,618
Wikipedia article count (numbers only)
Show replies (1)

Hypothes.is vs Tippanee

Main Page **Talk** > **Public**      





Welcome to Wikipedia,
the free encyclopedia that anyone can edit.
5,977,069 articles in English

From today's featured article

 **Chesapeake** was a 38-gun

Annotations ² **Page Notes** **Orphans** ⁴

victor_013 *(edited Jan 28) Sep 4, 2018*
Public
~~5,710,618~~
Wikipedia article count (numbers only)

Main Page

Talk

Read

View source

Welcome to **Wikipedia**,

the free encyclopedia that anyone can edit.

5,977,069 articles in English

- Arts
- Biogra
- Geogr

From today's featured article



Chesapeake was a 38-gun wooden-hulled, three-masted heavy frigate of the United States Navy. She was one of the original six



October 29, 2018 4:33 PM

5,742,333 articles in English

abc@abc.com

Wiki

Oct 30, 2018 9:32 AM

Hi there

Oct 30, 2018 9:49 AM



Add a note...



Similarity Index: 0.75

LSS



Tippaneé's Features i.e., Contributions

- Novel anchoring approach
 - stable and robust
 - works both online-offline
- End-user oriented features
 - personalized archival of textual content
 - social knowledge management
 - Linking and visualizing annotated content (*i.e.*, [knowledge graph](#))
 - enriching web content with semantic metadata
 - allows for creation of new semantic vocabularies* work in progress

Preliminary Evaluation

- Experiment 1:
 - replicated *735 (Hypothes.is) annotations from more 650 different websites*
 - observed annotations over 3 months (*expecting some web page decay*)
 - *91.41%* annotations were successfully attached
 - *12.41%* over Hypothes.is' expected success ([79%](#))
- Experiment 2:
 - presented the tool to 25 candidates
 - found the tool useful and easy to use
 - users preferred the tool for social interactions, expression of opinion and information sharing
 - helped identify bugs and suggested additional UI features

Ongoing Work

- Anchoring approach test bench:
 - 50 websites
 - 120 webpages
 - 9 annotations per page
 - 96 variants per annotation
 - *103,680 data points for evaluation*
- Support knowledge creation & exchange in organizations
[using Nonaka's SECI model]
- Develop a generic model for Collective Intelligence
[recently accepted @ ACM CSUR]
- Develop a novel user reputation model *[less prone to bias]*

“I think people’s fear of bad things happening on the internet is becoming, justifiably, greater and greater...”

“...If we leave the web as it is, there’s a very large number of things that will go wrong. We could end up with a digital dystopia if we don’t turn things around. It’s not that we need a 10-year plan for the web, we need to turn the web around now.”

- Tim Berners-Lee, the inventor of the web

[@ launch of “Contract for the Web”](#) 24th November’2019

Questions?

and

Feedback!