

I Annotate - Annoto Ergo Sum - Agenda
San Francisco, CA
April 23-24, 2015

LINK TO UPLOAD PICTURES [\[HERE\]](#)

Lehre von der Wahrnehmung bzw. vom sinnlichen Anschauen.



I Annotate 2015 on Ponga

A collection of Ponga pictures has been created from all pictures collected and shared. For write access to the pictures, see Barbara Tien (barbara@ponga.com), or Peg Fowler (pfowler@hypothes.is).

PONGA.COM

[Pictures](#) added into an IANNO 15 account on Ponga, sorted and accessible, [full list of pictures and links to Ponga annotated versions here](#).

Twitter hashtag [\[#ianno15\]](#) / **twitter account** [\[@i_annotate\]](#)

IRC [\[irc.freenode.net / #ianno15\]](#)

[Etherpad](#)

Abstracts [\[abstracts\]](#)

Open Sessions [\[ideas\]](#)

Conference hotel is the [Comfort Inn By the Bay](#) / 2775 Van Ness Ave @ Lombard, SF [\[map\]](#)

VENUE: Fort Mason Center / Golden Gate Room - all days

[\[map\]](#) The Golden Gate Room is in the building that the circle is on]

Attendee list:

<https://docs.google.com/spreadsheets/d/17vtrYyb9LLm9nkFZbiFBfDN-wfm9CEI-WWS7yaSBkpE/edit#gid=0>

AGENDA

Wed, April 22- Day 0

W3C Web Annotation Working Group Face to Face Meeting

[Minutes: <http://www.w3.org/2015/04/22-annotation-minutes.html> **]**

Challenges and Solutions in Scholarly Information Management: The changing face of research and the resources evolving to support it.

This is a free all day event run by [Digital Science](#).

Food and drinks included.

Lunch to be catered by [Yukol Thai](#).

To register [click here](#)

All I Annotate attendees

7:00pm Gather the night before! Cocktails at the [Brixton](#) [\[map\]](#)

Thu, April 23 - Day 1

ANNOTATION PLATFORMS: OVERVIEW AND STATUS

7:30 Coffee, tea and light breakfast.

8:15 **Opening remarks** [\[https://beta.etherpad.org/p/ianno15\]](https://beta.etherpad.org/p/ianno15) & notes

8:30 **Morning activities**

MISC Links:

- <http://annotatorjs.org/>
- <https://hypothes.is>
- <http://crowdtruth.org>
- <http://tagasauris.com>
- <https://www.historypin.org/>
- <http://jonreeve.com/projects/annotags/>
- <http://disciplineoforganizing.org/>
- <http://hyperstudio.mit.edu/>
- <https://www.wordnik.com/>
- <http://www.sketchengine.co.uk/>
- <https://www.diffbot.com/>
- <http://restory.swtr.us>
- <http://ponga.com>
- <http://mitan.in/bcp/raika>
- <http://factlink.com>
- <https://robot-reviewer.vortext.systems> / <https://annotate.vortext.systems/>

A definition of annotation would be helpful: annotation is done differently and has a different role in audiovisual (sound / film / moving image) archives than in other kinds of archives for example.

Morning Session 1: statements

- All ebooks should be browser-based
 - If it's not online, it doesn't exist
- There's no way to make money on annotation
 - Advertisement is like a bacteria and will always mutate
 - Engage users
 - Annotation brings added value to content
- Anonymity is the barrier to annotation at scale
 - Why do we need to know the identity of annotators
 - Comments in books are anonymous and scale
 - There's no such as anonymity
 - There is information overload
 - People don't take responsibility for their comments
- Annotations disrupt the experiences on some types of content
 - Annotations themselves don't bother it is the way they are visualized
 - Spoiler problem
 - Pop-up video

- You can add annotation identifying an anonymous author
- Experts are better in annotating
 - All the shared knowledge is destroyed by experts
 - Experts are rigid and not open minded
 - Experts assume single truth and single interpretation
- Copyright clarity is essential for annotations
 - If I don't know how to use it I can't use it
 - Publishers
 - For realtime analysis
 - Copyrights are ambiguous
 - There is no clarity in copyright
 - If you can't talk about things because they copyrighted
- Annotation should not be done with a browser plugins
 - Plugins should always be an option
 - Plugins should not be a requirement
 - To be useful annotations should be done natively in the browser
- Collective intelligence is stupid
 - Problems decomposed in small problems
 - Individuals don't have always the right answer but together yes
- Mobile annotation sucks
- Annotation must be decentralized
- Annotation will replace journalism
- People prefer to share annotations privately
- We are not ready for large scale annotations

Morning Session 2: taking annotations at scale

- what does annotation at scale means, e.g. in technology, standards, society terms
- what does it take to take the annotations at scale, e.g. roadmap, ecosystem

12:30 LUNCH: FREE Tacos by [La Tonayense](#) -- via their truck in the patio outside.
(1 truck with 3 servers means that 100 people will get their tacos < 40 minutes)

1:30 Speedgeeking

Wordnik (Erin) <https://www.wordnik.com/>

Papero (Jacob)

Ponga (Barbara) - picture annotation <http://ponga.com>

Checkdesk (Chris)

Lacuna stories (Michael) lacunastories.com

Historypin (Jon) <http://historypin.org> & [Panama Pacific International Exposition project](#)

Sweet web (Dinesh) <http://janastu.org> @ <http://wiki.janastu.org/AnnotationUseCase>

MX (Mark)

Annotags (Jonathan)

2048 Books (Bob) <http://people.ischool.berkeley.edu/~glushko/IAnnotate-GlushkoFinal.pdf>

Patents, contracts, agreements (Sumeet)

? (Drew)

Hypothes.is (Jake, Jeremy) <https://hypothes.is>

Annotation visualization (Andy Stuhl | HyperStudio) - <http://web.mit.edu/akstuhl/www/annotation-vis/> - <http://hyperstudio.mit.edu/>

SocioNet, an example of annotated relationships, etc. -

<https://socio.net.ru/publication.xml?h=repec:rus:mqijxk:31&l=en> (Sergey)

Check out also: <http://crowdtruth.org> & <http://tagasauris.com> for crowdsourcing annotations

3:00 **Working sessions (PLEASE ADD LINKS TO PADS OR UPLOADS PICTURES)**

[Attribution and anonymity](#) (Laura)

Federation for annotation (Dinesh, Benji)

Attribution in education (Phil)

[Ecosystem of interoperability](#) (Jon)

[What is annotation](#) (TS)

Deliberative reasoning (Che): [hackpad](#)

Semantics (Rob)

Future of AnnotatorJS (Michael) - Pictures uploaded with prefix "AnnotatorJS"

<https://drive.google.com/file/d/0B9QoL7Nb-di1YlpXMEFmUnlcnM/view?usp=sharing>

LINK TO UPLOAD PICTURES OF SESSION OUTPUT [[HERE](#)]

5:30 **Close**

6-9pm Cocktails at [East Side West](#) - 3154 Fillmore St @ Greenwich

Fri, April 24 - Day 2

COMMUNITIES: IMPLEMENTING AND USING ANNOTATION

7:30 Coffee/tea and light breakfast

8:15 **Opening remarks** [pad]

8:30 **Agenda TBD**

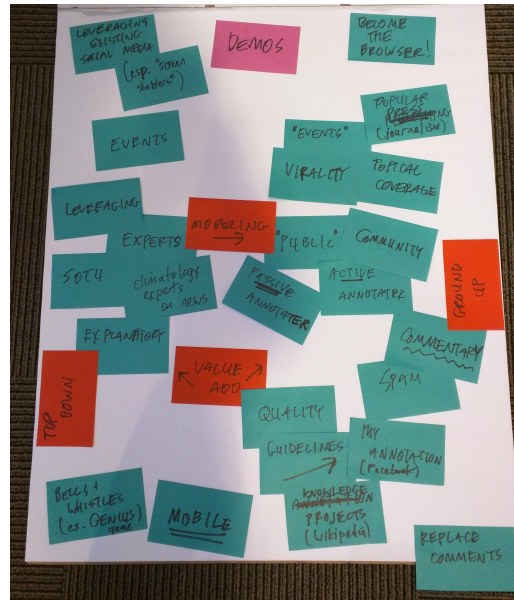
Morning Session 1: Inspiration & Motivation for Annotations

- Tim Berner-Lee (web is annotation)
- Douglas Engelbart
- Fermat
-

Morning Session 2:

Working Group1: How to Publically Demonstrate Annotations

- <http://www.politifact.com/>
- <http://www.factcheck.org/>
- <http://genius.com/>
- <http://fiskkit.com/>
- <http://climatefeedback.org/>
- <http://thestore.swtr.in> like annotation portability to feed into social networks



- How do you demo 'annotations'
 - Annotation events - work on annotation as performance - with expert annotations, using their tools
 - Twitter as an annotation tool (fiskkit)
 - Using famous people to show what did they annotated (genius)
 - Climate experts go to websites and annotate
 - Show people what's in it for them - show how annotations help them (getting it across immediately, and achieving a viral effect)
 - you can annotate small pieces, and not necessary the whole document - low-requirements for time investment
 - mobile
 - we can get the value of annotation - larger than twitter - through building a community
 - being able to toggle back and forwards between the annotations by me - and the rest
 - education
 - active outreach, like hypothes.is does
 - show people that annotation is a better version of what they're already doing. show them what's in it for them. make sure you engage them.

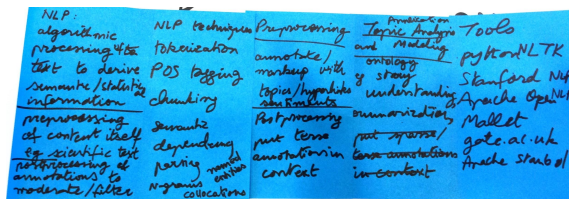
Working Group 2: User privacy and rights

- [Poster board image #1](#)
- [Poster board image #2](#)
- [Poster board image #3](#)

Working Group 3:

[Natural Language Processing](#) Notes

[NLP working group photos](#) , summary photo:



NLP afternoon working group notes:

- Discussed more tools eg LDA/LSA analysis - maps vectors of words into abstract math spaces and classes, but it is hard to ensure semantically meaningful classes
- One specific application example - how to automatically identify related press releases about a given topic, and persist user-curated comments/annotations into new press releases for continuity
- Question on maturity of NLP tools in NodeJS - NLTK and Apache OpenNLP are still much better
- Design comment - hard to design annotation questions well, to extract meaningful responses. Need input from mathematical psychologists.
- good beginner tutorial : <http://www.nltk.org/book/>

Overall NLP Working Group Assessment: group discussed this for the first time, and is in learning phase. Next steps are to come up with :

- Use cases for NLP in annotation
 - At input (preprocessing of content) or output (postprocessing of annotations)
 - Offline, real time or near real time
 - Application verticals (social, news/publishing, government, education, science, ecommerce, enterprise, ...)
 - Task categories (topic classification, clustering, named entity extraction, search, summarization, ...)
- Potential implementations in annotator
 - what should be 'integrated' with annotator, which version of annotator, what APIs, what NLP options (training sets, testing sets, dictionaries, NE datasets, languages, ...)
- Interested leaders and collaborators (please enter your names/email addresses if absent, or remove if not interested, or remove '?' if interested):
 - TS tswaterman@gmail.com
 - Phil Desenne pdesenne@gmail.com
 - Jamie Folsom jamiefolsom@gmail.com
 - Robert Casties r0bcas7@gmail.com
 - Chris Chapman chris@pentandra.com
 - Laurian Gridinoc laurian@gmail.com
 - Sarah Luger yo@sarahluger.com
 - Jonathan Reeve jon.reeve@gmail.com
 - Ray Denenberg rdn@loc.gov
 - Barbara Tien barbara@ponga.com ?
 - Christian Kleinfelchner ?
 - Emily Schneider elfs13@stanford.edu ?
 - Alex Kosorukoff allkos@yahoo.com ?
 - Kristof Csillag csillag@hypothes.is ?
 - Gergely Ujvari ujvari@hypothes.is ?
 - Erin McKean erin@wordnik.com
 - Sumeet Sandhu sumeet.k.sandhu@gmail.com (WG wildcat-herder)

Data Sources Group

Nick (leader), Frederick (Notes), Dan, James , (please add)

- **Goal:** Identify existing sources that can be used to pre-populate an annotation system to create a compelling demonstration for users as well as enabling development, testing and other uses.

- Group enumerated possibilities (see [Data Sources Poster](#)) and categorized on two axes : value and ease of re-use.

Value: quality of data, quantity of data, and interest

Ease of Reuse: technical ease of access, such as having existing APIs etc, legal ease, e.g. compatibility of copyright etc, business ease of relationship etc.

- **Result**: Many sources identified, there is more than one might expect. Group agreed that highest value and ease would come from well-curated annotations that can be imported - the skeptics annotated bible was given as an example. Other sources like the IMDb for movies, wikipedia/dbpedia and maps were also promising high value sources with relative ease of re-use.

The group also discussed how data sources can be refined and categorized further, for example Wikipedia has both relationships and citations.

12:30 LUNCH: Catered by [Mehfil Indian Cuisine](#)

1:30 Afternoon Session 1:

Extracting Meaning (<https://www.wordnik.com/>):

- frames with recurring elements, e.g. 'X was elected', 'Y crash' : <https://framenet.icsi.berkeley.edu/fndrupal/>
- word disambiguation
- computer vs. human errors
- REVS digital Library: <https://revslib.stanford.edu/>
- tools:
 - Corpora-L <http://www.hit.uib.no/corpora/>
 - SketchEngine <http://www.sketchengine.co.uk/>
 - NLTK python <http://www.nltk.org/>
 - Standford NLP <http://nlp.stanford.edu/>
 - GATE: <https://gate.ac.uk/>
 - diffbot → returning text from URL, <https://www.diffbot.com/>
- weekly wikipedia - weekly changes and additions → <http://weekly.hatnote.com/>
- Macro-Etymological Analyzer: <http://jonreeve.com/etym> (uses etymological WordNet: <http://www1.icsi.berkeley.edu/~demelo/etymwn/>)
- psychology prototype databases: link????
 - <http://www.cs.utexas.edu/users/mfkb/related.html/> ??
- affective lexica (for sentiment analysis): http://neuro.imm.dtu.dk/wiki/A_new_ANEW_evaluation_of_a_word_list_for_sentiment_analysis_in_microblogs
- topic signatures
 - → <http://ixa.si.ehu.es/ixa/resources/sensecorpus>
 - → <http://www.cs.upc.edu/~nlp/papers/cuadros05.pdf>
- Multi-word expression toolkit: <http://mwetoolkit.sourceforge.net/>, also Babelfy.org: <http://babelfy.org/>

epub and Web Annotation gap review

- picture
- epub 3.1 will include the Web Annotation data model which is a major aspect of convergence for portability and interoperability
- Robust Anchoring important both to epub 3.1 and Web Annotation;

UI/UX Session: [some raw notes](#) there is a also a list of emails of people interested in going further on this

Publisher Session

[Poster Board Image 1](#)

[Poster Board Image 2](#)

Annotation Federation Infrastructure

Summary:

The most important parts we need are

- motivation (to federate)
- a protocol
- identities (for users/agents and resources/targets)

The next steps can be informed by existing solutions in other fields (see red cards in poster image).

[Poster image 1](#) green: necessary parts, red: existing solutions in other fields.

[Poster image 2](#)

[Poster image 3](#)

[Poster image 4](#) (from first federation session)

[Poster image 5](#)

[Poster image 6](#) (from first federation session)

ePub and Web Annotation Session

Frederick (moderate/notes), Tim Cole, Rob Sanderson, Takeshi Kanai

The group did a gap analysis of ePub 3.1 and Web Annotation to determine where additional effort may be required for convergence as well as reviewing the differences. This is outlined in the [ePub 3.1 and Web Annotation chart](#)

Summary:

- [ePub](#) 3.1 will include support for the [Web Annotation](#) data model, a major aspect of interoperability.
 - Adding an example to the Web Annotation Data model for multiple renditions in ePub would be useful
- Both ePub 3.1 and Web Annotation require work on Annotation Sets; this can be coordinated.
- Robust Anchoring is important for all annotations, however there are differences in the details for ePub 3.1 and Web Annotation
 - ePub can fallback to a page if an annotation cannot be linked; web annotations do not have pages.
 - A special ePub robust anchoring use case is the common case of a book being updated without having the id changed.
- ePub provides URLs to access annotations (CFI), web annotations use selectors. There may be need to coordinate this work.
- hasScope in Web Annotations needs clarification

- Offline use and synchronization are essential to ePub (and may become important to Web Annotation) but are currently out of scope.

5:30 **Close**

6-9pm Cocktails at [Stock In Trade](#) - 2036 Lombard St @ Fillmore

Sat, April 25 - Day 3

Annotation Hack Day - Hands on, cross-"genre" collaboration

[Register at Meetup.com](#)

[What do you want to work on?](#)

[Annotator at I Annotate 2014 Ideas](#)

8:30-6pm Open space / hackathon format

(Coffee & lunch will be provided - [Little Star Pizza](#))

>>> [More Detailed Schedule](#) <<<

6-9pm Congrats, you're still here!

BBQ at 3 Fort Mason, SF

<https://www.google.com/maps/place/3+Fort+Mason/>

Sun, April 26 - Day 4

Annotation Hack Day - Hands on, cross-"genre" collaboration

9-5pm Open space / hackathon format

(Coffee & lunch will be provided - [The Sandwich Spot](#))

6-9pm Drinks nearby.