

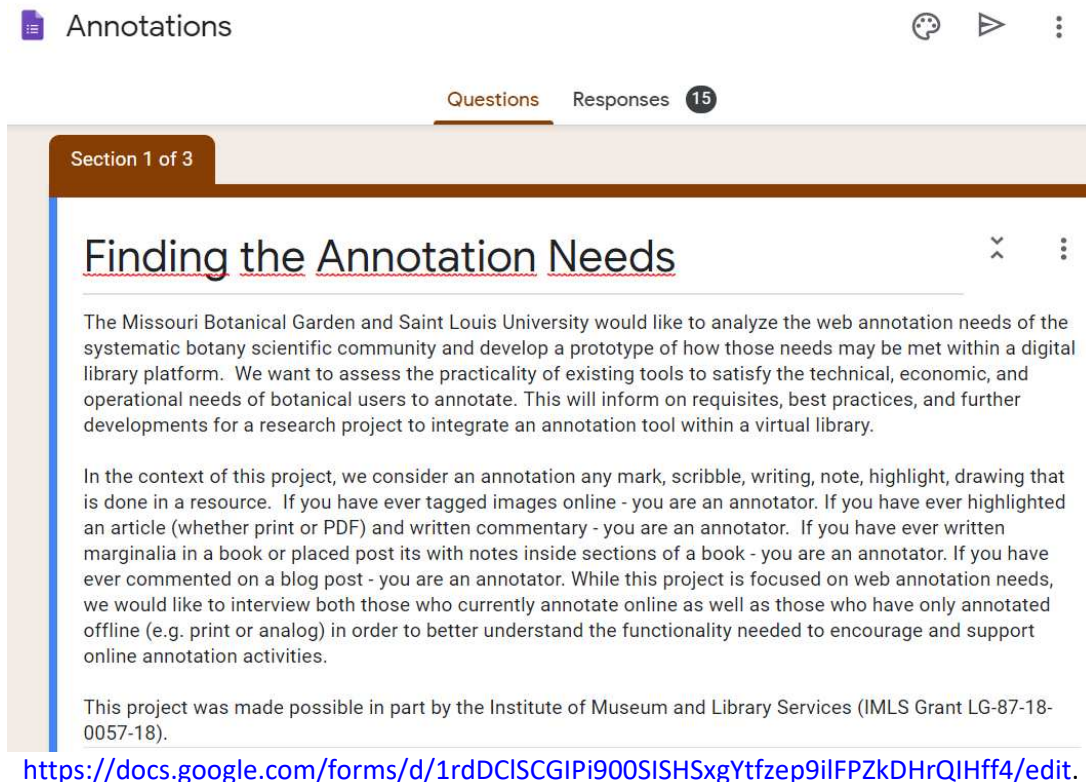
Appendix 3.

Assessment of Consumers and Creators Interviews

William Ulate and Trish Rose-Sandler

An interview was conducted with 14 individuals. The interviewed were asked to look at the questionnaire online and reply in advance if they wanted to. The interviewers will look at the answers before a personal (or online) interview so the interviewers would try to ask for more clarifications about the responses given and confirm by cross-referencing some of the answers provided.

The interview was done with a Google Form.



The screenshot shows a Google Form interface. At the top, there is a navigation bar with 'Annotations' on the left and icons for chat, play, and a menu on the right. Below this, there are tabs for 'Questions' and 'Responses' with a '15' indicator. The main content area is titled 'Section 1 of 3' and 'Finding the Annotation Needs'. The text of the form reads: 'The Missouri Botanical Garden and Saint Louis University would like to analyze the web annotation needs of the systematic botany scientific community and develop a prototype of how those needs may be met within a digital library platform. We want to assess the practicality of existing tools to satisfy the technical, economic, and operational needs of botanical users to annotate. This will inform on requisites, best practices, and further developments for a research project to integrate an annotation tool within a virtual library. In the context of this project, we consider an annotation any mark, scribble, writing, note, highlight, drawing that is done in a resource. If you have ever tagged images online - you are an annotator. If you have ever highlighted an article (whether print or PDF) and written commentary - you are an annotator. If you have ever written marginalia in a book or placed post its with notes inside sections of a book - you are an annotator. If you have ever commented on a blog post - you are an annotator. While this project is focused on web annotation needs, we would like to interview both those who currently annotate online as well as those who have only annotated offline (e.g. print or analog) in order to better understand the functionality needed to encourage and support online annotation activities. This project was made possible in part by the Institute of Museum and Library Services (IMLS Grant LG-87-18-0057-18). <https://docs.google.com/forms/d/1rdDCISCGIPi900SISHSxgYtfzep9iIFPZkDHRQIHff4/edit>

These results were presented at the iAnnotate Conference 2019:

<https://www.slideshare.net/wulate/finding-the-annotation-needs-of-the-botanical-community-in-a-digital-library>

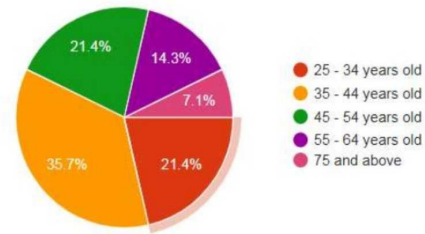
Needs analysis report

Survey

- 14 members of the botanical and scientific communities from 10 different institutions from 9 countries.
- Got a diverse representation sample:
 - 50% described themselves as non-white,
 - 2/3 identified themselves as females.
- Half of the people had a botanical background
 - The rest included entomologists, ecologists, a librarian and even an ex-lawyer.
- Included both, those who currently annotate online and those who don't.

Needs analysis report

Age group of respondents



Needs analysis report

Diverse profiles from the botanical and scientific community were represented in our sample:

- | | |
|--|-----------------------------------|
| • Herbarium Curatorial Assistant | • Curator and Associate Professor |
| • Assistant Professor | • Program Coordinator |
| • PhD Student | • Post Doctoral Researcher |
| • Curator | • Biodiversity Informatician |
| • Mother/citizen scientist/GLAM volunteer/Wikimedian | • Project Manager |
| • Outreach and Communication Manager | • Professor |
| | • Research specialist |

Needs analysis report

What do you annotate?

- Specimens
- Specimens determinations
- Original Descriptions
- Orthographic variants and synonyms
- Illustrations or plates
- Highlight, underline, write in the image
- Mostly annotate in Flickr, add species
- On an article or a long text, I usually annotate on the margins the topic of each paragraph of interest, so that I can find it easily in the future
- Adding of metadata, using vocabularies, fixing things in Wikipedia

A summary of the answers gathered from the interviews have been included below.

3.3. Where do you annotate?

Printing out an article and annotating tends to happen more when need is personal and not shared. Many folks still like printing articles and highlighting them – easy to quickly browse, take with you, etc. People tend to annotate more online when they have a need to share the annotations – with publishers, editors, colleagues. Then online is more of an advantage.

3.4. How fine is the object you annotate?

Most people highlight text but sometimes regions (images or parts of text)
Annotation tool need to handle both highlighting of texts (processable) and highlighting regions (image)

3.5. Why do you annotate?

Work

Tend to share more

Personal

Tend to keep to themselves

General

Comprehension

Highlight an important idea

Specific to field

Georeferences

Batch specimen re-identifications

Note morphological features

Habit descriptions (categorization)

Correcting names in IPNI

Memory recall
Corrections
Improve access for others – findability
Helps in building an article or topic
Generate online discussions & dialog
Manage & share information
Collocate similar info (e.g. bringing together data from single author or collector)
Linking
Peer Review
Refining ideas – annotate my papers as I write it.
Citation
Image tagging
To create a linked network of knowledge

3.6. At what stage of the process do you annotate?

Happens at every stage – beginning of the research cycle, middle and end. Beginning when gathering information and reading. Some read first then come back to annotate later, categorizing. Have to be able to add annotations at every stage.

Questions

- do our functional specs need to allow for validation of a target if it changes? (e.g. if the page or sentence on page changes, as in replaced or deleted, do we need to account for that and build in functionality to address that?)
- Do we need the functionality to allow for overlapping text and overlapping regions? (e.g. 2 different users highlight the same text or image but with slightly different boundaries)

3.7. How frequently do you annotate at each stage?

Variable depending on task

- Daily – most common response
- Hourly – at least two people annotate hourly
- Weekly – some did weekly

3.8. What methods or tools do you use to annotate?

General

- Print out on Paper –pencil, colored coded highlights, post-its
- PDFs
- Email
- MS Word review tools
- Google docs, Google Keep
- Documents in a shared drive – Google Drive
- Online – tagging or hashtags, comments (adding information or links)
- Screenshots (would consider those targets – people usually take a screenshot then put it in a doc or email with some explanatory text)
- Kindle
- Wikipedia, Flickr, Disqus, Wordpress, Pinterest, Zotero, Wordpress, Google Refine

Specific to field

- Physical specimen labels

- I classify specimens in a folder as an indication of the new identification needed. I reorganize lines in identification keys and sometimes add additional species.
- Leica Application Suite (proprietary software) to make measurements and annotations on microscopy photographs.
- Trove, Digital New Zealand, Smithsonian Transcription Center, Notes from Nature, VertNet, EOL, iNaturalist, AnnoSys, Tropicos, ADAM
- Vocabularies/checklists – The Plant List, WORMS, Catalog of Life, ITIS
- Normalize data with google Refine then pushing data (indexing) back to original source or pushing into new platforms

3.9. Why do you prefer these methods or tools indicated?

- Quick comprehension – “Color highlighting helps me differentiate the type of annotation”
- Way you were taught
- Interoperability – integrates well with other technologies
- Habit, comfort
- Flexibility,
- Simplicity
- Shareability
- Easy to quickly write thoughts on paper
- Pragmatic – limited time to learn new tools

Printed

I like to have "reference works" printed so that I can bring them to the herbarium, easily compare them to other texts I am reading (on screen or in print) and simply have them on hand at all times.

If I don't have access to Adobe to annotate online, I do it on paper. It is easy to quickly write thoughts on paper. I tend to not carry my laptop or ipad with me.

Online

For things I won't use as often, things I want on the cloud (books too heavy to carry on travels for example), and things I plan on sharing with coworkers, I would rather annotate the file digitally. The proprietary Leica software I use for microscopy images has more advanced functions (especially in terms of measurements) than the currently available free counterparts have.

I use sites that have the functionality if they don't allow annotation or tagging I don't do it. It depends on how site is designed. I don't know if institutions really think about that. Traditionally the trend has been to just provide content rather than interact with it but now some sites do. Allowing people to engage with content. Once you limit how people can engage with your content you limit number of people who do engage with your content.

Not a preference but a requirement

- No other option is available
- only option
- don't have a choice
- They're currently my only option.

Because no other option. For publishers they prefer our comments be done directly in PDFs so we do our annotations there. I might print out the PDF first to read and make notes but then I will add those

comments to the PDF before sending back to the publisher. Don't use tracking changes in Word because it's too easy to "accept all" changes and then all indications of where I made changes will be lost. Because I do most of my work on the computer my annotation method tends to be limited to the functionality of the website. For example in Flickr I'll use machine tags and the comment section. In Wikipedia I can add external links, citation links as well as wikilinks. It depends on what the website allows.

3.10. Do you use or refer to pre-existing lists (vocabularies) to annotate? If so, which ones?

Bolivia and Ecuador check lists

Check Lists

Stearn's "Botanical Latin" for any morphological term (it is a commonly used, widely accepted and easily available reference). IPNI for all plant names and author names. "Taxonomic Literature" (Stafleu and Cowan) for author names and journal title abbreviations (as requested when submitting an article to "Taxon", for example).

Morphological terms

Self-built vocabs

VIAF

Specific to field

OBO Foundry, Plant Phenology Ontology, FLOPO, PO, Gene Ontology, marineregions.org, Marine Species Traits, WWF Ecoregions, habitat ontologies, Atlas Living Australia, EOL, IPNI, Index Herbariorum

3.11. How do you use your annotations?

For explaining where you found information – rationale for the info you added.

Temporary vs long term

To return back to in the future or use only for a short time

Private vs public

Keep private unless someone challenges what I added

Always share publicly

Researcher vs Citizen Science motivations – researchers do not want others to steal on what they spent time. Citizen Scientists are motivated by the greater good and contributing to science. They want to save other researchers' time.

3.12. Who do you share your annotation with?

Responses were a balance of public vs private.

4 out of 14 chose "no one, only yourself" and no other options

Conclusion – need functionality of being able to 1) keep private 2) share with a group 3) share with everyone

Would there ever be an instance where annotations that are marked “private” would need to automatically be made public at a later date? We think no for the prototype but will leave this question open for discussion

e.g., raw data is required to be shared 1 year from date of report. LTER

3.13. Do you read or see other people's annotations?

Yes, they read it but they are not methodically looking at other people’s annotations as a way to enhance their own research or understanding of a topic

Why? They are not always aware it is available. Usually find through happenstance

Hard to understand the full history of annotations

Re-identifications are not published

No online way to link one annotation to multiple specimens (i.e. one body linked to many targets)

Annotations need to be discoverable outside of the place where they were added (i.e. separate from the website or target)

Would we have a functional need to overwrite an annotation? Yes if someone wants to fix his or her own data. If we allow editing, we need to link annotations to the different versions e.g. in Google docs “resolve” functionality hides the old comment but it is still there, not deleted.

Questions for RERUM

- can we determine who has access to our annotations (i.e. public vs private?)
- Can we have public annotations or are they all private?
- Does your API allow for synchronization with local storage?

3.14. Do you have a review or vetting process before sharing your annotations?

Majority say no, especially if their annotations are private. However, if annotations are public then sometimes review what they are writing before making it public.

In some cases with group editorial review, they are using annotations as support for the editorial process. There is a lot of vetting back and forth by editors and review of these annotations before they are shared with the author.

We would need to have the option for annotations to be made private or just visible to a group.

Annotations should be public by default. In each account you can specify if you want private by default or you may want it to ask you each time you annotate whether public, private or group

3.15. Do you sort or filter your annotations?

Majority do sort or filter them (65%) and many of those are by author. Some filter by species, date, status (e.g. resolved/not resolved), and creating institution. Some create customized hashtags to find them later.

Whether they sort or filter depends on what they are interested in

Most are filtering rather than sorting. Sorting can only be done once data is filtered first and sorting is done outside of the book viewer.

Filtering by author within a book is functionality needed at a minimum for the prototype (e.g. I want to see only John’s annotations in this book) as well as ability to filter and then export out of dig lib platform.

Filtering and sorting across books within the digital library platform would be desired for the future. Implicitly, filtering should be done by book or page in the viewer. Searching could produce a list of the annotations in table format where sorting and filtering can be done.

3.16. What information do you add in an annotation?

Lots of different types of info - specimen name, habitat types, corrected text, geographic locations, authors (artist, collector, dates, determined by), notes, reviews, links (URL, URI, DOI, barcode), customized categorization, personalized vocabularies or (hash)tags (“Interesting”, “evolution”, “new method”, “lacks documentation”, “lacks analysis”), bibliography (citation), ratings.

When we create annotation functionality should we allow for rich text or just text? I.e. text only or formatted text with images. For a basic prototype have text only but rich text have in a future platform. What about automatic hyperlinking? Maybe not for the basic prototype but definitely for the next version.

For the prototype, at minimum we need both target and annotation but for more advanced version we would, ideally, want target and its context (i.e. word “create” and sentence it belongs to) along with annotation. If the target were an image region, we would want a larger region it belongs to.

3.17. How can your current annotation process be improved?

No online way to link one annotation to multiple specimens (i.e. one body linked to many targets).
*Need to find out if the model supports this.

Make it easy to include annotations: 2-3 clicks process, have a dropdown list of controlled vocabularies, allow tagging with an URL.

Allow adding a tag in a specific place (region) within an image.

Implement search functionality by keyword or type (comments/descriptions/customized tags/categories)

Reuse previous annotations (add another target to an existing annotation?) Search and duplicate existing annotation and with a new different target?

Functionality for vetting annotations? Probably a good idea to have some sort of administrative vetting in basic prototype. This requires user roles

Autofill functionality (suggests words based on what you typed before) - we create it or the browser could handle

Users of our tool need to agree to make their public annotations available under a CCO license or something similar

Does the annotation model allow for annotations of annotations? (e.g., .replying to another annotation or identifying which best practice you used when you wrote an annotation.) In that case, an annotation would be both a target and body.

Is data in RERUM discoverable by other LOD resources?

Keep a simple tool integrable (with a click of a button) with Zotero, Hypothes.is and other tools and visible for non-users.

Annotations should be visible even without login in.

Provide feedback on the reading/modification/impact of an annotation. Is this supported by RERUM?

Being able to see my annotations, click on one and open there where I annotated previously.

3.18. Any comments related to annotations that you may want to add?

Allow for talk page. Wiki'mize more.

- **No anonymous login.** Every annotation with ID, Timestamp, Motivation
- **Privacy:** private, group, everyone.
- **Support workflow** (like editorial process of a publisher)
- **Where to store?** Agreement: Not stored locally only.
Interface. Default is to see all annotations, but being able to hide them or filter by author.