

iDigBio's Public Participation Platform Survey

Thank you for taking the time to consider and respond to the following questions. iDigBio will make your responses available to the natural history collections community, both as an online resource available to anyone, and as a reference specifically for participants in our upcoming Public Participation in Digitization of Biodiversity Collections course.

Please return your completed survey to Michael Denslow (michael.denslow@gmail.com). And thank you for your time!

Questions:

1. Name and email of person responding to this survey: [Ben Brumfield](#)
2. Name of public participation platform: FromThePage
3. Website: FromThePage.com
4. Company or group responsible for maintaining platform: Brumfield Labs, LLC
5. Long-term funding structure for maintaining platform (e.g., grants, membership, private): Since 2015, server hosting, technical support and software maintenance is funded through subscription fees paid by institutions running crowdsourcing projects; major feature development is largely crowd-funded by consortia of institutions or by individual institutions applying for grants.
6. Brief summary highlighting the market niche for this platform: FromThePage focuses on textual transcription, OCR correction and metadata description of digitized documents.
7. Types of biodiversity specimens most common on the platform: Field books and correspondence.
8. Types of tasks supported by the platform (e.g., transcription, measurement, georeferencing): Full-text transcription of letters or field books, structured data transcription of specimen labels, OCR correction from PDF or Internet Archive, spreadsheet transcription of ledgers, indexing entities within text, structured metadata creation/revision of sets of images.
9. Other disciplines that utilize the platform (e.g., camera traps): FromThePage is used by research libraries, special collections, state archives, and historians working on scholarly editions.
10. Example of the types of institutions/collections using your platform:
NYBG (John Torrey Letters)
Chicago Botanic Gardens (Orchidology sketch books)

Museum of Vertebrate Zoology (Joseph Grinnell field books)
San Diego Natural History Museum (Laurence Klauber field notes)
Missouri Botanical Garden (William Brewster field books)
Iowa State Library (Seed catalogs)
USDA National Agriculture Library

11. Any restrictions as to who can use the platform: no restrictions as to institution; projects may run public projects or private ones restricted to authorized users such as staff.

12. Primary language(s) supported: Any text supported by Unicode may be transcribed. (Additional development may be required for top-to-bottom scripts like Manchu or older Japanese.) Currently the software interface is in English, Spanish, and Portuguese. Interfaces in additional languages are available with minimal funding after discussion with the developers.

13. Primary volunteer base or core users of the platform: Institutions running projects recruit their own volunteers, but frequent users of FromThePage.com often discover projects on their own if they align with their interests. (Newsletter to transcribers on FromThePage.com can also help recruit volunteers.)

14. Brief description of method(s) for quality control/quality assurance available on the platform: Projects may be configured for optional review of contributions, required review (by any participant), or restricted review (only by authorized users). All transcriptions may be monitored by any user, so peer review/correction by volunteers is central to the process.

15. Types of media supported by the platform (e.g., image, video, 3D): Images

16. File formats supported by the platform. Indicate both input and output as appropriate: Input formats supported are PDF, ZIP files of images, and images in JPG, PNG, or TIFF. Digital collections systems that support IIIF have more robust integration, as does material hosted on the Internet Archive.

Output formats include plaintext, HTML, TEI-XML, DOCX, PDF and CSV. A IIIF-based API allows programmatic harvesting, and system-specific integrations exist for CONTENTdm and the Biodiversity Heritage Library.

17. Workflow customization available to the user:

For full-text transcription projects (appropriate to field books or correspondence), projects may define transcription conventions to be presented to users, determine whether entities may be indexed, and whether HTML/XML mark-up is supported. Individual documents may also support translation of the text in addition to transcription/OCR correction.

For structured-data projects, transcription fields are fully customizable by data type, label, and layout.

18. Costs to collections to use the platform: Hosting costs vary based on number of concurrent projects and number of images in active projects. Pricing starts at \$80/month for a single

project and up to 10,000 pages. The most popular option is \$3000/year for unlimited projects and a maximum of 50,000 images in active projects. Current pricing for hosting is at <https://fromthepage.com/pricing>

The FromThePage software is open source (licensed under the AGPL 3.0), so may be installed on an institution's own servers for free by IT staff with competency in Ruby-on-Rails and Linux.

19. Representative for potential users to contact: Ben Brumfield (ben@fromthepage.com) or Sara Brumfield (sara@fromthepage.com).

20. Best resources for additional information about the platform (e.g., presentations, brochures, recorded webinars, peer-reviewed papers): Presentations and project profiles are posted at <https://content.fromthepage.com/>. Monthly webinars to introduce crowdsourcing and the project are free to attend, listed at <https://content.fromthepage.com/webinars/>. Ben and Sara are always happy to meet via Zoom to discuss the platform, or brainstorm ideas for crowdsourcing projects.

21. Data ownership or usage policies (e.g., publication, sharing of data, public access): All data is owned by the institution or researcher running FromThePage projects, unless otherwise required by copyright.