



Media only: Liz O'Brien (202) 633-1522; [obriene@si.edu](mailto:obriene@si.edu)

June 9, 2015

## **Dartmouth's Tiltfactor Launches Games to Improve Access to Biodiversity Heritage Library Content**

### *Smorball and Beanstalk Allow Players to Transcribe Texts through Play*

Today, the Biodiversity Heritage Library, headquartered at the Smithsonian Libraries, announced the launch of two new crowdsourcing games, [Smorball](#) and [Beanstalk](#). The games, created by Dartmouth College's [Tiltfactor](#), will help improve access to books and journals online in the [BHL](#) collection by verifying the accuracy of text previously encoded by optical character recognition software.

“Cultural heritage institutions are increasingly benefiting from human computation approaches that have been used in revolutionary ways by scientific researchers. Engaging citizens to work together as decoders of our heritage is a natural progression, as preserving these records directly benefits the public,” says Dr. Mary Flanagan, founder and director of Tiltfactor. “Integrating the task of transcription with the engagement of computer games gives an extra layer of incentive to motivate the public to contribute.”

Smorball and Beanstalk are part of the “[Purposeful Gaming and BHL Project](#),” which is exploring how digital games can make scanned content more accessible and searchable for cultural institutions. Based at the Missouri Botanical Garden in St. Louis, Missouri, “Purposeful Gaming” was established in 2013 through an Institute of Museum and Library Services (IMLS) grant and includes partners from Harvard University, Cornell University and the New York Botanical Garden.

This project tackles a major challenge for digital libraries: full-text searching of digitized material is significantly hampered by poor output from Optical Character Recognition (OCR) software. When first scanned, the pages of digitized books and journals are merely image files, making the pages unsearchable and virtually unusable. While OCR converts page images to searchable, machine encoded text, historic literature, like that found in the Biodiversity Heritage Library, is difficult for OCR to accurately render because of its tendency to have varying fonts, typesetting, and layouts.

Smorball and Beanstalk, both quick and easy browser games, present players with phrases from scanned pages in BHL. After much verification, the words players type are sent to the libraries that store the corresponding pages, allowing those pages to be searched and data mined and ultimately making historic literature more usable for institutions, scholars, educators, and the public.

“The games provide a fun and engaging way for volunteers to help us with a task that we don’t have the staff to do ourselves,” explains Trish Rose-Sandler, Principal Investigator for “Purposeful Gaming.” “BHL benefits by having improved discoverability of its books and journals on plant and animal life. More importantly, benefits from the results of the project would extend to the broader digital library community. Any institution managing large text collections can learn from novel and more cost-effective approaches to generating searchable texts.”

### **About Smorball**

Players of the more challenging Smorball game are asked to type the words they see as quickly and accurately as possible to help coach their team, the Eugene Melonballers, to victory to win the coveted Dalahäst Trophy in the fictional sport of Smorball. Each word typed correctly defeats an opposing smorbot and brings the Melonballers closer to the championships.

### **About Beanstalk**

Players of the more relaxed Beanstalk game must type the words presented to them correctly in order to grow their beanstalk from a tiny tendril to a massive cloudscraper. The more words they type correctly, the faster the beanstalk grows. Players who accurately transcribe the most words will ascend to the top of the leaderboard as a result of their valuable contributions.

###

### **About Smithsonian Libraries**

The Smithsonian Libraries maintains a collection of more than 2 million volumes and serves as an educational resource for the Smithsonian Institution, the global research community and the public. The Libraries are located in Washington, Edgewater, Md., New York City and the Republic of Panama. For more information, visit the [Smithsonian Libraries website](#).

### **About the Biodiversity Heritage Library**

The Biodiversity Heritage Library is a consortium of major natural history, botanical and research libraries that seek to contribute to the global “biodiversity commons” by digitizing and

aggregating the resources housed within each of the participating institutions, providing free and open access to the legacy literature that underpins the work of the natural science community. For more information, visit the [Biodiversity Heritage Library website](#).

### **About the Institute of Museum and Library Services**

The Institute of Museum and Library Services is the primary source of federal support for the nation's 123,000 libraries and 35,000 museums. Our mission is to inspire libraries and museums to advance innovation, lifelong learning, and cultural and civic engagement. Our grant making, policy development, and research help libraries and museums deliver valuable services that make it possible for communities and individuals to thrive. To learn more, visit [www.ims.gov](http://www.ims.gov).

