The Digital Himalaya Project: Collection, Protection & Connection

Mark Turin

Published on: Feb 14, 2021
DOI: 10.21428/51bee781.028ec770
License: Creative Commons Attribution 4.0 International License (CC-BY 4.0)
Team

Project Lead: Mark Turin, Associate Professor, Institute for Critical Indigenous Studies and Anthropology, University of British Columbia

Project Team: The Digital Himalaya Project was established by Mark Turin, Sara Shneiderman, Alan Macfarlane, and Sarah Harrison in 2000. Since 2006, Komintal Thami has served as the Kathmandu-based data manager and digitization assistant. The project is supported by an active international Advisory Board and has benefited from skilled student interns and research assistants in Canada, Nepal, the United Kingdom, and the United States since its establishment.

Project URL

http://www.digitalhimalaya.org

Project Abstract

Digital Himalaya is a 20-year-old project established by anthropologists and historians to explore new methods for collecting, protecting, and connecting historical multimedia collections relating to the Himalayan region in ways that would widen access to the materials through emerging digital platforms.

The project set out to digitize older ethnographic data sets and collections in order to protect them from obsolescence and decay, forward migrate them as new standards emerged, and share the content with originating communities in the Himalayan region and with scholars everywhere online, on air, and in print.
The Digital Himalaya project was designed by Alan Macfarlane and Mark Turin as a strategy for archiving and making available ethnographic materials from the Himalayan region. Based at the Department of Social Anthropology at the University of Cambridge, the project was established in December 2000. From 2002 to 2005, the project moved to the Department of Anthropology at Cornell University and began its collaboration with the University of Virginia. From July 2014, the project has relocated to the University of British Columbia, Vancouver, Canada, and is engaged in a long term collaboration with Sichuan University.
Time Needed

*When did you begin this project? When did you complete this project?*

**Time Span:** January 1, 2000 - present

**Length:** 20+ years

Outcomes

*What is the outcome of the project?*

Project outcomes include a public-facing and frequently-updated website, a full data archive on DSpace, the University of Cambridge institutional repository with rich metadata, and a dedicated
Open Collection at the University of British Columbia library in active development. The Digital Himalaya archive includes over 200,000 pages of scanned text (most of which is machine-readable thanks to Optical Character Recognition technology), over 1,000 maps, hundreds of hours of video and audio and a large collection of assorted ethnographic content. Google Analytics provides helpful user information. Over the years, team members have written about the Digital Himalaya Project, and there are a number of public reviews of the website.

Resources

What tools, resources, programs, or equipment did you use for this project?

Making a virtue out of a necessity, Digital Himalaya has used exclusively open-source digital tools, MySQL, PHP, HTML, and other out-of-the-box tools and technologies. In it’s 20-year history, the project has been through three website redesigns and now conforms to W3C validation for accessibility purposes. The website also makes use of CC licenses.

Funding

Please describe any costs incurred for this project, and (if relevant) how you secured funding for these costs.

For the first five years while Digital Himalaya was still in active development, we had considerable success with competitive grants, all of which are listed on our website: http://www.digitalhimalaya.com/support.php.

Once the site was launched and most of the collections were digitized and accessible, it became more difficult to find ongoing resources to maintain our online collections. In general, funding agencies and research councils focus their calls on innovation and development, and not on digital maintenance and longevity. We have been fortunate to receive support in the form of donations from individuals and foundations who share our commitment to ensuring that unique cultural and historical content from the Himalayan region is shared more widely. We have also experimented, albeit with limited success, with modest referral commissions through Amazon.

Workflow

Please give an overview of the workflow or process you followed to execute this project, including time estimates where possible.
In our first five years, from 2000-2005, most digitization of film and scanning of texts and images was conducted at Cambridge and Cornell universities, in the UK and USA respectively. Thereafter, as project team members relocated to Asia, we moved our scanning operations to Kathmandu, Nepal, with some additional work conducted in Sikkim, India. Data were processed and analyzed, corrected and archived at Cornell University, the University of Virginia, Cambridge University, Yale University, and most recently at the University of British Columbia, as the project moved with the Director. A skeletal staff on short term contracts at Cambridge University, including Hikmat Khadka, were responsible for preparing much of the metadata that was ingested into the institutional repository.
Kathmandu-based data manager Komintal Thami holding an issue of Mulyankan, an important Nepali-language publication, the back issues of which Digital Himalaya hosts online. Kathmandu, Nepal, 2007. Photo by Mark Turin.
Challenges & Opportunities

What, if anything, changed between beginning your project and its current/final form?

While Digital Himalaya began as a strategy for collecting and protecting the products of colonial-era ethnographic collections on the Himalayas—for posterity and for access by local communities—the project has now become a collaborative digital publishing environment, welcoming new collections of contemporary academic and cultural content, with close to half a million web visitors since its establishment in 2000. The website has grown from being a static homepage with occasional updates to a dynamic content delivery platform for over 40GB of archival census data, maps, films, audio recordings, photographs, journals, and scanned books.

We have also witnessed dramatic changes in the profile of visitors to our site: from (overwhelmingly) scholars at European and American universities in the early years, to a well-networked, global, and digitally literate user community, in particular from South and East Asia.
Is there anything specific you wish you had known when beginning your project that might help other people to know?

Yes! When we launched the project, we did not anticipate the extraordinary development and global penetration of smart phones with data services and powerful web browsers. We expected some form of technological leapfrogging, to which end we were preparing for users to access Digital Himalaya content on ever-faster dial up connections and eventually on broadband. Our initial website was not designed for mobile use. We used advanced compression codecs for our audio-video collections to ensure that the files were not above 5 MB and could be easily downloaded, but the viewing quality was low. For colleagues actively developing a digital project now, it’s important to think ahead about who will be visiting your site and how they will be accessing the content.

**Next Steps**

*Do you have any plans to follow up on this project or work on something similar in the future?*

Together with some of the same colleagues, and building on the success of Digital Himalaya, I established the [World Oral Literature Project](http://www.digitalhimalaya.com/publications.php) using similar data structures and web tools.

**Publications & Presentations**