

Multilayered Collaboration: Documentation & Archival Construction in the Aftermath of the 2015 Nepal Earthquakes



Kristine Hildebrandt, Tanner Burge-Beckley, Jacob Sebok
Southern Illinois University Edwardsville
Slides available at <https://tinyurl.com/yxkkobj3>

Archiving Documentation Materials: The Common Scenario

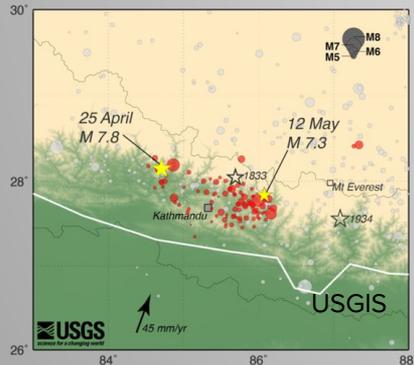
- The researcher and/or community activist records language data from a single language, representing different genres/contexts
- Once documentation and data preparation are completed, the materials are sent to an existing (external) repository, which “ingests” the materials for long-term storage and preservation
- Access can be controlled and restricted according to community concerns and considerations
- **Our project:** collect as many instances as possible of survivor narratives from speakers across myriad **Tibeto-Burman** ethnolinguistic communities representing multiple geographical regions of Nepal at a single point in time



Recording a Nubri narrative in Gorkha District

The Aftermath of the 2015 Earthquakes

- The first earthquake (4/25, M 7.8) killed 9,000, with an epicenter in Gorkha
- The second (5/12, M 7.3) killed an additional 200
- Our immediate concern (after learning about survivors, welfare of friends and relatives): capturing and preserving survivors' nuanced and diverse experiences of living through a natural disaster, in their own languages, before these narratives were generalized and altered for a global mass media audience
- We secured funding from the NSF RAPID mechanism (BCS 1547377: Narrating Disaster: Calibrating Causality & Responses to the 2015 Earthquakes in Nepal)



- We hoped survivors & responders would recount their reactions at the time of the events, and also describe what they know about earthquakes more generally, based on lifelong experiences & belief systems



Kathmandu

Gorkha District



The Linguistic Situation in Nepal

- Home to over 100 languages representing 5 families
- For many languages, the first major collection of transcribed, translated, and annotated materials with situational metadata and access to original audio and video.
- The materials center on theme of narration of survival & impact, in languages for which there is otherwise very little available description or data
- Exploration of a discourse genre that has barely been examined (but cf. Cox et al 2008 for media accounts of disaster narration; Baker et al 2008 & Hansen 2006 for refugee and foreign policy narratives) and allows further research on how survivor experiences map onto the physical world, and how this is linguistically encoded in terms of source, emotional experience, reactions to that experience, and also priorities identified by experiencers

Language, Glottocode, ISO 639-3	Sub-grouping	District	EGIDS Scale (Lewis & Simons 2010)
Nubri, nubri1241, kte	T-B: Bodish, Gyalsumdo-Nubri-Kyirong	Gorkha	6a Vigorous
Tsum, tsum1240, tz	T-B: South-Western Tibetic	Gorkha	6a Vigorous
Kutang, kuta1241, ght	T-B: Bodish, Kaikhe-Ghale-Tamangic	Gorkha	6a Vigorous
Gurung, west2414, gvr	T-B: Bodish, Tamangic	Manang	6b Threatened
Manange, mana1288, nmm	T-B: Bodish, Tamangic	Manang	6b Threatened
Nar, nar1239, npa	T-B: Bodish, Tamangic	Manang	6a Vigorous
Phu, nar1239, npa	T-B: Bodish, Tamangic	Manang	6a Vigorous
Lowa, lowa1242, loy	T-B: South-Western Tibetic	Mustang	6a Vigorous
Tibetan, tibe1272, bod	South-Western Tibetic, further classification uncertain	Mustang	Unknown
Baragaon Tibetan, bara1356	South-Western Tibetic, closely related to Lowa	Mustang	Unknown
Dokkeh/'brogs skad Tibetan, no known identifiers	South-Western Tibetic, further classification uncertain	Mustang	Unknown
Nepali, nepa1254, npi	Indo-European	Kathmandu	1 National

Languages in this Project

RAPID Project Organization

- We recruited 16 local residents who spoke the languages and who also had basic computer skills
- We trained them on A-V recording techniques, back-up, oral informed consent, and we co-constructed prompts to gather narratives (and in Mustang: more formal interviews)
- A day-long “crash course” in field methods in Kathmandu
- The residents returned to their homes over the coming weeks (as roads & footpaths became accessible); they also collected narratives from diaspora households in Kathmandu (also severely impacted)



Preparing to collect a narrative from a Manange woman in Kathmandu



Field methods crash course in Kathmandu



A Lowa interview in Mustang District

Archival Workflow

- In Nepal, recordings transcribed, translated (Nepali, English), and a sub-set had morpheme interlinearization
- We ended up with 96 fully or partially transcribed narratives (varying from 3-15+ minutes in length) across 8 languages
- Participants ranged from local residents (often agricultural or local civil servants) to members of the Buddhist religious community, to returning “diaspora responders”
- Data transfer to home institution in subsequent trips to Nepal
- UVA A-V and transcript proof-reading, cleaning-up, and materials upload done by PI and SIUE volunteer RA’s



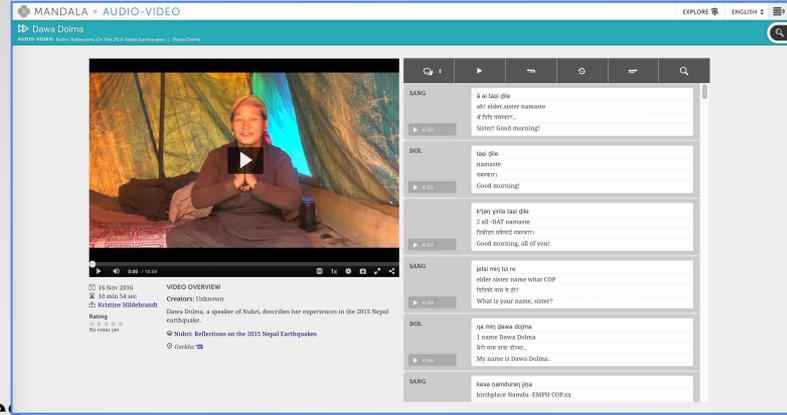
Recording a Tsum narrative in Gorkha

Project researcher Bhoj Raj Gautam with Tenzin Kunchok, transcribing a Nubri narrative in ELAN



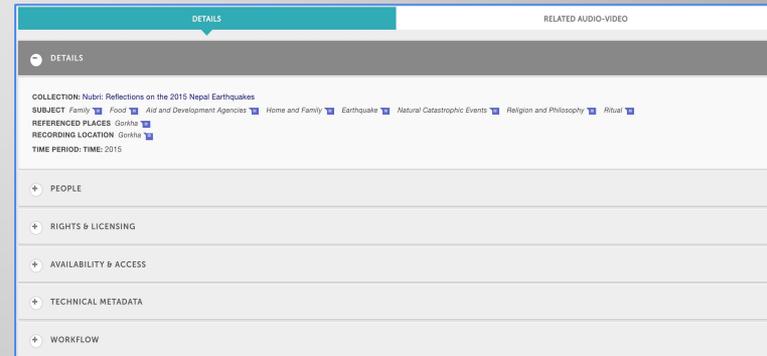
University of Virginia's Tibetan & Himalayan Library

- Hosts video narratives and transcripts into six collections: Nubri, Tsum, Kutang, Lowa, Nar-Phu, Manange
- Developed with a CMS designed by THL librarians
- Platform based on Drupal, an open source software for websites & other digital applications
- The metadata encoding worked with both universal standards, and standards specific to the “THL-verse”
- Each language is a unique collection; each collection houses narratives plus transcripts.
- Within each collection, items catalogued with item and content metadata, as allowed by oral consent.
- By clicking on the item: A thumbnail view of the video & time-aligned transcript. The video can be played in its entirety, or else individual transcript lines can be clicked on to hear that utterance in isolation



The screenshot displays the MANDALA AUDIO-VIDEO interface. At the top, it shows the title 'Dawa Dolma' and the collection 'AUDIO-VIDEO'. Below the title is a video player showing a woman sitting in a tent. To the right of the video player is a transcript with time-aligned lines. The transcript is organized into sections labeled 'SANG' and 'DOL'. The transcript text includes phrases like 'I at last die', 'old sister name', 'of this river.', 'Name: Good morning!', 'Good morning!', 'What is your name, sister?', 'My name is Dawa Dolma...', and 'My name is Dawa Dolma...'. The interface also includes a search bar and a play button.

Audio-video and technical-content metadata



The screenshot shows the 'DETAILS' page for the audio-video item. The page is titled 'DETAILS' and 'RELATED AUDIO-VIDEO'. The main content area is titled 'DETAILS' and contains the following metadata fields:

- COLLECTION:** Nubri: Reflections on the 2015 Nepal Earthquakes
- SUBJECT:** Family, Food, Aid and Development Agencies, Home and Family, Earthquakes, Natural Catastrophic Events, Religion and Philosophy, Ritual
- REFERENCED PLACES:** Gurkha
- RECORDING LOCATION:** Gurkha
- TIME PERIOD:** TIME: 2015

Below the metadata fields are several expandable sections:

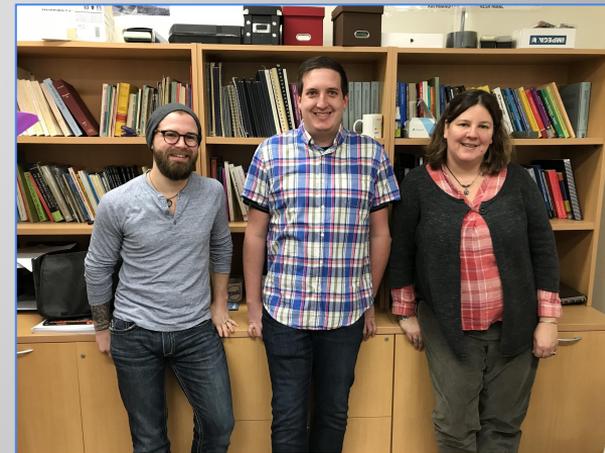
- PEOPLE
- RIGHTS & LICENSING
- AVAILABILITY & ACCESS
- TECHNICAL METADATA
- WORKFLOW

The Need For Another Repository, And an Idea....

- This project was assembled quickly (3 months from RAPID application to equipment purchase, travel to Nepal, recruitment & training of resident field assistants)
- We ended up with a variety of other materials that needed cataloguing & storage, but these were not in line with what we had originally agreed to deposit in THL: many still images, over 100 formal Mustang interviews, derivatives
- SIUE Interdisciplinary Research & Informatics Scholarship (IRIS) Center
- We have an ITS issuance of Omeka, and guarantee of regular backups
- But: Omeka is an “out of the box” application, requiring on-site construction
- We also wanted an online web exhibit interface that would link the two collections (UVA THL & our Omeka collection) along with other information about Nepal (language situation, post-earthquake updates)



REU Team: 2017-2018



What Now?

1. Requirements for Archive and Exhibit
3. Choose Platforms to Facilitate Requirements
4. Construct Site Architecture (Shells)
5. Populate the Site with Data

Functional Requirements

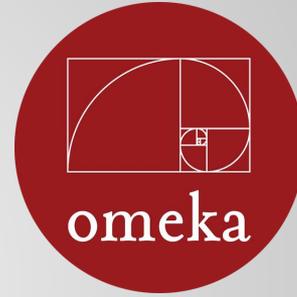
ID	Requirement Statement	Need	Comments
FR1.0	A persistent menu from which all other pages are accessible shall exist on all pages of site	Must	Perhaps this menu is different on homepage to account for different layout
FR1.1	A persistent footer from which contact form, institution info, copyright information, perhaps NSF info is accessible should exist on all pages of site	Want	There are many solutions for this issue but this solution eliminates the need for an extra "Contact" tab
FR1.2	Site shall be laid out as a museum exhibit, offering the ability to progress linearly through the narrative	Want	This requires a visual prompt (perhaps at the bottom of each page) to show the progression to next portion is available
FR1.3	Each page, excluding the homepage, shall include a localized page outline for easy navigation	Want	Think Wikipedia's hyperlinked table of contents
FR1.4	If user is on a specific page, they should be presented with the ability to advance to the next section without navigating to bottom of page	Want	Perhaps, a user has opted to navigate through site "as exhibit" but has no interest in phonology page. They should be allowed to slick a right arrow to progress to "Stories" page.
FR1.5	If user navigates to homepage for first time, page should display a welcome message	Want	Perhaps a brief wizard describing the exhibit layout? "Have you been here before? Start here."
FR1.6	Content pages will contain relevant metadata information (according to a developed standard) including captions on all images	Must	The standard will be developed and discussed as we move into further design stages
FR2.0	Site shall possess a home/welcome page that creates a context for the project	Must	The homepage should exemplify the navigation method of the site by requiring them to enter site manually
FR2.1	Homepage may possess an interactive graphic	Want	Located below a full-screen graphic?

Choosing the Platform(s)?



WORDPRESS

- Aesthetic Appeal
- Highly Customizable
- User-Friendly Interface
- Graphics-Driven Experience



- Respected Platform
- Dublin Core Metadata Standard
- Built to House Archives
- Infrastructure Already in Place

What is Omeka?

- Free
- Open Source
- Designed for Exhibits
- Dublin Core Metadata Standard

The screenshot shows the Omeka website for "Nepal Earthquakes". At the top, the title "Nepal Earthquakes" is on the left, and navigation links for "Items", "Collections", and "Exhibits" are on the right. Below the title is a search bar with the placeholder text "Search our digital archives..." and search icons. A banner image of a mountainous landscape is visible behind the search bar. Below the search bar, a status message reads: "Nepal Earthquakes contains 486 files and 479 items in 21 collections." The main content area is titled "Featured Exhibits" and contains four exhibit thumbnails: "Derivatives" (a group photo of people), "Manang" (a person sitting on steps outside a brick building), "Gorkha" (a person sitting on rocks with a laptop), and "Mustang" (people sitting on the grass in a field).

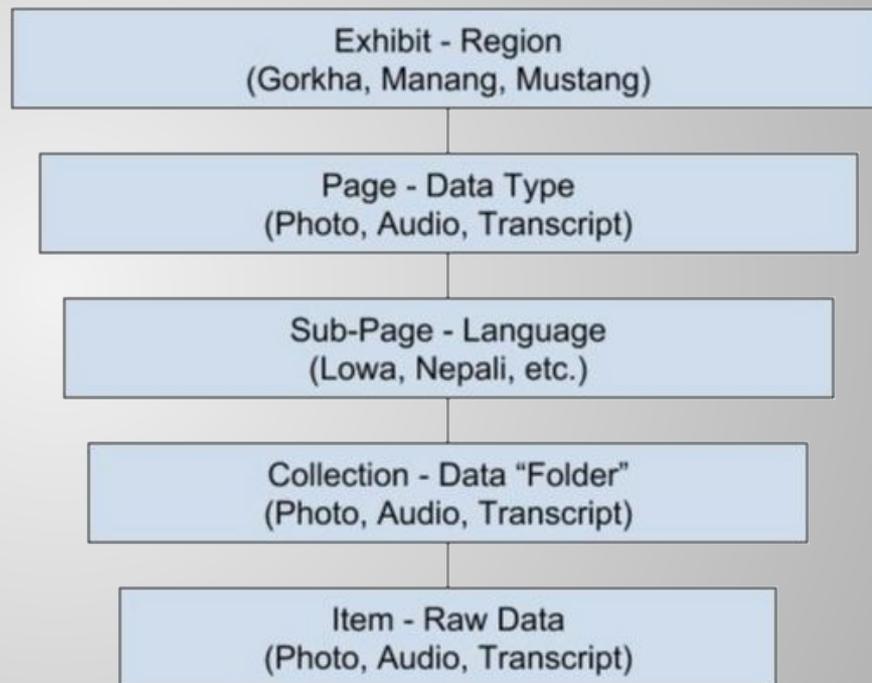
Information Architecture

We needed to organize

- Images,
- Audio Files
- Transcripts

According to

- Region
- Language



CSV Import - A Miracle

1) Import data to spreadsheet

- Mass-manipulation of data
- Sorting / Filtering
- Error Checking

2) Export spreadsheet as .CSV

- Single Location

3) Upload .CSV to Omeka

- Minimizing Potential User Error

Title	Subject	Description	Creator	Source	Publisher	Date	Contributor	Format	Language	Type	Identifier
Lowa_001_AUD	Effects of 2015 E	Interview of a cor	Jacob Sebok	Sienna Craig	National Science	24-Jun	Pema Tsering Gu	.mp3	Lowa	Audio Interview	130101_0004.%.M
Lowa_002_AUD	Effects of 2015 E	Interview of a cor	Jacob Sebok	Sienna Craig	National Science	28-Jun	Yangdorma Gurur	.mp3	Lowa	Audio Interview	130104_0014.M
Lowa_003_AUD	Effects of 2015 E	Interview of a cor	Jacob Sebok	Sienna Craig	National Science	28-Jun	Damchoe Bhutik	.mp3	Lowa	Audio Interview	130104_0016.M
Lowa_004_AUD	Effects of 2015 E	Interview of a cor	Jacob Sebok	Sienna Craig	National Science	28-Jun	Pasang Gurung	.mp3	Lowa	Audio Interview	130104_0018.M
Lowa_005_AUD	Effects of 2015 E	Interview of a cor	Jacob Sebok	Sienna Craig	National Science	28-Jun	Kunga Tenzin Gu	.mp3	Lowa	Audio Interview	130104_0019.M
Lowa_006_AUD	Effects of 2015 E	Interview of a cor	Jacob Sebok	Sienna Craig	National Science	29-Jun	Tashi Doma Gur	.mp3	Lowa	Audio Interview	130105_0021.M
Lowa_007_AUD	Effects of 2015 E	Interview of a cor	Jacob Sebok	Sienna Craig	National Science	29-Jun	Kasang Lhamo C	.mp3	Lowa	Audio Interview	130105_0022.M
Lowa_008_AUD	Effects of 2015 E	Interview of a cor	Jacob Sebok	Sienna Craig	National Science	29-Jun	Sherap Gurung	.mp3	Lowa	Audio Interview	130105_0023.M

Import Items Status

Step 2: Map Columns To Elements, Tags, or Files

	Example from CSV File	Map To Element	Use HTML?	Tags?	Files?
Title	Lowa_001_AUD	Title	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Subject	Effects of 2015 Earthquakes in the Mustang Regi...	Subject	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Description	Interview of a community member on the 2015 Nep...	Description	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Creator	Jacob Sebok	Select Below	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Source	Sienna Craig	Select Below	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Publisher	National Science Foundation / Southern Illinois...	Contributor	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Date	24-Jun	Coverage	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Contributor	Pema Tsering Gurung	Creator	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Format	.mp3	Date	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Language	Lowa	Description	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Type	Audio Interview	Format	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Identifier	130101_0004.%.%20MUL-2015-1	Identifier	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Coverage	2015 Nepal Earthquakes	Language	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
File	http://iris.siue.edu/nepal-earthquakes-archive/...	Publisher	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Relation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Rights	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Source	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Subject	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Title	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Type	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Item Type Metadata	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Bit Rate/Frequency	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
		Duration	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

The Archive

Each language had both **Audio Files** and **Transcripts**.

In order to facilitate a productive user experience, we adopted a **Trinomial** naming schema that identified:

- Interview Language
- Reference ID
- Media Type

The screenshot displays a digital archive interface. At the top, there is a search bar with the placeholder text "Search our digital archives..." and a magnifying glass icon. Below the search bar is a breadcrumb navigation path: "Home / Exhibits / Mustang / Audio / Transcripts / Lowa". The main content area is titled "Lowa" and is divided into two sections: "Interview Audio" and "Interview Transcripts".

The "Interview Audio" section features four audio file icons, each represented by a speaker icon. Below each icon is a label: "Lowa_025_AUD", "Lowa_026_AUD", "Lowa_027_AUD", and "Lowa_028_AUD".

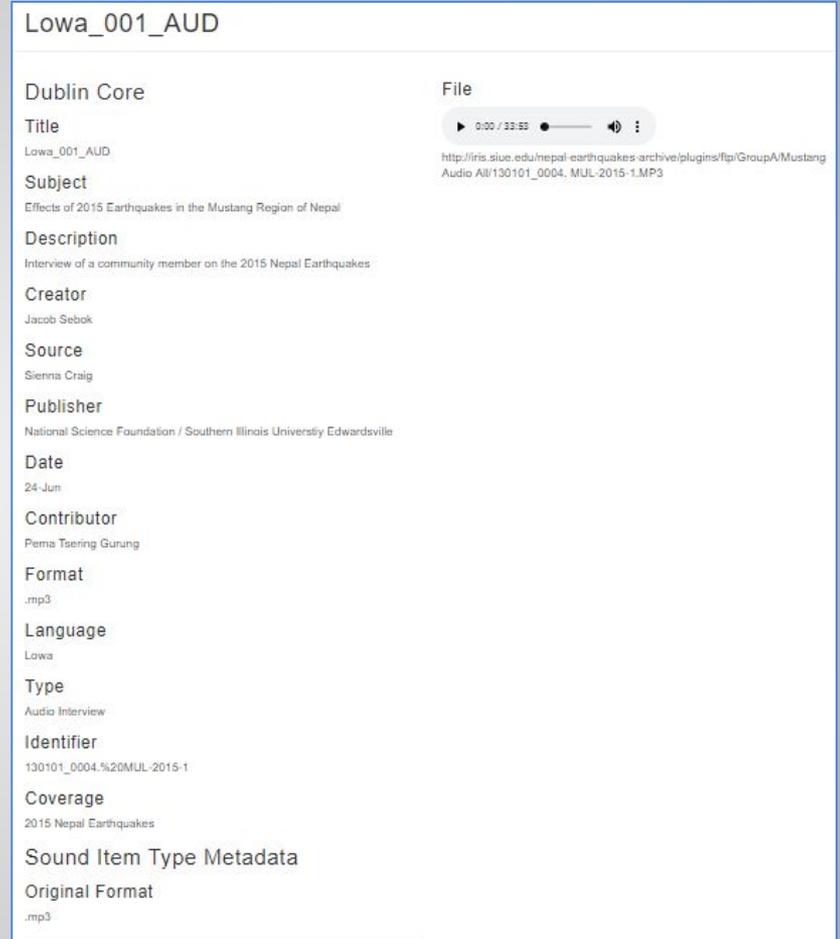
The "Interview Transcripts" section features four document icons, each represented by a square with a folded corner. Below each icon is a label: "Lowa_021_TRANS", "Lowa_027_TRANS", "Lowa_028_TRANS", and "Lowa_032_TRANS".

On the right side of the page, there is a sidebar menu titled "Mustang" with a list of categories: "1. Images", "2. Audio / Transcripts", "3. English", "4. Lowa", "5. Nepali", and "6. Tibetan". The "Lowa" category is highlighted in blue.

The Finished Product

With Omeka, a great deal of Metadata is visible to every user with the click of a button.

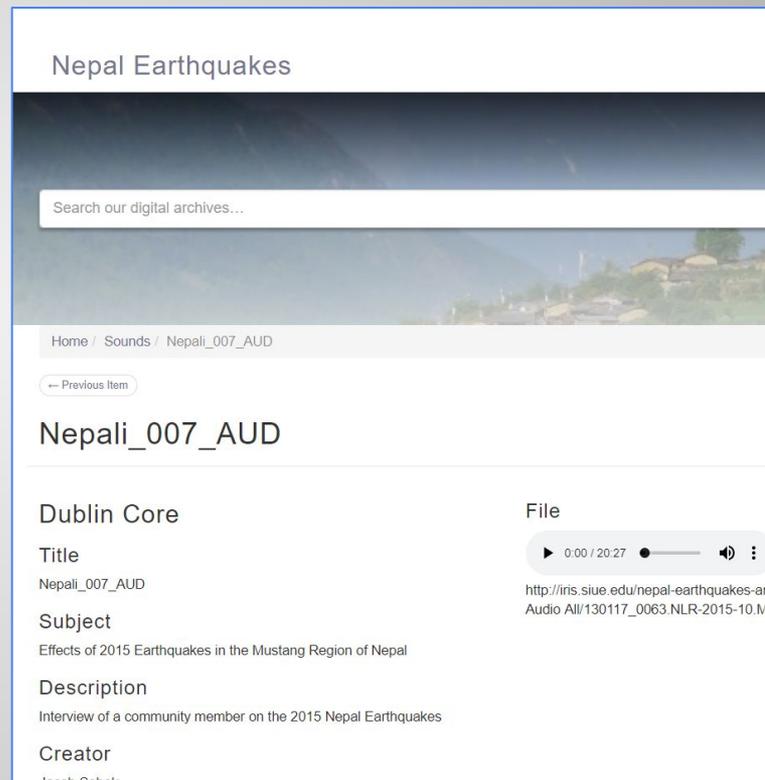
But how to make this accessible to a wider audience?



The screenshot displays an Omeka audio player interface. At the top, the title 'Lowa_001_AUD' is shown. Below it, a 'Dublin Core' metadata section lists various fields: Title (Lowa_001_AUD), Subject (Effects of 2015 Earthquakes in the Mustang Region of Nepal), Description (Interview of a community member on the 2015 Nepal Earthquakes), Creator (Jacob Sebok), Source (Sienna Craig), Publisher (National Science Foundation / Southern Illinois University Edwardsville), Date (24-Jun), Contributor (Perma Tsering Gurung), Format (.mp3), Language (Lowa), Type (Audio Interview), Identifier (130101_0004.%20MUL-2015-1), Coverage (2015 Nepal Earthquakes), Sound Item Type Metadata (Sound Item Type Metadata), and Original Format (.mp3). On the right side, there is a 'File' section with a play button, a progress bar showing 0:00 / 33:58, a volume icon, and a settings icon. Below the progress bar, the file URL is provided: http://iris.siu.edu/nepal-earthquakes/archive/plugins/fpi/GroupA/MustangAudio/All/130101_0004_MUL-2015-1.MP3.

Archive vs. Exhibit

- Not all of the research data is approachable to a wider audience
- A exhibit provides an accessible portal of contextual information for those unfamiliar with the research.



The screenshot displays a digital archive interface. At the top, the title "Nepal Earthquakes" is visible. Below it is a search bar with the placeholder text "Search our digital archives...". The main content area features a video player with a play button and a progress bar showing "0:00 / 20:27". To the right of the video player, there is a "File" section with a URL: "http://iris.siu.edu/nepal-earthquakes-ar/Audio/All/130117_0063.NLR-2015-10.M". Below the video player, there is a metadata section with the following fields:

- Dublin Core**
- Title**: Nepali_007_AUD
- Subject**: Effects of 2015 Earthquakes in the Mustang Region of Nepal
- Description**: Interview of a community member on the 2015 Nepal Earthquakes
- Creator**: Jacob Sabok

Wordpress and Omeka: Two Sides to a Coin

- For general information presentation, Wordpress is an open source, user friendly choice to build a complementary exhibit
- Like Omeka, is built on PHP as a server side script development language
- Provides a different form of CMS than Omeka



Cohesive Presentation

- Both Omeka and Wordpress provide free themes to customize color and arrangement presentation of page information
- With two complementary installations, we chose themes that were built on the same core Bootstrap framework
- Themes were then customized directly to appear more similar



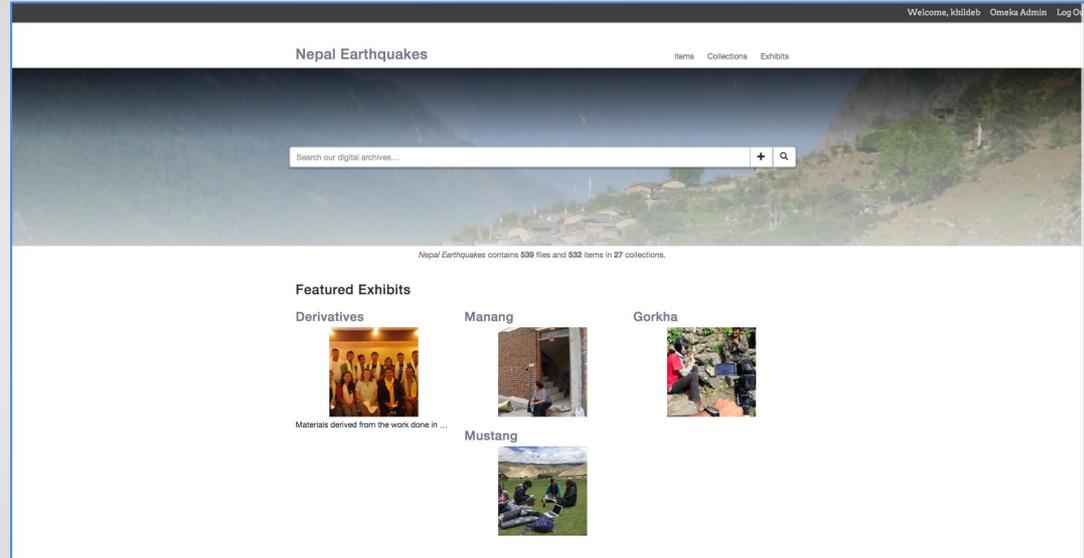
PHP Customization

- Theme customization involves PHP source editing
- Color and style customizations were made in CSS modification
- CSS and PHP knowledge can be employed from a knowledgeable Computer Science student or professional web developer

```
<?php endif; ?>
<div class="row">
<?php foreach(loop('collections') as $collection): ?>
  <div class="col-sm-6 col-lg-4">
    <div class="row">
      <div class="col-sm-4">
        <div class="collection-img">
          <?php
            $collectionImage = record_image('collection', 'square_thumbnail', array('cla
            $noFile = ' 'image'));
            else:
              echo link_to_collection($noFile, array('class' => 'image none'));
            endif;
          ?>
        </div>
      </div>
    </div>
  <div class="col-sm-8">
    <h2><?php echo link_to_collection(); ?></h2>
    <?php if (metadata('collection',array('Dublin Core', 'Contributor'))): ?>
    <div class="element well well-sm well-md well-lg">
      <p><span class="glyphicon glyphicon-user"></span><strong><?php echo __('Collect
      <div class="element-text">
        <p><?php echo metadata('collection',array('Dublin Core', 'Contributor')); ?>
      </div>
    </div>
  <?php endif; ?>
</div>
</div>
```

Benefits of a Self Hosted Archive

- Uses university resources
- Customization of archive design
- Set and uphold metadata standards
- Opportunity for cross-department collaboration



<https://iris.siue.edu/nepal-earthquakes-archive/>

We Also Faced Challenges

- We had to construct informed consent protocols to not compromise the dignity of participants in overwhelmingly complex and difficult situations
- We had to be sensitive to local administrative policies
- Our field researchers worked with a variety of equipment types, and often had to share resources
- They also struggled with aligned file naming conventions, and many transcript files needed to be transcoded
- Some of this stemmed from different disciplines coming together, and some of it came from the different field research team choices, issues of timing, issues of material access, and the constant tectonic, population, and political shifting that followed in the weeks and months after the earthquakes
- We also had challenges in terms of pairing mis-matched transcripts, with audio and video files, or A-V with no transcription whatsoever



Final Thoughts

- This presentation does not imply that documentation initiatives should attempt home-designed and locally-housed archival repositories instead of other, well established archives
- Established archival repositories have an advantage of a knowledgeable steering committee, trained staff, established standards for data formatting, encryption, preservation/protection, and of course a great deal of storage space
- We had residual materials and the chance to make use of local resources as a teaching and learning opportunity, and we have home institutional support
- We turned to established archives for inspiration and guidance
- We hope that by documenting our challenges and our solutions in the construction and customization of these two digital tools, other users will be able to adopt and adapt in their own endeavors

<https://tinyurl.com/yxkkobj3>



Thank you and Questions

- NSF BCS 1547377 funding, including two REU supplements
- We are grateful for the participation of: Geoff Childs, Sienna Craig, Mark Donohue (co-I's); Dr. Dubi Nanda Dhakal, Mr. Bhoj Raj Gautam (managers); Lhakpa Hritar, Lhakpa Norbu, Jangchuk Sangmo, Nyima Samdrub, Tenzin Kunchok, Namgyal Ngudup, Pema Norbu, Tenzin Gyaltzen, Ngawang Tsering Gurung, Yangjin Bista, Karma Choedon Gurung, Tsewang Gyurme Gurung, Sophiya Lama, Chimi Lama, Tsering Topden Gurung, Anjana Ghale (field researchers); Sarah Song, Kristin Kaskeski, Ishu Jha, and Pratik Lamsal (volunteer SIUE R.A.'s who assisted with UVA data transfer)

<https://tinyurl.com/yxkkobj3>



Video training in Kathmandu (July 2015)



Co-I S. Craig at project team check-in (October 2015)