The AV section
Wiki-based media categorization
Agenda: The AV section

1. Introduction
   a. Archiving, annotating, and classifying media
   b. Types of online classification tools, examples
   c. Why wikis?

2. Towards a taxonomy of media
   a. General notes
   b. Prior art: Baha’i Media Bank, Wikimedia Commons

3. Bahai.media
   a. Contents
   b. Classification scheme
   c. Content discovery & navigation
   d. Sub-projects and related projects

4. Conclusion and Q&A
Introduction

Archiving, annotating, and classifying media

Online classification tools

Why wikis?
The importance of archiving

Time is of the essence, and the sooner a systematic effort is undertaken to collect material such as tablets to the early believers, diaries, old photographs properly identified, and correspondence with the prominent early believers, the more productive it will be. Many of the important historic documents may have already been destroyed by the descendants of early believers who do not realize their importance.

(From a letter dated 6 August 1985 written on behalf of the Universal House of Justice to a National Spiritual Assembly)
Archiving media

By media we mean: photos, audio, video, charts, diagrams, art...

- The importance of archiving media is undeniable: every Baha’i community has a rich audio-visual heritage
- Effie Baker photographed unique sacred places; her photos were printed in the Dawn-Breakers
- Photographs and plans remain for destroyed holy sites, allowing for them to be rebuilt in the future
- Archiving media preserves Baha’i history for future researchers, readers, listeners and viewers

Now, although a picture is worth a thousand words…

…sometimes a picture needs words to give it context!
Annotating media

He thinks that when things are collected for the archives as much information as possible should be given about them. … every attempt should be made to amplify local records and accumulate historical data.

(On behalf of Shoghi Effendi, from a letter dated 25 November 1948 to an individual)
Annotating media (cont’d)

To help preserve the historical value of media, we often need to record information about context—What? When? Where? Who?

- Printed photographs were commonly annotated: on the back of a group photo, we could write down the name, date and location of an event and write down who was there.

- In the 21st century we have many ways of describing images, including totally new ones, e.g. tag a face on Facebook, add hashtags on Flickr or Instagram (social tagging / folksonomy).
Digital archival

Creating and maintaining digital archives has a lot of benefits over simply keeping things in boxes.

- Improve ease of access
- Reduce damage that can come from frequent handling of physical collections
- Enable complex research based on metadata

Nothing will replace physical archives, but the nature of media search makes digitization especially helpful for researchers.
What are the criteria by which media can be classified?
Organizing and classifying media

Keeping an online collection of media clean and organized is essential to making research easy. In considering how to organize and classify media, we have to ask: What scheme should we use?

Some general criteria include:

- **Subject**: What, when, where, who…
  - containing $x$, related to $y$, during the $z$-year plan…
- **Media-specific metadata**: Source, author/creator, medium…
- **Archival metadata**: Physical collection, processing status…

We will look at classification practices shortly, with some real-life examples.
Many types of solutions exist to organize and classify media online. Naturally, all of these tools allow for collaboration!

- **Cloud providers** offer space for storing media with basic organization tools (folders, albums)
  - Google, Office 365, Nextcloud...

- **Social media** sites offer channels for publishing audio/video and albums for photos
  - Facebook, Instagram, YouTube, Soundcloud...
  - These often feature **folksonomies** that offer a basic, user-generated system of classification: e.g. hashtags

- **Wiki-based systems** tend to offer hierarchical category trees, or taxonomies
  - Wikipedia, Wikimedia Commons...
Why wikis?

MediaWiki, one of the most popular wiki software applications, has built-in support for hierarchical categories.

- Categories can be organized from general ones to more specific ones
- Easy to navigate and discover content in many ways
  - Built-in category tree allows quick top-down browsing—drill down “vertically” to find appropriate categories
  - Category pages display all contained files, pages, and subcategories—browse “horizontally” along the same level of the category tree
  - Navigation templates can be created to allow quick thematic browsing

These features make wikis very powerful tools for organizing and classifying media into a taxonomy, enabling advanced research.
Towards a taxonomy of media

General notes
Prior art
How is the classification of textual works different from that of other media?
Initial notes on classification

Quite a lot has been said about how to classify Baha’i materials, especially sacred writings, publications...

- Phelps’ considerations on subject-based classifications of the Baha’i Writings
- Collins’s collected resources on library classifications (WABLA)

Questions on classification of Baha’i-related media:

- How do we effectively classify a wide variety of dissimilar media (images, audio, video…) on different subjects?
- Which overarching topics can best serve as the roots of a media taxonomy (the “main” or “top-level” topics)?
- How should our scheme express different types of characteristics and conceptual relations (is, has/contains, comes from, happened at/in/during, presented/was in attendance at…)
How to classify media

What are some best practices we can follow when trying to create a taxonomy of media?

- Place media in categories that meaningfully describe their subject's **characteristics**
  - A photo of a seasonal school in Cameroon will be filed under “Seasonal schools in Cameroon”; a photo of an event that took place in 1970, in “Events in 1970”
  - Naming conventions reflect different types of conceptual relations

- Categories should be **as specific as possible**, to avoid creating “catch-all” categories
  - As we saw, “Notable Bahá’ís” on bahaí.media contains more specific categories for Hands of the Cause, Counsellors, and many more
  - This makes categories more useful, and searching easier—as long as you know specific details about a subject you’re looking for, you can find media about it
Prior art #1: Bahá’í Media Bank

The Bahá’í Media Bank (media.bahai.org) is an example of a website that strives to classify media (photos, video) by subject

- Gives an idea of important top-level categories for a Baha’i project: activities, institutions, buildings, history, gatherings and events
- Further divided into more specific subcategories
  - Not a very deep taxonomy as this site is neither meant to house a large nor a very complex collection
- Suited for the most basic research needs—easy to look up photos of holy places and important figures
Prior art #2: Wikimedia Commons

Wikimedia Commons (commons.wikimedia.org) serves the needs of Wikipedia and its family of websites

- Over 80 million media files of all types
- Extensive, complex and interconnected taxonomy provided by MediaWiki software
- Database connections store myriad amounts of information about each category
- Looked after by an army of volunteers and supported by the Wikimedia Foundation
- Potentially usable for complex and specific research needs
Bahai.media

Contents
Classification scheme
Content discovery & navigation
Sub-projects and related projects
Bahai.media is an independent, community built repository for images, pdf files, audio talks and other media related to the Bahá’í Faith. It was started in September 2007 by David Haslip to meet the media needs of Bahaiwiki and its sister projects. The files hosted here mostly represent content derived from periodicals and other historical works published over the last 100 years. Images from these sources are cropped, uploaded and categorized with related pictures. Each category fits into a hierarchy of topics and related media types or files.

https://bahai.media/Bahaimedia:About
What kinds of content?

Bahai.media stores documents, photos, images, audio, video, and more…

- **PDFs**: Periodicals & publications for Bahai.works
  - See [David Haslip’s talk](#) on how these are used

- **Images**: Photos, drawings, charts, diagrams, maps…
  - Many images are digitized from periodicals and publications, for use as illustrations on Bahaipedia articles

- **Audio**: Audio talks and books, pronunciation keys

- **Video**: Movies, filmstrips

Most content is sourced from websites or digitized collections that are licensed or donated by individuals and institutions.
Copyrights & licenses

On copyright, we take a similar stance to Wikimedia—we aim not to use non-free works without some kind of permission.

- **Copyrighted media:** Arrangements have been made with some Bahá’í institutions/individuals to license and host certain media
  - The Bahá’í World Centre and Bahá’í International Community offer terms of use that permit reuse for certain applications
  - National Spiritual Assemblies: USA, Canada, Germany…
  - Individual licenses / “donated media”

- **Creative Commons:** Attribution (BY) and Share-Alike (SA) licenses
  - Unlike Wikimedia, Bahai.media accepts images licensed for non-commercial (NC); however, no-derivative (ND) licenses are generally not suitable

- **Public domain:** Older documents (prior to 1923) are generally in the public domain and can be used freely
  - Creative Commons Zero (CC0) is roughly equivalent to public domain
Self-made media

Of course, we welcome contributions of self-made media!

Anyone who, spontaneously and of his own free will, donates material to the Archives of his National Spiritual Assembly - whether this be Tablets, books, pictures, objects or the like... will have performed a highly meritorious act in the sight of God, and his name will be perpetuated in the records of the Spiritual Assemblies and his memory enshrined in the Archives for ever.

(From Shoghi Effendi, July 1925 to the Bahá’ís of the East, translated from the Persian)
Self-made media (cont’d)

- Photos or videos of local or regional Bahá’í activities can be used to illustrate articles
  - Both recent and historical photos and video are helpful!
- When uploading self-made media, contributors can choose a suitable license to release their work
- All media should be annotated with descriptions, including the real-life context (the event, time and date, notable individuals depicted…)
Classification on Bahai.media

Bahai.media uses MediaWiki’s categorization functionality to sort media into a taxonomy of categories.

- Top-level content categories are contained in the “Topics” category
  - Activities, objects, organizations, people, places, society…
- Each of these “topics” can have many sub-topics
  - Ex.: “People” contains sub-topics for Notable Bahá’ís and Bábís, Central Figures of the Faith, etc.
- Metacategories also exist to keep track of media properties, taxonomic structure, and project maintenance and administration

A list of major topics and subtopics is featured on the Main Page of Bahai.media.
Intersectional categories

Different topics can be combined to produce intersectional categories. To explain, let’s return to one of our previous examples. The photo at right shows a seasonal school in Cameroon in 1970. It has been filed under three different intersectional categories:

1. **Seasonal schools in Cameroon**
   Combines seasonal schools (activities/events) with Cameroon (places/countries)

2. **Events in 1970**
   Combines events (activities) with 1970 (history/years)

3. **Group photos of seasonal school participants**
   Combines group photos (images/photos) with seasonal schools (activities/events)

Intersectional categories on Bahai.media are listed under Category:Categories.
Navigating the system

Apart from the topics on the Main Page, here are some other ways to discover content:

- Built-in **category tree** allows quick top-down browsing to drill down “vertically”
- Category pages allow users to browse “horizontally” across the same level of the category tree
- Many navigation templates have been created to allow quick thematic browsing
- Search, of course!
Navigating the system (cont’d)

To illustrate, let’s return to our earlier example:

Category: Notable Bahá’ís
Sub-projects and related projects

The Baha’i Atlas, a sub-project of Bahai.media, is a front-end to the collection of maps stored on the site. It presents maps of different parts of the world, with a Bahá’í focus.

Other related projects that have been discussed:

- **BahaiData / Semantic Wiki**: Data store; collects semantic metadata about article & media
- **Inline annotation**: Adding Facebook-style tags, or annotations, to collected media
- **Bahaipedia**: Most of what has been said in this presentation can also apply to classifying articles
  - Collaboration: WikiProjects exist to focus on various topics (geography, biography, chronology, translation)
Conclusion/Q&A

Thank you for your attention!