I.   Background, Motivations, and Research Question

The mahua tree *(Madhuca longifolia)* isn't a mere botanical entity or economic resource.

Rather, it forms the very foundation of tribal religious life in central India. In the spirit of data

visualization, which allows researchers to "design" and "compress" vast knowledge into

digestible formats,[1] the following infographic[2, 3, 4, 5] conveys key research insights from my

thesis regarding the wide-ranging applicability of mahua in local ecology, ethnomedicine, social,

spiritual, and economic practices, and so on.

The mahua tree is a **keystone species** that supports hornbills, elephants, bats, deer, flying foxes, langurs, and other native biodiversity in the monsoon forests of central India.

For near-impoverished tribal populations in central and eastern India, mahua is a valuable and **versatile cash crop.** "They eat the flower, barter it for grain, and sell it for cash," recorded independent reporter Monica Jha in 2018.

An estimated 90% of flower production goes into **making alcohol,** and, in many areas, mahua flowers contribute up to 30% of a **family's annual cash income.**

India produces **more than a million tons of mahua flowers** every year, and **over 7.5 million people,** mostly women — this figure representing more than 75% of India's indigenous population — engage annually in the **collection of its flowers.**

Liquor distilled from the flowers of the mahua tree is used in the *tonda* (birth), *manda* (wedding), and *konda* (death) **cultural functions** of the Gond, a prominent central Indian tribe.

Various parts of the tree are believed to have **healing properties.** Thus, mahua products are widely used in **ethnomedicine** as a tonic for flu, skin diseases, rheumatism, and headaches.

The Gond tribals imagine **hell** as "miles and miles of **forest without any mahua** trees" recorded British anthropologist Verrier Elwin in his published diary, *Leaves from the Jungle,* in 1936.

Among other indigenous populations of central India, like the Oraons, Mundas, and Santhals, the mahua tree is revered as the sacred **'tree of life'.**

As illustrated above, the various roles and significances of mahua underscore its

indispensability to forest-dwelling communities residing in central India. However, once

ubiquitous across the landscape, mahua trees are growing endangered due to poor regeneration,

[1] McCandless, "The beauty of data visualization," 2010.

[2] Mukherjee, "Mahua: The Indian liquor the British banned," *BBC,* November 2022.

[3] Sekar, "Heaven Is Where There Are Mahua Trees – and Their Bat Friends, *Mongabay India,* June 2018."

[4] Mishra and Padhan, "Madhuca Longifolia: A Review of Traditional Uses and Nutritional Properties," *International Journal of Humanities and Social Science Invention,* 2013.

[5] Jha and Vittalamurthy, "The Mahua Story," *Fountain Ink Magazine,* May 2018.

overexploitation, and climate change.[6] My senior thesis historically analyzed the state-led suppression of the mahua tree and its products, particularly mahua liquor, in late colonial and early postcolonial central India. I conducted source-critical readings of state legislation, colonial ethnographies, district gazetteers, excise records, and anthropological surveys to understand the various roles of mahua within tribal society as well as its governance and marginalization.

However, the state and colonial archives, which were largely written and curated by elites, failed to capture the tribal perspective. This motivated me to conduct oral histories in rural central India, giving tribal participants the chance to retell their life stories in their own words and to weigh in on synchronous historical debates. In this way, I used oral histories to access the unrecorded memories of indigenous community members who had lived through the period I was researching. They revealed experiences that may have otherwise remained hidden. For instance, I learnt in an interview with Chinnaiah Jangam, who grew up in Nizamabad, Telangana, where mahua was heavily policed, that "state governments, because they wanted to sell their own alcohol, started cutting mahua trees and punishing distillers." He had continued:

> I remember when I was very small, someone was employed by the Excise Department to detect where mahua was being produced in tribal villages. Adivasis used to hide a layer of mahua flowers under the bottommost layer of grain in pots, so the detector was given a stick to poke through the grain and reveal the flowers.

My thesis ultimately argued that despite the efforts of colonial and nationalist governments efforts to wrest tribal communities from their traditional cultures surrounding the mahua tree, it remains deeply a part of the social, spiritual, and economic lives of forest-dwelling populations central India. While conducting field research on either end of Madhya Pradesh, from its westernmost border with Gujarat to its eastern frontiers neighboring Chhattisgarh, I not

---

[6] Garai et al., "Climate Change and Habitat Adaptability of (J. Madhuca Longifolia Koenig Ex L.) J.F. Macbr, 1062."

only recorded oral histories but also photographed incidences of mahua in daily life – as a prayer offering in the Bhilala temples of Kathiwada or a recurring motif in the songs of the *bana* or traditional three-stringed fiddle of the Pardhan Gonds, and much more.

Stories about mahua abound in the landscape. I was unexpectedly and delightfully overwhelmed by the richness of the oral history interviews that I conducted, and I believe that it would be worthwhile to create a larger online archive of oral histories, artifacts, photographs, etc. relating to mahua to not only unveil its cultural meanings in various contexts, but potentially also aspects of its governance or policing that have gone unrecorded or may have been suppressed.

This project will ask: *What might oral histories, alongside other unrecorded historical and contemporary sources, reveal about indigenous engagements with mahua, as well as state governance and policing of the mahua tree and its products?* Implicit in this research question and broader project framing is my conviction in the paucity of the colonial and nationalist state archives, as well as the treasure troves of data that can be made available through personal narratives and local traditions.

II.   <u>Dataset Identification</u>

I plan to travel through the state of Madhya Pradesh and continue conducting oral history interviews, as well as collect other promising source like photographs, news clippings, artifacts, etc., making sure to record the metadata methodically and uniformly for each source as I go along. I would like to adopt an existing metadata standard to ensure that my metadata is comprehensive, interoperable, and aligns with broader digital and archival practices. However, some of these, like the institutionally preferred MARC standard, don't support the incorporation of local languages and dialects or nuanced cultural markers.

Dublin Core might be the easiest option to adapt for a diverse range of digital resources, and the Dublin Core fifteen main metadata items are "generic, widely understood" terms with "simple descriptions" that are "easy to translate."[7] Manžuch had discussed the emergence of "community archives" as a democratizing force in archival studies, specifically noting how "[i]ndigenous communities have been using digitization to regain control over traditional cultural knowledge that is excessively commercialized and used without the consent of the communities where it originated."[8] I would like to involve as wide a range of contributors to my archival project as possible, honoring the knowledge and expertise held by local community members and tribal populations on their own terms.

However, the need for "bibliographic literacy" to contribute to and navigate within digital collections risks perpetuating historical divides in knowledge access, reinforcing barriers that privilege those with formal education or technological proficiency. Since I am keen on actually carrying out this project, I clarified the exact metadata requirements for each of the Dublin Core items as it pertains to the six source types I intend to work with, namely, oral histories, photographs, news clippings, documents and manuscripts, artifacts, and artworks. In the following table, I have organized simple guidances and clarifications for each Dublin Core item as it pertains to each source type. I will share this with prospective contributors to the archive to maintain uniformity all collected materials.

---

[7] *DCMI Today (2021-10-04).*
[8] Manžuch, "Ethical Issues in Digitization of Cultural Heritage," 3.

	element	definition	oral history	photograph	news article	document	artifact	artwork
1	contributor	an entity responsible for making contributions to the resource	interviewer, translator, transcriptionist, and any other key figures involved	assistant, photographers, editors, and any other key figures involved	columnists or commentators if applicable	editors, translators, and any other significant contributors	individual or organization responsible for finding or preserving the artifact	artists, patrons, or others involved in creating or restoring the artwork
2	coverage	the spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant	geographic location and time period discussed in the oral history	location and date depicted in the photograph	geographic focus of the news article and/or time period covered	geographical and temporal scope of the document's content	geographic origin and historical period of the artifact	cultural and temporal context or geographical focus depicted
3	creator	an entity primarily responsible for making the resource	interviewee or the person giving the testimony	photographer	author of the article	author or originator of the document	maker or culture that produced the artifact	artist or creator of the artwork
4	date	a point or period of time associated with an event in the lifecycle of the resource	date of the interview	date on which the photograph was taken	publication date	date of creation or publication	date of creation or discovery	date or period of creation
5	description	an account of the resource	summary of the interview content and major themes	context and content of the photograph	summary of the article, including key points	overview of document's content and purpose	physical description, use, and/or significance	description of work (i.e., style, themes, subject, etc.)
6	format	the file format, physical medium, or dimensions of the resource	audio or video format (e.g., MP3, MP4, AVI)	digital or physical format (e.g., JPEG, PNG, printed photo)	format of the clipping (e.g., PDF, HTML)	medium of the document (e.g., PDF, paper, scroll)	material and dimensions of artifact	medium used (e.g., oil on canvas, ink on paper, marble sculpture, etc.)
7	identifier	an unambiguous reference to the resource within a given context	unique ID or URL where the oral history can be accessed	unique ID or URL where the photo can be accessed	unique DOI, ID, or URL where clipping can be accessed	catalog or accession number, DOI, or URL	inventory number or unique identifier	unique accession number or URL where the artwork can be viewed
8	language	a language of the resource	language(s) spoken in the recording	language(s) of any text or inscriptions in the photo	language(s) of the article	language(s) in which the document is written	language(s) of inscriptions or any associated texts	language(s) of inscriptions or any associated texts
9	publisher	an entity responsible for making the resource available	entity responsible for making the oral history available (e.g., archive)	agency or individual distributing the photograph	newspaper or magazine that published the article	organization, publication, or person that issued or printed the document	museum or institution housing the artifact	gallery or museum exhibiting the artwork
10	relation	a related resource	relation to other sources (e.g., part of a series)	relation to other sources, photos, or collections	relationship to other articles or events	connection to other documents (e.g., series)	relationship to other artifacts or works	relation to other pieces (e.g., part of a series)
11	rights	information about rights held in and over the resource	copyright and usage rights, including permissions given by the interviewee	copyright or usage terms	copyright and reprinting rights	intellectual property rights and usage conditions	ownership and access rights	copyright and reproduction rights
12	source	a related resource from which the described resource is derived	reference to any prior versions or archival collection it belongs to	collection or original source of the photograph	name of the newspaper or database where it was obtained	collection, archive, or individual from whom the document was obtained	provenance of the artifact	an entity responsible for making contributions to the resource
13	topic	the topic of the resource	keywords relevant to interview content	main subjects or themes depicted in the photo	main topics covered in the news article	key topics or themes addressed in document	themes or cultural significance	where the artwork was acquired or donor name
14	title	a name given to the resource	title of the interview or a descriptive title	name or caption of the photograph	headline of the news article	title of the document or a descriptive title	common name or identified title of artifact	the main themes or subjects depicted
15	type	the nature or genre of the resource						

III.   <u>Methodology and Tools</u>

This stage will involve digitizing the sources and once again ensuring that the metadata is uniformly recorded. Here, it is critical to address the prevailing lack of transparency in digitization processes, especially considering how colonial legacies have, in various other instances, permeated the digital archiving of cultural heritage.[9] If not carefully managed, this mahua-centered online archive risks overlooking or appropriating indigenous narratives in the documentary process. To mitigate this risk, it will be essential to collaboratively develop the archive with continuous input and feedback from the indigenous communities whose histories and cultures are being documented.

As mentioned, I do plan on informing prospective contributors beforehand about the 'why' and 'how' of standardizing metadata to maximize transparency and streamline analysis. However, if there are compelling sources that I haven't personally collected but would like to include in the archive, or if contributors fail to record metadata in accordance with Dublin Core standards, I will use OpenRefine to clean and standardize the metadata to ensure consistency and reliability across the entire archive.

Alongside metadata management, I intend to create an interactive map that hosts each source at its location, adding a spatial dimension to the archive. To analyze the dataset, I will first sort by source type, period, or whichever analyzable grouping seems appropriate based on the distribution and diversity of the sources collected, such as their geographic origins, time periods covered, and the type of media they represent. I will apply tools from the digital humanities like distant reading and topic modeling. According to digital humanist Nina Tahmasebi, the primary advantage of computational methods lies in their ability to enable

---

[9] Varner and Porter, "Dream Lab Podcast: Digital Surrogates."

researchers to "zoom out" and efficiently process vast bodies of text.[10] Distant reading enables analysis across various authors, large time spans, and even different languages – in a way that would be impossible to do manually. I could apply Voyant Tools to oral history transcripts, machine-readable documents and newspaper clippings, etc. to perform exploratory data analysis. For instance, I could use TermsBerry to visualize interconnections between themes like 'tradition', 'spirituality', and 'governance', or employ the Word Trends panel to trace the frequencies and contexts of specific words like *"daru"* or liquor, *"sarkar"* or government, "police," "excise," etc. that recur. This might help me discern overarching patterns and thematic consistencies within the discourse, yielding fertile grounds for future inquiry.

Additionally, I could apply topic modeling to illuminate how discussions of mahua are embedded within broader themes such as indigenous rights or national identity. By applying topic modeling to analyze large datasets like legislative records, news or magazine articles, ethnographic accounts, etc., we can track shifts in associations with mahua over time.

Furthermore, by analyzing the frequencies and contexts of terms related to mahua, we might, hypothetically, discover a transition from seeing it solely as a "tribal drink" to, more recently, as a "heritage liquor." My advisor had proposed exploring the ways in which the ritual uses and understandings of mahua have shifted and evolved – rather than remaining a static "tradition" – in the last hundred years, as a question future researchers might take up. Topic modeling could quantitatively document these changes in discourse, revealing patterns in how mahua is discussed and thereby signaling shifts in its ritualistic use.

IV.   Presentation and Dissemination

---

[10] Tahmasebi and Hengchen.

I was inspired by the [Hiroshima Archive](#) to create a map that serves as a repository for mahua-related archives, locating the origins of various contents. For longer oral history videos or audio recordings, the map will feature pop-ups that provide a brief description of the content and link to a complete recording on YouTube.


I did find the Hiroshima Archive slightly overwhelming, so to minimize visual clutter, I have assigned a unique symbol to each type of source, as summarized in the following graphic:

	source type	description	symbol
1	oral histories	audio or video recordings of firsthand accounts, interviews, and testimonies	🎤
2	photographs	historical, and contemporary images reflecting the social, cultural, spiritual, environmental, and economic uses and significances of mahua in central India	📷
3	news clippings	digitized articles and features from newspapers, magazines, and other periodicals pertaining to mahua	📰
4	documents + manuscripts	letters, diaries, journals, official records, legislation, and other written or printed documents that reference mahua	📖
5	artifacts	scanned and photographed three-dimensional objects like distillation equipment, ceremonial items, etc. for example, the sacred three-stringed fiddle known as the *bana* played by Pardhan bards, who drink mahua while performing and reference it in the *bana* songs	🏺
6	artworks	digitized images of paintings, drawings, sculptures, and other forms of visual art representing, or referencing, mahua	🎨

Furthermore, I will assign a unique color to each 50-year period from 1850 to the present. I used the 'Square' Color Harmony setting on Adobe Color to identify visually distinct and thus easily distinguishable colors, and I have listed their respective hex codes and time periods underneath each color block in the palette below:

| #C8D5FA | #C6FAD8 | #F9B2B9 | #F9E9C1 |
| 1850-1899 | 1900-1949 | 1950-1999 | 2000— |
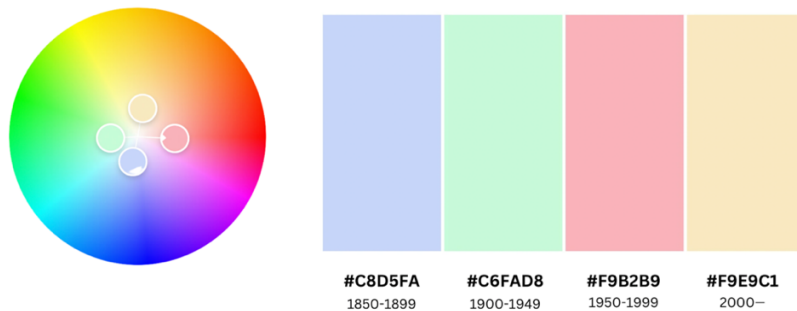
Combining these, I created the following mockup of my map:
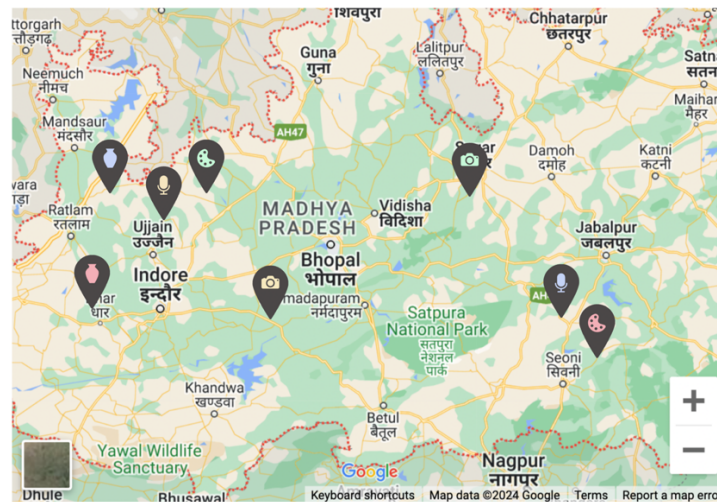
This project is not just about preserving history; it is about creating a space where the living traditions of these communities can be explored and appreciated. The ultimate goal is to ensure that the mahua tree and its various meanings to the people of central India are recognized and preserved in the face of modern challenges like climate change, deforestation, and cultural assimilation. Through careful collaboration, transparent methodologies, and respectful

engagement with local communities, this archive might come to truly reflect the complex

interrelations between people, their environment, and their history.

V.   Works Cited

DCMI Today (2021-10-04), 2021. https://www.youtube.com/watch?v=G-qY0nd9CyU.

Garai, Sanjoy, Ashish Ranjan, Mujibar Rahaman, Masjuda Khatun, Raj Kumar, Basant Kumar, Shambhu Mishra, and Sharad Tiwari. "Climate Change and Habitat Adaptability of (J. Madhuca Longifolia Koenig Ex L.) J.F. Macbr.: An Endangered Tree Species of High Socio-Economic Significance to Tribal Community of Eastern India." *Indian Journal of Ecology* 48 (September 15, 2021): 1062–67.

Manžuch, Zinaida. "Ethical Issues in Digitization of Cultural Heritage." *Journal of Contemporary Archival Studies* 4, no. 2 (2017): 1–17.

McCandless, David. "The Beauty of Data Visualization." Accessed April 1, 2024. https://www.ted.com/talks/david_mccandless_the_beauty_of_data_visualization.

Mishra, Sunita, and Sarojini Padhan. "Madhuca Longifolia: A Review of Traditional Uses and Nutritional Properties." *International Journal of Humanities and Social Science Invention* 2, no. 5 (May 2013): 30–36.

Mukherjee, Sugato. "Mahua: The Indian Liquor the British Banned." Accessed November 20, 2023. https://www.bbc.com/travel/article/20221121-mahua-the-indian-liquor-the-british-banned.

Sekar, Sandhya. "Heaven Is Where There Are Mahua Trees – and Their Bat Friends." Mongabay-India, June 19, 2018. https://india.mongabay.com/2018/06/heaven-is-where-there-are-mahua-trees-and-their-bat-friends/.

Tahmasebi, Nina, and Simon Hengchen. "The Strengths and Pitfalls of Large-Scale Text Mining for Literary Studies." *Samlaren* 140 (March 30, 2020): 198–227.

"The Mahua Story." Accessed November 20, 2023. https://fountainink.in/interactive/the-mahua-story/.

This paper represents my own work in accordance with University regulations.