

Wikimedia Wikimeet
India 2021

Presenting

SERVANTS OF KNOWLEDGE INITIATIVE

FOR DIGITAL ARCHIVING KANNADA & OTHER INDIC LANGUAGE CONTENT

ಓಂಶಿವಪ್ರಕಾಶ್ - @Omshivaprakash

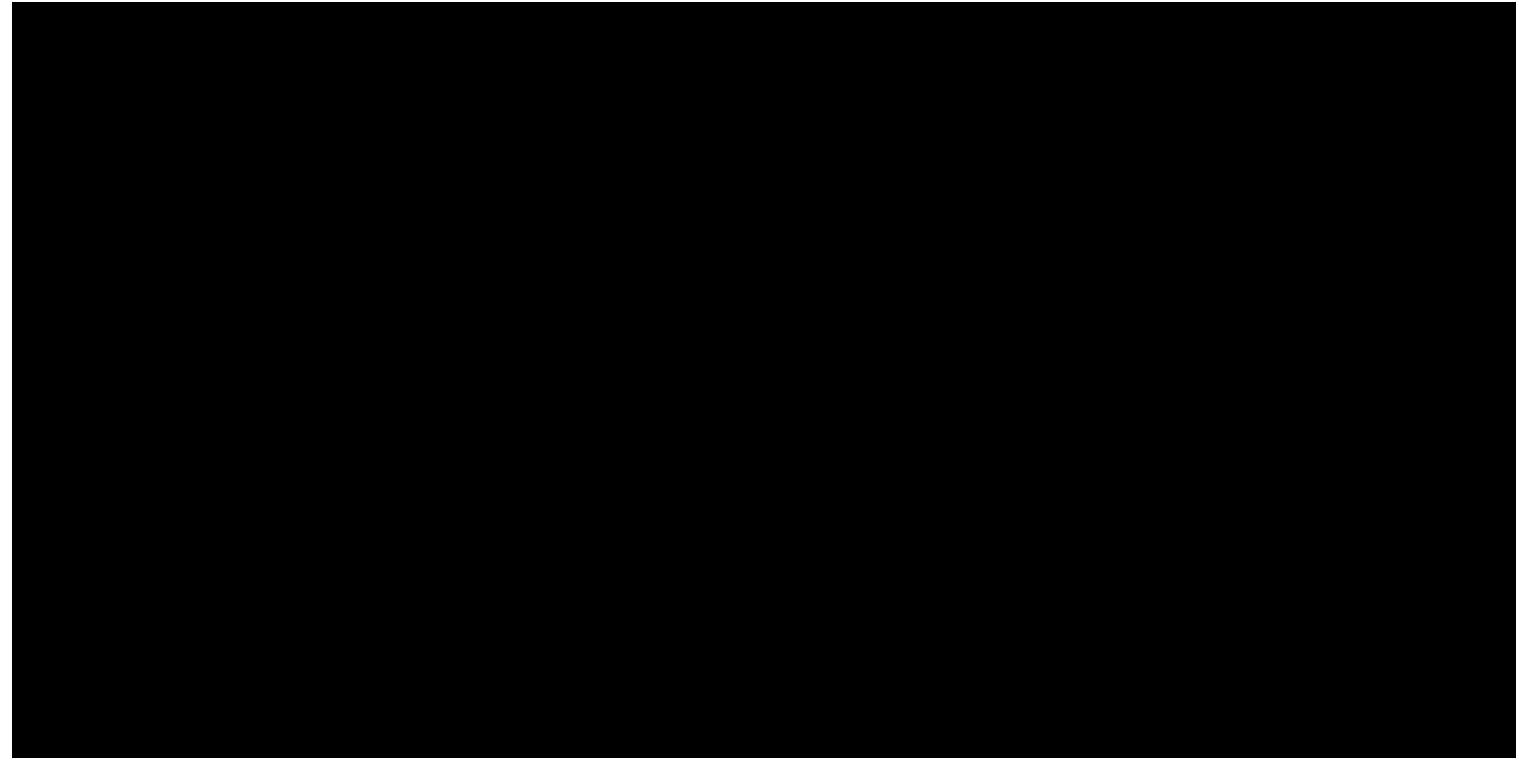
About myself



**OMSHIVAPRAKASH-
WIKIPEDIAN, MOZILLAN, CC CONTRIBUTOR,,
L10N, STARTUPS, FLOSS, LANGUAGE
TECHNOLOGY – KANNADA, DIGITAL
ARCHIVIST, OPEN DATA ETC.**



A QUICK HISTORY OF SOK

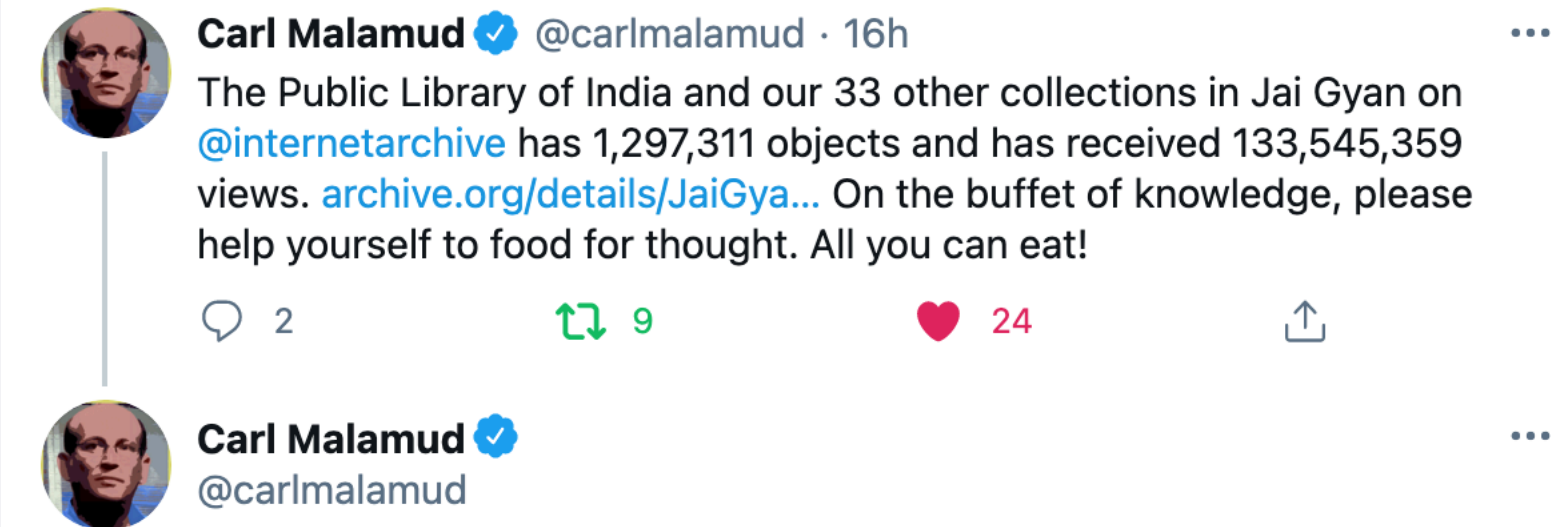
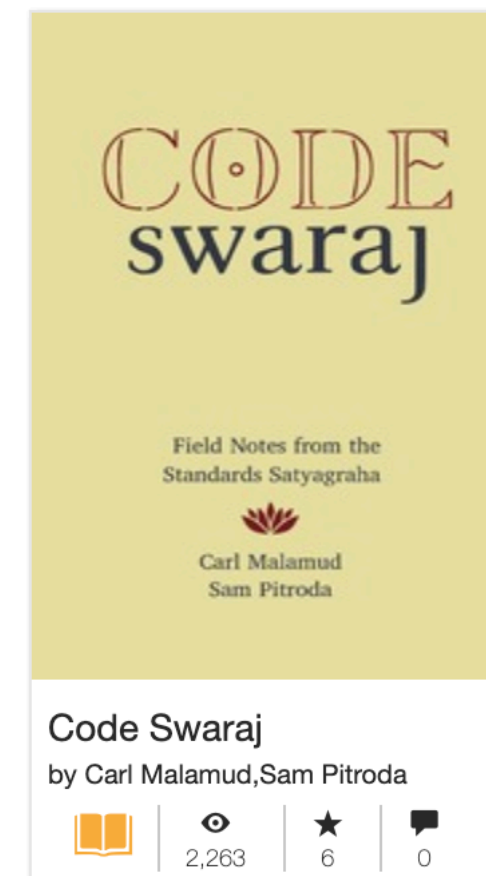


https://en.wikipedia.org/wiki/Carl_Malamud

{{cc-by-2.0}} {{flickrreview|Nishkid64|18:22, 27 October 2007 (UTC)}}

American technologist, author, and public domain advocate, known for his foundation Public.Resource.Org

Social connect -> Building a Community project
-> Sourcing books via partnerships -> Metadata
-> Scanning -> Collaboration



Replying to @carlmalamud and @internetarchive

In addition to the India materials, my 11 other collections on @internetarchive have 48,203,789 views. We have 9,837,894 views on Flickr, 83,879,373 on YouTube. Not enough to be considered mainstream. Most of the material I pick because it is hard, not because it will be popular.

1:20 AM · Feb 21, 2021 · TweetDeck

“

- Scanning is the new spinning... *Carl Malamud*
- public.resource.org

OUR PARTNERSHIPS GREW WITH THE SUPPORT OF ORGS AND GENERAL PUBLIC

SERVANTS OF KNOWLEDGE INITIATIVE



<https://archive.org/details/ServantsOfKnowledge>

- We start scanning books in IASc Bengaluru since March 2019
- ~2 years
- 4250 Books & growing
- More than a dozen associations including authors/ authors families, publishers
- 10+ magazines archived (Kasturi 50 years corpus)
- More than 60% local language books
- ~600K Pages (~6 Lac)
- Kannada, English, Tamil, Malayalam, Konkani and more...

WHY I PURSUED THIS PROJECT FOR OPEN RESOURCES

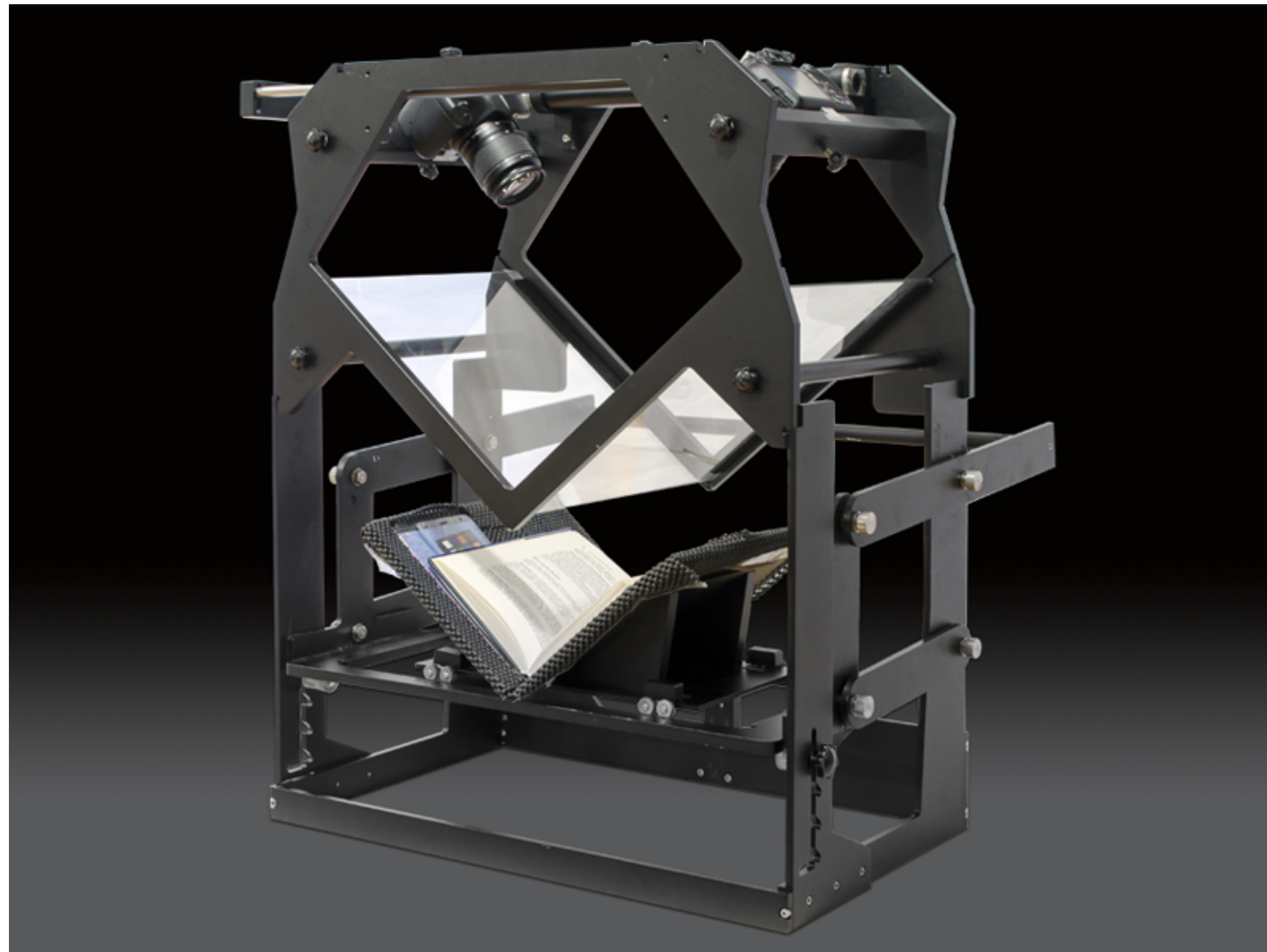
- No verifiable resources on Internet for my language
- Lack of Open content
- Lack of understanding about licensing & copyrights
- Need for standardisation
- Need for collaboration

This limited my Wiki contributions at one stage...

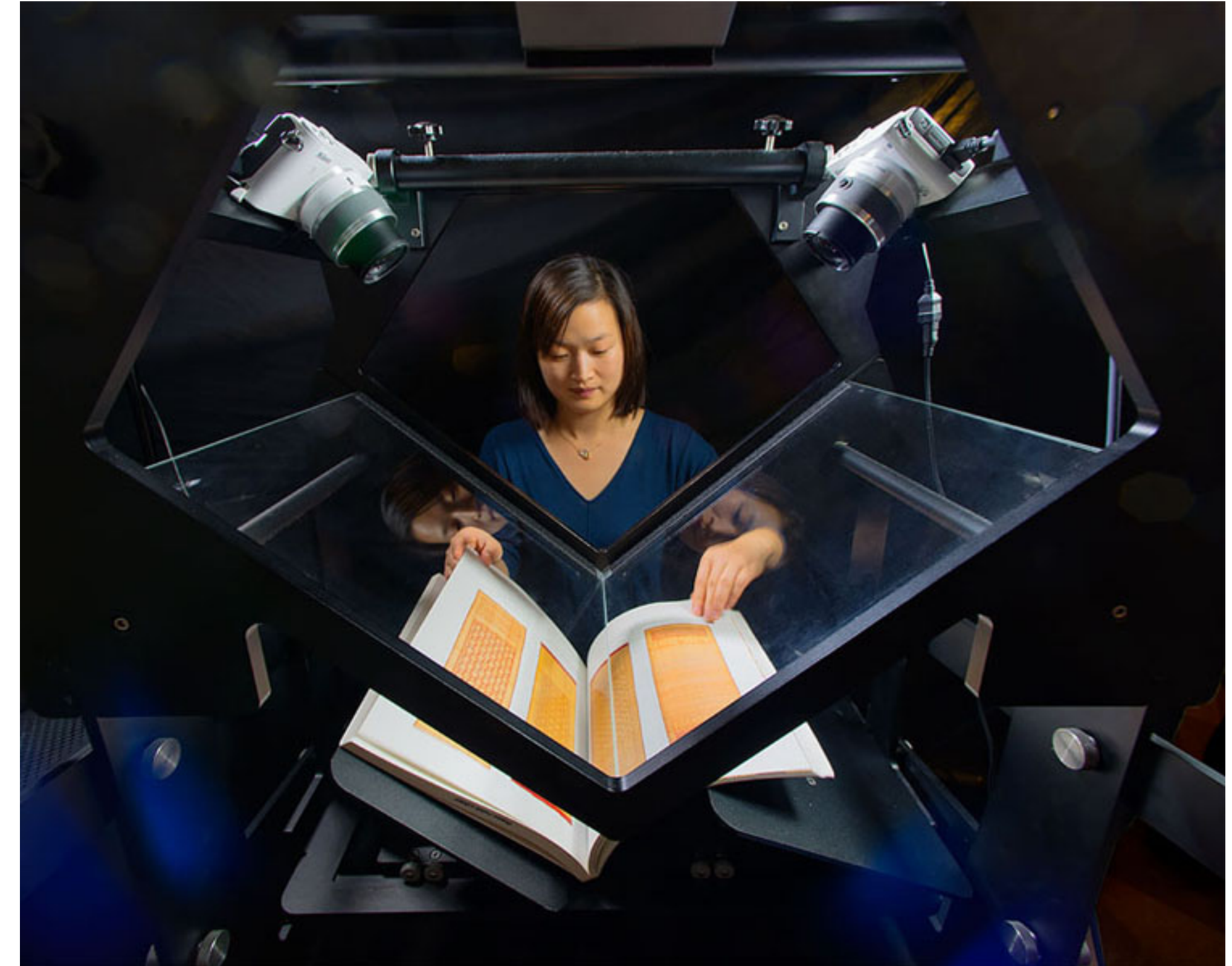
And I started chasing references through digitiaztion projects

POWER HORSE BEHIND THE PROJECT

- TTScribe by Internet Archive



OUR SCAN ASSISTANTS

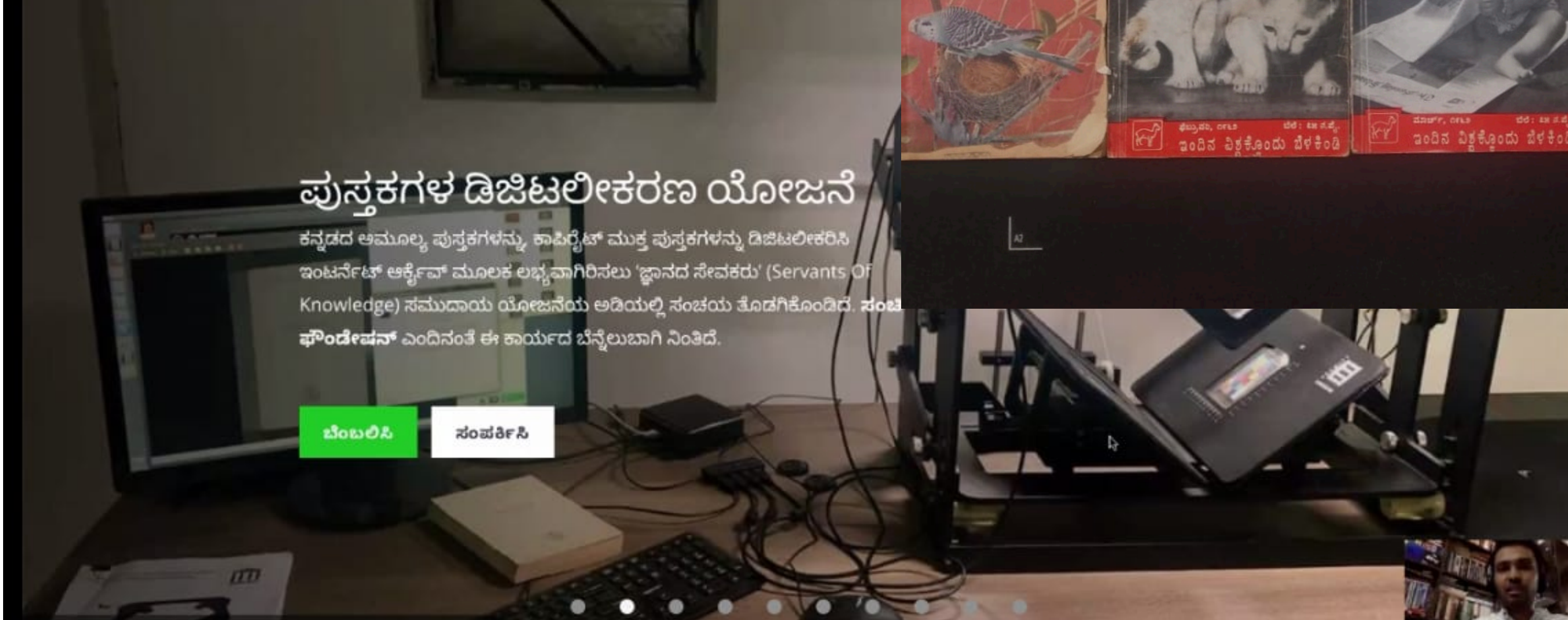
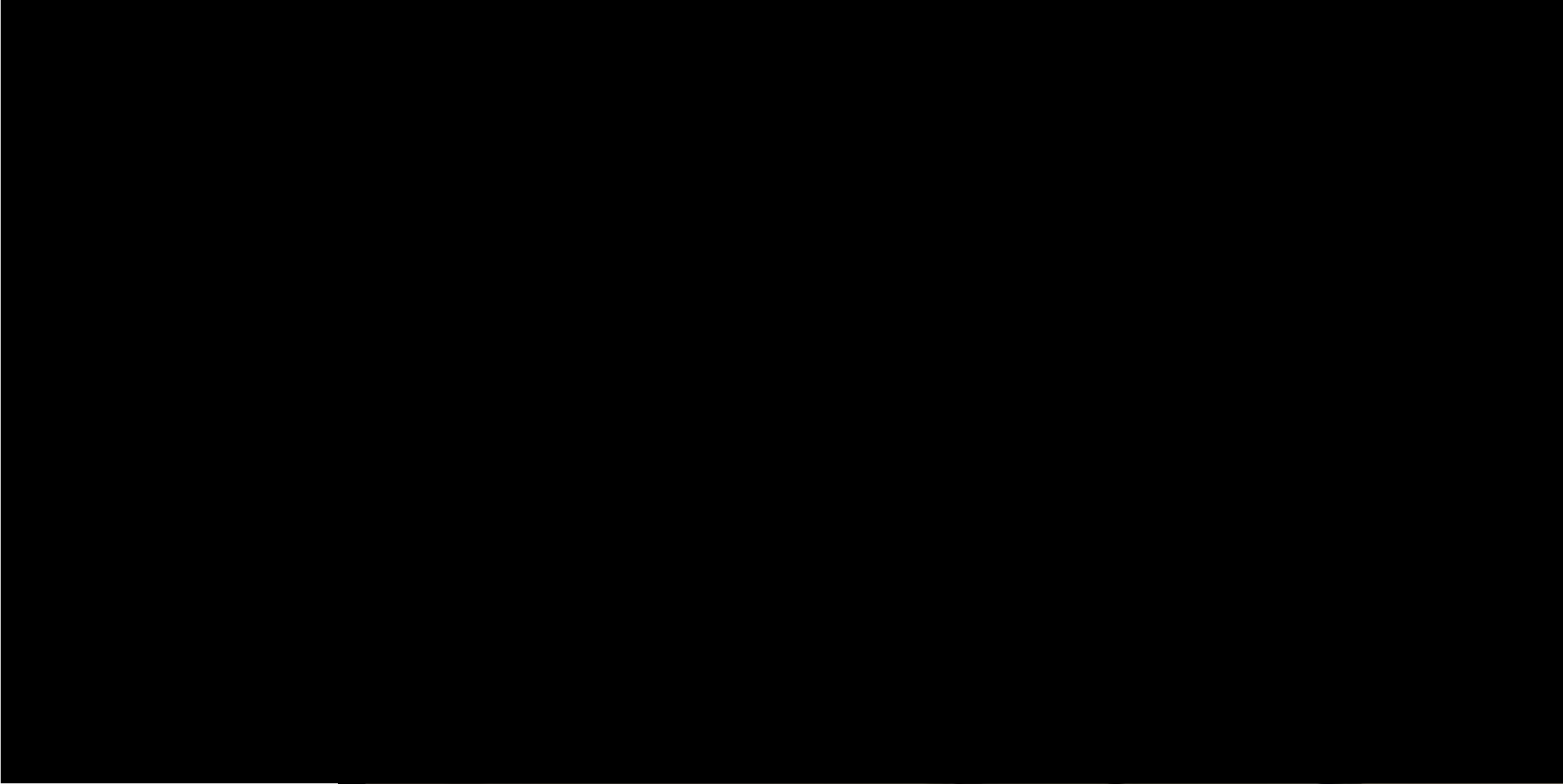


& INTERNET ARCHIVES BACKEND TEAM

HOW WE ARE SOURCING BOOKS

- Individual contributions
- Collaborate with Authors/Publishers to release content under Creative Commons
- Source public domain content via libraries/universities
- Requesting weeded out books from libraries

OUR DIGITIZATION WORKS



ಪುಸ್ತಕಗಳ ಡಿಜಿಟಲೀಕರಣ ಯೋಜನೆ

ಕನ್ನಡದ ಅಮೂಲ್ಯ ಪುಸ್ತಕಗಳನ್ನು ಕಾಪಿರೈಟ್ ಮುಕ್ತ ಪುಸ್ತಕಗಳನ್ನು ಡಿಜಿಟಲೀಕರಿಸಿ ಇಂಟರ್ನೆಟ್ ಆರೈವ್ ಮೂಲಕ ಲಭ್ಯವಾಗಿರಿಸಲು 'ಜ್ಞಾನದ ಸೇವಕರು' (Servants Of Knowledge) ಸಮುದಾಯ ಯೋಜನೆಯ ಅಡಿಯಲ್ಲಿ ಸಂಚಯ ತೊಡಗಿಕೊಂಡಿದೆ. ಸಂಚಯ ಫೌಂಡೇಷನ್ ಎಂದಿನಂತೆ ಈ ಕಾರ್ಯದ ಬೆನ್ನೆಲುಬಾಗಿ ನಿಂತಿದೆ.

ಬೆಂಬಲಿಸಿ

ಸಂಪರ್ಕಿಸಿ

DIGITIZED CONTENT ACROSS FOLLOWING CATEGORIES

digital.sanchaya.net

- Science
- History
- Arts
- Literature
- Classical Text
- Newspapers
- Magazines
- Dictionaries
-

ಸಂಚಯ

ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಹಾಗೂ ಡಾ. ಜ್ಯೋತ್ಸ್ನಾ ಕಾಮತ್ ಅವರ ಡಿಜಿಟಲ್ ಕನ್ನಡ ಪುಸ್ತಕಗಳ ಸಂಚಯ
ಕಾಮತರ ಕಲಸೋಗರ | ಸಂಚಯ | ಡಿಜಿಟಲ್ ಸಂಚಯ

	ಮಧ್ಯಪ್ರದೇಶದ ಮಡಿಲಲ್ಲಿ ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಕನ್ನಡ ಸಾಹಿತ್ಯ ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಅವರ ಈ ಪುಸ್ತಕವನ್ನು ಸಂಚಿ ಫೌಂಡೇಶನ್ (https://sanchifoundation.org) ಹಾಗೂ ಸಂಚಯದ (https://sanchaya.org)		ನಾ ರಾಜಸ್ಥಾನದಲ್ಲಿ ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಕನ್ನಡ ಸಾಹಿತ್ಯ ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಅವರ ಈ ಪುಸ್ತಕವನ್ನು ಸಂಚಿ ಫೌಂಡೇಶನ್ (https://sanchifoundation.org) ಹಾಗೂ ಸಂಚಯದ (https://sanchaya.org)		ಕರುನಾಡಿನ ಜನಜೀವನ ಡಾ. ಜ್ಯೋತ್ಸ್ನಾ ಕಾಮತ್ ಕನ್ನಡ ಸಾಹಿತ್ಯ ಡಾ. ಜ್ಯೋತ್ಸ್ನಾ ಕಾಮತ್ ಅವರ ಈ ಪುಸ್ತಕವನ್ನು ಸಂಚಿ ಫೌಂಡೇಶನ್ (https://sanchifoundation.org) ಹಾಗೂ ಸಂಚಯದ (https://sanchaya.org)
	ಕೊಂಕಣಿ ಭಾಷೆ - ಸಾಹಿತ್ಯ ಮುರಳೀಧರ ಉಪಾಧ್ಯಾ ಹಿರಿಯಡಕ ಕನ್ನಡ ಸಾಹಿತ್ಯ ಈ ಪುಸ್ತಕವನ್ನು ಸಂಚಿ ಫೌಂಡೇಶನ್ (https://sanchifoundation.org) ಹಾಗೂ ಸಂಚಯದ (https://sanchaya.org) ವತಿಯಿಂದ ಇಂಡಿಯನ್ ಅಕಾಡೆಮಿ ಆಫ್		ನಾ ರಾಜಸ್ಥಾನದಲ್ಲಿ ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಕನ್ನಡ ಸಾಹಿತ್ಯ ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಅವರ ಈ ಪುಸ್ತಕವನ್ನು ಸಂಚಿ ಫೌಂಡೇಶನ್ (https://sanchifoundation.org) ಹಾಗೂ ಸಂಚಯದ (https://sanchaya.org)		ಕೀಟ ಜಗತ್ತು ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಕನ್ನಡ ಸಾಹಿತ್ಯ ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಅವರ ಈ ಪುಸ್ತಕವನ್ನು ಸಂಚಿ ಫೌಂಡೇಶನ್ (https://sanchifoundation.org) ಹಾಗೂ ಸಂಚಯದ (https://sanchaya.org)
	ಕಮಲಾದೇವಿ ಚಟೋಪಾಧ್ಯಾಯ ಡಾ. ಜ್ಯೋತ್ಸ್ನಾ ಕಾಮತ್ ಕನ್ನಡ ಸಾಹಿತ್ಯ ಡಾ. ಜ್ಯೋತ್ಸ್ನಾ ಕಾಮತ್ ಅವರ ಈ ಪುಸ್ತಕವನ್ನು ಸಂಚಿ ಫೌಂಡೇಶನ್ (https://sanchifoundation.org) ಹಾಗೂ ಸಂಚಯದ (https://sanchaya.org)		ಪಶು ಪಕ್ಷಿ ಪ್ರಪಂಚ ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಕನ್ನಡ ಸಾಹಿತ್ಯ ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಅವರ ಈ ಪುಸ್ತಕವನ್ನು ಸಂಚಿ ಫೌಂಡೇಶನ್ (https://sanchifoundation.org) ಹಾಗೂ ಸಂಚಯದ (https://sanchaya.org)		ಪರಿಸರ ಪರ್ಯಟನೆ ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಕನ್ನಡ ಸಾಹಿತ್ಯ ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಅವರ ಈ ಪುಸ್ತಕವನ್ನು ಸಂಚಿ ಫೌಂಡೇಶನ್ (https://sanchifoundation.org) ಹಾಗೂ ಸಂಚಯದ (https://sanchaya.org)
	ಸಸ್ಯ ಪರಿಸರ ಡಾ. ಕೃಷ್ಣಾನಂದ ಕಾಮತ್ ಕನ್ನಡ ಸಾಹಿತ್ಯ		ಕರ್ನಾಟಕದ ಶಿಕ್ಷಣ ಪರಂಪರೆ ಡಾ. ಜ್ಯೋತ್ಸ್ನಾ ಕಾಮತ್ ಕನ್ನಡ ಸಾಹಿತ್ಯ		ಮಾನವನ ಅವತಾರ ಪ್ರಕಾಶ್ ಬುರ್ಡೆ ಕನ್ನಡ ಸಾಹಿತ್ಯ

RESULT – IT'S NOW SEARCHABLE & ACCESSIBLE

* We are revisiting our literature

* Learning about our authors

* Learning about publishers from various regions

* Learning about the lexicon, print history

* And a never ending new possibilities opened up

ಪುಸ್ತಕ ಸಂಚಯ alpha ವಿಭಾಗಗಳು ನಮ್ಮ ಬಗ್ಗೆ ಸಹಾಯ ಸಂಪರ್ಕಿಸಿ

1 ಸಂಚಯ ಬೆಂಬಲಿಸಿ Secured by Razorpay

ಅಂತರಜಾಲದಲ್ಲಿರುವ ಡಿಜಿಟಲ್ ಲೈಬ್ರರಿಗಳಲ್ಲಿನ ಕನ್ನಡ ಪುಸ್ತಕಗಳನ್ನು ಹುಡುಕುವುದನ್ನು ಸುಲಭವಾಗಿರುವುದು ಪುಸ್ತಕ ಸಂಚಯದ ಮುಖ್ಯ ಗುರಿ. ಡಿಜಿಟಲ್ ಲೈಬ್ರರಿ ಆಫ್ ಇಂಡಿಯಾ , ಓನ್ಮಾನಿಯ ವಿಶ್ವವಿದ್ಯಾಲಯದ ಡಿಜಿಟಲ್ ಲೈಬ್ರರಿ , ಹಾಗೂ ಕಣಜ ಡಿಜಿಟಲ್ ಲೈಬ್ರರಿಗಳಲ್ಲಿರುವ ಪುಸ್ತಕಗಳನ್ನು ಈಗ ನೀವಿಲ್ಲಿ ಹುಡುಕಬಹುದು.

ಹುಡುಕು

ಪುಸ್ತಕ ಸಂಚಯಕ್ಕೆ ಅವಶ್ಯವಿರುವ ಮಾನವ ಸಂಪನ್ಮೂಲವನ್ನು ಸಮೂಹ ಸಂಚಯದ ಕ್ರೌಡ್‌ಸೋರ್ಸಿಂಗ್ ಮೂಲಕ ಪಡೆದುಕೊಳ್ಳಲಾಗಿದೆ. ನಮ್ಮ ಯೋಜನೆಗಳಲ್ಲಿ ಭಾಗವಹಿಸಲು ಇಚ್ಛಿಸುವವರು ಸಂಚಯ ಸಂಪರ್ಕಿಸಬಹುದು.

ಕಾದಂಬರೀಸಂಗ್ರಹ W

ಈಗ ಪುಸ್ತಕ ಸಂಚಯದಲ್ಲಿ ಪುಸ್ತಕಗಳನ್ನು ಹುಡುಕುವುದರ ಜೊತೆಗೆ ಕನ್ನಡ ವಿಕಿಪೀಡಿಯದಲ್ಲಿ ಪುಸ್ತಕಗಳ ಮಾಹಿತಿ ಪುಟದ ಸಂಪಾದನೆಗೂ ತೊಡಗಿ, ಕನ್ನಡಕ್ಕೆ ಪುಟ್ಟ ಕೊಡುಗೆ ನೀಡಿ (ಮತ್ತಷ್ಟು ತಿಳಿಯಿರಿ) .

ಹಂತ ೧

ನಿಮ್ಮ ವಿಕಿ ಬಳಕೆದಾರ ನೊಂದಣಿ ಖಚಿತಪಡಿಸಿ.

ಹಂತ ೨

ವಿಕಿಪೀಡಿಯಾದಲ್ಲಿ ಸಂಪಾದನೆಗೆ ತೊಡಗಿ.

ಹಂತ ೩

ನಿಮ್ಮ ವಿಕಿ ಬಳಕೆಯ ಮಾಹಿತಿ ದೃಢೀಕರಿಸಿ.

ನಿಮ್ಮಲ್ಲಿ ವಿಕಿಪೀಡಿಯದ ಬಳಕೆದಾರ ಖಾತೆ (ಅಕೌಂಟ್) ಇದೆಯೇ?

● ಹೌದು ● ಇಲ್ಲ

6 3150 2587 10660

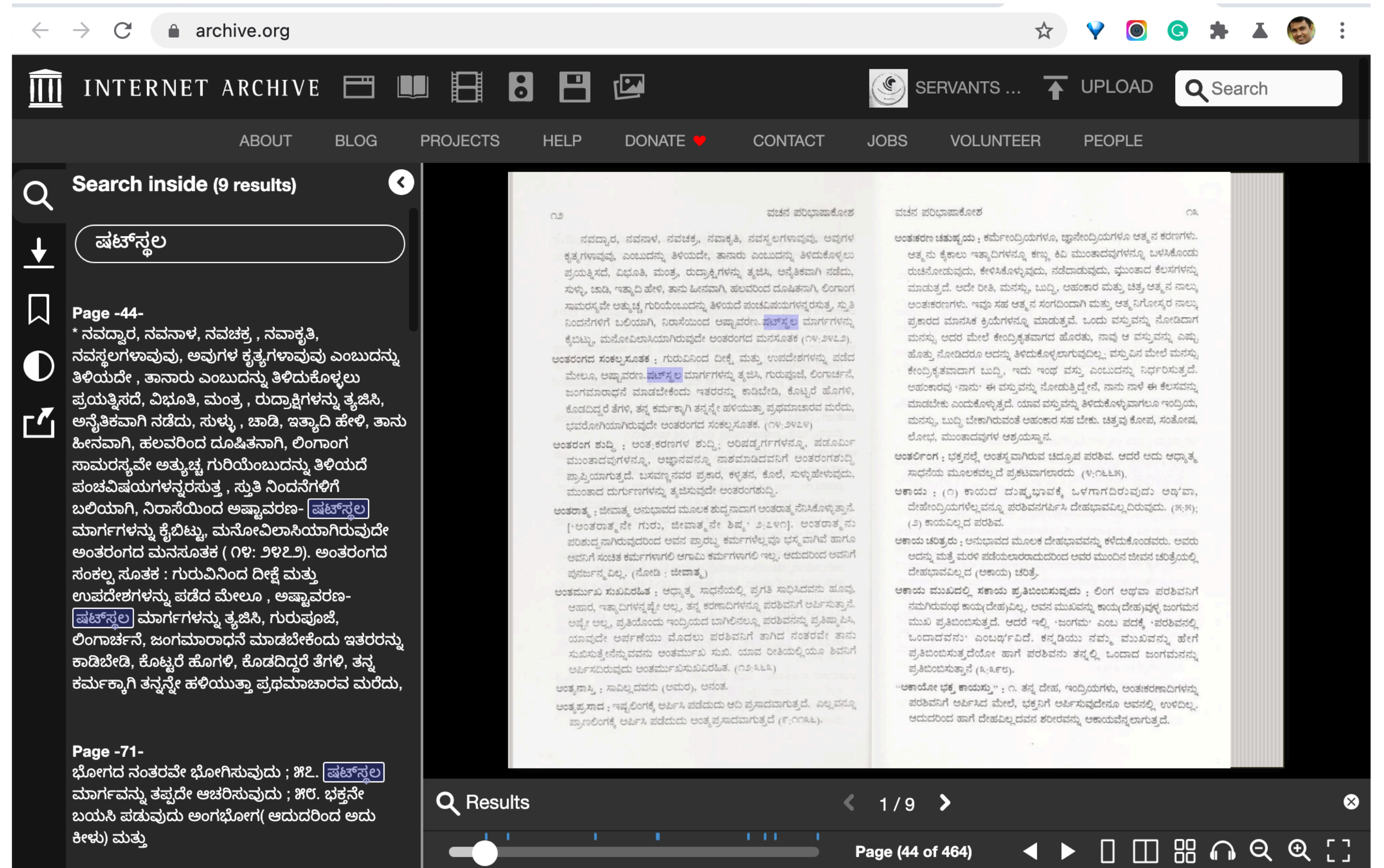
METADATA MANAGEMENT

- Ensure book info is added in both source language & English (Author, publisher, Title)
- Add the Copyright/Licensing / any contributions from the community members

OCR & SEARCHABLE INDIC CONTENT

Thanks to Sushant from IndiaKanoon

<https://github.com/sushant354/egazette>





We are here...

Servants Of Knowledge is currently operating 3 scanning centres

- * *Bengaluru - Indian Academy of Sciences*
- * *Mangaluru - World Konkani Centre*
- * *Chennai - Roja Muthaiah Research Library*

CHALLENGES & NEXT STEPS

- Make use of the reference material largely for Indic wikis - Kannada to start with
- Contribute the metadata to Wikidata
- Finding PD books
- Finding Copyright holders
- Convincing to release under CC-by-SA
- Testing OCR & contributing back to tesseract etc
- Expansion in terms local tech, collaboration

MY OTHER AFFILIATIONS SANCHAYA.ORG & SANCHIFOUNDATION.ORG

The screenshot shows the Sanchaya website interface. At the top, there's a navigation bar with the site name 'ವಚನ ಸಂಚಯ' and various menu items. Below that, there's a search bar and a list of categories. The main content area features a post creation form with fields for title, author, and content. A sidebar on the left displays a bar chart of categories. At the bottom, there's a footer with social media links and a 'MAKE A DONATION' button.

ವಚನಕಾರರು	ಅಂಕ
ಅಕ್ಕಮಹಾದೇವಿ	7
ಅಕ್ಕಮ್ಮ	7
ಅಂಗಸೋಂಕಿನ ಲಿಂಗತಂದೆ	1
ಅರಿವಿನ ಮಾರಿತಂದೆ	2
ಅಲ್ಲಮಪ್ರಭುದೇವರು	9
ಆದಯ್ಯ	3
ಆನಂದಯ್ಯ	3
ಆಯಸ್ತಿ	0

- We built projects to create content to preserve art, history & culture through audio & visual content - intent was to make a rich-reusable-archive for educational purpose.
- Released them under Creative Commons - cc-by-sa license on Youtube, Portal & Internet Archive (350+ hrs of visual content)
- Created partnerships with repertory such as Ninasam, Kamat Foundation etc.

ಧನ್ಯವಾದಗಳು
THANK YOU

Wiki/Twitter/fb/linkedin - @omshivaprakash

omshivaprakash at gmail dot com