Integrated Search and Analytics Platform

For House Proceedings In Printed Books, Video, Newspaper Clippings
For Vidhan Sabha Secretariat, Madhya Pradesh

SeMT- Madhya Pradesh
1st May, 2019
Topic Covered

• Need Assessment
• Solution Drive
• Ease of process
• Technology Strength
  • Digitization solution
  • Bi-Lingual intelligent search engine
  • Crawling (Newspaper) and e Books
  • Data analytics
Need Assessment

• Document-Hardcopies more then 40+ years old records
  • Problem searching records
  • Indexing and searching of house proceedings
  • Search and arrange the questions raised in Vidhansabha Secretariat
  • Relevant information dispersal and transparency.
  • Audio/Video content indexing and searching.
  • Record maintenance.
  • Workflow and tracking issues
  • Organized structure of Question and its replies.
  • Tracking and indexing of questions raised over a period of time.
  • Reporting and Data Analytics
Solution drive

- Intelligent digitization, Annotation, Indexing and Image processing of the House proceedings, Video cassettes and news Paper clippings and various other important books
- Articulating a Data Architecture that meets the storage, archival, retention, maintenance and exchange of data across multiple systems
- Defining a security architecture that stems the proliferation of fragmented and overlapping identity and authorization mechanisms with an organization.
- Platform for OCR and Document Classification
- Platform for information analysis and retrieval.
Ease of process

- Search and indexing Platform
- Data Analytics Platform
- Workflow Management system
- Interactive eBooks
- Crawling (Newspaper)
- Content Management and digitization platform
- Document classification
Digitization Solution

Books, Proceedings + Newspaper Clippings + Video Recordings

Intelligent Search Solution
Digitization Process- Books & Newspaper Clippings

Scan
- Scan Physical Copies
- Scan Newspaper Clippings

OCR
- Image Enhancement
- Multi Format Conversion
- Storage in secure vault

Quality Check
- Correction - Skew, low brightness, ink blots
- Annotate Image for further use

E-book Generation

Request Proceeding Book

Perform quality checks

Scan Image Scan QA

Annotation Annotation Review Image Annotation VS Review
Digitization Process - Video

**Request Video Cassette**
- Digitize Video Digitization
  - Using state-of-the-art techniques we convert Video cassettes to Digital videos

**Video Resolution**
- Performing Video Cleansing and Noise Removal
- Splitting of video and merging
- Storing the Video in a secure digital vault

**Annotation**
- Bookmarks created for the video and attached
- Attaching image annotations to the video to the corresponding topic

**Audio Muting**
- Removal of un-parliamentary words

**Final processed Video**
Bi-lingual Intelligent Search Engine
Search Engine Features

- Simple GUI Interface
- Support for Search by String or Key Persons name or Department or Topics
- On-screen keyboard (INSCRIPT & Phonetic) and Transliteration
- Search query Auto Suggestion
- Advanced search options like Phrase Search, Boolean Search, Date Range
- Dictionary based Spell Correction and alternate spellings in search keywords
- Dynamic facet filtering
- Intelligent Ranking Algorithm which will rank the results according to Year, View Count
- Latest annotated files are auto-indexed and searchable
Search Page

Madhya Pradesh Vidhan Sabha

Search in Proceedings, News crawler or News by Member Name, Key word, Topic etc.

Matches on: All Words Exact Word Any Words

Facet Filter

Relevancy

Advance Search

1 of total result(s)

Search

Reset

Vidhansabha Number:
Select Vidhansabha

Proceeding Date:
Select Date

Section Type:
Select Section Type
Search Result Page

Page, e-book, Video and Audio displayed for the searched discussion/headline

Madhya Pradesh
Sample E-Book Of Proceedings

- E-Book will be created for all the Proceeding books
- Available at a click of a mouse
- Will have options like flip to next page, return back to last page, zoom full screen, print etc.
Search Engine Features

- Auto Suggest words in Hindi, auto completes your words according to popular searches.

- Quickly populates a list of different searches done in the past & saves as a link for effective time savings.
Search Engine Features

- **Synonym**
  - Search for word or phrase that means exactly or nearly the same as the typed search query
  - फसल का नुकसान where नुकसान is matched to क्षति, हानि and relevant results are retrieved

- **Bi-lingual Translation**
  - Search system translates the English search queries in Hindi and retrieves results
  - Madhya Pradesh amendment is translated in Hindi as मध्य प्रदेश संशोधन and searched

- **Homophone**
  - Search system matches the words having same pronunciation and retrieve results for probable homophone matches
  - हिंदी and हिंदी both word will searched

- **Facets**
  - Key persons, locations and department are added as filter
  - multiple filters applied to drill down to specific discussion

Madhya Pradesh
Newspaper Clippings

- Scanning of the available News paper Clippings of various news papers
- Image cleaning and skew correction
- Support for dictionary based tagging.
- Indexing annotated content and tags to search engine.
बैतूल में हाईवे पर धाड़ हवाई जहाज

*Headline*: बैतूल में हाईवे पर धाड़ हवाई जहाज

*Section*: अन्य

*Classification*: 03 : बुराईमति

*Sub-Classification*: "भ.ज.स.क.,/अ.व.म.भ."

*Date of News*: 01-01-2014

*Vidhan Sabha*: बैतूल

*Newspaper Name*: सांविधान

*News Type*: State

*Departments*: 

*Locations*: बैतूल
Newspaper Crawler (RSS - Rich Site Summary Feed)

- RSS feed based online newspapers to be crawled
- Pattern dictionary based parsing and cleansing of content for text extraction.
- Support for dictionary based tagging.
- Indexing parsed content and tags to search engine.
<table>
<thead>
<tr>
<th>News title</th>
<th>Summary</th>
<th>Newspaper Name</th>
<th>Published Date</th>
<th>Link</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>महागाई भंता सरकारी कर्मचारियों का कानूनी अधिकार : हाईकोर्ट</td>
<td>नायाप्रीष्ठ देशवासी कर्मचारियों एवं नायाप्रीष्ठ शेखर बेबी सरफक की खंडपीठ ने ढींटी की दर तप करने का फैसला सेट पर ही होता है।</td>
<td>Jagran Hindi News - news:national</td>
<td>2018-09-01</td>
<td>Local URL not available</td>
<td>Not set</td>
</tr>
<tr>
<td>Rahul wanted China role in his Mansarovar trip: BJP</td>
<td>Seeking to undercut what it sees as Rahul Gandhi’s “Hindu outreach”, BJP claimed on Friday that the Congress chief sought a “see-off” from Chinese diplomats for his Kailash Mansarovar trek. It said the Chinese embassy had asked for three passes for the IGI Airport lounge.</td>
<td>India News</td>
<td>2018-09-01</td>
<td>View news</td>
<td>Not set</td>
</tr>
<tr>
<td>JAGRAN TOP TEN: दीप मुख, आज इन बड़ी सूचनाओं पर बनी हुई है नजर</td>
<td>प्रसिद्ध कैन मूले तरह सागर महाराज का शानिवार सुबह दिल्ली में निभाया दिखा। प्रमुखमंत्री मोदी अब पूरे देश में एक सभा इंडिया पोस्ट पेमेंट बैंक को लोंग करेंगे।</td>
<td>Jagran Hindi News - news:national</td>
<td>2018-09-01</td>
<td>Local URL not available</td>
<td>Not set</td>
</tr>
<tr>
<td>भारत v/s पाकिस्तान: जल संधि से रार तक, आवश्यक है निदेशक के बंदरगाह का नाम</td>
<td>इसके तहत सिंह सन्तान की सहायक नंदियों को पूर्वी और पश्चिमी नंदियों में बंटा गया है। सततताव, स्वास्थ्य व रासी को पूर्वी नंदी में रखा गया है, जबकि ब्रेकज, सेनाब और सिंह को पश्चिम की नंदिया हैं।</td>
<td>Jagran Hindi News - news:national</td>
<td>2018-09-01</td>
<td>Local URL not available</td>
<td>Not set</td>
</tr>
</tbody>
</table>
Data Analytics Platform – ShareInsights
MPVS Analytical Dashboards

- **विभाग विश्लेषण**: Dashboard is useful to understand most important discussions around the department.
- **वार्षिक गतिविधियाँ**: This dashboard display snapshot what happened in a year.
- **सदस्य विश्लेषण**: will give quick glance of the MLA participation in various discussions.
- **प्रश्न विश्लेषण**: shows all the questions, type of question, question count, participant details.
- **ध्यानाकर्षण विश्लेषण**: shows - Headline details, Department, location, number of people participated in the discussion.
- **विधेयक विश्लेषण**: shows Bills statistic- passed, returned, proposed count, People participation, Topics etc. for the referred bill.
- **स्थगन प्रस्ताव विश्लेषण**: shows- Number of Adjournment Motion in a session.
- **अनुदानों की मांग विश्लेषण**: shows Headline details, Demand Number, Demand proposal accepted/rejected, Deduction Proposal accepted/Rejected, number of people participated in the discussion.
- **सारांश**: can give quick glance of the activities conducted in a session.
- **खोज विश्लेषण**: It shows top 10 searched headlines, sessions and on which topics.
MPVS Analytical Dashboards

Madhya Pradesh
Analysis report displays, participation of the members under different sections along with number of occurrence of a keyword in headline. This will give a quick glimpse on what are the popular topics discussed. Example- कार्यवाही is the most popular key word, Indicates that more discussion happened on the topic related to कार्यवाही in Vidhan Sabha
Thank You!
SeMT Madhya Pradesh
Vishal Bangard, Shivam Tripathi
Madhya Pradesh Vidhansabha
Bhopal, Madhya Pradesh