

A Data Visualization Tool for Heritage Buildings in India and Bangladesh: The Project of a MediaWiki-based Platform for ID-SCAPES

Giuseppe Resta, Sidh Losa Mendiratta, Tiago Trindade Cruz

Abstract: The ID-SCAPES research project studies medieval and early modern religious architecture of Christian minorities in India and Bangladesh, often representing contested heritage. A key output will be *Herichurch*, a MediaWiki-based platform designed to present historical entries on significant churches, making information accessible to a broader audience. The Wiki will integrate texts, photographs, 3D visualizations, and drawings for effective cultural heritage dissemination. Drawing lessons from successful case studies, *Herichurch* Wiki intends to explore the connection between a MediaWiki encyclopaedia and a structured database to produce accurate and accessible content about heritage buildings.

Keywords: Digital Humanities, heritage buildings, India, churches, preservation by record.

1. Introduction

In late 2023, the European Research Council approved funding for the ID-SCAPES project, which aims to study and document the medieval and early modern religious architecture of Christian minorities in India and Bangladesh (Fig. 1). These religious sites are often multi-layered and contested heritage, and some buildings have suffered from increasing neglect, erasure and even effacement during recent decades. Furthermore, following the end of colonial rule in those two countries, many within the Church hierarchies sought to distance themselves from colonial legacies, and many of the older churches were radically transformed or completely rebuilt with contemporary designs. Taking into consideration the risks arising from these processes and the progressive erasure of cultural heritage, ID-SCAPES aims to produce a social history of the built environment of India and Bangladesh's medieval and early modern churches and sacral landscapes (built before ca. 1800), including functioning, ruined, and disappeared buildings.

ID-SCAPES challenges the European-centric historiographical framework that has traditionally been applied to these topics. Through extensive

Giuseppe Resta, University of Porto, Portugal, gresta@arq.up.pt, 0000-0001-8489-5291
Sidh Losa Mendiratta, University of Porto, Portugal, smendiratta@arq.up.pt, 0000-0003-2960-8100
Tiago Trindade Cruz, University of Porto, Portugal, tcruz@arq.up.pt, 0000-0001-8792-5142

Referee List (DOI 10.36253/fup_referee_list)

FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Giuseppe Resta, Sidh Losa Mendiratta, Tiago Trindade Cruz, *A Data Visualization Tool for Heritage Buildings in India and Bangladesh: The Project of a MediaWiki-based Platform for ID-SCAPES*, © Author(s), CC BY 4.0, DOI 10.36253/979-12-215-1010-2.22, in Elena Marangoni, Camillo Carlo Pellizzari di San Girolamo (edited by), *Wikidata e la ricerca. Condividere esperienze. Atti del Convegno, Firenze, 5-6 giugno 2025 / Wikidata and Research. Sharing Experiences. Conference Proceedings, Florence, 5-6 giugno 2025*, pp. 175-183, 2026, published by Firenze University Press, ISBN 979-12-215-1010-2, DOI 10.36253/979-12-215-1010-2

fieldwork and the analysis of previously unexplored visual and written documents, this research project promotes a new methodological approach that acknowledges the complex histories of these buildings. It addresses issues such as cultural encounters, caste, religious accommodation, indigenous agency, and local spatial and artistic traditions, revealing how these factors have influenced church architecture.

One of the principal outputs of the project will be *Herichurch.org*, a MediaWiki-based platform intended to provide entries with historical information on a selection of significant churches, accessible to a wider audience. Following the idea of “preservation by record”, the project’s Wiki will combine texts, photographs, 3D visualizations and CAD drawings as fundamental assets for cultural heritage dissemination, conservation and research.

In this paper, we highlight the typologies of data that our project will address. Additionally, we review several comparable projects in the field of digital humanities that have utilized the Wikidata database. Furthermore, we examine the common challenges faced by these projects in relation to ID-SCAPES. Finally, we outline a potential roadmap for integrating the project outputs into an indexed collection of church architectures, organized in a MediaWiki-based platform, with a focus on how they can be published to reach a broader audience.



Figure 1 – Church of Our Lady of Divine Providence, Goa, India (credits: Sidh Losa Mendiratta, 2020).

2. ID-SCAPES structure of knowledge: from photo archives to CAD drawings

The ID-SCAPES project will create a comprehensive collection of datasets focused on historical, architectural, and geographical research gathered from archival sources and fieldwork (Fig. 2). These datasets will facilitate analysis and interpretation of various subjects. A significant portion of the data generated regarding the religious architecture and sacred landscapes of Christian minorities in India and Bangladesh will be synthesized and made accessible to a broad audience through the *Herichurch.org* Wiki. The online resource will serve as a semi-structured georeferenced database in open access that can be collaboratively edited, ensuring interoperability and access to the information sourced from all the project's datasets. *Herichurch.org* and other digital tools created by ID-SCAPES can impact heritage management interventions, especially in the case of endangered or contested sites.

During the project's timeline (2024–29), ID-SCAPES will curate six datasets that will contribute to the *Herichurch.org* entries:

- 1) Photographs (~12000 project photographs taken during previous and forthcoming fieldwork in India and Bangladesh and ~18500 historical photographs acquired from private collections in Europe, South Asia and the USA),
- 2) Cartographic / Iconographic documents (~250 manuscripts and ~100 printed documents acquired from private collections in Europe, South Asia and the USA),
- 3) CAD drawings (~120 two-dimensional visuals and ~30 three-dimensional visuals generated by the project),
- 4) Project map (1 map of georeferenced sites linked with the individual entries on *Herichurch.org*),
- 5) Bibliographic resources (~500 manuscript sources, ~500 printed sources, and ~50 published outputs collected and generated by the project, organized with a reference management software),
- 6) Oral Sources (~50 field work interviews).

Datasets 1 and 2 contain image files in PDF, JPEG and TIFF formats; dataset 3 vector files in DWG and DWF formats; dataset 4 Geopackage files for geospatial information; dataset 5 both ENL library format and PDFs; dataset 6 audio files in MP4 format.

Published outputs and *Herichurch.org* entries will be assigned persistent identifiers (DOIs). Data gathered from archival and private sources will have restricted access, while the Creative Commons Attribution 4.0 International (CC BY 4.0) license will apply exclusively to the project photography, allowing for broad reuse. Datasets 1 and 2 are provided with ~15 metadata fields to describe each item.

For the functioning of the *Herichurch.org* Wiki, each entry is associated with one ID-SCAPES inventory number; geographic information (coordinate location, country, state, district, sub-division, city, locator map image); religious

significance (ethnolinguistic group, parish, past parish, diocese, dedicated to); architectural quality (church in use, conservation status, original typology, present typology, built area, heritage designation); historical facts (past dedication, inception date, rebuilding date, Christian liturgical rite, original religious order); contact information (official website, present religious order, current priest); and licensing information. All these fields will use standardized Wikidata property IDs. Also, the web layout will take into consideration that most internet connections in India and Bangladesh are accessed via mobile devices (Waghmare 2025).



Figure 2 – Sample photo from a fieldwork at the church of Our Lady of Mount Carmel, Goa, India (credits: Sidh Losa Mendiratta, 2017).

To illustrate the *Herichurch.org* entries, we will use photographs, 3D visualizations and CAD drawings generated by ID-SCAPES, and visual documents already existing in Wikimedia Commons. Regarding visual documents from institutions and private collections, the licensing will be negotiated on a case-by-case basis, image by image. This process is coherent with the activities postulated for standard data curation: description, annotation, collection/aggregation, storage, and migration (Sabharwal 2015, 20).

ID-SCAPES intends to organise this knowledge structure through Wikibase, the software for managing structured data developed by Wikimedia Deutschland. The Wikibase Data Model is particularly significant because it allows for the enrichment of the basic semantic triple with references (sources) and qualifiers, which better represent ranges or contrasting perspectives on historical

facts (Bacchi and Bergamin 2018). This applies, for example, when the foundation date of a church can only be defined within a specific interval, as reported by certain authors. Additionally, embedded maps allow for Wikibase projects to be searched both in georeferenced and alphanumeric methods. These methods are currently employed in digital humanities in order to make data as interoperable as possible and knowledge dissemination as wide as possible (Zhao 2023).

3. Wikidata and georeferenced history

The preservation by record approach is preferable where “preservation *in situ* (keeping a site in an unchanged state) is impossible or impractical and involves making a detailed and comprehensive record before the material is lost” (Janik 2014, 112). The research theme of ID-SCAPES is in a geography different from that of the host institution, focusing on a body of cultural heritage that is often associated with colonialism in South Asia. There is a sense of urgency in documenting these sites, given that many of these structures have disappeared, and that present government policies in South Asia, generally speaking, don’t prioritize the research and preservation of minority heritage (Sahgal et al. 2021; Thames and Chaudhry 2014). Thus, studying early modern religious architecture using an innovative digital platform that captures and engages with the complex histories of these significant cultural sites is more effective and fitting.

The project faces typical challenges associated with research practices in digital humanities. Sabharwal (2015, 31–35) reviews the criticism that was voiced after the publication of the first *Digital Humanities Manifesto* in 2008, addressing the renewed importance of archives for the humanities. As shown with ID-SCAPES’s structure of knowledge, the project’s archive spans across multiple formats and modes of systematization.

Flanders and Muñoz (2011) outline three methods for data treatment in the humanities: interpretive layering, data capture and preparation, and capturing scholarly agency, each reflecting different levels of data curation. These methods emphasize the importance of documentation, curatorial decisions, and the contextualization of original works to encourage new interpretations of sources. Focusing on historical phenomena, access to sources has evolved from being limited to a few scholarly publications to now offering ubiquitous access to extensively digitized collections. However, such digital ephemera cause “implications for archives and libraries” that warn “about data loss due to crises, lack of control, or emerging political landscapes intolerant of open knowledge” (Sabharwal 2015, 51). The cyberattacks on the *Internet Archive* that rendered the digital library unavailable between May and October 2024 (Wikipedia contributors 2025), and on the British Library in October 2023 are two examples of such crises.

In the archive sector, Wikidata has proven to be an essential tool for expanding the infrastructure of linked archived data. Cultural heritage institutions, such as the Museum of Modern Art, the Library of Congress, and the Smithsonian, have successfully established a two-way integration of their authority data with Wikidata. This integration benefits both their institutional websites and allows

for open access to a portion of their collections through Wikipedia (Hawkins 2022). The same drive is observed in recent research projects.

As shown in the analysis by Zhao (2023), Wikidata for Digital Humanities “is evolving as a linking hub of data” in a direction that can integrate original datasets with existing ones. For instance, the COURAGE project faced the issue of matching pre-existing entities of people, groups, and events pertaining to political opposition in the Socialist era, with the possibility to link project-specific individuals or organizations to a wider network of authority IDs (Faraj and Micsik 2019). Reducing data redundancy was one of the main issues that led to the release of Wikidata in 2012. ID-SCAPES will similarly face the problem of integrating some existing data on Indian and Bangladeshi churches with more granular information found in private and public collections. Rossenova, Duchesne, and Blümel (2022) analyze two ways of linking Wikidata and independent Wikibases in cultural heritage research projects. One can prefer the networked, standardized option of Wikidata, as in the case of the *DigAMus goes Wikidata* project (von Heyl, Thiel, and Schlott 2025), where items or properties are already present in Wikidata and only need curation. Alternatively, a private Wikibase containing customized historical information can be interlinked with Wikidata entries, as in the *Corpus der barocken Deckenmalerei in Deutschland* project (Vogel and Moises 2021), in which archival material, 2D drawings, photos, and 3D models datasets complement standard Wikidata items such as places and historical figures. This second approach better describes how the ID-SCAPES structure of knowledge could be ideally rendered.

Another relevant nucleus of reflection for the project is the mapping of geo-referenced information linked to the project’s Wiki. Many of the buildings under research have disappeared, are in ruins, or have been deeply transformed, and there is a sense of urgency in documenting whatever structures remain. This leads to four possibilities:

- The contemporary building, in use or ruins, is the same as described in sources from Datasets 2 and 5, and its location is certain,
- The contemporary building represents a significant alteration of the original structure that can be found in sources from Datasets 2 and 5, with its location also being certain,
- The original building left no traces, but Datasets 1, 2, and 5 allow geographical cross-examination to pin an approximate location,
- The original building left no material traces and only generic positional information in Dataset 5, i.e. the name of the parish. The location can only be identified with an area and referenced to literary or oral sources.

The work involving spatial approximations and uncertainties aims to balance scientific reliability with the need to provide consistent information across the set of ID-SCAPES architectural example. The authors have previously experimented with methods from Historical GIS to map buildings that no longer exist (Resta 2025). This was done in parallel with the cross-examination of historical visual documents to support fragmented information from written sources

(Mendiratta, Resta, and Cruz 2025; Mendiratta, Resta, and Reis Leite 2024). The integration of GIS with the semantic web in a web GIS suite interlinks archival documents in a mapping environment. The *Ghent Mapped* project clearly illustrates the challenges mentioned earlier. It provides “a public history project that spatially aggregates and maps cultural heritage records—predominantly movable heritage—from seven Ghentian heritage organizations using a custom web GIS” (Ducatteeuw, Danniau, and Verbruggen 2025). Diverse datasets, depicting both national landmarks and everyday places with material provided by the communities, are organized in a web map featuring a timeline that spans from 1550 to the present day. The authors have represented those places that have disappeared using an approximate polygon shape derived from historical documents. Boundaries are defined as the spatial domains that interact with each specific building (Ducatteeuw, Danniau, and Verbruggen 2025). In ID-SCAPES, one church without a specific location can be coupled with the corresponding religious jurisdiction that is usually available in registries. Alternatively, this spatial domain can correspond to an area whose offset reflects the degree of certainty that sources allow in each case, constituting one additional parameter to be associated with the geocoordinates.

4. Herichurch roadmap

At present, the development of the *Herichurch.org* Wiki is in its initial stage. The research team is considering the best strategy for making the project’s findings easily accessible. The goal is to create a practical tool that can be used on-site and visualized on a smartphone. As previously discussed, two main concerns for the forthcoming Wiki are how to link the project’s data with existing Wikidata entries and how to georeference historical records. These decisions pass through a thorough study of best practices in the sector of DH that ID-SCAPES initiated in March 2025. All the datasets presented in this paper are being consolidated in the first years of the project, while the elaboration of the Wiki will start in 2026. This allows time to define the appropriate metadata structure that better integrates with the tools to be employed.

Some degree of connection with Wikidata entities is necessary to ensure interoperability and the reusability of the Wiki output by third parties such as research institutions, administrations, and associations.

The analysis of *eViterbo*, which is the MediaWiki-based research output of the *TechNetEMPIRE* project, provides significant insights. This project focuses on the network of individuals—such as engineers, architects, cartographers, carpenters, and masons—who contributed to the formation of the Portuguese colonial empire. For this reason, it is adjacent to ID-SCAPES both in content and methodology. The limitations and challenges addressed in *eViterbo*’s global report (Faria and Correia 2022) could be a starting point to set the *Herichurch* Wiki in the right direction. We believe that several issues should be addressed in the initial phases of the project. These include the project’s infobox template script, the alignment of the project’s properties with Wikidata, redundant prop-

erties in data ontology, the implementation of *WikibaseQualityConstraints*, and proper data modeling.

Entries will likely link to existing entries in Wikipedia, *eViterbo*, and *Heritage of Portuguese Influence* (HPIP). HPIP is an online encyclopedic resource that maps and describes the architectural and urban heritage of Portuguese influence in various regions around the world. This project was initiated in 2007 by the Calouste Gulbenkian Foundation, and published as a four-volume work in 2011.

In summary, *Herichurch.org* Wiki intends to explore further this connection between a MediaWiki-based encyclopedia and a structured database to produce accurate and accessible content about India and Bangladesh's medieval and early modern churches and sacral landscapes. In this initial phase, significant efforts are devoted to designing three aspects of the authoring and publishing platform: integrating a selection of existing data on Indian churches with more granular information found in private and public collections; setting a privately hosted Wikibase containing customized historical information interlinked with Wikidata entries; mapping georeferenced historical records linked to the project's findings. To align with the spirit of Wikimedia projects, analyzing successful case studies and setbacks from similar experiences was crucial for taking a more informed and strategic approach to develop the *Herichurch.org* project.

References

- Bacchi, Cristian, and Giovanni Bergamin. 2018. "New Ways of Creating and Sharing Bibliographic Information: An Experiment of Using the Wikibase Data Model for UNIMARC Data." *JLIS: Italian Journal of Library, Archives and Information Science* 9 (3): 35–74. <<https://doi.org/10.4403/jlis.it-12458>>.
- Ducatteuw, Vincent, Fien Danniauw, and Christophe Verbruggen. 2025. "Mapping Ghent's Cultural Heritage: A Place-Based Approach with Web GIS." *International Journal of Digital Humanities*. <<https://doi.org/10.1007/s42803-025-00099-4>>.
- Faraj, Ghazal, and András Micsik. 2019. "Enriching Wikidata with Cultural Heritage Data from the COURAGE Project." In *Metadata and Semantic Research*, edited by Emmanouel Garoufallou, Francesca Fallucchi, and Ernesto William De Luca, 407–18. Cham: Springer International Publishing.
- Faria, Alice Santiago, and Rute Correia. 2022. *eViterbo Global Report. Outlines on the Creation of a Research Open Platform Using MediaWiki and Wikibase*. Zenodo. <https://doi.org/10.5281/zenodo.7380114>.
- Flanders, Julia, and Trevor Muñoz. 2011. "An Introduction to Humanities Data Curation." *DH Curation Guide: A Community Resource Guide to Data Curation in the Digital Humanities*. <<https://archive.mith.umd.edu/dhcuration-guide/guide.dhcuration.org/glossary/intro/index.html>>.
- Hawkins, Ashleigh. 2022. "Archives, Linked Data and the Digital Humanities: Increasing Access to Digitised and Born-Digital Archives via the Semantic Web." *Archival Science* 22 (3): 319–44. <<https://doi.org/10.1007/s10502-021-09381-0>>.
- Janik, Liliana. 2014. "'Preservation by Record': The Case from Eastern Scandinavia." In *Open-Air Rock-Art Conservation and Management: State of the Art and Future Perspectives*, edited by Timothy Darvill and Antonio Batarda Fernandes, 112–24. New York-London: Taylor and Francis.

- Mendiratta, Sidh Losa, Giuseppe Resta, and Antonieta Reis Leite. 2024. *Rediscovered Visual Documents of the Portuguese Estado da Índia: The Philippe Nery Xavier Collection*. Coimbra: Universidade de Coimbra, Centro de Estudos Sociais.
- Mendiratta, Sidh Losa, Giuseppe Resta, and Tiago Cruz. 2025. "The Fall of Portuguese Jaffna in a Rediscovered Visual Document." *The Ceylon Journal* (3).
- Resta, Giuseppe. 2025. "Mapping the Invisible: Problems of Interpretation and Representation of Hormuz Island, Iran." *Athens Journal of Architecture* 11 (2): 201-32. <<https://doi.org/10.30958/aja.11-2-5>>.
- Rossenova, Lozana, Paul Duchesne, and Ina Blümel. 2022. "Wikidata and Wikibase as Complementary Research Data Management Services for Cultural Heritage Data." *CEUR Workshop Proceedings*.
- Sabharwal, Arjun. 2015. *Digital Curation in the Digital Humanities: Preserving and Promoting Archival and Special Collections*. Waltham, MA: Chandos Publishing.
- Sahgal, Neha, Jonathan Evans, Ariana Monique Salazar, Kelsey Jo Starr, and Manolo Corichi. 2021. *Religion in India: Tolerance and Segregation*. Washington, DC: Pew Research Center.
- Thames, Knox, and Sahar Chaudhry. 2014. "Shrinking Religious Freedom in South Asia." *Foreign Policy*, 30 April, 2014. <<https://foreignpolicy.com/2014/04/30/shrinking-religious-freedom-in-south-asia/>>.
- Vogel, Evelyn, and Jürgen Moises. 2021. "Blick nach oben." *Süddeutsche Zeitung*. <<https://www.sueddeutsche.de/muenchen/bayern-virtuelle-sehenswuerdigkeiten-dek-kengemaelde-1.5184036>>.
- von Heyl, Anke, Sonja Thiel, and Zahia Schlott. 2025. "Datenlage gut – Kooperation zwischen NFDI4Culture und dem DigAMus Award." NFDI4Culture. <<https://nfdi4culture.de/de/nachrichten/good-data-situation-cooperation-between-nfdi4culture-and-digamus-award.html>>.
- Waghmare, Abhishek. 2025. "Access to Phones and the Internet." *Data for India*. <<https://www.dataforindia.com/comm-tech/>>.
- Wikipedia contributors. 2025. "Internet Archive." *Wikipedia, The Free Encyclopedia*. <https://en.wikipedia.org/w/index.php?title=Internet_Archive&oldid=1282670714>.
- Zhao, Fudie. 2023. "A Systematic Review of Wikidata in Digital Humanities Projects." *Digital Scholarship in the Humanities* 38 (2): 852–74. <<https://doi.org/10.1093/llc/fqac083>>.