Setting the Stage: Enabling Spaces for Dialogues with Ibali Digital Collections UCT

Sanjin Muftić
University of Cape Town
sanjin.muftic@uct.ac.za

Abstract

In 2021, the Digital Scholarship Services (DSS – formerly Digital Library Services) department at the University of Cape Town Libraries launched a university-wide platform for showcasing digital collections: https://ibali.uct.ac.za. The site, called Ibali (Xhosa for story), runs on a set of semantic web technologies Omeka S and IIIF. Ibali is part of the UCT Libraries drive to nurture an Open Access space where digital collections can be created, curated, published, and showcased. It is a collaborative, flexible, and future-oriented repository that supports Digital Humanities projects. Since its launch, Ibali has hosted diverse sites, including heritage resources on endangered South African languages, a theatrical research video archive, a repeat landscape photography database, student-led curations, and selected library collections.

Omeka S, through its structure and modules, enables interactions between showcase sites, creators, and visitors. Semantic technologies within metadata allow conversations and feedback to be integrated into the site, enhancing archives and enabling multiple voices and interpretations. The digital showcase becomes a stage for stories, embracing the transactional, transformative, and migratory nature of images, events, and recordings of our collective memories.

This paper outlines Ibalis approach to fostering dialogue and participation, opportunities in Omeka S metadata infrastructure, and sustainability challenges,

1 Ibali and Omeka S

The main architecture of Ibali is the open-source software Omeka S. Omeka S is a web publishing platform for GLAMs (Galleries | Libraries | Archives | Museums), designed to create relationships between objects in collections as well as describe them through linked open data resources on the internet. Omeka S is a piece of open-source software developed and maintained by the Corporation for Digital Scholarship and Roy Rosenzweig Center for History and New Media. It is a web-publishing platform built specifically for those seeking to showcase heritage collections online. It is maintained by a dedicated team who provide regular updates and further supported by a large and active user community. Omeka S aligns itself with the Centres mission to use digital media to democratize history (Roy Rosenzweig Center for History and New Media, n.d.). The S in Omeka S stands for semantic, as in connecting to the semantic web, where data in web pages is structured and tagged. Its primary focus is on organising elements of a collection such that the links in between items and the greater elements of the internet are strengthened, allowing for much more relevant searches and deeper explorations.

In simple terms, Omeka S requires that any object uploaded to the platform is described using customizable metadata fields (Omeka Team, n.d.). This metadata is stored in a database and linked to associated media files, enabling structured display when visitors navigate a collection site. The structured display is suitable for both humans viewing the site and "machines processing the data. Building on this foundation, Omeka S offers several appealing features for showcasing collections. One is the balance between required and flexible metadata, allowing collection curators to create templates suited to different item types or entire collections. Another is the clear separation between the ingesting layer, where items and media are uploaded and described, and the presentation layer, which focuses on designing website pages to host the curated content.

Omeka S is also structured in a modular way so that there is a core offering of the software, which provides the basic functionality. This is then en-

hanced by a variety of modules[1] which can be installed on top of the base installation and integrated easily with the main software. In going along with the Open-Source model, all these modules are free and developed by either the core Omeka S team or various other developers around the world who make it accessible for others to download and install. The modules offered range from advanced search functionality, customized design, enhancing online discoverability, and even, more recently, modules that directly leverage AI tools for metadata generation from uploaded media.

The UCT Libraries instance of Omeka S, under the Ibali banner, has been implemented through a partnership with the Universitys Information Technology (ICTS) division and has been branded Ibali. While one of the motivations for Ibali was to showcase some of the Libraries own collections, we also saw a demand from various of the universitys research groups and researchers who wanted to deposit and showcase their own collections. Our aim was to provide an institutional home for these, together with curation support, while also eventually being able to leave the regular updating of the showcase site in the researchers own hands.

2 Building a Collection

The process for building a collection site on Ibali is interactive, as it engages both the curator and DSS in a series of consultations to prepare the data and its metadata for the platform. In this period of consultation, several meetings are held with the purpose to understand how the collection of data can best be showcased through Omeka S.

A questionnaire kicks off proceedings that asks the curator to consider some of their desired output:

1. What kind of media items do you have to upload in your collection?

2. What other kind of non-media "items" are frequently referred to by your items?

3. How do all the items connect to each other?

4. For each of the item types, how would you describe them?

5. How would you group your items to assist with managing them?

6. Are your descriptive terms shared across your items?

7. Describe how you would want your site to look/work?

Notice that only the final question is about the look and feel of the final site. However, it is possible to work backwards and start with question seven and trace back the curator to their data and metadata, unpacking the organisation in a way that assists the presentation of the collection.

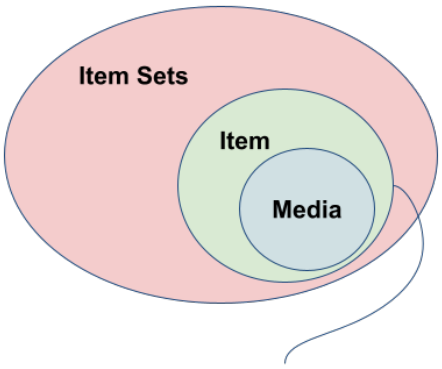
Question two, around the presence of non-media items, can make a huge difference in how the collection is showcased. It also connects to the subject of the collection and the importance of their place on the site. For example, if a collection consists of several documents relating to relationships between people, places, and events, it would be prudent to consider those non-media items as items deserving of their own entry into the Omeka S database. This way each of those items would get its own page and metadata structure, allowing themselves not only to be described but also linked to from the various documents that reference it. A site that showcases works of art on display at a university would benefit from treating artists as items, so they get their own item page and link out to their artworks.

3 Metadata Consultation

This leads to another major consultation point in the building of the showcase site, the discussion around the metadata fields and schemas to use to describe the various types of items. While on Ibali we provide certain generic templates that use mostly Dublin Core metadata schema to describe media objects, there is an incredible amount of freedom in how the rest of the metadata can be set up. As Omeka S can be connected or have several recognized ontologies imported, the properties from these ontologies become selectable to the template a curator wants to set up for a particular item. This way, a template can comprise of fields from Dublin Core, schema.org[2], and others to capture metadata in a structured format. No matter what template is used, the metadata remains in a structured format that can be exported and mapped out later. The consultation process continues with discussions around this metadata, especially if the curator who

[1] The official module list with downloads can be found on the official Omeka Site, while a more thorough one is found on the developers GitHub.

[2] While Dublin Core was developed for libraries and archives, schema.org is a vocab for structured data on the internet.



Resource Template

Metadata Property_1	value
Metadata Property_2	value
Metadata Property_3	value
...	...

Figure 1: Omeka S breaks up resources into media, item and item-sets, with each being able to be described with customized metadata templates based of standard onotologies

is asking to build a showcase site wants to use the metadata to draw a particular argument in their curation. Most significantly, it is to ensure that within the collection there is a consistency around the use of the various metadata templates to allow visitors to navigate the new showcase collection in a way that makes it easier to discover connections.

Within the metadata consultations, there is also a discovery of what possible modules can be used to highlight these collections, and as such, this gets closer to looking at the overall design of the site. For example, formatting dates in a standardized YYYY-MM-DD format allows for an Omeka S module to build timelines, which could highlight the temporal nature of the events or materials in the collection.

Through the consultations with DSS staff, the intricacies of the collection are made visible and then applied into the nature of how the collection will be described and showcased.

The next phase, after the templates have been set up and the collection mapped out, is to begin the process of ingesting items. This can be done in one of two ways either through using spreadsheets pre-formatted with the metadata property names as headings[3], or through the manual creation of items

[3]The CSV Import module allows curators to batch ingest metadata and links to media through the import of a spreadsheet.

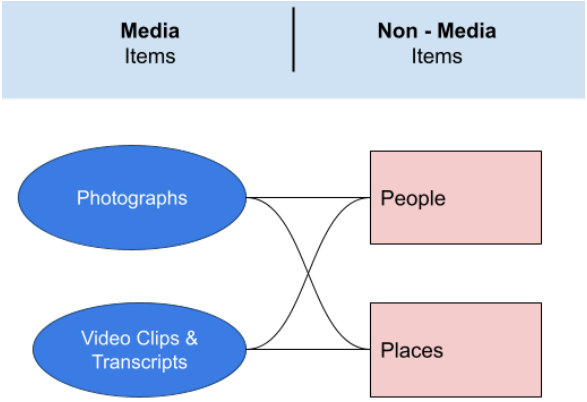


Figure 2: A sample collection Metadata Map that demonstrates the links made between media items and non-media items

on the platform itself. Each one has its own characteristics that are beneficial. While the manual process of adding items is time-consuming, it can allow the curator to become more comfortable with the platform and build the connections between the items in real-time. Working in spreadsheets on the other hand, can make the ingest process quicker because many items can be imported at once. However the order of import becomes important as the various connections between items need to be hardcoded into the import of various item types. For example, in the case of where certain documents reference individuals, it would be necessary to import the individuals first, retrieve their unique Omeka S identifier, and use that in the field that connects the two. This allows the software to build the connections automatically. It is only once the items have been loaded that the phase moves to the design of the site, looking at the presentation elements. On top of this the curator might want to add more general narrative, context, or argument pages that can be constructed in the same way as most other web publishing platforms. What separates Omeka S is that through the pre-loading of the items, any reference can connect to the item that is sitting within the collection now on the database.

4 Metadata Mapping

Within the curation process of each consultation, translating a collection diagram into how the collection will be practically set up on the platform is a significant task, one that determines the engagement with the site and showcases the desired

durational slant the curator would like to have. Simultaneously, for the custodians of the platform working at Digital Scholarship Services, our vision is to enable visitors to search all the collections at the same time, collating all the various digital items. For this to happen, the metadata fields that are used across the various collections need to be close enough or sticking to structure so that a thread of consistency runs through the database, allowing visitors to filter, facet and search fields that contain similar types of information. In the case of the Dublin Core field for Type, it would be appropriate if every collection that uses that field to identify the type of resource according to the class element of Dublin Core - a controlled vocab. They can still use another metadata field if they want to be more specific about the sub-type (Miller, 2022).

To achieve these, DSS has set up certain key metadata templates which have rules for what should be captured in them. These templates are shared with any collection curators so they can decide how they want to describe their items, which of these fields they can use to be beneficial both for them and the entire platform, and which extra fields need to be added to further describe the collection as they see fit.

For further description, Omeka S allows the import of various ontologies onto the system. These ontologies then add their respective metadata fields as available fields to use on the platform. While Omeka S comes with a few standard ontologies, we have setup a few extra ones ourselves, including the previously mentioned schema.org which offers a smorgasbord of fields to be used that fit within how the web shares its data, and Darwin Core which is used when describing flora and fauna.

When building templates for their items, Omeka S curators are also allowed to rename the property for their site. So that one can pick a property that might be out of practical use for the system and rename it into a property that is more relevant to them. This preserves the structure of the metadata, and allows the curator, through the template, to be aware of these adjusted fields to make searches and issues easier.

5 Metadata Consistency

Being consistent within metadata is another ingredient in a successful collection site that will make it easier for visitors to navigate the resources. For this purpose, Omeka S module of Custom Vocab allows a curator to predefine a set list of terms or items that can be used as a dropdown when filling out metadata to ensure the consistency of what is being entered. This bears further fruit down the line because it allows the system to recognize the shared keywords and allow visitors to jump across to similar items which might share those same key words. This is done through the Metadata Browse module, and this ensures that out of any page there are links to take you to another section without having to go back and forth.

In the building of the showcase site of the ǂKhomani | Hugh Brody collection, we have leveraged the use of the custom vocab, to assist the researchers involved in the collection to set up their keywords. While we uploaded a range of images and videos with some initial keywords, based on previous curation work, we have given the researchers edit access. However, we have also set up specific drop downs for different categories of keywords, to enable them to select directly from a controlled vocab list one dedicated to generic keywords, but others also including the names of the people and places as well as cultural events that might be mentioned in the relevant media.

For the standardization processes, when setting up metadata templates, it is possible to specify that type of metadata that will be going into the system. Among being able to indicate if the type is going to be a text, or date, it is also possible to connect directly to Linked Open Vocabularies[4], so that the metadata gets aligned to these internationally recognized standards. For example, taking the type metadata field and linking it to the Dublin Core terms, or taking a Spatial Coverage field and allowing it to dropdown from the Geonames[5] database which further encourages standardization.

6 Working with Languages

Including the perspectives and stories of languages is part of working towards greater inclusion. The ability to present information in various languages is almost second to the fact that the information should also be initially presented and foregrounded in the metadata and site creation. On the African continent, dealing with legacies of colonization and

[4]Linked Open Data vocabularies are published term sets with semantic links that make data machine-readable, for example a birth date for a person in schema.orgs would link to their birthDate field.

[5]GeoNames is an example of a Linked Open Data vocabulary that provides structured information about geographic locations worldwide.

disrespect of indigenous languages this takes on even greater importance.

Within Omeka S, every metadata field gives an option to indicate the language being used. This allows for multiple entries within a field, and each one with a corresponding language tag. These tags become useful when building multi-lingual showcase sites, allowing the presentation, and filtering of only the metadata in the requested language thus only presenting one language at a time at the press of the button.

However, there are times when the multiplicity of languages needs to be showcased. Such was the case in the work of the San and Khoi Centre within UCT who had digitized *Of the Same Breath: Indigenous animal and place names* compiled by Lucy A. Möller, which captures the names of local animals in the multitude of indigenous languages of the Southern African region. This was requested to be presented alongside each other to highlight the rich heritage of the language. The names of these animals were broken down based on language groups according to regional distribution identified by Dorothea Bleek in 1956. What the Custom Ontology module allowed us to do was to capture the breakdown of the various language groups into a metadata ontology that we could load into the system, creating metadata properties for each of the various languages. This indeed allowed us to create items for each of the animals and present their metadata of their names in the various languages in a structured manner to possibly allow others to work with it, but at the same time allow a visitor to view the information in one go. We could then also use these fields to search for all the words in particular languages.

While the presentation of this information could be better served in a digital dictionary, at the same time, Omeka S has the flexibility to do something similar within the digital showcase presentation, and yet also keep the information structured for future reuse. The experiments with the language ontology led us to set up something similar for languages in the greater South African region.

7 Interactions with Forms

The Collecting module encourages dialogue by enabling curators to set up pages to collect objects and metadata from visitors through forms. When creating such a form, the site curator is required to immediately link the various fields of the form to align with metadata fields in the Omeka S system, and specifically their own metadata schema. What the visitor enters is then immediately mapped across and placed as items in the database, with the metadata fields of the form correlating with fields in the system. The item that is submitted via the form does not become public, but sits on the backend of the items of the database waiting for the curator of the site to take further action.

While these forms can be used for simple comments and observations or questions, another way to leverage them is for public contributions to the showcase site. For example, if a visitor might have something that aligns itself with what is being presented and wants to have the curators include it on the site, the form would allow them to complete a lot of the desired metadata ahead of time, saving valuable curation time from the admin side in collecting certain information after the fact.

On the Ibali platform, this feature is active on several sites. The Climate Change Library is a site dedicated to showcasing the open access educational resources which can be used by school teachers to discuss climate change with their students. The original curators had set up a detailed metadata schema for capturing important metadata about each resource to make it easier for teachers visiting the site to find what they need for their classroom. Some of the metadata includes main things such as the target audience of the resource, as well as the type of resource, but then also goes on to include what sustainable development goal is most aligned with the resource.

When setting up the public facing form, which asks for submissions to be added to the library, the curators of the site simplified the form to include only the most basic information such as creator, target audience, etc. The workflow being that once a visitor has submitted a resource, one of the curators would go through the submission and add the extra metadata needed by reviewing the resource and then publishing it on the site in line with the rest of the items.

The Collecting module can thus be harnessed to create community dialogues. It is also possible to have the visitors select in the form, a particular item on the site that they want to make a comment about, or add metadata, or correct a mistake. This eliminates the need to send emails with weblinks to the curators but rather opens a more direct line of conversation that is hosted on the platform. A simple way that the Re-imagining Tragedy in Africa

and the Global South site on Ibali uses this is by crowd-sourcing performances of Greek Tragedy on the African Continent. As the ReTAGS project is eager to uncover the connections between Greek Tragedy and the African continent, it attempting to collate the various performances of adaptations of Greek Tragedy on the continent. It has thus set up a form open to the public where they can submit performances of Greek Tragedy that they are aware of, and which are not already on the site. When submitting information about the performance, the visitor is not required to type in the Greek text the performance they watched is based on but rather selects from a drop-down list of Greek classical texts. Through the mapping of the form with list of plays in the database, this connection is immediately made, and it simplifies the work of the curator in sorting out the metadata before making the item public and adding it to the list of Greek Tragedy inspired performances on the continent.

8 Annotations

To further enhance dialogue within archives, a core functionality of Omeka S is the value annotation functionality. This provides a mechanism for curators to layer additional context onto existing metadata values. This capability is significant for inclusion because it allows multiple perspectives to coexist within the metadata structure and stored within the database. Examples of this would be including data on where a particular piece of metadata originated from, or even indicating if something is an approximation or a best guess. However, it can also comment, disagree, question and add context to what is being presented. This is a big step in allowing each showcase to present different viewpoints, thus de-centralizing the archive and allowing a layering of points of view to sit on top and around each other so that those who view the showcase can appreciate the dialogue and complexity.

The use cases for annotations sit within qualifying and questioning the archive and its various elements, especially if the original metadata that is captured comes from more antiquated sources of information. From a provenance standpoint, annotations make it possible to trace the evolution of metadata over time, documenting sources, interpretive decisions, and any changes applied to the original record. In doing so, the annotation process decentralizes authority and opens space for collaborative enrichment, which is essential for transparent and inclusive digital collections.

If we take, for an example, a site which showcases a collection of musical instruments collected by a European in the 18th century, we can assume that much of that metadata might come from that period and carry a limited amount of information describing those instruments. Being able to annotate these records, would allow the visitor to see how the metadata has changed over time, and allow for subject experts to enhance what is there, while still preserving the information that was captured at the time for the purpose of historical records.

Curators building sites on Omeka S for academic purposes, such as part of their PhD or curation studies, can use this mechanism, among others, to comment on the more institutional metadata. While they can add their own metadata fields to related objects to include their own descriptions and connections, now they can also add annotations to pre-existing metadata. A few sites on Ibali have explored customising their metadata, notably The Medicine Chest, curated by Nina Liebenberg. As part of her PhD studies, Nina explored using Omeka S and its metadata powers to present her own curation style and showcase her own way of how different resonances between objects connect to each other to present an argument (Bode, 2018).

9 Presentation

The last part of the inclusion paradigm for the sites is how they are presented. Omeka S separates the creation of the showcase sites from the actual creation of the metadata so that one can focus on a particular design without impacting the data itself. It works through a combination of themes which are available for download and allows you, with a little bit of software building experience in PHP and HTML to build some of your own. The aspect of visual design of a site contributes to the overall sense of inclusion when experiencing the showcase collection sites.

On top of this there are specific modules that use the way the metadata has been ingested to present it in interesting ways. For example, by formatting all the fields for the dates in the YYYY-MM-DD the timeline module can be used to create an interactive timeline to allow visitors to the site to horizontally scroll through items. The timeline can take start dates and end dates to create a horizontal graph. As

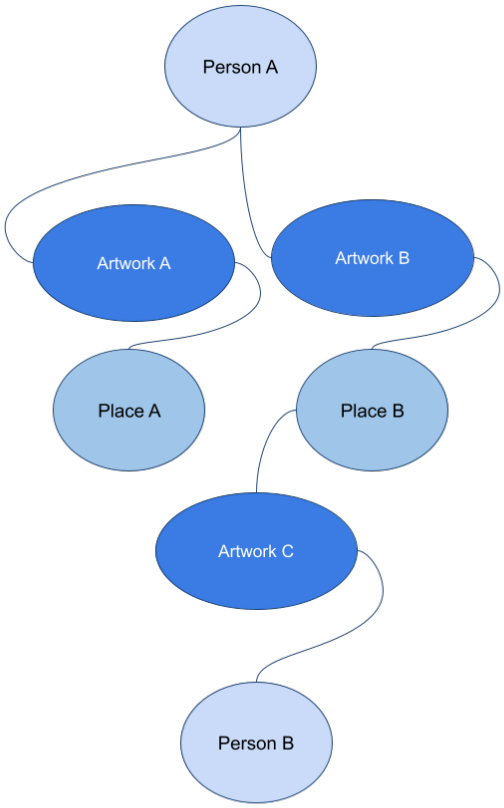


Figure 3: Item Connections from the Works of Art Collection

one navigates the timeline, a clickthrough allows for visitors to jump through to that specific item. On the Ibali site commemorating the Jagger Library fire we had used this to document the events surrounding the salvage and recovery.

The mapping modules allows for another round of visualization by using geographical coordinates that have been ingested to place those items on a map, allowing visitors to the site to explore maps and the items that are connected. The site dedicated to the Works of Art held at the University of Cape Town, showcases this tool by placing the various buildings that hold works of art on the map, and as the visitor clicks through to the building, they can then see all the works of art that are within.

More than anything else the separation between the metadata and the showcase site, encourages the creation of multiple sites inspired by the same digital objects and media. So, while it is possible to have a more official presentation of certain objects for people to navigate, a curator can also decide to build their own interpretive site that presents the same digital objects but presents and contextualizes them in a different way while also adding a completely different visual design.

10 Challenges and Considerations

The implementation of open-source software on local university infrastructure presents several challenges. While Omeka S benefits from a vibrant developer and user community, troubleshooting issues often relies on support from the core developers and contributions via the Omeka Forum. Locally, this necessitates close coordination with the universitys Information and Communication Technology Services (ICTS) to ensure ongoing functionality and maintenance of the server infrastructure that supports the Ibali platform.

In the age of AI, it is important to acknowledge that open access curatorial heritage platforms such as Ibali have become targets for automated bots and crawlers seeking data to train language models (Weinberg, 2025). This creates a double-edged situation. On one hand, the increased demand and crawling can slow down system performance and extract content from the platform without direct engagement. On the other hand, this process contributes to sourcing material for future AI models, which may help address biases and support efforts to decolonize these technologies.

One other emerging challenge is the limitation of storage capacity. As more media-rich collections are added, particularly those containing video files, the demand for storage increases significantly. This requires financial allocations for expanded storage, which places pressure on the Libraries and their supporting institutional structures to maintain the platform sustainably.

Site maintenance also requires ongoing effort. While it is beneficial to have overarching administrators with full access to the platform, it is equally important to cultivate a broader community of curators who are confident in using the system. This is achieved through experimentation, training, and the use of a dedicated university development instance of the software that allows curators to explore the platforms capabilities in a sandbox environment.

Digital Library Services (DSS) hosts regular Ibali Indaba webinars to share updates, showcase new developments, and provide training materials aimed at growing the curator community. The process of building a collection site inherently includes training and support, with the goal of enabling cura-

tors to manage their own collections independently. After the initial development phase, the responsibility for ongoing maintenance typically shifts to the curators who created the site.

However, with time, the curators of sites will move on, and few will be left to maintain the site from the collection side, placing further demand on library staff. To assist with this, Omeka S has developed a module for the export of sites into more static code, which will not require software maintenance. Although the Static Site Export module limits its dynamic functionality, it simplifies long-term maintenance and ensures that collections can remain accessible through digital preservation even if the core platform evolves or is retired. More importantly, because of the structured data infrastructure of Omeka S (through JSON-LD and API access), all of the data and metadata can be migrated easily to other systems.

Conclusion

With the growing number of digital collections, through the data that has been collected over many years by research individuals and groups, there is a drive to make this information available FAIR-ly[6] (Leon, 2025), to allow others to build upon this work, and to re-address the worldwide inequalities of academic output around the world. Using Omeka S as a showcase platform for such digital collections provides both the structure and the freedom to build collections to support this practice and receive recognition for the work that might be sitting on old hard drives. Of course, the process of curating it enables others to better navigate and to locate it, as well as for the curator a multitude of different ways of presenting. Not only that, but their items and ways of curating can offer different options. It is a way to include the collections, specifically from the African continent.

Omeka S is a suitable platform for this exercise of building and showcasing indigenous collections, as it allows for a good balance between structure and flexibility. More important however, have been the processes that Omeka S has encouraged us to do, in terms of engaging with the showcase developers to unpack their collection and figure out the metadata that works for them. It is in this process that the building of these sites becomes a community affair, promoting a sense of inclusion in the showcase that hopes the present itself the larger online community as well.

[6] A principle for ensuring that data online is Findable, Accessible, Interoperable and Reusable (FAIR). Read more about it on the Go Fair website.

Limitations

This work is subject to several limitations. First, Omeka S may pose challenges for users without prior experience in digital curation, requiring ongoing training and support. Advanced features, such as the use of API, may also require programming experience. Second, resource constraints particularly storage for media-rich collections and institutional staffing can impact scalability and long-term sustainability. Additionally, there are a lot more modules that have not been fully explored and implemented, which have potentially limited opportunities for collaborative metadata enrichment. Ongoing curation of Ibali items toward FAIR data practices can also be more fully integrated within workflows. Finally, although the paper acknowledges the potential role of AI in data curation, its practical integration is still at very early stage.

Ethics

The development and maintenance of Ibali raise important ethical considerations with respect to digital curation. As an open access platform, Ibali promotes transparency and inclusivity, but this openness also makes collections vulnerable to automated data harvesting by AI systems. While such activity can contribute to improving language models and reducing bias, it also raises concerns about consent, attribution, and the potential misuse of cultural heritage data. Equally significant is the responsibility to represent indigenous knowledge and multilingual resources with care. Decisions about metadata, language tagging, and interpretation must respect the communities whose histories and voices are being showcased. The platforms design encourages collaboration and layered descriptions, but ethical stewardship remains essential to ensure that these collections are not only accessible but also handled with integrity and sensitivity(Hughes, 2012).

Acknowledgements

The staff and management at University of Cape Town Libraries, especially those working with Digital Scholarship Services. Also, deep gratitude to the various researchers at UCT who make use of the Ibali platform and push its development through their explorations.

References

Katherine Bode. 2018. *A world of fiction: digital collections and the future of literary history*. University of Michigan Press, Ann Arbor, Michigan.

Lorna M. Hughes, editor. 2012. *Evaluating and measuring the value, use and impact of digital collections*. Facet, London.

Sharon Leon. 2025. Omeka and fair. https://omeka.org/news/2025/08/13/omeka-and-FAIR/. Accessed: 30 October 2025.

Steven J. Miller. 2022. *Metadata for digital collections*, second edition. Facet, London.

Omeka Team. n.d. Omeka s user manual. https://omeka.org/s/docs/user-manual/. Accessed: 30 October 2025.

Michael Weinberg. 2025. Are ai bots knocking cultural heritage offline? https://www.glamelab.org/products/are-ai-bots-knocking-cultural-heritage-offline/. Accessed: 30 October 2025.