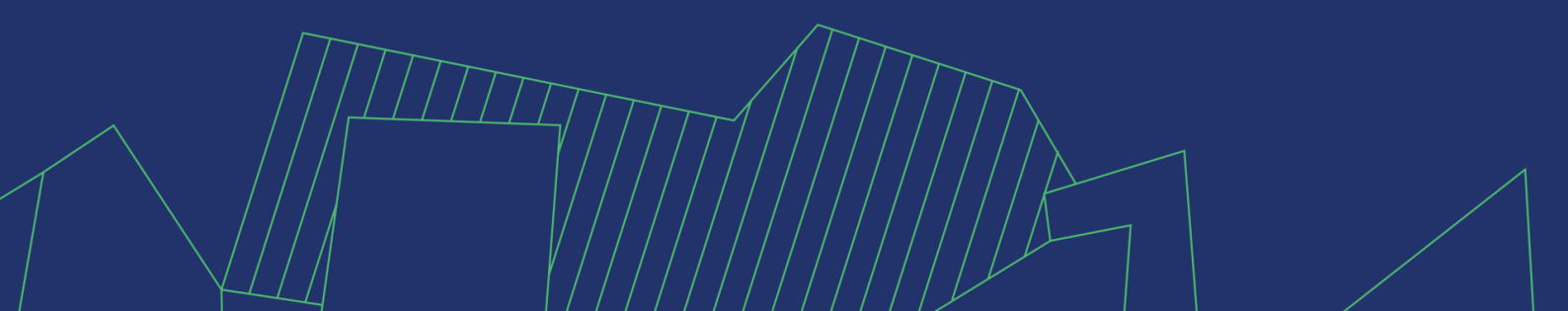Colonial Collections
Consortium

Report
Workshop *Against Opacity*

Working towards a digital platform that brings together, enriches and
provides insight into information on collections from colonial contexts

Colonial Collections Consortium
May 23-24, 2023

Workshop *Against Opacity* - Content

Project

The Dutch Ministry of Education, Culture and Science requested the *Cultural Heritage Agency of the Netherlands* (https://www.cultureelerfgoed.nl/), *Rijksmuseum Amsterdam* (https://www.rijksmuseum.nl/en), the *Institute for War, Holocaust and Genocide Studies NIOD* (https://www.niod.nl/en), *Bronbeek Military Museum for Dutch colonial history and the Royal Netherlands East Indies Army* (https://english.defensie.nl/topics/bronbeek/bronbeek-museum), and the *Nationaal Museum van Wereldculturen* (https://www.tropenmuseum.nl/en/about-tropenmuseum/mission-0) to form a consortium for the project "Against Opacity" in mid-2022. The institutions submitted a proposal to the ministry and formally established the consortium after the summer.

The National Consortium for expertise on the return of objects collected in the colonial period was established to:
1. support the implementation of the Dutch policy on return through the sharing of knowledge and expertise on collections acquired in the colonial context,
2. support the diverse caretakers of collections in the Netherlands in the process of provenance research on collections,
3. support and facilitate the proper research, documentation and (public) accessibility of these collections,
4. be a central access point (or a hub) for stakeholders from countries of origin for questions regarding colonial collections and provenance research.

The project has three work programmes: (1) building trust; (2) from data opacity to data transparency; and (3) co-creating knowledge of the past for the present and the future.

Data Hub

The engineering team responsible for the data hub (https://colonial-heritage.triply.cc/) started to work on the 9th of January 2023. During the first three months the team has converted the first datasets of the consortium partners to linked data, and modeled it according to the wishes of the respective musea:
- the object catalog of the Nationaal Museum van Wereldculturen
- the collection of objects with a colonial context from the Cultural Heritage Agency of the Netherlands
- the military personnel records of the Royal Netherlands East Indies Army (KNIL) of Bronbeek Museum

All the data in the data hub is linked data. It is browseable, visualizable, comes with APIs, and the data that is open is by definition FAIR (findable, accessible, interoperable, and reusable).

The data hub is ready for the inclusion of future datasets. Over the course of this year we expect e.g. the object catalog of Bronbeek Museum, and the collections of Rijksmuseum in Amsterdam. The focus in 2023 is on consortium partners, other datasets in the Netherlands, or perhaps abroad, will likely follow in 2024 and 2025.

Dataset Browser

To maintain an overview of the datasets included in the data hub the consortium has constructed a dataset browser (https://datasets.colonialcollections.nl). This application allows users to search for datasets and view their transparency rating. What opacity and transparency means, is subject to a lot of discussion in the project. You can read more about our current view on the topic in the section below.

Research Application

The first version of the dataset browser was delivered mid-April. The team has now switched to the development of the research application (https://app.colonialcollections.nl). The

difference between the dataset browser and the research application is that the latter allows users to work with the data contained in the datasets.

The team has developed a unified model for objects in all the catalogs that are registered in the dataset browser, and all persons found in these datasets. A unified data model makes it possible to search through all objects and all persons as if they were one dataset.

At the end of June 2023, a first version of the research application was made available containing all objects and persons. Users will notice that the applications are very basic – although they work, they contain few features, a lot of dirty data, and a lot of comments like 'not implemented', 'unknown' etc. This is entirely by choice. The team rather shows what data is available right now so that users can start giving feedback immediately, than to wait for improvements. Over the coming months more and more features will be implemented, and data will be cleaned. The order and priority of the work is highly dependent on user feedback.

View on Opacity and Transparency

Opacity is described in the Consortium Proposal: "*...while objects or collections may have been acquired through single acts of exchange, whether through purchase or military looting, they, like their archives, are spread across different institutions. As such, they are less accessible or legible to anyone outside the museum. Importantly, this inaccessibility leads to a problem of opacity: what is included in or excluded from the category of colonial collections, and how to access the archives, is not entirely clear, even for the experienced researchers. Indeed, we have often heard from researchers from countries from which objects originated that they do not know what we have. Such opacity is due to a number of factors, ranging from a lack of acquaintance with our data-management systems, to the fact that the language and terminology used in our documentation is not accessible to all.*"

Later in the same document other causes are set out: "*…accessibility may be hindered by the inadequate or incomplete documentation of collections, or by the lack of or limited digitization of collections. Especially in cases where collections are extensive, such as in ethnographic museums, large datasets – in both number of objects and available archival data – present challenges to the easy and efficient access to these collections. Objects collected within the colonial context, with their associated archives, can be distributed across different institutions. The dispersed nature of such data can contribute to its opaqueness. Access may further be curtailed by even more mundane yet important causes, such as the language in which data is available.*"

Based on these descriptions, 'opacity' is broadly interpreted as 'any hindrance that limits access to data'. Transparency is more than exposing a full museum catalog on the internet. Hannah Turner, in her paper Organizing Knowledge in Museums: A Review of Concepts and Concerns, critically analyses and ties together contemporary perspectives in information studies, science and technology studies, knowledge organization and indigenous postcolonial theory and defines the development of a field of thought for museum knowledge organization. She writes: "*…that concepts and categories such as facts and data and objects are localized and temporal, and that alternative overlapping worlds of knowledge organize the world in radically different ways… The study of the creation or stabilization of knowledge*

through looking at infrastructures such as standards, categories, and data arrangements can potentially explain how it is that multiple worlds become subsumed into one through normative practice." To realize that there are many perspectives – multiple worlds – possible, not only when interpreting data, but also in the production of data, makes it impossible to claim completeness. Opacity and transparency are not a toggle but a continuous scale, one cannot state that 'dataset X is opaque' or 'dataset Y is transparent'. Instead, the focus in this project will be on defining the hindrances that are relevant to the project and for which the team must implement solutions.

To define Opacity and Transparency, the data hub will establish metrics that relate to relevant and solvable hindrances. The combination of these metrics describes the current state of Opacity/Transparency for an object, collection, and even the sum of collections. Metrics for the evaluation of the FAIRness of data (Findable, Accessible, Interoperable and Reusable) will serve as a basis for this objective, as will the CARE principles for indigenous data governance.

At this point the dataset browser of the data hub includes the following metrics:
- *License*: measures whether a dataset is available with an open license.
- *Languages*: the dataset makes clear in which language(s) it is available.
- *Downloadable or accessible*: the dataset is digitally available online.
- *Structured format*: the dataset is available in a format like Excel.
- *Open format*: the dataset available in an open format like CSV (instead of Excel) or XML.
- *RDF format*: the dataset is available as linked data and therefore in a semantic web format, allowing it to be integrated with other datasets.

These metrics make Opacity/Transparency technically tangible. They can be measured manually through forms, or through automated processes. The results show the project's progress and are reported to data providers (museums, archives, etc.). Moreover, they help them to prioritize curation and enrichment activities that directly benefit the opacity of their datasets.

Workshop

Overview

The project members are working closely with Dutch researchers and Dutch heritage institutions. Through monthly workshops they have formed a clear understanding of what these user groups want from a data hub. However, the story of objects in a colonial context is not complete without at least as much input from the nations of origin. The workshop on May 23rd and 24th was the first conversation in a dialogue that will help us to understand what these users need from the online platform.

On Tuesday May 23rd we made a start learning what opacity and transparency means to people from the nations of origin. What can't they do now, that they would really like to? Or need to? The team gathered stories and compared those to the tooling in development. And asked the question: are we doing the right things? Or do we need different solutions?

On Wednesday May 24th we expanded these first thoughts for solutions and shaped them into practical ideas that the team can act on in the months ahead. Parallel to the session UI/UX designer Bas Doppen interviewed all participants to collect as much information on the circumstances and use cases as possible from the nations of origin.

Method

In the workshop Jauco Noordzij and Gertjan Filarski acted as facilitators using a set of methods known as Liberating Structures. These are intended to engage everyone in the room and gather the ideas, opinions, insights, and knowledge of the full group – instead of the most dominant speakers.

Many workshops, meetings, and conferences resort to one of five conventional "microstructures": Presentations, Managed discussions, Status reports, Open discussions, and Brainstorms. The problem with these is that for the far majority of purposes they are either too constraining (in the case of presentations, managed discussions, and status reports) or too loose (in the case of open discussions and brainstorms).

Liberating Structures are a selection of 33 alternative structures for facilitating meetings and conversations. They were curated by Henri Lipmanowicz and Keith McCandless and designed to embrace distributed control and include a fairer, larger number of people in shaping the outcomes of the meeting.

For more information:
- https://www.liberatingstructures.com
- https://medium.com/the-liberators/liberating-structures/home
- https://reimaginaire.medium.com/what-are-liberating-structures-de6f6d14c2c8the

Participants

For the workshop we chose consciously to limit the number of Dutch participants to the actual members of the team developing the software. Including more people from the consortium, like researchers and heritage specialists from the Netherlands, would have unbalanced the meeting and reduced the opportunity for a meaningful conversation with the guests from the nations of origin. Many Dutch participants would be familiar with the team and speak the same language. This gives the group such an advantage that it would be virtually impossible for them not to (subconsciously) dominate the session and start translating ideas and thoughts for the team.

We landed on a group of ten to twelve Dutch participants and about fifteen to twenty invited guests from the nations of origin. During the workshop guests participated coming from: Aruba, Curacao, Indonesia, Nigeria, South Africa, Sri Lanka, St. Eustatius, and Suriname. The balance was carefully maintained over both days and resulted in meaningful conversations and exchanges of ideas and opinions. Together we created a space in which people could have fun and enjoy working together, and also disagree and express frustrations, turning all of these into powerful ideas for the future.

Outcomes

Authority over cultural data

During the workshop we noticed that the participants from the nations of origin did not mention restitution at all. They explained that governments need the infrastructure, space, security and resources to handle returned objects. Purely pragmatically, many participants find that the Netherlands, or western nations in general, are much better equipped to conserve these objects. Restitution is considered an option for a limited number of objects with tremendous cultural or material value.

We explored what, if not the actual restitution of objects, concerned participants most. The one thing all objects have in common is that they were initially made within an original context different from their current existence within a western museum. They had a purpose for real people. This original context is now for a large part lost. The objects have become untouchable, are hidden in a glass box or a drawer, and the data descriptions focus on characteristics like size, material and colors. Museums and western heritage experts are fascinated by the history that the objects represent, but are much less aware of the role of those or similar objects in a living culture. Museums give another meaning to objects by taking them out of their cultural context.

Virtually all participants considered the restoration of the original context as much more valuable than the restitution of the actual object. Participants demand the ability to add their knowledge of the original purpose for the original people in the original location in the original language. They want to tell the story of the objects and preserve knowledge - not necessarily for museum visitors but primarily for themselves: to strengthen identification as a people. The suggested usage of this knowledge is highly practical: in local education, by local artists, or for the benefit of people living in diaspora.

Moreover, participants are adamant that this story cannot, and must not, be told by western experts. Participants passionately claim ownership of the stories of their objects. Western heritage experts now have the authority over the representation and description of the objects. This is a *scholarly* knowledge founded on an academic western philosophy that is, more often than not, very different from the original system of thought. Western academic principles have not excelled in their acceptance of these systems, and scholars often struggle to view them as equal. The workshop showed that the people in diaspora and in the nations of origin have a *native* knowledge. Participants want us to listen to that story without judgment and as equals. For example: their name for an object is the correct name (they made it) and their pronunciation is the actual pronunciation (it is their language). In addition, participants explain that also within the community there are many differences. These original contexts and the western colonial context of objects have to complement each other.

Participants are not so much interested in a software environment for provenance research to aid in governmental restitution procedures of museum objects. The application should enable them to add data to objects: the stories, original sources, names, descriptions, images and videos.

Local knowledge

The authority over cultural data provides a solid ground for communities in countries of origin to share local knowledge on objects in the data hub, according to the participants. The possibility of linking to the data hub will assist nations of origin moreover in sharing information and data themselves from sources that are only present in their scope. However, differences between countries exist in the state of datasets and the possibilities to develop them. Therefore, enrichment must exist on the level of dataset to dataset as well as contextualization of specific (collections of) objects. One of the main examples of what the participants thought of as highly important in regard to context, was the name of an object. It is often the case that objects are described in Dutch databases by generic terms. These can be insulting to the culture to which the 'object' belongs, as they are a reflection of appropriation by the Dutch. How can we create the ability for counties of origin to use their own names and tell their own stories? How are we able to return the authority over these matters to the communities?

A number of solutions were brought forward. First, there is a cultural change required in The Netherlands. At times, for example, an object label may tell that an object belongs to an 'indigenous' group, where in fact it is part of Maroon culture. The participants experience no space to correct or add to the information provided by a curator. Second, a possibility may be for users to create their own 'Rijksstudio'. The environment would be suitable to build your own collections of objects and information from the data hub. Third, crowdsourcing and citizen science should be used to complement, create and add information to the data hub. In regard to these functionalities and how to handle anonymous users, the participants urge us to look to Wikimedia Indonesia. Concerning these functionalities, new ontologies should be considered with codified native knowledge. Fourth, language is an aspect that the participants draw our attention to. The application and data should be available in local languages and the question is raised if unauthorized translations should be allowed within this scope. Moreover, a number of problems arise within the frame of language that could lead to 'no results'. For example, users will use street language to find objects or use country-specific characters. An option would be to allow requests for translation to be made to museums and institutions.

Furthermore, the vital point of adding stories and narratives can be applied to 'category pages'. As put forward by the participants, narratives most often do not concern a single object, but a category. Allowing these stories to come forward creates a bigger picture of a whole network of artifacts and how they relate to one another. Additionally, adding information per category is more manageable than on each object. Subsequently, a body of knowledge can be held together in which links to research papers and other publications can be made that reference an object or category. Moreover, part of the storytelling and sharing of narratives is organization and moderation of content. The participants stretched that this should change per nation of origin. For example, Wikimedia Indonesia is moderated locally and open to all. They solved the problem by making everyone able to contribute and these contributions are verified. Additionally, Wikidata provides the assignment of different terminology. However, questions still exist on whether this would be suitable for a separate community page within the hub.

Regarding ownership and narrative, participants moreover expressed that there are certain conditions to be met in order to have an equal footing. The first is ownership of knowledge and data - how do narratives and perspectives obtain a place in which they are represented? Moreover, there is the issue of Dutch classification standards, which might be incorrect and cause objects to be unfindable. The participants wish for their classifications to be visible and able to reflect more than mere legal ownership. Relating questions that are brought up regarding regional differences in classification standards, e.g. between South America and Indonesia. And existing colonial Dutch perspective on the data by means of classification. Examples are Anton de Kom, who is hard to find as he is classified as trouble maker; 'Rebellion' vs 'Resistance', and 'to run away' vs 'to seek refuge'. The data hub should allow for different perspectives and contexts on the data and create equality of authorship between Dutch institutions and countries of origin.

A point that comes up regarding ownership and narrative is furthermore if moderated feedback will stay on the platform of the data hub itself or if it should be forwarded to the source owner of the data. A community page might connect users from countries of origin with these museums and heritage organizations that hold objects of interest.

A place for inspiration

In the workshops, the participants expressed the wish for the data hub to be a place for inspiration, not just a research tool. Therefore, it is important to create a hub that is visually inspiring and displays virtual collections. Participants believe that images are crucial in this environment and that they should not only consist of frontal shots, but also bottom, top, back, and sides. The images should ideally be high-resolution with zoom-in functionalities and 360 view. If possible, 3D scans would be welcome. Participants would like an option to request this type of material from the data provider. We should use everything that museums and institutions have to offer in this regard. However, the availability of digital copies of an object should not restrict access to the original.

In this view, Digital Benin (https://digitalbenin.org) is a great example. In their digital space, the first thing you encounter is not an object, but a person. You can listen to the pronunciation of the names of objects, and when present, their categories. There exists a focus on the stories behind an artifact, giving it an identity. Seeing this can make one realize that narratives are an inherent part of these items. In museums, stories are often told through visuals, and a question could be if we can use this as a starting point. Last, this visuality can be brought forward through the use of maps. Proper geolocations that are presented in this way, are more inspiring than viewing the data as a list.

Audience and bridge function

The workshop showed that it is essential to redefine the audience of the data hub. We cannot view Dutch researchers as a separate user community from researchers in the nations of origin, nor should we treat heritage organizations in the nations of origin differently from their counterparts in the Netherlands.

Instead the data hub should function as a bridge between all these communities of researchers, heritage specialists, and the interested public. Features to stimulate this bridge

function include forums to exchange knowledge, shared curation of the data, helpful guides to show the way through Dutch policy and heritage collections, and simply adding an email address for questions regarding certain objects and collections.

The only way to accomplish the above is to support multiple languages in both the user interface and the available data. Besides western languages such as English, French, Spanish, and Dutch, these should include those spoken in the nations of origin. During the workshop participants explained that even they, as specialists, often do not have the original knowledge related to objects either. They know how to get that knowledge through various communities whose members do not always speak western languages. To gain trust it is essential to offer the user interface in their language they understand.

Technical requirements

In the workshops, several technical requirements were put forward by the participants. These relate especially to the accessibility of the hub and its data. Access through mobile phones should be facilitated by the design of the user interface and it should be adapted to suit small-screen devices. Moreover, not all users have a fast internet connection, which should be considered in the development of the data hub. Additionally, the participants expressed their need for an API to access and use the data of the hub in their own applications. The latter allows others in the nations of origin to do the same, and not be dependent on institutions in The Netherlands.

Limitations

In different countries, different topics are part of the momentum at different times. Not every country can give their input today because they may not have the topic on the agenda right now. The exchange of ideas and information can also be difficult because not everyone is proficient enough in English to be part of the conversation. Furthermore, a participant expressed that not all countries are as advanced as The Netherlands on various technical fronts, in which there is also a big difference between large and small institutions. Additionally, countries of origin and the participants themselves are not always able to provide feedback at the same pace as the team would like. It is good to recognize the limitations.

Posters

How might we make the data hub inspiring to people?

The broader public is possibly already interested due to the theme of the project and the fact that it touches many people, their cultures, and identities. However, there may also be many people who have no interest in the project yet. Therefore, the data hub needs to inspire people to participate. Whether it is for students, with research projects, or for personal use and research. The challenge is especially how to engage non-professional researchers; people who have a lot of knowledge to add to the hub. This engagement could be promoted through stories of items that were either returned or have gained back their context through the data hub, which was lost in the process of appropriation and collecting. Therefore, an attitude of openness and generosity is needed, which allows for various points of view. Items

need to be findable, and co-ownership should help build recognizability. The item in question should be allowed to gain back its identity. Hence, a storytelling space for each object should be facilitated. As a consequence of this process, a rich environment can emerge with information. A part of this can be data visualization that touches upon events. In this way, the result will speak to people and the digital space can be seen as a research hub with stories and perspectives, not necessarily as a data hub alone.

How might we receive and maintain input and feedback from the diverse communities around the world?

To receive and maintain input and feedback from diverse communities around the world the focus should be placed on crowdsourcing, regarding each object. However, this input should not be mixed with feedback, but separated. In this manner, the feedback can be responded to, and for example, accepted. The question of who should accept the feedback is still to be answered, but moderation is required. One option is a voting system in which feedback, input, and additions can receive votes, and therefore 'float' to the top of a comment section. Moreover, feedback that is added and changes or additions that are made should have sources that can be verified. Next to this, a section can be made for knowledge that does not have or require these sources, e.g. oral heritage. Furthermore, there needs to be a way to verify user accounts. To accomplish this, existing accounts on social media or academic platforms could be used. The verification allows users to e.g. add related objects with comments, and create collections of items. The latter can be enriched even further by enabling the addition of various types of information such as links to other objects and images. These verified user accounts could moreover be linked to specific objects on which they worked. As such, another user can find that person who might have different or new knowledge about a specific item. A database could be made to connect users who are interested in specific objects. Moreover, there should be networks and structures in place to invite communities for workshops, work groups, etc. In this way, the data hub also becomes a platform where people can communicate and organize. Last, clear definitions should be set for terms such as "receive", "maintain" and "diverse communities", for adequate communication with users.

How might we add indigenous perspectives to the data in the data hub (the voices of the nations of origin)?

To add the 'indigenous' perspectives to the data in the data hub (voices of the nations of origin) they should first be included in the team, as e.g. liaison. The advantages and necessity of this inclusion far outway the costs. Moreover, committees of experts should be created and facilitated, which not only emphasize academic knowledge, but also oral tradition, intangible heritage, and non-academic knowledge. These committees can potentially overlap with the collaboration partners from nations of origin, and be self-initialized. To focus the effort, the aim should be to do this country by country, not all at once. The committees of nations with which The Netherlands collaborates on a diplomatic level, such as Nigeria, Indonesia, and Sri Lanka, can be organized or facilitated first. Due to the international scale of the project, there should be an option to add local languages and remove language barriers. The intention would be to include three languages: English, the National official language, and a local language.

Furthermore, the question of access should be considered: should there be controlled content or open public access, such as Wikipedia, or both? Concerning this topic, a partnership can be set up with the proven experience of e.g. Wikimedia Indonesia. A solution could also be to think in terms of different levels of access. The first may be open comments. However, these may not be productive. The second may be edited or moderated text and oral interviews, such as with Digital Benin. It is important not to over-moderate since anger about the situation is understandable. Last, there could be an option of access to content after registration, to avoid bots, racism, commercial content, and unreferenced individuals.

How might we add datasets from outside the Netherlands to the data hub?

The response to the question of how we might add datasets from outside The Netherlands to the data hub evolves around equality, contextualization, and accessibility. First, full databases must be provided from Dutch collections to create an equal footing. Here, questions may arise about how to deal with hidden information and sensitivities that are part of those databases. This process can only evolve if a complete image of what is present in Dutch collections emerges, through full disclosure of collection data. Therefore, the intention should be that no curation is done beforehand in terms of data selection. Furthermore, multiple views need to be included in the data that is collected. On an object level, this can mean a 360-degree presentation of artifacts, so users can interact with them to a high level. Additionally, sound, video, film, folk music, and folklore can be part of the contextualization, which brings vital significance to the foreground of an interface. These features can be added through buttons next to photos of artifacts, e.g. an eye, an ear, a mouth, and its use.

Furthermore, to be able to add to and be part of the data process through local hubs, which would be country-based, is required. The local hubs are able to bring forward intangible heritage and indicate active views. The first of the two is an element that is very much lacking within the data that is present in Dutch collections and institutions. To aid the exchange between local hubs and the Consortium data hub, open international standards are required.

How might we facilitate communications between users (abroad) and museum curators (primarily Dutch)?

To answer the question of how we might facilitate communications between users (abroad) and museum curators (primarily Dutch ones), we should first consider if the Consortium and Against Opacity technical team inform the stakeholders from the countries of origin/former colonized countries. If there is no communication at the moment, the following is what is needed in an Indonesian context. Step 1) Reach out: first, a diplomatic route should be taken through the procedures of the Ministry of Education and Culture, and other governmental agencies. Next, there would be direct communication with museums, which could be facilitated by a participant of the workshop. Afterward, universities, pesantren (https://en.wikipedia.org/wiki/Pesantren), and their communities can be involved. Facebook groups are another audience that can be approached; there are very active users on the subject of the platform.

After the sub-question is answered, the focus can shift towards the network that emerges through the described process and other active users (e.g. researchers and communities) that are involved. From these groups, databases can be created that determine a network of active people, involved in the data hub. The goal is that these databases can aid communication between the users, and facilitate the exchange of ideas and knowledge. In this way, if one has a question, it becomes easier to find another who has knowledge on the topic of interest. The latter can also be linked to platforms such as Google Scholar, from which publications and researchers could be harvested to enlarge the network.

Subsequently, a second step may focus on getting the user in touch with a contact person through a button in the interface. The person answering the button will work through a triage system, to find the right person (e.g. museum curator) who has information about the request or object in question. The result is quick and direct communication. In this process, the database of users can be of assistance as well. Moreover, if the user can form their collections of interest, consisting of objects in the data hub, more points of information, contact, and expertise may emerge, which relates to the previous paragraph. The consequence of a well-connected and socialized network of users, that can find each other and museum curators for information, is that contextualization of objects and development of knowledge is advanced.

How might we make the data truly open?

To make the data truly open it should be freely used, re-used, and redistributed by anyone. An example can be seen in the Open Knowledge Foundation (https://en.wikipedia.org/wiki/Open_Knowledge_Foundation). Moreover, the FAIR (https://en.wikipedia.org/wiki/FAIR_data) principles should be adopted; to make the data findable, accessible, interoperable, and reusable. Standards for the data should also be considered within this context, making clear what one can expect when interacting with the content of the hub. In line with the principles is an API (https://en.wikipedia.org/wiki/API) which gives access to the data and makes it possible to download it, or make other connections. Security is an essential factor concerning e.g. an API and should be central to such connectivity. Furthermore, the openness of data is dependent on the collaboration between the various parties involved. To make it truly open, it is an essential point of attention. An example of such a collaboration effort, which leads to open data, is the selection of data that determines what is part of the hub. The decision-making process of what to include and exclude should be highly accessible. Preferably, the selection of data should be minimized to give stakeholders the decision on what is important. The openness of data is moreover determined by the user-friendliness of the interface. No technical jargon should be used and users should have access through their mobile phones. The multilingualism of the data hub relates to this user-friendliness, as it is essential for openness. Last, adding comments on records is required to have an open interaction with the content and context of the data hub.

Priorities, future work, and ongoing conversations

During the two-day workshop, the participants expressed many wishes, of which the majority is bundled into the sections on outcomes and posters. These outcomes and posters concern e.g. what type of digital space the datahub should be, what should be possible in this

environment, and above all, what it should aim to do. The implementation of these elements into the datahub is an elaborate exercise. With each aim for the implementation of x, new questions arise concerning how to approach x, how to manage x on the backend, and how to shape x in the user interface. Because of this, it is very challenging to implement all outcomes and poster points at the current stage. Priorities have to be made regarding what to do first and what to focus on in a later phase of the project. The priorities for 2023 and 2024 are described below, as well as the ongoing conversations.

2023

The main focus for this year is the implementation of features related to perspectives, collections, input data enrichment, communication, and open data(sets). The most important feature is the accommodation of multiple perspectives and contexts on the heritage of an object or collection of objects. It will also be possible to add multiple perspectives on the data points that are brought together from the data providers. Users will be able to make collections of objects. These 'category pages' allow stories to come forward by creating a bigger picture of a whole network of artifacts and how they relate to one another.

These collections can also be made by communities of people, which the data hub allows users to create. As a community, people can build their collections from all the objects in the hub and share them with others. Like an individual, communities can add comments, alternative data, sources, names, descriptions, and create provenance records that show the (probable) history of an object or a collection of objects. These additions are moderated through a voting system. Users and communities can add their support for existing or added information on the object pages of the application. In that environment, they can also contact the data providers (e.g. to request material), even when not logged in. In this case, access to the content of the hub is limited. To log in users can utilize their conventional accounts, such as Facebook and Gmail.

The combination of collection data enrichment and linked data creates new information. Users, data experts, and automated systems together will start to relate historical events, people, and locations to objects in the museums. The resulting data and knowledge can be used to create more visual features in the user interface. Engagement is also increased through stories of items that were either returned or have gained back their context through the data hub.

The open character of the data is central to the hub. Users can query and retrieve its data from a web-API and use the results for their own research and applications. Through endpoints, the user can utilize SPARQL for linked data exposure. Because of these features the data hub will follow the FAIR principles: to make the data findable, accessible, interoperable, and reusable. The functionalities are accompanied by documentation on the standards for the data, to make clear what the visitor can expect when interacting with the content of the hub. Whether data is open is also determined by the user-friendliness of the interface. No technical jargon will be used where it is not necessary. User friendliness and access to the hub will be optimized by the design of the user interface. The latter will be adapted to suit small-screen devices so use through mobile phones will be possible.

By the end of the year, the data hub will provide access to the object catalogs of the National Museum of World Cultures, Bronbeek Military Museum for Dutch colonial history and the Royal Netherlands East Indies Army (KNIL), Rijksmuseum Amsterdam, and the Cultural Heritage Agency of the Netherlands. Moreover, integrated into the data we will add a guide to colonial heritage provenance research developed by the Dutch Institute for War, Holocaust and Genocide Studies (NIOD), and the public registers of all Dutch military personnel serving in the Netherlands East Indies. The datasets will be subject to the integrated opacity measuring tool; it aims to provide insights on how accessible and transparent a dataset is.

2024

In 2024, the focus will shift to language, storytelling, collaboration through a forum, and more datasets. The intention is to have the application and data available in three languages: English, the national or official language, and a local language. This future multilingualism of the data hub contributes to a new application that focuses on storytelling.

The ambition is to create a 'storytelling application'. Features of this environment could be e.g. creation of stories and narratives with the help of uploading images and sound. Since we are still in the investigation phase of this idea, the goal is to come to a feature set and a possible version 0 in 2024. For more inspiration of functionalities, we look at the Digital Benin project (https://digitalbenin.org/).

Community Engagement Platform features are also options for the future. These would offer the possibility of the exchange of knowledge, shared curation of the data, and coming together in a digital environment. In this manner, users can find others who have knowledge about a specific item or collection of interest.

Further down the line, there are ideas about connecting to local hubs; creating a data hub network. However, the scope of this concept is enormous, and could easily be the entirety of work in 2024. The potential lies here in the exchange and storage of one's own enrichments to the data, which will shift ownership.

In 2024, we will shift our focus from bringing together the data from the consortium partners to adding datasets from data providers in The Netherlands. In 2025, this will possibly shift again, toward the international stage. We acknowledge the importance of the step aimed for in 2025. Doing this already in 2024 would be a big challenge; creating problems in relation to scope and scale.

Ongoing conversations

The focus described above is a start. However, certain topics are ongoing conversations. Enabling these dialogues between (Dutch) heritage organizations and users and communities (in the nations of origin) is an essential part of the project. For example, there are many outstanding issues related to the ownership of the data, who controls it, and who decides which objects are - and are not - part of colonial heritage and the hub. What about objects that seemingly do not relate to colonial heritage?

Once objects from museums in the Netherlands are restituted to the nations of origin, other questions arise. What, for example, happens to data "owned" by Dutch museums on objects that have been returned? Do we remove that if that is what the communities want? At the same time, we also have an obligation to researchers to guarantee sustainable and persistent data. What happens to the links in research papers and scientific journals, or the reproduction of results?

Furthermore, how do we create equality of authorship between Dutch institutions and countries of origin within the datahub? And how should we include voices of the nations of origin in the team itself? An option is to create and facilitate committees of experts and consultants, which not only emphasize academic knowledge but also oral tradition and intangible heritage.

These conversations provide us with continued input to improve the applications and take a position in these debates. Additional requested features will be added over the coming months and years.