Visualize and Communicate Tangible and Intangible Cultural Heritage with the QueryLab System

Maria Teresa Artese[0000-0001-8453-0807] and Isabella Gagliardi[0000-0002-3706-0993]

IMATI - CNR (National Research Council), Milan, Italy
{teresa,isabella}@mi.imati.cnr.it

Abstract. QueryLab is a platform designed to query different cultural databases at the same time in a transparent way. This paper presents the design of a new section, ICH Discovery, dedicated to the intangible heritage, where we tried to improve the structuring, querying and visualization of the results, alternative to the classically available ones, to facilitate the sharing, dissemination, and preservation of cultural heritage.

Keywords: Living Heritage, ICH Collection, Common Metadata Model, Inventory Integration, Data Visualization.

1 Introduction

Mapping cultural heritage, both tangible and intangible, is the first step towards safeguarding it, as established by UNESCO in its 2003 convention. Ratifying the convention, many countries are developing online inventories or even simple lists of tangible and intangible assets to be preserved, recognized as part of their culture. However, there are more virtuous nations, such as South Korea, Japan and China, which had already defined various strategies to enhance and safeguard their traditions before 2003.

Dynamic websites allowing the consultation of large databases are a standard we are used to, but unfortunately the mere presence in archives or inventories, even digital ones, does not guarantee the fruition and preservation of intangible heritage. Technological tools are needed to spread the knowledge of these assets to as many people as possible, involving communities in the protection and conservation action. Intangible heritage, due to the complex nature of the assets it describes and the variety of related multimedia documentation, requires more sophisticated query and visualization tools, as also highlighted by UNESCO which has recently added the "Dive in" section to its website (*https://ich.unesco.org/en/dive*), allowing innovative and dynamic searches and visualizations.

The QueryLab platform has been designed to allow the online querying of several cultural heritage databases all at the same time, with a single query and in a simple and transparent way for the user, thus enabling the knowledge and approach of different, but integrated, cultural heritages. We present here an extension of QueryLab, focused on intangible heritage, developed to offer new ways of querying and visualizing data coming from different inventories.

Digital Presentation and Preservation of Cultural and Scientific Heritage. Conference Proceedings. Vol. 11, Sofia, Bulgaria: Institute of Mathematics and Informatics – BAS, 2021. ISSN: 1314-4006, eISSN: 2535-0366

Simultaneous querying of different intangible heritage archives is possible if they share the same data model, which is why we introduced the ICH-Light model for QueryLab (*http://arm.mi.imati.cnr.it/querylab/*), created specifically to index intangible heritage, with the aim to record the minimum set of information needed to search and integrate data from different realities, while always ensuring the link to the original resource. The model studied can host and structure information at various levels, allowing for a more advanced and dynamic visualization, and can be the basis for a participatory system where contributions to the index can come either through an interchange format or through specific web services.

QueryLab new tools have been tested on different archives: Intangible Search, Sahapedia, ACCU static pages and German Nationwide Inventory. Preliminary results have demonstrated the simplicity of use, combined with the effectiveness of QueryLab, in displaying results from different archives in a single system and facilitating the comparison, dissemination and preservation of cultural heritage.

This paper describes the different layers of the ICH-Light host structure and the tools implemented for the ICH-Discovery section, which allow users to explore ICH assets in a new way, filtering and navigating through the results, in order to show the interconnections existing between different cultures and disseminate traditions and knowledge, favoring transmission to new generations.

The first results provide confidence in the ease of use and effectiveness of QueryLab in visualizing the results of different archives in a single system and thus facilitating the comparison, dissemination and preservation of cultural heritage.

The paper is structured as follows: after a short review of related work about data integration, an overview of the developed functionalities is presented detailing the inventories currently collected in the platform, specifying the data model designed and how it has been used. The ICH Discovery tools, its features, motivations and some screenshots are then presented, and the paper ends with conclusions and future work.

2 Related Works

Integrating data from different sources is a topical issue that has been studied from different perspectives, one of the most famous examples of integration is “Europeana”, the European digital cultural platform that includes, gathered in a single platform, the records of more than 10 million cultural artefacts presented in different ways and oriented to today's users. Another example, but in the field of books, is “HathiTrust”, a large-scale collaborative repository of digital content, offering a collection of millions of digitized titles from libraries around the world. The difficulty in the world of intangible assets lies in finding intangible heritage inventories available online. A good starting point is the “Map of e-Inventories of ICH”, written and coordinated by Memória Imaterial, a UNESCO-accredited non-governmental organization based in Portugal (Sousa & Imaterial, 2017), which shows the world map with several points representing the archives that have been identified and examined, giving a brief description and a web address for each of them. The updated version also introduces possibilities to search for archives by keywords or categories, but not on the contents.

For intangible heritage, UNESCO offers, on its website, good tools for querying and immersive navigation, but only related to the elements included in its Lists of Intangible Cultural Heritage and in the Register of Good Safeguarding Practices.

3 QueryLab and Intangible Heritage

The QueryLab platform originates from the need to query different archives in an integrated way, so it provides several databases that can be queried simultaneously in a simplified way using both web services and local queries (Artese & Gagliardi, 2020). This study focuses on the inventories specifically dedicated to ICH items, considered as "living goods", storing the memory of people and communities evolving over time. A new section has been developed, dedicated to them: new tools have been designed to query and visualize the results, integrating them with the other databases already available reached through the web services. The former “ICH-Light” metadata structure, the common data structure for intangible heritage used so far, has been reshaped in order to obtain a new extensible and customizable one, capable of storing data coming from different ICH datasets. The new section realized is called “ICH Discovery”.

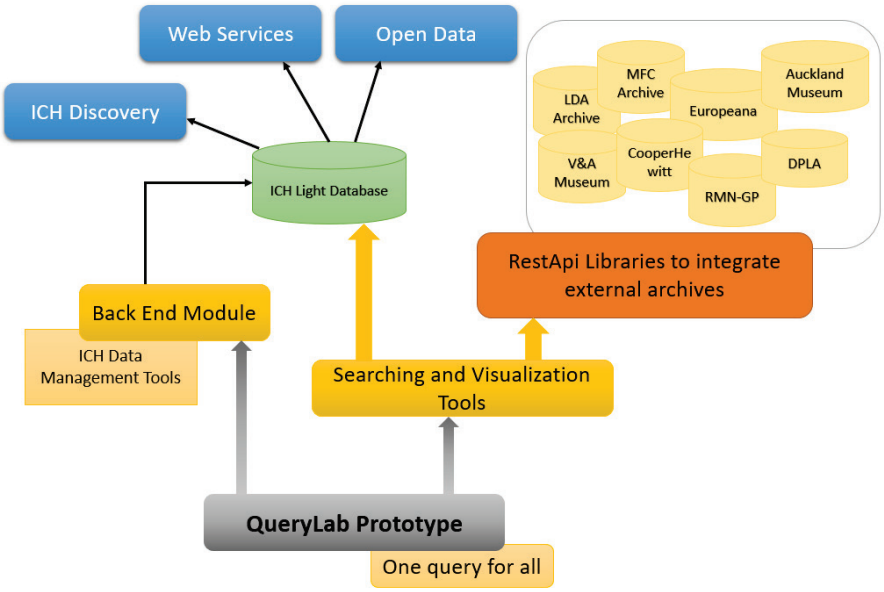

Fig. 1. QueryLab Structure

To provide a simple and integrated navigation of different archives collected by ICH Light metadata, it is necessary to obtain a general index of contents, modelled on the approach of the REST web-services integrated in QueryLab, which allows it to be easily integrated with the portal. Several archives available online were studied to model the general index: some with a dedicated website, while others static, consisting of simple web pages that do not allow dynamic querying.

We studied and applied the Cultural Heritage in Digital Environments Model (CHDE) proposed by Wijesundara & Sugimoto, 2018: "ICH does not exist as a single physical item but is usually instantiated in many occasions", so a performance is an *Instantiation* of the corresponding intangible entity, who is performed in a specific time and place by different people. Introducing the *Instantiation* concept it is possible to transform all digital resources describing an intangible entity into digital archive record instances, according to the One-to-One Principle of Metadata. The one to one principle is a long-standing principle within the Dublin Core and other cultural heritage metadata communities. It is «the principle whereby related but conceptually different entities, for example a painting and a digital image of the painting, are described by separate metadata records» (Miller, 2010). It allows to clearly identify physical and digital objects, it is a key issue to aggregate metadata for a single cultural heritage object and link item-based metadata to web resources. This is a basical step to study a metadata classification using the most suitable ontologies and to be able to produce Linked Open Data from the data collected using "ICH-Light" metadata.

4 Towards a Standard: ICH-Light Levels

Several problems were encountered during the creation of the data model that allows the simultaneous querying of very different archives: the main problems concern the different languages used to store the contents, which often do not provide the English translation, as well as the categories used which frequently do not correspond to the categories indicated by Unesco, the cataloguing methods which can be very different and finally the availability or lack of web services to access the contents.

The inventories used for the analysis and testing phase so far cover different geographic areas of the world and are specifically dedicated to ICH entities:

- From Italy: IntangibleSearch, the archive of the Lombardy region in Italy (*http://www.intangiblesearch.eu*)
- From Asia Pacific Area: ACCU Data Bank, the Asia Pacific Database on Intangible Cultural Heritage, including Australia, Cambodia, Fiji, Kyrgyzstan, Tajikistan, Tonga and Vanuatu (was at *http://www.accu.or.jp/ich/en/arts/arts1.html*, now available only from WayBackMachine).
- From India: Sahapedia (*https://www.sahapedia.org/*), an open online resource on the arts, cultures and heritage of India
- Form Germany: the German Nationwide Inventory of Intangible Cultural Heritage (*https://www.unesco.de/en/kultur/immaterielles-kulturerbe/german-inventory.html*)

The model building phase went through successive revisions and refinements, and the resulting data model is organized on three levels to allow for different phases of data storage:

Index Level: this is the basic level dedicated to the main data of an ICH asset, which are used to generate the main indexing because it provides the minimal set of information expressed in one language only: English or the native language if English is not

available. The metadata set at this level contains information about the archive of provenance, the domains of pertinence, the place where the asset is performed or live, a representative image, denomination, brief description, link where the original resource can be found and a set of keywords which are labelled using a glossary to associate, if possible, a translation in different languages (Italian, English, French and German till now).

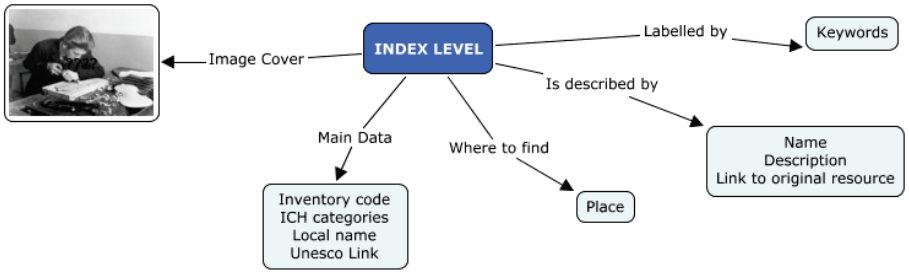

Fig. 2. Index Level

Catalog Level: this is the cataloguing level and is designed to extend and enrich the metadata index level, to better detail the catalogued assets expressing data in different languages and adding information about communities, valorization and safeguard measures. This level is especially useful for those archives that are currently stored as simple flat lists, without a real database to preserve data and is organized to allow different repetitions of each semantic group according to the different languages used to catalog items.
Instance Level: this level is dedicated to the different instances of the catalogued ICH assets, which can be different collection of data and multimedia contents useful to deepen and witness the evolution and adaptions of the ICH entity over time.
As shown in Fig. 3 Catalog Level and Instance Level are linked, because the instance level contains the physical objects representing the entity described in the linked catalogue level, organized according to the different representations or records of the asset. The catalogue level is also an expansion of the index level, as indicated by the green blocks and the blue specifications and is able to store data in multiple languages for the same entity.

The creation of the unique index for intangible assets, integrated in the QueryLab platform, allows to:

- realize search tools independent from the different languages used for contents.
- collect keywords to create guided tours, in automatic way.
- develop a participatory system dedicated to the loading of data, to contribute directly to the integrated index.

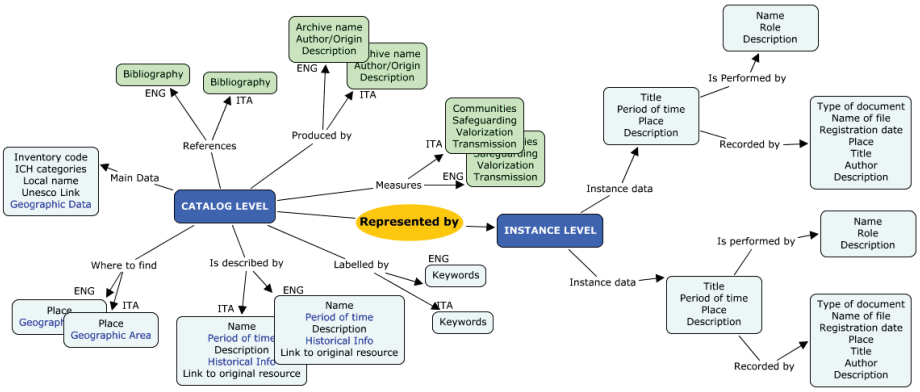

Fig. 3. Catalog and Instance Level

5 ICH Discovery Tools

The tools developed for the ICH Discovery section of QueryLab are focused on searching and visualizing ICH data from different inventories and stored by the ICH-Light model at the basic level. The available search keys combine both the data available for the Index Level and other keys, which can also be cross-referenced with specified keys entered directly: users can search for ICH resources using the proposed thematic paths and refine them using country, category, available keywords, source archive or free text.

In addition to the traditional mosaic of images, new visualization methods have been developed, which are less traditional but provide a visually simpler and clearer idea of the results obtained and their distribution within ICH categories, countries of data origin and inventories of provenance.

On the left-hand side of the page filters are available for searching while the results are displayed in the rest of the page. There are two blocks of keys for the filters: the first group offers the themed routes, while the second group contains all the other keys coming from the Index Level, by crossing them users can search or refine results.

The main part of the page is the output section, where the ICH assets extracted from the query entered by the user are displayed. There are several ways of displaying these items, for each available mode there are various sub-representations following the main available keys: category (domain), country, inventory of origin. The new visualization modes such as Icons, Sunburst and Force Graph enable access to the ICH item directly, while the others are useful to show ICH distributions according to the query executed. Users can switch from one to the other using the proper buttons at the top.

Sunburst: is a three-level interactive circle graph, where the innermost one is the main one and highlights the country, category or inventory of origin and each section can be explored in detail by clicking on it, enlarging and navigating the subsequent levels. This type of visualization provides useful feedback on the type of distribution of query results in terms of categories, countries and source inventories. The results can be browsed interactively by clicking on each section in the three circles.

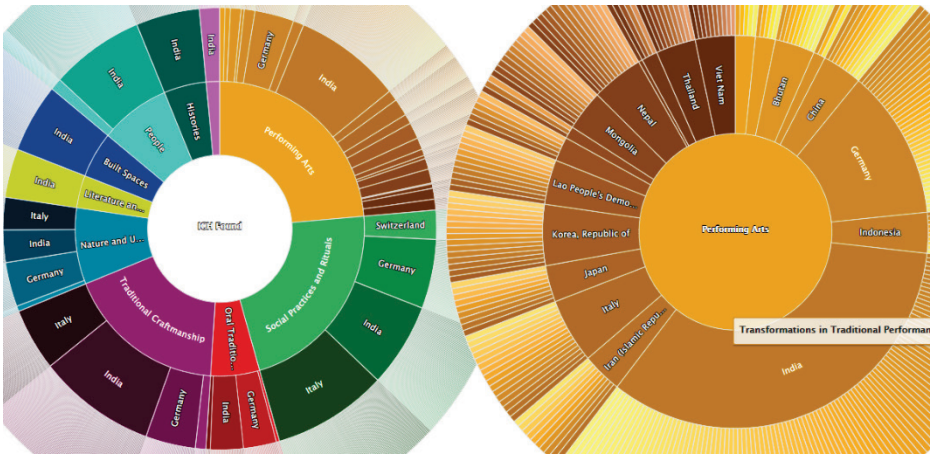

Fig. 4. Sunburst by Category on the left and "Performing Arts" expanded on the right

Force Graph: this type of graph shows the items resulting from the query linked to their domain, country or inventory, depending on the type of sub-graph chosen. The objects represented can be dragged to evaluate their links and each ICH, represented by its image, can be opened in detail using the Ctrl key on the keyboard. The legend, available on the left, shows the association between the color and the corresponding category/domain, also used to trace the connections between domain and country or sources. To identify countries their flags are used and for inventories their logos are used.

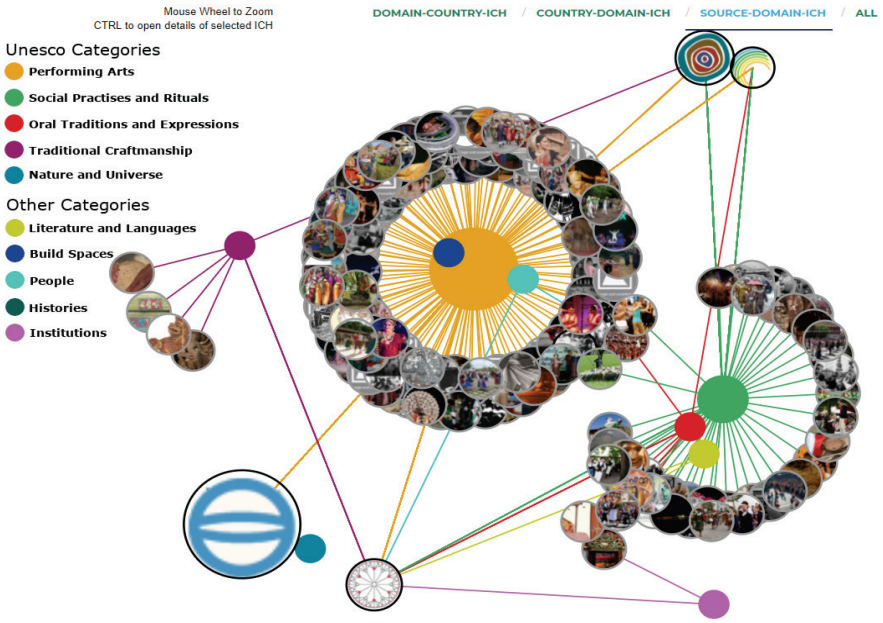

Fig. 5. Force graph for "Traditional Dances" query

Bubbles: this graph shows results using bubbles grouped together by a bigger one, representing the country, the category or the inventory to which items belong to, the size of each bubble reflects the corresponding number of items.

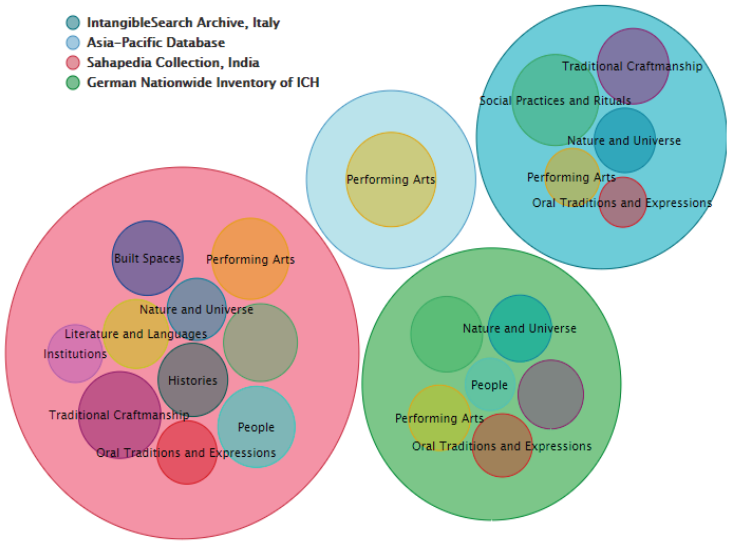

Fig. 6. Bubbles by Domain

Dependency Wheel and Wind Rose: both have query keys on the outer circle, in the first keys are connected each other by strings with different thickness depending on the number of objects resulting from the connected criteria. Moving the mouse onto the graph will highlight connections. Wind Rose instead show distributions of the number of ICH results of a query with respect to a main key (country, domain or inventory) identified using different colors while the secondary grouping key is around the circle. Moving the mouse onto the graph or legend the corresponding item is highlighted.

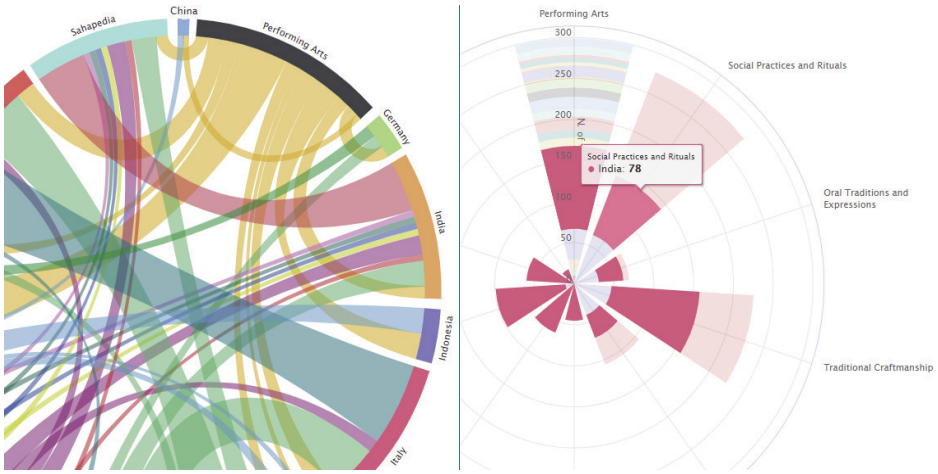

Fig. 7. Dependency Wheel and Wind Rose graph, where "India" is highlighted

Sankey: this graph is similar to the dependency wheel but is shown on a plane where the main key is placed on the central bar and on either side, there are two other secondary grouping keys, the connecting ribbons being proportional to the number of corresponding elements. This system highlights the connections of the key in the center with the two different keys on the left and right, simultaneously. In the example below, for the query "Wedding" the key in the middle is the domain, connected with countries on the left and inventories on the right.

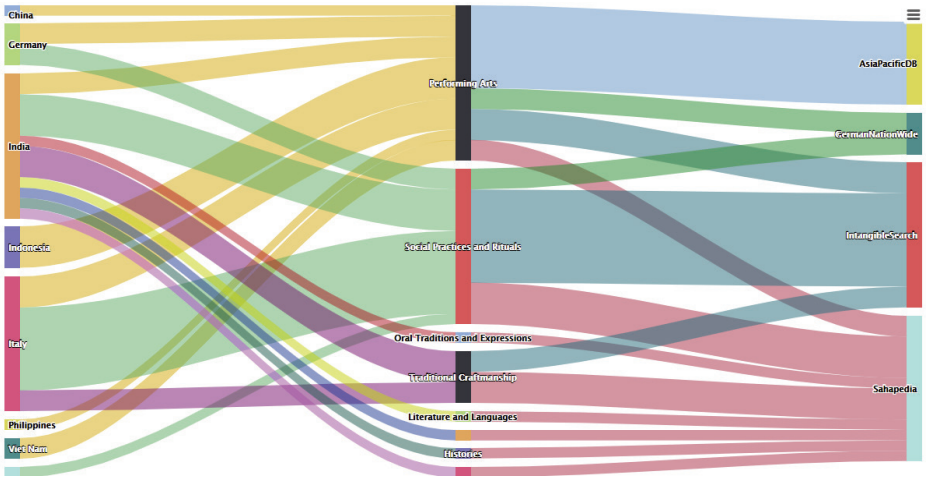

Fig. 8. Sankey graph with "Domain" key in the middle

When output results are shown using classic mosaic icons, force graph or sunburst it is possible to access each ICH item directly, opening details about it where the words used for query are highlighted, as shown in the figure below, where the active search key is "Wedding". Details come from the Index level content, the basic one, and the direct link to the original resource is always available by "Go to Source" button.

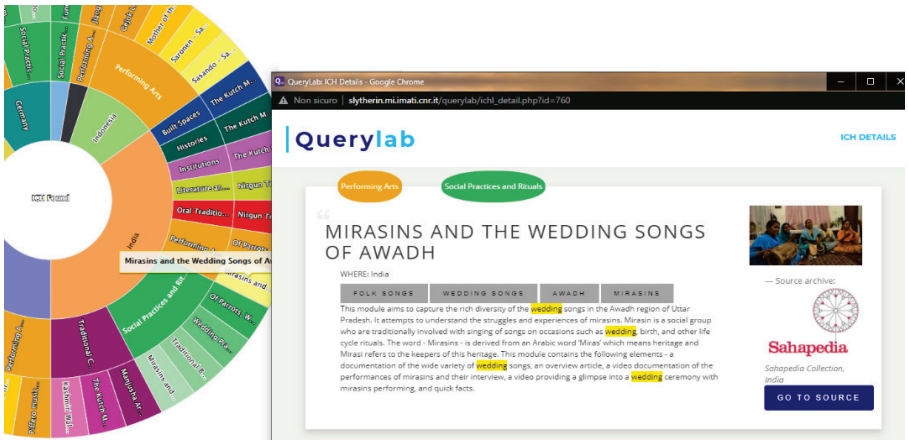

Fig. 9. Details opened from Sunburst graph: the query key is highlighted with yellow

6 Conclusion and Future Works

QueryLab has been provided with a new section designed to enhance and deepen the fruition of items coming from different ICH inventories. For the data analysis and testing phases of the proposed structure, several inventories containing data from Asia Pacific and India were identified, whose inventory contents, thanks to Unesco guidelines, are incredibly close to the Europeans' inventories in terms of approach and tradition.

Through the implementation of visualization tools alternative to the classically available ones, preliminary results have demonstrated the ease of use and efficiency of QueryLab new tools to visualize the results of several archives in a single scenario and thus facilitate the sharing, dissemination and preservation of cultural heritage.

For future developments, we propose to further improve the data query methods, expanding the inventories considered to better detail the proposed structure, and develop a back-end environment to directly manage the collected data, increasing the possibility of preserving and sharing intangible cultural heritage online.

References

Artese, M. T., & Gagliardi, I. (2019). A platform for safeguarding cultural memory: the QueryLab prototype. *Memoriamedia Review.* Review 4. Art. 2. 2019. https://review.memoriamedia.net/index.php/querylab-prototype.

Artese, M. T., & Gagliardi, I. (2020). A unified data model for ICH online inventories: a proposal. *Digital Presentation and Preservation of Cultural and Scientific Heritage.* Burgas (Bulgaria): Institute of Mathematics and Informatics at the Bulgarian Academy of Sciences.

Miller, S. J. (2010). The one-to-one principle: challenges in current practice. *International Conference on Dublin Core and Metadata Applications*, (p. 150-164).

Monika, W., Wijesundara, C., & Sugimoto, S. (2017, November). Modeling digital archives of intangible cultural heritage based on one-to-one principle of metadata. In Proceedings of the 8th Asia-Pacific Conference on Library & Information Education and Practice (A-LIEP 2017) (pp. 137-148).

Sousa, F., & Imaterial, I. R.--. (2017). Map of e-Inventories of Intangible Cultural Heritage. Memoriamedia Review. *Memoriamedia Review*.

Wijesundara, C., & Sugimoto, S. (2018). Metadata model for organizing digital archives of tangible and intangible cultural heritage, and linking cultural heritage information in digital space. *LIBRES: Library and Information Science Research Electronic Journal*, 58-80.

Received: June 28, 2021
Reviewed: July 17, 2021
Finally Accepted: July 22, 2021