Advancing the Interoperability of the GLAM+ and Cultural Tourism Sectors through KOS
Perspectives and Challenges

Abstract:
The possibilities and challenges of knowledge organization systems (KOS) to collaborate in the interconnection between the cultural heritage sector–galleries, libraries, archives, museums, publishers… (GLAM+)–and the increasingly important cultural tourism industry are explored, and a model for framing their interaction is proposed. Due to the diversity of KOS implied in GLAM+, this project is to be treated as an interoperability problem, thought a strong user-oriented purpose is also needed, based on a careful assessment of tourists’ segmentation and their needs. The main components of the model are five: the real phenomena that form up the potential world of interest, the universe of potential sources, a web taxonomy, a domain thesaurus and an interoperability hub. The St James’ Way is used as a source of examples. It is concluded that thesauri based on ISO 25964 offer a great potential for the simple, flexible, dynamic and distributed interconnection between the institutions of memory (GLAM+, research institutions in digital humanities and social sciences, transparency portals…) and the growing demand from the tourist sector for a more personalized and contextualized experience that can make a difference in an increasingly competitive international market.

1.0 Introduction: context and motivation

It has become a common experience that web information is changing tourism and travelling, but a new phase is developing in the last years. There is increasing evidence of “a growing “bifurcation” between traditional online travellers, i.e., those who use the Internet for standard travel products, and those who are beginning to adopt alternative channels and products in search of deeper and more authentic experiences”, with the first market entering into a relative stagnation and the second offering new opportunities for combining different products (Xiang et al. 2015). Some experts have even identified a ‘cultural turn’ in tourism (Dabbage 2018a, 55-56; 2018b).

In this context, relating the immaterial, material, artistic, bibliographical and archival heritage to competent touristic proposals and infrastructure to be used by people of different nationalities, languages and cultures is becoming a key strategical challenge for both sectors. On one hand, the tourism sector can profit from more contextualized, personalized and interesting information resources. On the other, Humanities and GLAM+ (acronym for galleries, libraries, archives and museums, and, in general, the heritage preserving institutions) can improve their visibility and relevance through an undisputable, applied and practical contribution to social and economic development.

Rich references to cultural artefacts–both before, during and after the visit (in webpages, VR applications, QR code support, augmented reality…)–can help tourists to opt for a particular tour; decide future activities outside the standard ones, widening choices and benefiting the local tourism market; improve their cultural, educational and life experience; promote word-of-mouth recommendation; and enhance the acquisition
of cultural and historical lessons and knowledge. Such a vision seems a gain-gain one both for tourists, destinations, and tourism agents and organizations.

Within this framework, this paper explores how experts from the field of knowledge organization and information architecture can collaborate on devising and proposing action lines to improve the feedback between both sectors (tourism and cultural heritage institutions). In a first stage, by facilitating the use of the huge amount of data and digital artefacts made available by the information and communication professionals, humanists and social scientists, which allow the contextualization, enrichment and personalization of the touristic experience in its relation to relevant cultural objects. Second and reciprocally, by looking for strategies to enhance the transfer of resources from the tourism industry to the field of basic research in the humanities and social sciences, promoting its economic sustainability by connecting it with its potential market uses. This last point seems specially important at a time when Humanities have become neglected by funding agencies because of the longstanding economic constrains following the 2008 crisis and the increasingly acrimonious cultural wars between global and identitarian political stakeholders.

2.0 Aims and research questions

The overall intention of this paper is to explore how KO research and development can contribute to the interconnection of the institutions of memory (libraries, archives, museums, documentation centres, research institutions in digital humanities and social sciences, transparency portals…) and the needs of cultural tourists, an increasingly important industry (Fang 2020); and to develop a model that can contribute to frame and guide future research on the field. Specifically, six research questions were addressed:

Which roles can the GLAM sector perform inside the digital information ecology of cultural tourism? Which are their implications for KO? How can the relation between the GLAM and cultural tourism sectors be modelled to reveal the central role of KOS in such interactions? What are the relevant characteristics of KOS used in the main GLAM subsectors? Are they interoperable with cultural tourism websites? On which terms? To answer these questions, in the following sections, the main agents and factors involved in the interoperability between the GLAM and cultural tourism sectors are identified and considered from the point of view of Knowledge Organization, that is, of the role that KOS might have in their successful interaction; and a model is proposed. Finally, the problems and opportunities for KO in this field are identified.

4.0 Analysis and discussion

In the next sections, we will consider the relation among cultural tourism and the GLAM+ sector in the digital age as a potential information ecology. In such an ecology, three main kind of agents and artefacts can be identified: tourism information mediators and their web sites; cultural heritage institutions (GLAM+ sector) with their information sources and databases; and end users (tourists and travellers) with their needs, which are partially known but must be partially disclosed.

In this emergent ecology, KO can contribute to close the gaps among their agents in three ways: modelling the connection among GLAM+ resources and user needs; identifying knowledge representation and organization technologies and methodologies that
can be useful to implement the model; and offering an operative model to start experimentation. Two contextual problems outside the model must be also stressed because of their current importance: communicational issues and legal concerns.

4.1 The information ecology of cultural tourism

The concept of ‘information ecology’ allows the modelling of information systems that have not been purposely designed, in contrast to those compact and well-differentiated from their environment, like libraries, archives and information centres. An information ecology (Hubermann 2001; Shim and Lee 2006; Sebastiá 2008) can be defined as a network of information organisms interacting among them and with their environment to form a complex system. This concept is very useful to think about situations where different information agents co-exist, cooperate and compete to fulfil an information need. This is the typical situation in the Internet; and, in our case, what users interested in cultural tourism will experience when trying to solve their information needs.

For example, Internet information on the St James’ Way is provided by a distributed network of independent agents: the non-governmental sector (associations, religious institutions, informal groups and individuals), private for-profit companies (travel services firms, publishing houses, consumer cooperatives...), and public institutions (council, national and regional governments...) (López del Ramo and García Marco 2018). Each of them has its own aims and provides specific information services to pilgrims and tourists, who, on their part, are also very segmented regarding their specific interests.

4.2 The tourist information mediators (providers and associations)

Though GLAM+ institutions sometimes address proactively the information needs of tourists, relevant information is generally vehiculated through the tourism industry, and more precisely by those institutions and departments that are specialized in connecting tourists and destinations. This is mainly a marketing activity, and their more specialized agents are the so-called destination marketing organizations (DMOs): institutional, council, regional, national tourist information offices, and their background marketing departments. DMOs can be the main mediator agents between GLAM+ institutions and tourists, though online travel agencies (OTAs) and social media have been gaining prominence (Xiang et al. 2015). Of course, it should not be forgotten that there are big GALM+ institutions that have great DMOs inside them, e.g., big museums and galleries. Associations of experts and fans related to the protection and dissemination of cultural heritage are also another important sector of mediators in the field of cultural tourism. Internet has brought a revolution to marketing, and now even the smallest organizations can have an Internet global presence.

From the point of view of tourist mediators, knowledge organization may have two different uses, one internal and the other external. The internal objective is related to knowledge management: the representation of knowledge into information; information preservation, retrieval and sharing; and the transformation of information into knowledge. For this, a corporative KOS is needed.
The external goal is to communicate the part of this information that is relevant to tourists, so that they become aware, transform it into knowledge and hopefully choose the proposed destinations for their travels. This is usually done through websites and social platforms, but also increasingly by more and more sophisticated mobile applications. In this regard, tourist information systems must be seen as integrated, with a core of organized information and a set of well-established distribution procedures. The main KO tools for these purposes are web taxonomies in the case of cultural tourism webs, and folksonomies and ad hoc taxonomies in the case of social networks and blogs.

Of course, knowledge organization experts now that, for proper functioning, both aims must be connected into a successful knowledge organization system, though both of them should be efficaciously addressed and easily differentiated. In fact, both functions are more or less separated in many organizations: sometimes one is absorbed or displaced by the other; other times website and corporate information databases are not well communicated. Frequently, the corporate taxonomy becomes the structure of a tourist website, and, as a result, the website reflects more the ontology of the organization than that of tourists (López del Ramo and García Marco 2018) or is mainly product-oriented; but there are also many well-designed sites from a user-oriented perspective (Table I).

<table>
<thead>
<tr>
<th>Table I. Main categories in the taxonomies of four St James’ Way websites</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>caminodesantiago.gal</strong></td>
</tr>
<tr>
<td>Galicia government tourist-oriented website</td>
</tr>
<tr>
<td>Discover</td>
</tr>
<tr>
<td>Get ready</td>
</tr>
<tr>
<td>On bicycle</td>
</tr>
<tr>
<td>Plan your trip</td>
</tr>
<tr>
<td>‘Turismo España’</td>
</tr>
<tr>
<td>On the way</td>
</tr>
<tr>
<td>Recyling</td>
</tr>
<tr>
<td>Knowl. and research</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Xunta de Galicia</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

4.3 The side of the information sources: GLAM, research centres and publishing

The information sources side is quite complex. Any approach to integrate GLAM resources in tourist information sites requires enhancing interoperability among very diverse systems.

Libraries are by far the more standardized GLAM subsector, though the gap is being closed very quickly because of the pressure for global and integrated access that Internet at the same time offers and requires. Their resource description and interchange standards are fully international and integrated (MARC21 family), and to a lesser grade also are their KOS, both systematic classifications (LCC, DC, UDC…) and alphabetic subject headings lists (LCSH, RAMEAU, EMBNE…). There has been even a strong work on integrating systematic and alphabetic KOS in libraries, some finished and other still
on course (LCC-LCSH, UDC and national subject heading systems in Europe...). Also, relevant mappings among competing systems are on course (e.g., Slavic 2011). These are strong points for the interoperability of the GLAM and cultural tourism sectors. But there are also weaknesses from the tourist or traveller perspective. Classifications frequently do not provide the level of specificity required to map sources to their needs; and topics are scattered among many classes, doing mapping projects really complicate.

For example, there is no class in the UDC for St James’ Way. The classification of works about the St James’ Way is usually done in the Regional Geography class (913), and less in Routes, etc. (656.022) or in the many (and scattered) classes available for Travel (e.g., 338.48-12, 656.022.33), as transport and tourism are in different trees. For the purpose of adding specificity, this selected class is frequently faceted in Spain by the two main countries: (44+460), that is, France and Spain (to denote the “Camino Francés”), though the actual countries may change (e.g. Portugal). The National Library of Spain has noted the lack of specificity of the expression, and its librarians have usually added “Camino de Santiago” between the brackets and after the country codes.

Subject headings are usually much more specific. Though the exact subject heading is not present in many international systems, it can be easily recognized in the strings:

- EMBNE: Peregrinaciones cristianas — Santiago de Compostela
- LCSH: (Christian pilgrims and pilgrimages-Spain-Santiago de Compostela)
- RAMEAU: [Pèlerinages chrétiens-Espagne-Saint-Jacques-de-Compstelle (Espagne)]

Therefore, there is a strong potential for an easy interconnection of libraries and cultural tourism websites. But a serious problem stands in the way: practical travel decisions require data, not systematized knowledge, and there is no easy method to transform books and other complex library materials into high quality data without the cooperation of the research, publishing and media sectors, as it will be discussed later.

Big museums, which rely on very analytical databases for the control of their collections and operations, are at the forefront of providing semantic data that can be linked by external sites. This data provides factual, contextual and bibliographical information on pieces of art and craftsmanship and reproductions that can be inserted in websites and A/VR applications. The CIDOC CRM ontology, originated in the museum field, provides a full frame for the interoperability of other controlled vocabularies and is being influential in the GLAM sector. Museums have also sound KOS: AAT, Iconclass, etc.

Archives, their fonds and documents are rarely used by tourists and their mediators in a direct way. There are interesting projects linking historical documents to heritage sites, monuments and pieces of art, but this kind of information is usually secondary for tourists: they will use it only out of curiosity and not for practical decisions. Archives are also very idiosyncratic in their intellectual organization, as their KO leitmotiv is the principle of provenance and respect for the original order, which is complemented by the macrolevel organizations based on function to allow the navigation of the multiple changes that organizations suffer along their life. Anyway, authorities have grown increasingly important for the multilevel description in archives, and there is a strong movement to make library, museum and archival headings compatible.

Outside GLAM but closely related, there is a huge industry that is also fully involved in knowledge preservation: the publishing sector. Though a first sight could suggest that
their role is incorporated through the library network, editorial organizations have a huge task ahead that is crucial for the successful interoperability among the cultural and scientific sectors: adapting the publishing practices to the promises and requisites of the semantic web and big data revolutions. Cultural tourists have two broad categories of information needs: a) gaining broad perspectives for travel planning and background knowledge development, and for this purpose books and journals are totally suitable; b) acquiring and accessing data throughout different sources, and it is here were the digitally-enhanced user has problems. Most books are oriented to the first need; and, as they are not usually automatically searchable and navigable through library systems, a big gap is opened between both worlds. Semantic web offers the technology to bridge this gap, but editions and library catalogues must be transformed to fulfil this promise.

What it has been said about the ‘traditional’ publishing sector must be applied to the current multimedia environment, including audio-visual industries and videogaming. Other emerging ones, like augmented and virtual reality, use semantic technologies by default.

Strictly, the GLAM acronym does not include the producers of information; and, as a great part of the task of producing semantic- and data-supporting documents rely on them, it is suggested that an extended superset should be considered to include publishers and other media producers, e.g. GLAM+.

4.4 The tourist side: end users first?

It has been observed that “huge discrepancies exist between the domain ontology derived from tourism Web sites and the one emerging from user queries”, mainly because the tourism industry uses a very specific terminology and categorization that is not simply connected with the terms that users actually search, and because, apart from a group of overrepresented categories, the “the overall domain is extremely rich and largely idiosyncratic […] with numerous destination specifics” (Xiang et al. 2009). Though consumer generated contents – e.g., blogs and reviews (Gretzel, Hwang, and Fesenmaier 2006) – can be used to learn about the language that travellers use, filling the terminological gap; there are problems that do not seem easy to solve, especially those derived from the intrinsic limitations of attention, processing memory, interface space, and the representation of a really complex domain.

Problems increase when taking into account that cultural tourism is a huge sector and that tourists themselves are indeed very segmented. Only in the field of cultural tourism, McKercher and Du Cross (2002) identified five types of tourists: purposeful, sightseeing, serendipitous, casual and incidental; and Csapó (2012) classified seven different kinds of cultural tourism products for them: heritage tourism; cultural thematic routes; cultural city tourism, cultural tours; traditions, ethnic tourism; event and festival tourism; religious tourism, pilgrimage routes; and creative culture, creative tourism.

On the other hand, this gap between the overwhelming worlds of information and tourists is a very interesting challenge that has attracted a lot of talent. In particular, modelling users through ‘ontological’ profiles, based on types and/or personality traits, is an effort that has been on course for more than a decade and now constitutes the established research topic of tourism recommendation systems (Grün, Neidhardt, and Werthner 2017). However, GLAM+ connection requires specific typologies for both
tourists and sources, based on their information needs: broad ones–planning, and con-
necting and developing their mind maps–and specific ones–solving problems with rel-
levant data, and navigating among sources in search of unanswered questions–.

In this situation, mediators and users can work through general-purpose search sys
tems like Google, or with the source providers’ KOS. In the first case, users will
obtain relevant selections, but not necessarily very precise when leaving the first results,
nor exhaustive or filtered by source. With the second approach, these aims could be
better served, but a difficult work of knowledge organization engineering will lie ahead.
The first evidence gained in any experience in KOS/GLAM+ interoperability for
cultural tourism is that it is no easy task, specially when leaving the basic operations
and data and getting into the realm of learning, sense-making and culture.

4.5 Bridging the gap (1): towards an ontological and epistemological model

Thus, can a model be developed to integrate the different information needs, struc-
tures and finally perspectives of GLAM+, DMOs and tourists so that a truly cultural
tourism information ecosystem may be born?

In our opinion, any proposal has to take in account four different layers: the cultural
world (artefacts, persons and organizations, sites, abstract realities), counting on the
KOS that are already functioning in the GLAM sector; cultural tourism ‘science’ and
operation, as expressed in some KOS, like the important WTO thesaurus (World Tour-
ism Organization 2001), in the tourism website taxonomies and some prospective ont-
tologies that are being developed (Li, Buhalis, and Zhang 2013); and the last one for
representing the user. To this last aim, Maslow’s (1954) pyramid of human needs seems
especially well suited to frame the classification and interrelation of so different layers
as basic needs (accommodation, eating and drinking, security…), social ones (identity,
relation…) and self-actualization (sense-making, high culture, wisdom-building…).

4.6 Bridging the gap (2): available standards and technologies

To model such a KO ecology, fully developed technologies and standards are avail-
able both in the field of KOS and in the semantic web realm. Regarding networking
knowledge organization systems (KOS), thesauri offer a great potential for the simple,
flexible, dynamic and distributed interconnection between the institutions of memory,
especially after their relaunch with the new ISO 25964 standard (Aitchison and Dextre
Clarke 2004; Dextre Clarke 2012). In particular, Part 2 of ISO 25964 deals with the
 interoperability of all these different KOS implied in the information ecology of cultural
tourism, and proposes specific mapping models and specific devices (EQ, ~EQ, +, |,
Bm, NM, RM). In our project, we are experimenting with a hub architecture, but a lot
of system-to-system projects are on course, so the environment is changing quickly.

Dextre Clarke (2011) has made a very clear diagnostic of what are the expected re-
results of mapping thesauri and the main kinds of KOS, and Soergel (2010) has proposed
a clear conceptual framework that also considers facet-based search. In the particular
case of cultural tourism, more positive results can be expected from interoperability with
authority and subject lists, because they accurately represent cultural artefacts (through
titles), persons and places; though the problems will persist for higher level concepts.
On the part of the semantic web research, W3C has by now greatly completed the standards that deal with its lower and middle layers; thesauri have been successfully expressed in RDF, OWL and SKOS; and tourism recommender systems based on semantic web and ontologies have become one the main research fronts in information science and tourism: “Recommender systems based on semantic web and ontology technologies are an effective method and tool to improve the quality of internet service through personalization and customization.” (Li, Buhalis, and Zhang 2013).

4.7 Bridging the gap (3): a methodology

KO interoperability is difficult when working with already well-seasoned KOS, but it becomes a mess when trying to connect ad hoc web taxonomies and folksonomies—on the side of DMOs and other traveller-oriented websites—with GLAM+ KOS on the other. In a current project on the Aragonian part of the St James’ Way, we have divided our work in two lines of action.

On one hand, the different websites are being classified in homogeneous groups and their taxonomies studied: council, regional and national DMOs, mainly from the public sector; pilgrim associations; and business providing information (publishing houses, consumer cooperatives, hotel chains, transport firms...). In this way, it is possible to obtain some sort of a ‘least common multiple’ of all these empirical taxonomies, including all their concepts without repeating them. Equivalent terms, which are extremely important in the “language of tourism” (Xiang, Gretzel, and Fesenmaier 2009), are then controlled, so they can be used to resolve searches; taxonomy-incompatible polyhierarchical relations are resolved into BT/NT and RT; and the relations among the taxonomy categories (now thesaurus concepts) are expressed through RT. As a result, a special kind of thesaurus emerges, which could be called a Common Compact Taxonomical Thesaurus (CCTT). Each of its concepts is to be a node (page) in a trial CMS-supported website. Thinking in future interoperability, a SPARQL point should be developed, with an independent system behind to manage prospective and external KOS links.

On the other hand, the most used KOS and knowledge representation systems in the different GLAM+ sectors are being studied to explore their potential interoperability. Prospective analysis has shown that mapping each KOS to the others would be a difficult, uncertain and huge task; and that it is also unfeasible to take one of them as a hub. Therefore, the idea is to connect the subsets of them that offer greater potentiality (because they are related to the concepts of the CCTT) to the nodes of the website.

Though the best solution for actual interoperability would be to establish an independent KOS hub, for the moment we intent to build a trial DMO website that will incorporate the pilot CCTT as an extended taxonomy; and then link the external resources to each node of the CMS-based website, based on a case-by-case analysis of their correspondences with the controlled tags in the sources. Complementarily, relevant Wikidata nodes and their relations could be mined to produce a set of mappings.

All these steps are based on a presupposition: that DMOs KOS really express user needs, but this extreme should be further researched, because actual evidence is limited. This gap can only be filled in the future by doing user studies: both subjective (satisfaction surveys, etc.) and objective (eye tracking, log analysis, search analysis...). Certainly, the strong need to connect both subdisciplines of information science (i.e., KO
and user studies) becomes even greater when approaching multidisciplinary and multi-platform fields, like cultural tourism and GLAM+ integration. Only in this way, the whole cycle of evidence-based KOS design and research could be effectively closed.

4.8 Beyond KOS and linked open data: legal concerns and communication issues
Although the paper is focused on the analysis of networking knowledge organization systems, there are other aspects that are essential for their operations in the real world: in particular, communicational issues (communication security, effectiveness of the informational message) and legal concerns (privacy, data protection and intellectual property, basically), especially considering the problems created by the massive data processing that is becoming inherent to the Internet, which requires the use of algorithms that respect the legislation and standards on human rights in the use of open data.

5.0 Conclusion
In this paper, it has been shown that KO, cultural tourism, GLAM+ and humanities can develop a fruitful alliance as disciplines. In the digital realm, cultural tourism needs humanities to support sound and personalized products, incorporating GLAM+ resources and data to enhance tourists’ experiences. To integrate all this information, it has a networking KO gap. On the other hand, cultural tourism is a fascinating subject for KO. First, it is a truly transdisciplinary field, while far from compact. Second, it brings forward the problem of end users’ segmentation with their different information needs, and therefore diverse conceptual maps.

In practice, a KO-enhanced operative model for the networking of GLAM and cultural tourism websites has been outlined, generalizing from an on-going project on the St. James’ Way. As a conclusion, it can be affirmed that networking KOS— in particular thesauri following its relaunch after the new ISO 25964 standard (Aitchison and Dextre Clarke 2004; Dextre Clarke 2012) —offer a great potential for a simple, flexible, dynamic and distributed interconnection between the institutions of memory (libraries, archives, museums, documentation centres, research institutions in digital humanities and social sciences, transparency portals…) and the growing demand of the tourism sector for a more and more personalized and contextualized experience that can make a difference in an increasingly competitive market (Abrahams and Dai 2005).

Acknowledgments
This research work was carried within the CSO2015-65448-R (MINECO/FEDER) project. We are also grateful to the anonymous referees for their suggestions.

References