310

Knowl. Org. 51(2024)No.5
E. Daga. Citizen Experiences in Cultural Heritage Archives: A Data Journey

Citizen Experiences in Cultural Heritage Archives:
A Data Journey

Enrico Daga

The Open University – Walton Hall, Milton Keynes MK7 6AA, United Kingdom

enrico.daga@open.ac.uk

Enrico Daga is a Senior Research Fellow at the Knowledge Media Institute (KMi) of The Open University in the UK, where he obtained a PhD in 2018. After developing intelligent systems for Ontology Engineering (NeOn) and Smart Cities (MK:Smart), his current research explores the application of knowledge graph technologies in data-intensive socio-technical environments: cultural heritage, smart cities and robotics, and healthcare. His current focus is developing novel knowledge graph construction and curation methods, including policies and methodological information. Recently, he engaged in two EU-funded projects in Cultural Heritage and Music: SPICE -- as co-investigator and Polifonia -- as Technical Director and Principal Investigator for the OU. Both projects established knowledge graphs as the primary solution to integrate and make sense of heterogeneous data for supporting frontier humanities scholarship and cultural engagement. OrcId: 0000-0002-3184-5407

Abstract: Digital archives of memory institutions are typically concerned with the cataloguing of artefacts of artistic, historical, and cultural value. However, experiencing cultural heritage requires engaging with the so-called cultural background (historical, social) but also, and possibly more importantly, relating the objects (artefacts, artworks, ...) *to our own experiences* and, eventually, the experiences of others, which cannot be forced into a unique, objective meaning. Recently, new forms of citizen participation in cultural heritage have emerged, producing a wealth of material spanning from visitors' experiential feedback on exhibitions and cultural artefacts to digitally mediated interactions like those on social media platforms. Citizen curation promotes the adoption of intelligent, extended technologies for cultural heritage engagement that mediate the production, collection, interpretation, and archiving of people's responses to cultural objects, favouring the emergence of multiple, sometimes conflicting viewpoints and motivating the users and memory institutions to reflect upon them. There are good reasons for institutions to archive people's responses to cultural objects. Therefore, here we focus on the impact of integrating citizen experiences in cultural heritage archives. As we rely on complex systems to support the management of cultural heritage collections and digitally mediated systems to enable innovative engagement applications, it becomes vital to equip underlying infrastructures with means for monitoring, capturing, and explaining what users do with those systems. By analysing the case studies of the EU-funded SPICE project, we argue that a knowledge organisation system for "data journeys" can help disentangle problems that include distribution, sense-making, ownership, sensitivity, privacy, and rights management.

1.0 Introduction

Digital archives of memory institutions are typically concerned with the cataloguing of artefacts of artistic, historical, and cultural value. In 2015, The Warwick Commission found that "the gap in participation between the white and BAME population is widening" and reported that UK residents from higher socioeconomic groups accounted for 87% of museum visitors. The challenge was not only accessibility but also "a mismatch between the public's taste and the publicly funded cultural offer" (Neelands et al. 2015). Such concerns are part of the debate since the Faro convention on the value of cultural heritage for society (Council of Europe 2006) where it was declared the need to "involve everyone in society in the ongoing process of defining and managing cultural heritage" and that active participation to

Knowl. Org. 51(2024)No.5
E. Daga. Citizen Experiences in Cultural Heritage Archives: A Data Journey

311

cultural heritage is one fundamental right of citizens. This vision has been echoed more recently by the International Council of Museums (ICOM)[1], whose definition of a museum highlights the value of diversity and inclusion of multiple perspectives[2]. Museums aim to be participatory and collaborate with diverse communities to contribute to the collection, preservation, interpretation, and understanding of our heritage. Experiencing cultural heritage requires engaging with the so-called cultural background (historical, social) but also, and possibly more importantly, relating the objects (artefacts, artworks, ...) *to our own experiences* and, eventually, the experiences of others, which cannot be forced into a unique, objective meaning (Whitehead 2011). Thus, new forms of citizen participation in cultural heritage have emerged, producing a wealth of material from visitors' experiential feedback on exhibitions and cultural artefacts to digitally mediated interactions like those on social media platforms.

Crowdsourcing initiatives have been considered by cultural heritage archives as a way, for example, to enrich the library metadata (Ridge 2013), targeting users who are typically happy to volunteer instead of focusing on creating new spaces of engagement. Enabling multiple voices in museums is better reflected in initiatives to decolonise museums that introduce new perspectives to challenge dominant narratives (Coombes and Phillips 2020). Grassroots projects have emerged intending to document and preserve the experiences of communities outside the mainstream, but those are developed independently from cultural heritage institutions (Flinn et al. 2009). A useful direction sees the direct involvement of artists to imagine new modalities. In the European project GIFT, artists, museum professionals, and researchers design new types of hybrid experiences, such as the Gift App, where users use their smartphone to create a digital object to capture the cultural heritage artifact as a "gift" for someone they care about (Back et al. 2018).

1.1 Citizen Curation

Citizen curation (Bruni et al. 2020) promotes the adoption of intelligent, extended technologies for cultural heritage engagement that mediate the production, collection, interpretation, and archiving of people's responses to cultural objects, favouring the emergence of multiple, sometimes conflicting viewpoints and motivating the users and memory institutions to reflect upon them. In the EU project Social Cohesion, Participation, and Inclusion through Cultural Engagement (SPICE), museums and researchers experiment with the citizen curation idea to foster participation and inclusion, targeting communities of users whose voices are often left unheard, such as asylum seekers, people living with illnesses that prevent physical visits, and people in secure environments (Stoneman et al. 2021). An intermediate *linked data layer* supported the applications by mediating between the collections' metadata archives and the end user applications: the SPICE Linked Data Hub (LDH)[3]. In what follows, we will use the SPICE project to learn about citizen curation and use its findings to reflect on how citizen curation impacts data infrastructures. SPICE case studies can be briefly summarised as follows, reflecting the leading cultural heritage organisation that hosted the co-design activities:

- Design Museum Helsinki (DMH)[4], developed by Aalto University in collaboration with the Design Museum Helsinki. The Pop-Up VR Museum aims to bridge the physical accessibility gap, making it easier for people to experience art and culture (Vishwantha et al. 2023).
- *Galleria d'Arte Moderna (GAM)*[5], developed by the University of Turin in collaboration with
- *Fondazione Torino Musei*[6], *GAM*. Gam allows visitors to create stories to document their moods and reactions to the contents they encounter during the visit (Lieto et al. 2023a).
- Hecht Museum (HM) led by the University of Haifa. The Hecht's Museum[7] case study focuses on engaging with students of secondary school(s) where historical artifacts are linked to a historical event (the Galilee rebellion); participants elaborate short essays expressing diverse opinions regarding historical and national issues (SPICE 2022).
- Irish Museum of Modern Art (IMMA)[8]. The Deep Viewpoints system, developed by The Open University, is based on the 'slow looking' methodology, where the experience of the artworks is mediated through prompts and questions, and user responses are collected and confronted, and visitors propose their own prompts for other users, effectively scripting their own citizen curation (Muhlholland et al.2022).
- *Museo Nacional Ciencias Naturales (MNCN)*[9]. Padaone games develops "serious games" for cultural engagement. In this case study, a treasure hunt game mixes puzzles, quizzes, and questions linking objects in the Natural History Museum to themes of environmental sustainability (Gutiérrez-Sánchez et al. 2023).

1.2 A general workflow in citizen curation systems

A recent survey on the topic (Daga et al. 2022) covers requirements, state-of-the-art technologies, and infrastructures for citizen curation. It characterises citizen curation from the view of user roles and devises a general workflow[10]. Figure 1 illustrates a typical citizen curation scenario. It starts from a cultural heritage digital asset belonging to some original author or copyright holder, then moves to how it is collected and curated by a cultural heritage in-
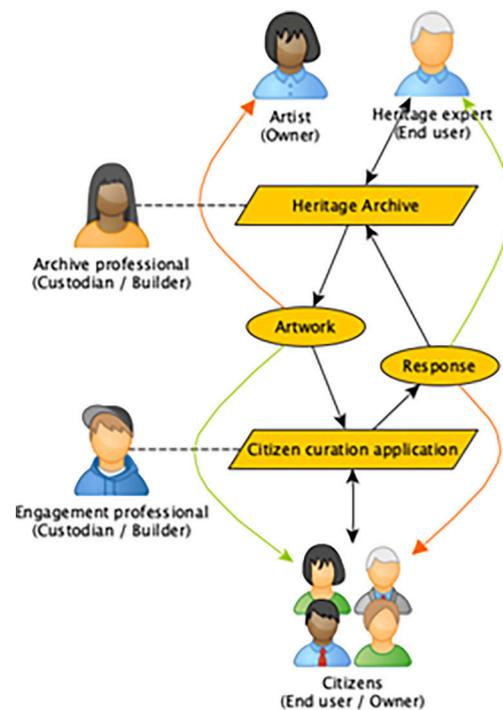
312

Knowl. Org. 51(2024)No.5
E. Daga. Citizen Experiences in Cultural Heritage Archives: A Data Journey

Figure 1. Citizen curation scenario, abstracted from the case studies of the SPICE project.

stitution, and finally, how it is used in a citizen engagement system, typically produced by a third-party organisation, for example, a company active in the tourism sector. We can identify four major roles: the 'owner', the 'custodian', the 'builder', and the 'end user'. The owner is the copyright holder of the cultural heritage asset; the custodian is the intermediate organisation (a museum); the builder is the company that produces the system that processes the digital object; and, finally, the visitor is the end user.

It is interesting how citizen curation turns this model upside down. In a citizen curation application, all the above is still valid, but, in addition, the roles apply in reverse. Visitors produce initial responses to the artworks and may be acknowledged as authors (copyright owners) who delegate their content management to a platform provider. Thus, such novel, digital-born cultural heritage is handled by an intermediate organisation (e.g. the tourism company) that has to oversee the content produced and, for example, ensure its quality and compliance with regulations (e.g. the content itself does not violate any law). This organisation acts as a 'custodian', with a relationship to the citizen like the one that ties museums to artists. The application provider then passes the newly acquired content to the museum professionals that select and curate citizen responses, thus 'building' on the content received and archiving it in the collection management system (or its extension tailored to citizen contributions). Finally, curators, researchers, and histo-

rians are the 'end users' of this journey following citizen contributions into cultural heritage archives (Daga et al. 2022).

The black arrows indicate the movement of data objects; the dotted lines indicate the role of intermediaries; the orange line ownership (and copyright); while the green arcs point to who is the target end user.

It is straightforward to recognise how such a scenario raises several problems with information management that go well beyond what is typically supported by collection management systems.

– Sense-making. These systems generate a wealth of information that is strongly dependent on the method used for their collection. To make sense of this information, capturing the modality of their production is fundamental, as well as explaining citizen curation applications alongside the generated content.
– Ownership. Users produce original content, and there is a question of whether they should be recognised as authors and whether such recognition should be promoted by the cultural heritage institution, for example, with attribution statements.
– Monitoring sensitive content. The produced content has the same characteristics as social media data, which can include inappropriate or sensitive information. Thus, monitoring and moderation is essential before it is included in the archive or repurposed. In addition, the con-

Knowl. Org. 51(2024)No.5
E. Daga. Citizen Experiences in Cultural Heritage Archives: A Data Journey

313

tent may include personally identifiable information that may violate privacy regulations. Crucially, museums could be seen as being responsible for or endorsing opinions found in user-generated content.

– Terms of use. Terms and conditions may restrict the use of digital images, and users should be informed of any restriction when designing the experiences (e.g. including artwork in a co-design workshop) and when the contributed content is collected and associated with the original artwork. In addition, these systems aim to support citizens in sharing their contributions with the museum and each other, opening essential issues regarding rights and terms of use of the generated content.

Considering this research program, it is an open question of what type of knowledge representation could support citizen curation. As we rely on complex systems to support the management of cultural heritage collections and digitally mediated systems to enable innovative engagement applications, it becomes vital to equip underlying infrastructures with means for monitoring, capturing, and explaining what users do with those systems. However, what does this mean from the 'knowledge organisation' standpoint?

2.0 Data Journeys

2.1 Background

Provenance is a well-established notion in museum curatorial practice, where it is related to ensuring the quality and lineage of an object as part of the acquisition management phase. This idea has been borrowed by information science research, which reformulates it as the problem of describing how a certain information object has been produced, who is responsible for it, and associated usage requirements. Digital library research stresses the importance of understanding the context in which catalogue metadata is being produced and the impact that such background has on how the catalogued items are perceived (Chowdhury 2010). Provenance plays a key role in the web literature (Moreau 2010), considering the people as content creators and advocating for integrating this feature in the semantic web (Harth et al. 2007). In this declination, provenance becomes a relevant concept for us, where the assets produced by citizen curation activities are supposed to be managed as first-class objects in museum archives. This line of research is being resurged recently in the context of data studies, with the notion of 'data journeys', defined it as the *"movement of data from their production site to many other sites in which they are processed, mobilised and re-purposed."*(Leonelli 2020, 9). The work in data studies emphasises the difficulty of empirically understanding data journeys because of many perspectives[11]. Daga and Groth (2023) introduced a layered seman-

tics perspective to the definition of data journeys: "a Data Journey is a multi-layered, semantic representation of a data processing activity, linked to the digital assets involved (code, components, data)".

2.2 The Data Journeys Ontology

The Data Journeys Ontology (DJO) capitalises on semantic web research on workflow representation and reasoning (Daga et al. 2015; Garijo et al. 2014; Lieto et al. 2023b). DJO identifies two abstraction layers: a data node graph linked to data objects (resources) and their changes within the process and an activity graph connecting high-level operations. Figure 2 shows the hierarchy of classes defined by the ontology. Apart from the type Data Journey, the ontology defines two main top-level types: activities and data nodes. Activities represent operations performed on the data, while data nodes qualify roles of resources in the process, either files, variables, or temporary objects.

Thus, a multi-layered data journey allows a multiplicity of perspectives that can be overlaid to describe the process. This multiplicity can help capture the context around a data journey while still allowing for computational analysis. Layering representations allow linking them to the concrete assets involved (e.g., with URLs) but also defining incremental intermediate abstractions.

Fundamentally, here we argue that the journey a citizen curation object goes through, its *lineage* or *provenance*, is a powerful way of describing citizen curation applications. We pose two main questions: (1) how do DJO components relate to citizen curation? (2) what additional representational layers are needed to capture the complexity of citizen curation?

3.0 Data journeys in the cultural heritage

We can use the five case studies of the EU project SPICE to reflect on the applicability (and utility) of data journeys. Table 1 shows a summary of our analysis of the five SPICE case studies. Although DJO specifies both classes and relations, here we mainly focus on discussing activity types, which seems the most reasonable way to approach the problem and leave an analysis of data node relationships to future work.

Thus, we can look at how DJO activity types relate to citizen curation:

– Retrieval. This activity is performed in all applications scrutinised, who need to access items in different ways; it can relate to finding an artwork within a collection or collecting user responses.
– Preparation. Most scenarios require a content curation phase, where developers (builders) setup additional

314

Knowl. Org. 51(2024)No.5
E. Daga. Citizen Experiences in Cultural Heritage Archives: A Data Journey
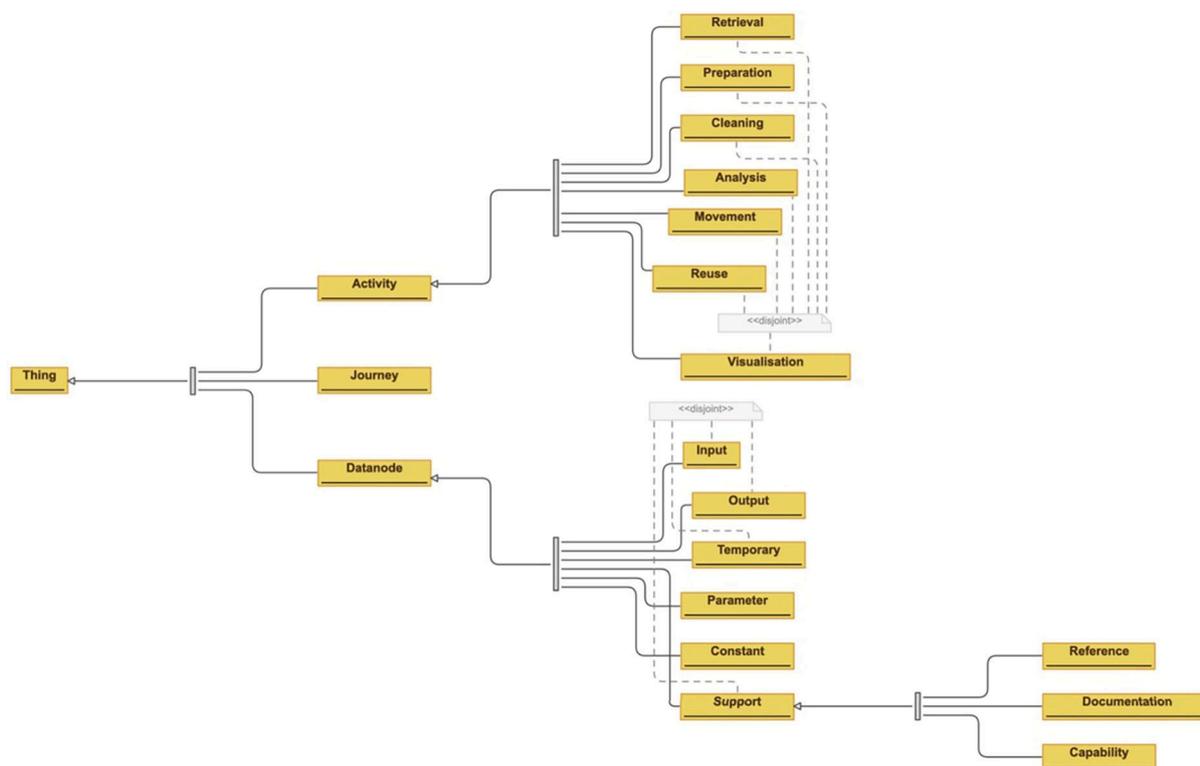

Figure 2. Class hierarchy of the Data Journeys Ontology (DJO), from (Daga et al. 2022).

metadata required by the application, which can be different from what is typically considered by cultural heritage archive metadata (it is the case of VR objects that are newly acquired). In some cases, like the Deep Viewpoint system, museum practitioners or even visitors engage in a curatorial activity, preparing new experiences for future users.

– Cleaning. In computational data science, cleaning refers to filtering out unwanted data points. This notion relates well with the case of citizen contributed content who can be considered harmful, for example, for including aggressive or hateful speech, or for disclosing personal information. Citizen contributed content is monitored, moderated, and possibly removed.

– Analysis. Analyses can be quantitative, such as statistical methods to visitors' responses. More complex analytics involve automatic or semi-automatic content analysis algorithms (e.g. to identify content that requires monitoring, e.g. detecting hate speech), or automated reasoning (e.g. for emotion classification (Lieto et al. 2023a). All these methods somehow augment the data. Content analyses in citizen curation applications span different methods but the most striking difference with data science is the role of human intervention. This is not surprising but opens the opportunity of expanding data journeys to cases where the workflow is more complex than a input/output data flow, and multiple user interventions occur during the process.

Crucially, new input in the form of annotations can be produced, in search for similarities or differences between citizen responses (compare) and are typically mediated by user interfaces. However, qualitative analyses may be also performed offline with the aid of questionnaires, surveys, or in focus groups.

– Movement. Citizen curation applications transfer data across different systems. In SPICE, these include one (or more) engagement interfaces (e.g. via tablets or mobile phones) often sustained by Web APIs of intermediate, "headless" applications which, in turn, access catalogue metadata collected and augmented on a common 'data hub'. Such data hub acts as mediator between application and cultural heritage archival systems and takes care of publishing the collections' digital items (and their metadata) with the applications and, in turn, support the sharing of such metadata across applications, and the sharing of collected responses with the museums archive (Carvalho et al. 2023).

– Reuse. In citizen curation applications, the most widely form or reuse relates to cultural heritage digital artefacts being picked by museum professionals at the beginning of the process (select). Digitization processes can be also considered forms of reuse. Visitor responses can be analysed by dashboards and other tools that are separated from the systems where they originate from.

Knowl. Org. 51(2024)No.5

315

E. Daga. Citizen Experiences in Cultural Heritage Archives: A Data Journey

Case Study	Re-trieval	Preparation	Cleaning	Analysis	Movement	Reuse	Visualisation
DMH	Access Collect	VR/AR digital object setup (Curate) Metadata preparation (Curate)	-	Annotate (Human medi-ated analyses / Augment)	Share Publish Data hub (Mediate)	Select Scan artifact (Acquire)	Engage Sense-making Multi-Modality Immersive VR-Interaction AR-Interaction
GAM	Access Collect	Metadata preparation (Curate)	-	Find similarities or differ-ences (Compare) Annotate (Augment) Recommend	Share Publish Data hub (Mediate)	Select	Storytelling Explore Receive recommendations
HECHT	-	-	-	Find similarities or differ-ences (Compare) Statistical (Quantitative Analysis) Thematic analysis (Quali-tative Analysis) Effect (Impact Assess-ment)	Share Publish Data hub (Mediate)	-	Explore
IMMA	Find Access Collect	Design script (Cu-rate)	Monitor Moderate Delete responses (Remove)	Find similarities or differ-ences (Compare) Thematic analysis (Quali-tative Analysis) Statistical (Quantitative Analysis)	Share Publish Data hub (Mediate)	Select	Explore Storytelling Wooden model (Cyber-phisical)
MNCN	Access Collect	Metadata preparation (Curate)	-	Find similarities or differ-ences (Compare) Annotate (Augment)	Share Publish Data hub (Mediate)	Select Scan artifact (Acquire)	Games, Puzzles, Treasure Hunt, Cyber-physical

Table 1. Thematic analysis of the five SPICE case studies.

– Visualisation. This activity is the one that differs signifi-cantly from stand-alone data science pipelines to more interactive, user-intensive applications in this domain. Users of all types (curators, citizens, researchers) interact with the data in many ways, engaging with the digital ob-jects via viewing, listening, and sometimes touching, or themselves in a virtual or augmented reality. Citizens en-gage with the artifacts via mediated experiences that in-volve in-presence storytelling as well as hybrid modali-ties, for example, combining physical reconstructions of objects (e.g. a wooden or 3D-printed replica) with digital representations in 'cyber-physical' spaces.

Another set of classes in DJO refer to roles that data can have: input, output, temporary, etc... (see Fig. 2) These are certainly valid in any software engineering setting, but they don't tell much about the types of resources (assets) in-volved. Citizen curation data journeys require a tailored characterisation of the types of (digital) objects involved, that can be in-turn mapped to several different nodes in the data flow. The concept of 'citizen curation object' is used

here to refer to any digital resource used or created through the 'citizen curation process'. It includes (i) digital represen-tations of artworks and their metadata and museum labels, (ii) resources that guide the citizen curation activity (e.g. quizzes, interpretation exercises) and (iii) the results of the activity (e.g. citizen answers, stories, interpretations). The data and metadata associated with the results of citizen cu-ration activities (e.g. citizen answers to questions plus metadata associated with the author (e.g. their identity and community membership) and content of the activity (e.g. the text and extracted features such as its sentiment and val-ues) is what is essential to making sense of these journeys. Table 2 shows a list of resource types derived from analysing the five SPICE case studies (but many more could emerge in a broader survey).

We have seen how the backbone components of compu-tational data journeys are compatible with citizen curation applications. We also observed how the ontology would re-quire additional components to express the richness of these systems, both in terms of activities and types of resources involved. However, to reflect further on the opportunities

316

Knowl. Org. 51(2024)No.5
E. Daga. Citizen Experiences in Cultural Heritage Archives: A Data Journey

Case Study	Resource types
DMH	Artifact, Design Objects, Stories, Text, Audio, Video, 3D objects
GAM	Artwork metadata, Images, Comments, Emoticons, Emotions, Stories, Characters, Focus groups, Online survey, Ethnographic observation
HECHT	Dilemma (Prompt), Stories (Autoethnographies), Photos
IMMA	Artwork metadata, Artwork images, Interview, Survey, Question, Prompt, Stories (Autoethnographies), Scripts, Response, Text, Choice
MNCN	Images, Essay, Game, Puzzle

Table 2. Types of resources used in the SPICE case studies.

derived from such a perspective it is worth considering the key issue of 'capturing' data journeys and reflecting on the additional representational layers needed.

In the case of the SPICE Linked Data Hub (Carvalho et al. 2024), the project developed an activity monitoring layer that has the purpose of recording events from connected citizen curation applications, linking catalogued artifacts with citizen responses, and make them reusable for analysis. The backbone representational layer is the established W3C Prov-O data model (WC3 2013). The model can be further extended covering the specificity of citizen curation artifacts and activities. It is worth noting how such representational layer is agnostic with respect to the underlying technology. Events described as such could be stored in a traditional relational database, in a graph database, or a blockchain (Anagnostakis 2018).

Artworks, metadata, and responses travel across various systems whose competences vary from hosting data, monitoring activities, and providing extended user interfaces. Citizen responses may be generated by users interacting with a mobile application, when the underlying system (the app itself) generates a new event referencing the artifact, the activity performed, and the response. End-user systems can be seen as operating in a diversified social media network. Another citizen may receive a notification, via another citizen engagement system in SPICE, asking to react to a newly generated response. The new user comments with an emoticon, and the underlying infrastructure record the new event.

An extended set of activities and object types as well as tracing user operations in event graphs can help in answering the issues mentioned before, for the benefit of both citizens and museum practitioners.

- Sense-making. By representing citizen curation applications as data journeys, we can potentially support analytics covering multiple dimensions of interest to museum professionals and researchers, including emotions, preferences, and choices, and finding similarities and differences across communities of users. Curators can use the data journeys to explore the responses to a given artifact in specific interaction contexts, responses of a given community, or how different activities relate. Data journeys can be leveraged by an analytics dashboard able to support curators in exploring the contributions from a multiplicity of perspectives, independently from the diversity of user-facing systems. Event graphs allow for analysing users' behaviours and contextualising the content within a rich interaction context. Data journeys can be leveraged for observing reception by difference community groups, sensitivities, and cultural backgrounds, and explore ways of characterising it in the archive without imposing one view over others.

- Monitoring sensitive content. New content produced by end-user systems can be monitored thanks to the events graph before it reaches the archive. Data managers (curators, developers) can review the collected information, as well as establish moderation processes or make use of intelligent systems for content analysis, to automatically flag content that can be inappropriate or potentially violate the privacy of users. Some content can be flagged as sensitive but still considered of value to be included in the archive to preserve the authenticity of the original response.

- Ownership. Requesting a specific citizen response, the citizen curation application will be able to also know who produced that response, whether it was authored and by whom, accessing the full lineage of the resource. Similarly, when a citizen curation script uses an artifact image, the data journey will describe how that metadata record was created, the link to the original source, being it the museums' Website or a collection management system. Citizens can claim ownership of contributed content and ask for specific attribution statements to be included.

- Terms of use. Data journeys provide a high-level representation of how a certain asset (artifact, image, etc...) is being used, by whom, and for what purpose, in end-user applications. Such representation could be fine-grained and leverage existing standards for digital rights management such as W3C ODRL (Renato 2007). Information can include the usage policies applicable to that context so that

Knowl. Org. 51(2024)No.5
E. Daga. Citizen Experiences in Cultural Heritage Archives: A Data Journey

317

applications can adapt and mediate intelligently with their users, relying on a standard protocol for adapting user interfaces to content, such as IIIF (Snydman 2015).

We can summarise what has been discussed so far about citizen curation data journeys in five layers:

- Resources: resources used in the data journey such as artwork images, metadata records, data sources, licencing information, and terms of use, each one identified by a Linked Data entity URI, and leveraging a rich set of types (extending Table 2).
- Event graphs: events occurring in different systems, such as a citizen curation activity that generates a user response or a curator who selects a response to be included in the archive.
- Data nodes graph: a graph of data-to-data relationships, such as reused tools and resources manipulated by the process, abstracting from the event graphs, that characterise the data flow of a given application, focusing on its design (abstracting on actual events).
- Activity graph: a graph of high-level activities. In the context of citizen curation, these can be specialisations of the general scenario introduced before (but expanding on the activities listed in Table 1).
- Policies graph: a graph of metadata about ownership, licences, and terms of use, to support the computational analysis of terms of use.

4.0 Discussion and conclusions

Before concluding, it can be useful to reflect on the implications of having these different layers altogether and on possible, concrete use cases that can benefit from such a holistic view.

- Tracing sensitivity: heritage institutions are typically strongly characterised both geographically and culturally. Some institutions may find sensitive certain content because of how it triggers local, unsolved issues. Other content can be considered sensitive by museum curators at a certain point in time but this may change in the future. The heritage archive is the perfect place where this knowledge can be preserved. Crucially, data journeys may support the study of how reception changes with changing cultural norms.
- Granularity of policies: licences (terms of use) may be associated to a whole collection of entities (such as catalogue metadata). However, catalogue-level terms of use may not be applicable to all items in the same way, as specific photograph of an artwork, for example, may have special ownership and terms of use. Linked data and related graph technologies can help in specifying terms

with a high degree of granularity, pointing to collections, items, or their parts.
- Composite objects: citizen curation applications may generate composite objects, including images of artefacts, curators' notes (e.g. questions of a slow looking activity), and citizen contributed content. Data journeys allow to capture information about rights and terms of use and reason upon the compatibility of rights when joining content in composite objects.
- Compatibility of terms of use: applications should make users aware of the difference in terms of use associated with each one of them. Potentially, an intelligent system could raise issues in relation to intended use (using a simulated workflow to verify agreement with current policies).
- Visibility and access control: when terms of use affect access control, relevant users should be notified and instructed on what type of actions are needed to ensure a continued availability of resources.
- Time-dependent information: it is not uncommon for museums to have limited control on the terms and conditions associated with the artworks, and often negotiate with artists (or rights management agencies) terms bound to specific contexts (e.g. a festival). Data journeys allows to capture and preserve the usage policies, ownership and terms of use associated with assets involved in a specific citizen curation scenario 'at the time' of the events occurred.
- Changes of terms of use: when an owner changes the terms of use of an image, they should be notified that there are applications having rights to access that image for a purpose that should not be allowed anymore. In this case, the owner may decide to either revoke the permission or restore the original policies.
- Revoke consent: similarly, curators shall know if a citizen does not want their content to be used anymore, and such changes should be propagated to citizen curation applications.

In this article, we argued that a knowledge organisation system for 'data journeys', such as the DJO, can help in disentangling problems that include issues of sense-making, ownership, sensitivity, privacy, and rights management. Data journeys can help govern the complexity of citizen curation applications. Here, we observed citizen curation by abstracting it as a data-intensive system. However, to realise this vision, many problems need to be solved concerning infrastructure, technologies, and methods that could support the implementation of data journeys. To that end, stakeholders should privilege open standards and distribution rather than offering end-to-end systems in isolated silos.

318

Knowl. Org. 51(2024)No.5
E. Daga. Citizen Experiences in Cultural Heritage Archives: A Data Journey

Endnotes

1. https://icom.museum
2. "A museum is a not-for-profit, permanent institution in the service of society that researches, collects, conserves, interprets and exhibits tangible and intangible heritage. Open to the public, accessible and inclusive, museums foster diversity and sustainability. They operate and communicate ethically, professionally and with the participation of communities, offering varied experiences for education, enjoyment, reflection and knowledge sharing." See https://icom.museum/en/resources/standards-guidelines/museum-definition/
3. http://spice.kmi.open.ac.uk
4. https://www.designmuseum.fi/
5. https://www.gamtorino.it/
6. https://www.fondazionetorinomusei.it/
7. https://mushecht.haifa.ac.il/
8. http://imma.ie
9. https://www.padaonegames.com/
10. Here, we focus on analysing how to describe those systems as data journeys and why this is useful. In Daga et al. (2022) we conducted in-depth analysis of the technologies that could support the integration of citizen experiences in cultural heritage archives.
11. It is interesting for us to note how in the book, the one chapter touching on cultural heritage is focused on the issues of forgery and attribution in the arts. Provenance is indeed the conceptual ancestor of data journeys.

References

Anagnostakis, Aristidis G. 2018. "Towards a Blockchain Architecture for cultural Heritage Tokens." In: *Transdisciplinary Multispectral Modeling and Cooperation for the Preservation of Cultural Heritage, First International Conference, TMM_CH 10 13 October 2018, Athens, Greece*. 541-551.

Back, Jon, Benjamin Bedwell, Steve Benford, Lina Eklund, Anders Løvlie, William Preston, Paulina Rajkowska, Karin Ryding, Jocelyn Spence, Emily Thorn, Annika Waern, and Tim Wray. 2018. "GIFT: Hybrid Museum Experiences through Gifting and Play." In: *Proceedings of the Workshop on Cultural Informatics, co-located with the EUROMED International Conference on Digital Heritage, 3 November Nicosia, Cyprus* edited by Angeliki Antoniou and Manolis Wallace. 31-45.

Bruni, Luis Emilio, Enrico Daga, Rossana Damiano, Lily Diaz, Tsvi Kuflik, Antonio Lieto, Aldo Gangemi, Paul Mulholland, Silvio Peroni, Sofia Pescarin and Alan Wecker. 2020. "Towards Advanced Interfaces for Citizen Curation". In: *AVI2CH 2020: Workshop on Advanced Visual Interfaces and Interactions in Cultural Heritage, 29 Sep 2020 Ischia, Italy*. URL: http://avi2ch-20.di.unito.it/index.html.

Carvalho, Jason, Enrico Daga, Luigi Asprino, Aldo Gangemi, Mark Maguire, Paul Mulholland, Adam Stoneman, and Chukwudi Uwasomba. 2023. "SPICE Linked Data Hub – Data infrastructure and tools to support Citizen Curation". In: *ExICE – Extended Intelligence for Cultural Engagement, 15 Mar 2023 Bologna, Italy*. URL: https://spice-h2020.eu/conference/.

Carvalho, Jason, Enrico Daga, Paul Mulholland, Luigi Asprino, Chukwudi Uwasomba, Daquino Marilena, Aldo Gangemi, Mark Maguire, and Adam Stoneman. 2024. "SPICE Linked Data Hub – Data infrastructure and tools to support Citizen Curation". Manuscript submitted for publication.: ACM Journal of Computing and Cultural Heritage (JOCCH).

Chowdhury, Gobina. 2010. "From Digital Libraries to Digital Preservation Research: The Importance Of Users And Context". *Journal of Documentation* 66, no.2: 207-223. doi:10.1108/00220411011023625.

Coombes, Annie B., and Ruth B. Phillips, eds. 2020. *Museum transformations: Decolonization and Democratization*. John Wiley & Sons.

Council of Europe. 2006. "Framework Convention on the Value of Cultural Heritage for Society: Faro, 27.10. 2005". Council of Europe treaty series.

Daga, Enrico, Mathieu d'Aquin, Aldo Gangemi, and Enrico Motta. 2015. "Propagation of Policies in Rich Data Flows". In: *Proceedings of the 8th International Conference on Knowledge Capture (K-CAP 2015), October 7-10 2015 Palisades, NY, USA*, 1–8.

Daga, Enrico; Luigi Asprino, Rossana Damiano, Belen Diaz-Agudo, Aldo Gangemi, Tsvi Kuflik, Antonio Lieto, Anna Maria Marras, Delfina Martinez-Pandiani, Paul Mulholland, Silvio Peroni, Sofia Pescarin and Alan Wecker. 2022. "Integrating Citizen Experiences in Cultural Heritage Archives: Requirements, State of The Art, And Challenges". *Journal on Computing and Cultural Heritage* (JOCCH) 15, no.1: 1.35.

Daga, Enrico and Paul Groth. 2023. "Data Journeys: Explaining AI Workflows Through Abstraction". *Semantic Web* 1-27. https://doi.org/10.3233/SW-233407.

Flinn, Andrew, Mary Stevens, and Elizabeth Shepherd. 2009. "Whose Memories, Whose Archives? Independent Community Archives, Autonomy and The Mainstream". *Archival Science* 9: 71-86.

Garijo, Daniel, Alper Pinar, Khalid Belhajjame, Oscar Corcho, Yolanda Gil and Carole Goble. 2013. "Common Motifs in Scientific Workflows: An Empirical Analysis". *Future Generation Computers Systems* 36: 338-341.

Gutiérrez-Sánchez, Pablo, Irene Camps-Ortueta, Pedro Gómez-Martín, and Pedro Antonio. 2023. "Co-design of Treasure Hunts for a Natural History Museum." In

Knowl. Org. 51(2024)No.5

319

E. Daga. Citizen Experiences in Cultural Heritage Archives: A Data Journey

ExICE – Extended Intelligence for Cultural Engagement, 15 March 2023, Bologna, Italy.

Harth, Andreas, Alex Polleres and Stefam Decker. 2007. "Towards a Social Provenance Model for The Web." *Workshop on Principles of Provenance* (PrOPr).

Leonelli Sabina and Niccolo Tempini, eds. 2020. *Data Journeys in The Sciences.* Springer Nature.

Lieto Antonio, Manuel Striani, Cristina Gena, Enrico Dolza, Anna Maria Marras, Gian Luca Pozzato, and Rossana Damiano. 2023a. "A Sensemaking System for Grouping And Suggesting Stories From Multiple Affective Viewpoints In Museums." Preprint, submitted April 27, 2023. http://arXiv:2304.14117I.

Lieto, Antonio, Gian Luca Pozzato, Manuel Striani, Stefano Zoia, and Rossana Damiano. 2023b. "2.0: A Diversity-Seeking, Explainable, And Affective Art Recommender for Social Inclusion." Cognitive Systems Research 77: 1-17. https://doi.org/10.1016/j.cogsys.2022.10.001

Moreau, Luc. 2010. *The foundations for Provenance on The Web.* Now Publishers Inc.

Mulholland, Paul, Adam Stoneman, Naomi Barker, Mark Maguire, Jason Carvalho, Enrico Daga, and Paul Warren. 2023. "The Sound of Paintings: Using Citizen Curation to Explore the Cross-Modal Personalization of Museum Experiences". In: *Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization*, 26-29 June, 2023, Limassol Cyprus, 408-418. https://doi.org/10.1145/3563359.3596662.

Neelands, Johathan, Eleonora Belfiore, Catriona Firth, Natalie Hart, Liese Perrin, Susan Brock, Dominic Holdaway, and Jane Woddis. 2015. *Enriching Britain: Culture, Creativity and Growth.* Coventry: University of Warwick.

Renato, Iannella. 2007. "Open Digital Rights Language (ODRL)". In: *Open Content Licensing: Cultivating the Creative Commons*, edited by B. Fitzgerald, J. Coates, and S. Lewis Sydney University Press.

Ridge, Mia. 2013. "From Tagging to Theorizing: Deepening Engagement with Cultural Heritage Through Crowdsourcing." *Curator: The Museum Journal* 56, no. 4: 435-450. https://doi.org/10.1111/cura.12046

Snydman, Stuart, Robert Sanderson, and Tom Cramer. 2015. "The International Image Interoperability Framework (IIIF): A Community & Technology Approach for Web-Based Images." In: *Archiving Conference,* vol. 2015, Society for Imaging Science and Technology 16-21.

SPICE. 2022. "D7.4 Socio-Technical roadmap with project management tool". URL: https://spice-h2020.eu/document/deliverable/D7.4.pdf.

Stoneman, Adam, Jason Carvallo, Enrico Daga, Mark Maguire and Paul Mulholland. 2021. "Uncomfortable Revelations: can citizen curation widen access to museums?" *Museum Ireland* 28: 64-71.

Vishwanath, Gautam, Lily Diaz-Kommonen and Leena Svinhufvud. 2023. "Citizen Curation of Digital Cultural Heritage through the Co-design of a Virtual Reality Museum". In *ExICE – Extended Intelligence for Cultural Engagement*, Bologna, Italy. doi:10.5281/zenodo.7973634.

WC3. 2013. "PROV-O: The PROV Ontology". URL: https://www.w3.org/TR/prov-o/

Whitehead, Chris. 2011. *Interpreting Art in Museums and Galleries.* Routledg

Acknowledgements