HACKING CULTURAL HERITAGE

THE HACKAHTON AS A METHOD FOR HERITAGE INTERPRETATION

Leonardo Moura de Araújo

Submitted to the Faculty 3
(Mathematics and Computer Science)
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Engineering (Dr.-Ing.)
at the University of Bremen

2018
Supervisors:
1. Prof. Dr. Frieder Nake
   FB3 Mathematik und Informatik
   Universität Bremen

2. Prof. Dr. Karsten Wolf
   FB 12 Erziehungs- und Bildungswissenschaften
   Universität Bremen

Date of Oral Examination:
September 24th, 2018

The present work was carried out with the support of the National Council of Scientific and Technological Development - CNPq - Brazil
ABSTRACT

Hackathons were originated from both the evolution of and revolution caused by personal computers. Initially, they have been implemented as a collaborative method for solving computer-related problems or conceptualizing new possibilities based on specific infrastructures. Only later on, when Cultural Institutions had undergone intensive digitization, Hackathons started to be part of their repertoire. Because of the special nature of Cultural Institutions, Hackathons for Cultural Heritage cannot be understood in the same way as their counterparts happening in a purely engineering domain. Problem solving and conceptualization through collaborative programming are entangled with the significance of the content matter they intend to deal with: the institutions’ collections. Based on these considerations, this thesis aims at explaining the underlying principles, interactions, and infrastructures of the Hackathon as a method for Heritage Interpretation. Moreover, the thesis also proposes a Fast-speed IT Platform, which was designed within the context of the Two-speed IT infrastructure, where a foundational, stable, and slow infrastructure is complemented by an additional creative, experimental, and agile infrastructure, which is capable of promptly responding to the needs of communities. The platform is an effort to implement strategies for interpreting, recontextualizing, and telling stories with Digital Collections. In addition, the platform aims at mitigating problems concerning technical knowledge that is usually required for taking advantage of the affordances of Digital Collections as a creative material.
ACKNOWLEDGMENTS

To begin with, I would like to express my sincere gratitude to my advisor Prof. Dr. Frieder Nake for the continuous support of my Ph.D study. His motivation, energy, and immense knowledge have been sources of personal inspiration and have enormously contributed to the conclusion of this work. Likewise, my sincere thanks go to my second advisor Prof. Dr. Karsten Wolf and the members of my thesis committee: Prof. Dr. Andreas Hepp and Prof. Dr. Sara Hofmann.

To all fellow labmates, colleagues, test participants, and team members that provided valuable insights over the course of this research project. I would especially like to thank Prof. Dr. Heidi Schelhowe, Prof. Dr. Ruth Schilling, Prof. Dr. Thomas Schneider, Stefanos Trialonis, Doris Caliz, Iris Bockermann, Hestiasari Rante, Luis Carlos Chaparro, Roger Meintjes, Nina Hentschel, Adrienn Kovács, Nicole Mayorga, Tossawat Mok, Michael Lund, Axel Menning, Philipp Geisler, Gregor Rohmann, Christiane Weber, and Wolfgang Thomsen.

Last but not the least, I would like to thank my family for supporting me spiritually throughout writing this thesis. To my old and new parents, and to my sister who also provided very valuables insights during the research. A very special thanks goes to my wife Hanna Błońska for the patience, insights, and endurance for putting up with me and my research. It was certainly not an easy task!
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>MIT’s Tech Model Railroad Club</td>
<td>15</td>
</tr>
<tr>
<td>Figure 2</td>
<td>First Europeana Hackathon</td>
<td>19</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Structures of collaborative and competitive Hackathons</td>
<td>22</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Motivations to take part in Hackathons according to [160]</td>
<td>24</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Age groups of hackers according to survey A.2.1.1</td>
<td>26</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Participants of the Kultur-Hackathon Coding da Vinci</td>
<td>28</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Zeitblick (right) and Midiola (left)</td>
<td>30</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Alt-Berlin (right) and DNB - Data Explorer (left)</td>
<td>31</td>
</tr>
<tr>
<td>Figure 9</td>
<td>The virtual reality app Skelex</td>
<td>32</td>
</tr>
<tr>
<td>Figure 10</td>
<td>From the module Images of Recursion - Forever Programs</td>
<td>46</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Scratch Visual Programming Environment</td>
<td>49</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Curator Table - Google Arts and Culture Project</td>
<td>50</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Homebrew Computer Club’s action call</td>
<td>55</td>
</tr>
<tr>
<td>Figure 14</td>
<td>The Altair 8800</td>
<td>56</td>
</tr>
<tr>
<td>Figure 15</td>
<td>System/360 Model</td>
<td>60</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Merging conflicting code versions with a Git GUI</td>
<td>64</td>
</tr>
<tr>
<td>Figure 17</td>
<td>“Collecting” artifacts with the Cooper Hewitt Pen</td>
<td>67</td>
</tr>
<tr>
<td>Figure 18</td>
<td>Treasure hunt-like game Collect &amp; Connect</td>
<td>73</td>
</tr>
<tr>
<td>Figure 19</td>
<td>File cabinets at the Oakdale Workmen’s Institute</td>
<td>75</td>
</tr>
<tr>
<td>Figure 20</td>
<td>Analog entries generated using the British Museum rules</td>
<td>76</td>
</tr>
<tr>
<td>Figure 21</td>
<td>Expressing temporal information with the CIDOC-CRM</td>
<td>79</td>
</tr>
<tr>
<td>Figure 22</td>
<td>The influence of datasets in the design process</td>
<td>86</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>23</td>
<td>The need for easily extending and modifying datasets</td>
<td>87</td>
</tr>
<tr>
<td>24</td>
<td>The utilization of digital collections beyond search and retrieval</td>
<td>89</td>
</tr>
<tr>
<td>25</td>
<td>UI of The Museum System</td>
<td>91</td>
</tr>
<tr>
<td>26</td>
<td>Card catalogue from the Denver Art Museum</td>
<td>92</td>
</tr>
<tr>
<td>27</td>
<td>ITS’ physical archive</td>
<td>94</td>
</tr>
<tr>
<td>28</td>
<td>Central Name Index</td>
<td>95</td>
</tr>
<tr>
<td>29</td>
<td>OuSArchiv’s UI</td>
<td>96</td>
</tr>
<tr>
<td>30</td>
<td>OuSArchiv’s UI components</td>
<td>98</td>
</tr>
<tr>
<td>31</td>
<td>Hierarchy of the archival objects as described in [179]</td>
<td>99</td>
</tr>
<tr>
<td>32</td>
<td>Typo3’s Administrative User Interface</td>
<td>111</td>
</tr>
<tr>
<td>33</td>
<td>Artfacts’ within the Two-speed IT Infrastructure</td>
<td>117</td>
</tr>
<tr>
<td>34</td>
<td>Artfacts’ architecture overview</td>
<td>118</td>
</tr>
<tr>
<td>35</td>
<td>Artfacts’ User Interface</td>
<td>118</td>
</tr>
<tr>
<td>36</td>
<td>Conceptualization of an exhibition supported by graph organizers</td>
<td>120</td>
</tr>
<tr>
<td>37</td>
<td>Storytelling as a graph organizer</td>
<td>121</td>
</tr>
<tr>
<td>38</td>
<td>Porphyrian Tree</td>
<td>122</td>
</tr>
<tr>
<td>39</td>
<td>Tree (left) and networks (right) diagrams</td>
<td>125</td>
</tr>
<tr>
<td>40</td>
<td>A Mind Map depicting topics related to “Time Management”</td>
<td>126</td>
</tr>
<tr>
<td>41</td>
<td>Snippets displaying objects from Google’s Knowledge Graph</td>
<td>129</td>
</tr>
<tr>
<td>42</td>
<td>Hierarchical contextualization of concepts with SKOS</td>
<td>131</td>
</tr>
<tr>
<td>43</td>
<td>Components of the Artfacts’ GUI</td>
<td>135</td>
</tr>
<tr>
<td>44</td>
<td>Research questions displayed on the walls of the exhibition</td>
<td>137</td>
</tr>
<tr>
<td>45</td>
<td>Coding with Atlas.ti</td>
<td>138</td>
</tr>
<tr>
<td>46</td>
<td>The Knowledge Map</td>
<td>140</td>
</tr>
<tr>
<td>47</td>
<td>A tagged statement on the Content Panel</td>
<td>141</td>
</tr>
<tr>
<td>48</td>
<td>Definitions of relationships</td>
<td>143</td>
</tr>
<tr>
<td>49</td>
<td>Rhizomatic relationships</td>
<td>144</td>
</tr>
<tr>
<td>50</td>
<td>Hierarchical structure for classes</td>
<td>145</td>
</tr>
<tr>
<td>51</td>
<td>Hierarchical structure for schemes</td>
<td>146</td>
</tr>
<tr>
<td>52</td>
<td>Architectural Plan and Collection Gallery</td>
<td>147</td>
</tr>
<tr>
<td>53</td>
<td>Visualization Reports</td>
<td>148</td>
</tr>
</tbody>
</table>
ACRONYMS

AAT  Art and Architecture Thesaurus
ADM  Artfacts’ Data Model
CCO  Cataloging Cultural Objects
CdV  Kultur-Hackathon Coding da Vinci
CDWA  Categories for the Description of Works of Art
CIDOC  International Council of Museums
CIDOC-CRM  CIDOC - Conceptual Reference Model
CMS  Collection Management Systems
CNI  Central Name Index
CntMS  Content Management Systems
DC  Dublin Core
DCMI  Dublin Core Metadata Initiative
DIA  Digital Interpretive Artifact
EDM  Europeana Data Model
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESE</td>
<td>Europeana Semantic Elements</td>
</tr>
<tr>
<td>FOAF</td>
<td>Friend-of-a-Friend</td>
</tr>
<tr>
<td>GLAM</td>
<td>Galleries, Libraries, Archives, and Museums</td>
</tr>
<tr>
<td>GRIPHOS</td>
<td>General Retrieval and Information Processor for Humanities Oriented Studies</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical user interface</td>
</tr>
<tr>
<td>HCI</td>
<td>Human-Computer Interaction</td>
</tr>
<tr>
<td>ITS</td>
<td>International Tracing Service</td>
</tr>
<tr>
<td>KGLink</td>
<td>Knowledge Graph Link</td>
</tr>
<tr>
<td>KGNode</td>
<td>Knowledge Graph Node</td>
</tr>
<tr>
<td>KMLink</td>
<td>Knowledge Map Link</td>
</tr>
<tr>
<td>KMNode</td>
<td>Knowledge Map Node</td>
</tr>
<tr>
<td>KOS</td>
<td>Knowledge Organization System</td>
</tr>
<tr>
<td>MARC</td>
<td>MAchine-Readable Cataloging</td>
</tr>
<tr>
<td>MCN</td>
<td>Museum Computer Network</td>
</tr>
<tr>
<td>OCLC</td>
<td>Online Computer Library Center</td>
</tr>
<tr>
<td>OOP</td>
<td>Object-Oriented Programming</td>
</tr>
<tr>
<td>OOUI</td>
<td>Object-Oriented User Interface</td>
</tr>
<tr>
<td>QDAS</td>
<td>Qualitative Data Analysis Systems</td>
</tr>
<tr>
<td>RDF</td>
<td>Resource Description Framework</td>
</tr>
<tr>
<td>RDFS</td>
<td>Resource Description Framework Schema</td>
</tr>
<tr>
<td>RDMS</td>
<td>Relational Database Management System</td>
</tr>
<tr>
<td>RnED</td>
<td>International Tracing Service’s Research and Education Department</td>
</tr>
<tr>
<td>SKOS</td>
<td>Simple Knowledge Organization System</td>
</tr>
<tr>
<td>TGN</td>
<td>Thesaurus for Geographic Names</td>
</tr>
<tr>
<td>UI</td>
<td>User Interface</td>
</tr>
<tr>
<td>ULAN</td>
<td>Union List of Artist Names</td>
</tr>
<tr>
<td>VPL</td>
<td>Visual Programming Language</td>
</tr>
</tbody>
</table>
INTRODUCTION

In November of 2012, the Museum of Modern Art\(^1\) (MoMA - New York, USA) sparked heated debates in the art world when it decided to include fourteen video games, among them Pac-man, Tetris, and SimCity, in its Architecture and Design collection. The move was considered outrageous by some. "MoMA claims these games belong in its collection because they are art. Really? Is that so?" \(^{118}\), wrote the art critic Jonathan Jones, in an article in the Guardian\(^2\). In another article, Liel Leibovitz declared MoMA had mistaken video games for art, because "they are, quite thoroughly, something else: code." \(^{134}\)

Does code hold any cultural significance to be considered as art? And should it be preserved by being part of a museum collection? According to Paola Antonelli, the senior curator of MoMA’s Department of Architecture and Design, the answers to both questions are yes. "They sure are" \(^7\), wrote Antonelli on MoMA’s Blog describing video games as a new category of artworks. For the curator, "this acquisition allows the Museum to study, preserve, and exhibit video games as part of its Architecture and Design collection." \(^7\)

The relevance of this little anecdote for this dissertation does not lie in the discussion whether video games can be considered artworks or not. Instead, this story sheds light on how to situate, compare, and understand the concept of Cultural Heritage, as something that holds value and must be preserved, in an increasingly digitized world, where the significance and pervasiveness of digitality can be found on all levels of society.

1.0.1 Cultural Heritage

Marilena Vecco \(^{232}\) points to the French Revolution as a historical turning point that helped to shape the notion of common heritage, in that the goods and properties of the king were made public through a process of nationalization, which "was a sort of public appropriation" \(^{232}\). In earlier times, the notion of public good was inexistenable, and so was the notion of a common Cultural Heritage.

\(^{1}\) https://www.moma.org/
\(^{2}\) https://www.theguardian.com/
According to Vecco, the term Cultural Heritage *(patrimoine culturel* in French) was firstly adopted by André Malraux, France’s first Minister of Cultural Affairs, in the decree from 1959 [232]. The usage of the term was then restricted to the realm of fine arts, but with the consolidation of international bodies, such as The International Council on Monuments and Sites³ (ICOMOS) and the United Nations Educational, Scientific and Cultural Organization⁴ (UNESCO), it has been gradually expanded. Nowadays, the term is used not only to define monuments and collections of objects that, because of their significance, represent a tangible support to social memory [63], but also traditions or living expressions, such as “oral traditions, performing arts, social practices, rituals, festive events, knowledge and practices concerning nature and the universe or the knowledge and skills to produce traditional crafts.” [238] In 2002, ICOMOS provided a concise but general definition for the term:

> Cultural heritage is an expression of the ways of living, developed by a community and passed on from generation to generation, including customs, practices, places, objects, artistic expressions and values. Cultural heritage is often expressed as either intangible or tangible Cultural Heritage. [228]

### 1.0.2 Digital Cultural Heritage

There is little doubt that Picasso’s *Les Demoiselles d’Avignon⁵*, as a traditional tangible Cultural Heritage, alongside the collection of video games acquired by MoMA, as partially tangible and partially intangible Cultural Heritage⁶, fall into the ICOMOS definition described above. However, in regard to its digital existence, it is not *game over* for the *Les Demoiselles d’Avignon*, since the physical artwork, through digitization, gains almost similar properties as some born-digital artifacts through the (re)production of its digital surrogate(s).

The digital representation of Cultural Heritage as well as the use of digital media for management, research, interpretation, and preservation is the concern of initiatives done in the field of  

---

³ [https://www.icomos.org/](https://www.icomos.org/)
⁴ [https://en.unesco.org/](https://en.unesco.org/)
⁵ The artwork is part of the permanent collection of MoMA
⁶ Computer applications are made of a formal description (code) that is executed to become electrical processes (running code).
Digital Cultural Heritage. Information Systems play a key role here, because they are important tools for creating collections of digital representations (data) of both born-digital and traditional assets. Digital collections/archives, long understood as digital extensions of the original collections, have increasingly gained recognition for their significance by their unique qualities and properties as code.

Pac-Man and Picasso sharing the same museum space invite us indeed to think about the digital in different ways. Authors, such as Matthew Kirschenbaum et al. [122], Fatih Ozenc et al. [181], Anna Vallgarda et al. [231], Verena Fuchsberger et al. [89], and Paul Leonardi [135] suggest a certain kind of materiality embedded in the digital. Digital artists, such as Aaron Koblin [224] and Julie Freeman [87] promptly declare the digital, more specifically digital data, as an art material. “Data is no longer just in the domain of engineers and scientists” [87], declares Freeman et al.. To think about data as a raw, artistic, and elementary material one can construct with, requires a different epistemology that stresses design instead of engineering. A design perspective provides both flexibility in information modeling and intuitive reasoning about the limitations and affordances of data. New, complex and diverse meanings are formed from the modular connection of different data elements. According to Freeman et al., “in their raw format they are sets of individual values which can be manipulated, reconfigured, and transformed. This highly flexible, malleable substance is an ideal art material.” [87]

1.0.3 Hacking Cultural Heritage

MoMA’s wish-list for future video game acquisitions includes Spacewar! [7], which was not only one of the very first video games produced, but also the outcome of a series of guiding principles shared among MIT hackers. As a hack, Spacewar! was the result of the “deviant” use of the computer, a machine that during and

---

7 This term does not stand for digital artifacts as it may suggest at first glance (see [37]).
8 The term affordances, widely used in this thesis, should be understood under the J. J. Gibson’s standpoint. An affordance is an actionable property provided by the characteristics of an environment or entities of an environment in relationship to the characteristics of an actor (person or animal). Therefore, an affordance is a relative relationship that provides an outcome and exists independently from the actor’s ability to perceive it and act upon it [93].
9 This topic is further discussed in Chapter 5.
10 See Chapter 2.
after the Second World War was seen as a strategic asset for military and scientific research. In addition, not only the computer was being used to create something outrageously unimportant, but also, and likely because of that, it was developed collaboratively and distributed freely to any one willing to play or improve the code. This hacking approach to technology was the precursor of what later would be known as Open Source. The question whether video games belong in Cultural Institutions is also a question to ponder where hacking belongs, because video games are historically bound to it. However, while Spacewar! is still in the wish-list for future acquisitions, and therefore a hack to be preserved and exhibited in MoMA’s collection, hacking as a method guided by principles has been already used in the Cultural Heritage Sector.

Owning vast stockpiles of data, a number of cultural institutions, especially the so-called Memory Institutions or GLAMs (Galleries, Libraries, Archives, and Museums), have organized Hackathons in order to explore the creative possibilities of their digital collections. The word Hackathon comes from a combination of two other words, namely hack and marathon, and is commonly used to describe an event where a group or groups of individuals from a wide variety of backgrounds engage in collaborative computer programming. As the PDP-12 offered the material conditions so that hacking principles could be manifested to create Spacewar!, digital collections are one of the most important means that enable Hackathons in the Cultural Heritage Sector.

In its origin and essence, hacking as a practice can be situated in the intersection between engineering and humanities, science and art, proprietary and open. It is a limbo state of trial and error, and unrealized full potential, but it holds nevertheless great promises. The earlier hackers were the first to find new applications for the computer beyond their usual military employment13. Although taken for granted today, the idea of playing games, producing art, or even writing a dissertation such as this one on a computer were unthinkable before. First a phenomenon in Computer Science,
hacking has now spread across the cultural landscape. This thesis is an attempt to understand this phenomenon in this very special context, where Digital Collections have become a creative material for the construction of new meanings and interpretations, rather than the means for fast search and retrieval of information only.

1.1 Research Goals and Questions

At the beginning of this research project, it was popular among Cultural Institutions the idea that online social networking services had to be adopted in order to engage with their audiences. Social media platforms appeared in the horizon even as promising and powerful tools for providing more leverage for communities over institutions. The hypothesis was that the multiple conversations and discussions happening through these social networking services would result in a participatory interpretation of heritage happening digitally, because of the active participation of communities and the contrast of points of view there expressed. The museum edition of the Horizon Report of 2010 [149], classified social media as one of the key emerging technologies for their potential impact on and use in education and interpretation within the museum environment. According to the report:

Social media tap an entire world of user-generated content created by new technologies and applications, readily available on our smart phones and computers, and easily disseminated and accessed via the Internet. Social media use video, audio, and other media as a catalyst to encourage, facilitate and provoke social interaction. [...] Social media engage others in conversations and interactions with, about, and through media. Collectively, social media are above all the voice of the audience, endlessly expressive and creative. [149]

Indeed, Facebook15, Twitter16, Instagram17 and others have offered additional communication channels that can be beneficial for institutions and communities, but it has become increasingly clear

14 Or democratic, as another commonly used word to refer to some aspects of social media.
15 https://www.facebook.com/
16 https://twitter.com/
17 https://www.instagram.com/
that a truly genuine digital interpretation of Cultural Heritage is
difficult to be accomplished through social networking services
that operate as black boxes, in the sense that no one knows
exactly what happens inside, or are offered the means to modify
these platforms so that they can better adapt to the needs of
communities. On the contrary, communities need to conform to the
terms and conditions of social media platforms. Neither are the
algorithms open for scrutiny or modification by the community, nor
are the organizational practices transparent in regard to e.g. the
utilization and analysis of the data generated by users while on
the Web. The fundament of these social networking services is their
financial survival through a business model that exchange personal
information for advertisements. The manipulation of user behavior
through interface and procedural strategies are also used. The goals
of such practices are unclear, besides keeping users “hooked”\(^\text{18}\). All
of these factors limit the creative and educational potential that such
powerful technologies are able to offer.

Other projects and movements in the Computer Science history,
such as Wikipedia\(^\text{19}\) and the Open-Source Movement\(^\text{20}\), operating
on the principles that have offered considerably higher amount of
leverage to the individual, who is seen as a co-creator, free to
reuse and profit from the contributions produced by the others
in the community, and take part in the decisions that shape the
ecosystem he or she is a part of - not only content-wise, but
also structurally. Such projects offer a more suitable foundation
for a participatory interpretation of heritage happening digitally.
These projects are more ambitious and have demonstrated their
significance through their contributions, which were capable of
delivering undeniable benefits to society\(^\text{21}\). Openness, inclusiveness,
freedom, and transparency together with the high quality outcomes
produced by such projects were the key qualities that pointed
this research to truly participatory models that have been able to
handle multiplicity, evolve according to the needs and wishes of the
affiliated communities, and therefore be potentially more suitable
to offer a deeper digital interpretation of Cultural Heritage that
accounts for multiple voices and points of view. This direction
soon took this research project to pay attention to hacking as a

\(^{18}\) See Section 3.1.3.2.
\(^{19}\) https://www.wikipedia.org/
\(^{20}\) See Section 4.2.
\(^{21}\) See Chapters 2, 3, and 4.
creative and impetuous attitude that was behind the origin of great transformations in the computer science history, and the so called Hackathons, a more recent phenomenon happening in the Cultural Heritage Sector that utilizes the same attitude, placing the community at the center of curatorial and interpretation processes. Therefore, with hacking and Hackathons as cornerstones, this thesis presents the following research goals and questions:

1.1.1 Research Goals

- **RG1**: To understand the underlying principles, interactions, and infrastructures of Hackathons as a method for heritage interpretation.
- **RG2**: Based on the results of the analysis, to conceptualize, implement, and test an Information System that incorporates principles, enables interactions, and supports hacking as an interpretive method.

1.1.2 Research Questions

- **RQ1**: What are Hackathons and their particularities in regard to the Cultural Heritage Sector?
- **RQ2**: How can Heritage Interpretation be understood in the context of Hackathons?
- **RQ3**: What are the technologies and their characteristics that enable the occurrence of such events?
- **RQ4**: Considering the vital importance of Collection Management Systems as one of the main enablers of digitization (i.e. producers of digital assets) in Cultural Institutions, the following question is posed: are the currently used Collection Management Systems suitable for the requirements of Hackathons as a method for the interpretation of heritage? Why?

1.2 Research Methods

Research concerned with Information Systems has traditionally been conducted through quantitative research methods [217]. Recently these systems are understood as holistic entities that are highly depended on the universe of discourse they are inserted (see [217], [172], [12]). The complexities and subjectivities present in the communities to which these systems are designed for can in many cases only be grasped by qualitative methods. Therefore, the research
questions are addressed here through mixed methods, qualitative and quantitative, that comprehend both theoretical and practical research.

In order to provide a solid theoretical foundation, an extensive literature review was carried out throughout the entire duration of this research project in all topics that concern the research questions, such as the particularities and historical context of Hackathons, Heritage Interpretation, Digital Platforms, Information Systems for Cultural Heritage, and so on. On top of the literature review, a series of semi-structured interviews (with curators, researchers, hackers, and archivists) were conducted in order to understand the context of their practice, identify key characteristics and requirements in regard to their activities, and obtain specialized reviews on the prototype developed.

The practical research approach and the derived knowledge produced from it come in great part from the iterative design of digital prototypes (namely the different versions of the Artfacts Platform and its modules). Prototyping, as a practice-based method, has been discussed and used since the 1970s in the development of Information Systems [20]. The kinds of problems this method is particular beneficial are the ones of uncertain nature [94] that cannot be resolved through analysis only, because they were not sufficiently explored and therefore present many uncertain and unsolved issues. Due to the lack of literature on the topic covered in this thesis, this is certainly the case here. The development process of Information System in general “is embedded within a search for knowledge about the universe of discourse” [94]. Therefore, in particular to this research project, prototyping was used as a catalyst for the generation of insights and questions before, during and after the development process. In this sense, the prototype design process has served not only for collecting data, evaluating and validating requirements through dialogs with the target groups, but also to organize conceptually the complexities of the topic dealt in this thesis. In this sense, the platform that was developed as a product of this practice-based method is in close connection with the theoretical foundations here presented.

In parallel to interviews and prototyping, focus group workshops were conducted in order to gather collective insights and feedback on the software in issues related to its role within heritage interpretation processes and usability. As a qualitative research method, focus groups are group interviews, typically composed of 6 to 10 individuals with similar backgrounds [183], that is employed to
provide a social setting for arguments to be constructed collectively [128]. Discussions happening in this setting are ideal to gather “high-quality data in a social context where people can consider their own views in the context of the views of others” [183, p.386]. This kind of method, provides also a rich environment where data about a phenomenon that is difficult to measure quantitatively can be systematic collected, organized, and interpreted. Furthermore, an usability test, as an additional user-centered technique approach in iterative design, was employed in order evaluate the software especially in regard to attributes, such as usefulness\textsuperscript{22}, effectiveness\textsuperscript{23}, satisfaction\textsuperscript{24}, and learnability\textsuperscript{25} [199]. Specifically, it was been applied a summative usability test, which is conducted in a middle stage of the product development cycle in order to assess “how well a user can actually perform full-blown realistic tasks and in identifying specific usability deficiencies in the product” [199, p.35].

Case studies were also conducted for examining closely the context in which a specific Information System was used in order to organize and support institutional workflows within a Cultural Institution. This method was also employed to gather insights and evaluate the usage of the second version of the prototype in a real case scenario (Hackathon). As a research method, case studies are particularly useful, because they enable the investigation of “contemporary real-life phenomenon through detailed contextual analysis of a limited number of events or conditions, and their relationships” [241]. Case studies therefore enable the researcher to understand the object of research in its own context, providing a range of different kinds of data for analysis, such as contextualized and naturally occurring data in e.g. group dynamics. Finally, online surveys were conducted in order to obtain quantifiable results for not only supporting some findings of the qualitative research, but also obtaining new insights in topics related to the research questions. An in-depth overview of the empirical research, methods, results and findings conducted during this research project can be obtained in Appendix A.

\begin{itemize}
\item \textsuperscript{22} Measured by the number of problems the tool helps to solve.
\item \textsuperscript{23} Evaluated by the efficacy of methods the tool utilize for solving problems.
\item \textsuperscript{24} Measured by the user’s perceptions, opinions and feelings towards the software
\item \textsuperscript{25} Measured by the amount of training for reaching a certain level of competence.
\end{itemize}
1.3 Organization of the Dissertation

The dissertation is organized as follows: Chapter 2 provides the foundations of hacking as a mindset capable of empowering communities by appropriating technology. In the 1950s, the first hackers and their technological innovations were driven by a set of principles that influenced many cultures, movements, philosophies, and initiatives that were still to come into existence, such as the Open-source Movement and Wikipedia. Besides giving a historical context to hacking, the chapter also examines the origins of Hackathons as events capable of aggregating communities and intensively focusing their abilities to come up with ingenious solutions to technological problems. Not only hacking, but also Hackathons are based on the same set of principles that enable them to occur, such as free information, decentralization, meritocracy, and the belief that the computer could be used as a tool to change old and create new worlds. Currently, the Cultural Heritage Sector has also appropriated from hacking principles and Hackathons as a way to regaining relevance in a fast-paced and increasingly digitized society, because Hackathons are powerful strategies to advance innovation not only, but also in the Cultural Heritage Sector. Hackathons invigorate the digital and participatory strategies of Cultural Heritage Institutions. Furthermore, these events offer numerous advantages not only to institutions, but also to affiliated communities and the institutions’ audiences. The chapter not only discusses about these advantages, but also provides a closer look at the structures of Hackathons and presents concrete examples of these events and their outcomes in the Cultural Heritage Sector.

Chapter 3 examines the historical motivations of Cultural Heritage Institutions behind their efforts in sharing their assets and opening up to their audiences. At the core of memory institutions is Heritage Interpretation, a complex concept that encompasses not only the relationship between current cultural values, moral judgments, and emotional factors with collections of tangible and intangible objects, but also the connection between institutions and their audiences. In this chapter, it is explained the influences of technology in the Cultural Heritage Sector especially in concern to how the Internet, in particular Web 2.0 principles, changed Heritage Interpretation from a one-directional to a multi-faceted process made by multiple voices and opinions. The Open-Source Movement, Wikipedia, and a variety
of projects and business models that are built on top of the sharing economy have one common characteristic - communities are given the means to co-create. Cultural Institutions have adopted the same principles giving their audiences the opportunity for co-producing experiences, narratives, and artifacts that tell stories about heritage. Hackathons for Cultural Heritage appear in this context. Audiences are invited to construct digital artifacts that are algorithmic interpretations of the institutions’ collections. These applications are defined here as Digital Interpretive Artifacts. This kind of Heritage Interpretation happens through digital fabrication and uses datasets as a creative material. The employment of digital fabrication done through the interpretation of Cultural Heritage Datasets in the context of Hackathons is defined at the end of this chapter as Constructionist Heritage Interpretation. In essence, Constructionist Heritage Interpretation is justified in that the interpretation of heritage is operationalized by the manipulation of construction tools, such as programming languages, and digital materials, especially datasets representing collections. Digital Collections are used to intermediate the design of computational artifacts that externalize understandings (stories) about heritage, which can be then shared with others.

Chapter 4 shows the close connection between technology and communities. It analyzes especially certain characteristics of technologies that enable communities of hackers to form and create. At the beginning of the chapter, a closer look is taken at the particular features of the MITS Altair 8800 and how it was able to pull like-minded individuals together to form the so called Homebrew Computer Club, a milestone in the Computer Science history that is considered as the root of not only Hackathons, but also a series of innovations that created new meaning to the computer and consequently changed the society as a whole. The chapter shows that these powerful technological principles capable of driving innovation and aggregating communities have been present ever since. The concept of platforms, rooted in modularization, affords extensible and open-ended systems, which invite individuals to think about new possibilities given a certain set of standardized components and constraints. In addition, open access to the technology and its architecture is essential to enable the evolution of technology by dedicated ecosystems. Open access to the architecture of the technology, by means of open-source code, open schematics, open data, and so on, enables anyone to contribute
with adaptations to the system. By adapting a certain technology to solve a particular problem, the developer might be potentially contributing to the solution of similar problems encountered by numerous other individuals, if open access is granted. The chapter finally presents a series of examples of Cultural Institutions that have adopted an open attitude by making their Digital Collections freely available as online repositories, and engaging with and releasing new open-source projects in order to profit from vibrant digital ecosystems.

Chapter 5 discusses the implications of the format and content of Digital Collections for the implementation of Digital Interpretive Artifacts, since not only the design process, but also the capability of these applications are defined by Digital Collections. In this sense, the applicability of Digital Collections as a creative material depends on a number of intertwined factors and the entangled relationship between Conceptual Models, Data Standards, and Collection Management Systems. One of the cornerstones of the chapter is the visibility and invisibility of representations, which, in case of memory institutions, are crucial for the success of their social mandate. The question of visibility is the basis of interpretive processes. That is because the non-existence of formalized memory (as e.g. what is represented by Digital Collections) hinders any kind of collective discussion that is supposed to provide understanding about a topic. The interpretation of Digital Collections is therefore bounded to the flexibility and speed in which conceptual models of Collection Management Systems are able to adapt. In addition to that, the design approaches used to implement Collection Management Systems are defining factors for their institutional role and interpretive possibilities. The enforcement of workflows due to issues such as compatibility with legacy systems, current engineering and curatorial practices, and end-user expectations arise in the discussion by presenting both advantages and disadvantages. On the one hand, rigid workflows provide consistency to institutional work. On the other hand, content that do not conform to the infrastructure in place cannot be described by these systems. The end of the chapter presents a case study that discusses how these issues are dealt with in a real case scenario. It is evident the appearance of parallel institutional infrastructures that are able to better accommodate educational and research needs, which, because of their elasticity, do not fully conform to core conceptual models and standards.
Chapter 6 presents the conceptualization, implementation, and evaluation of a Fast-speed IT platform called *Artfacts*, which was designed within the context of the two-speed IT infrastructure, where a foundational, stable, and slow infrastructure is complemented by an additional creative, experimental, and agile infrastructure capable of promptly responding to the needs of communities. The platform is an attempt to digitally incorporate strategies for making sense and reusing Digital Collections and mitigate problems concerning specialized knowledge required for profiting from their affordances as a creative material. In this sense, through the cartography of information, the platform aims at widening the participation of individuals with no technical background in the development and maintenance process of Digital Interpretive Artifacts, no matter whether within cultural institutions or events, such as Hackathons for Cultural Heritage. The Artfacts Platform intermediates the reinterpretation of cultural datasets and the fabrication of Digital Interpretive Artifacts by means of a flexible, general, and interoperable data model that is able to adapt to the demands of storytellers, and an open-ended Object-Oriented UI that enables analysis and experimentation by arranging and rearranging data elements into digital narratives.

Finally, chapter 7 discusses the main results and contributions offered by this research project. A complete overview in regard to Hackathons as constructionist methods for the interpretation of Cultural Heritage is detailed and situated within the Two-speed IT Infrastructure. In addition, the chapter also discusses the main benefits and limitations of the Artfacts Platform within the context of these institutional infrastructures and affiliated and non-affiliated communities.
Steven Levy, in his book *Hackers: Heroes of the Computer Revolution* [137], traced the roots of the *Hacker Culture* back to members of MIT’s *Tech Model Railroad Club* (TMRC - see Fig. 1) in the 1960s. Located in an unattractive building erected during World War II for the development of radar technology, the TMRC members enjoyed enough space and freedom to use the location as they wished. The TMRC was organized mainly around a large-scale model of a railway system powered by a complex matrix of wires, relays, and switches. All the members had to contribute to the model somehow. A few were interested in experimenting with electronics and computers, especially the IBM 704² and later the PDP-1³, in order to improve the scale model. The TMRC members saw the possibilities of computers beyond the military, which was the main sector where these machines were used at the time. MIT, however, possessed a small number

---

1 Picture source: https://www.wired.com/2014/11/the-tech-model-railroad-club/
2 See https://www-03.ibm.com/ibm/history/exhibits/mainframe/mainframe_PP704.html
3 See http://history-computer.com/ModernComputer/Electronic/PDP-1.html
of these expensive machines. Although being denied access to the computers, the club members found always ways through the strict safeguard. The lack of authorization to use the computers was only one of the problems students had. The club also lacked resources, such as electronic parts necessary to make the scale model run. In order to bypass the difficulties, they had to work with donated and discarded devices. The only way to keep on building and improving the railway model was to take these devices apart, reuse and remix their components in order to create the pieces they needed. According to Levy, “the most productive people working on Signals and Power called themselves ‘hackers’ with great pride” [137, p.7].

This anecdote shows clearly that the origin of the terms hacker and to hack had under no circumstances a negative connotation. As stated by Jay London [141], the term hack, whose ancient meaning relates the act of cutting something with rough movements, appeared for the first time on the minutes of TMRC in April of 1955 in the passage: “Mr. Eccles requests that anyone working or hacking on the electrical system turn the power off to avoid fuse blowing” [141]. The hacker was not someone who caused cyberattacks, stole credit card information, or used computers to manipulate democratic elections, as the term is partially used today. In the 1960s and 70s, a hacker was someone who, driven by necessity, had to make use of what was available around to come up with ingenious solutions to a technological problem. Someone, whose creativity was a driver for finding usefulness to things that seemed completely useless. Hackers got together not to destroy things, but to build them. According to Levy, these pioneers believed that “essential lessons can be learned about the systems — about the world — from taking things apart, seeing how they work, and using this knowledge to create new and even more interesting things” [137, p.28]. This creative building process was fulfilled of a very special kind of ethics, that ranged from dismissing authority and resisting bureaucracy to envisioning computers as machines capable of deeply changing society. The main tenets were:

- “Access to computers — and anything that might teach you something about the way the world works — should be unlimited and total. Always yield to the Hands-On Imperative!” [137, p.28]
- “All information should be free.” [137, p.28]
- “Mistrust Authority — Promote Decentralization.” [137, p.29]
- “Hackers should be judged by their hacking, not bogus criteria such as degrees, age, race, or position.” [137, p.31]
• “You can create art and beauty on a computer.” [137, p.31]
• “Computers can change your life for the better.” [137, p.34]

Before anyone else, the first hackers were dealing with principles that later on shaped many of the cultures, movements, philosophies, and initiatives that followed. Meritocracy and decentralization, key principles for current projects such as Wikipedia and the Open-Source Movement, already dictated human relations among MIT hackers. Technology as a driver for empowerment, disruption of authority, and subversion of established rules showed its potential in a small, but nevertheless vigorous way at MIT long before the Web 2.0 transformed society. A hands-on attitude, which is today the cornerstone of the DIY Culture and the Maker Movement, was a must-have mindset to bypass challenges. The belief in the computer as a transformative social tool, capable of changing and being changed by traditional human disciplines, made those old pioneers to come up with not only the first space shooter computer game ever created, but also the first music compiler, and first word processor. They shifted the function of the computer from a mere number crunching machine to a device to create, learn, and play with. This though was hugely influential and generated the seeds of disciplines such as Digital Humanities, learning theories such as Constructionism, and a variety of industries ranging from computer industry itself to digital gaming.

The powerful ideas and transformative possibilities afforded by technology, deeply affected the Cultural Heritage Sector. As a young audience grew accustomed to interaction, instant feedback, and dynamic communication provided by new media, Cultural Institutions were pushed to take an audacious turn towards openness in order to avoid the eminent collapse in visit numbers. Openness, for Cultural Institutions, meant not only implementing interactive technology within exhibitions, having an online presence,
and communication channels with their visitors, but also and most importantly opening up to public as an active co-contributor in the construction of cultural narratives\textsuperscript{14}. Institutions were compelled to decentralize and share their authority with communities and other fields of knowledge. Cultural content has been shared extensively on e.g. social media and specialized platforms such as GitHub\textsuperscript{15}. The inevitable and uncontrollable usage of cultural content lead by the Internet and the ubiquitous presence of smartphones equipped with cameras forced more flexible copyright agreements \textsuperscript{5}. The visitor has been called to take part in reinventing institutions\textsuperscript{16}. Communities, made of experts in areas that the institutions do not detain the know-how are asked to hack their content. Not only long-term or permanent educational-oriented programs, such as Makerspaces, but also short-term professional-targeted events, such as Hackathons, are now part of the repertoire of institutions.

2.1 The Hackathon

On the 4th of June 1999, a group of ten software engineers got together at a private location in Calgary, Canada in an event called the Hackathon (see \textsuperscript{60}, \textsuperscript{176}). The event was organized by the Canadian Theo de Raadt, the founder of the open-source project OpenBSD\textsuperscript{17}, and had as objective working collaboratively in order to fix bugs and further develop the network protocol IPsec\textsuperscript{18}, which is a standard that ensures packages are sent securely over a network \textsuperscript{176}. The term Hackathon is derived from a combination of two other words, namely hack and marathon \textsuperscript{32}. In this specific case, however, hacking did not stand for gaining unauthorized access to someone else’s computer, but for engaging in experimenting with technology to come up with a concrete solution to programming challenges in short amount of time\textsuperscript{19}. Coincidently or not, some days later, a second Hackathon took place. This time, the event was organized by Sun Microsystems during the JavaOne Conference in San Francisco, USA,
which took place between June 15-19 of 1999. In an almost similar but competitive format, developers took part in the Most Visionary App contest to develop innovative applications programmed in Java and running on the portable Personal Digital Assistant (PDA) called Palm V [36]. This second event resembles a lot more the format of some of the Hackathons employed today, which are not only well-organized and competitive, but also backed by well-established companies, institutions, and sometimes even governments. Notwithstanding, no matter if past or present, collaborative or competitive, the limited time that force individuals to immerse themselves into an intensive work-flow, barely stopping for eating or sleeping, and the exploratory nature towards the possibilities of what one can achieve with a certain technology are among the characteristics that define Hackathons.

2.1.0.1 Hackathons also for Cultural Heritage

Hackathons have largely popularized across many different areas besides the obvious ones, such as Computer Science. Take the example of the Science Hack Day20, which is a two-day-all-night event that gathers a variety of individuals from many different branches of science in order to prototype their ideas. According to the organizers, 50 events took place in 19 different countries in 2015 [207]. In 2017, the number grew to 92 events taking place in 45 cities in 27 countries [208]. According to the agency BEMYAPP21, which

20 http://sciencehackday.org/
21 http://agency.bemyapp.com/
manages a platform\textsuperscript{22} that helps to list and organize Hackathons, in 2016, more than two hundred thousand individuals took part in Hackathons, producing thirteen thousand prototypes in across more than a hundred countries [131]. However, in the Cultural Heritage Sector there are neither precise numbers on the occurrence of Hackathons nor their exact origin.

A post from April 2011\textsuperscript{23} on the Europeana Blog\textsuperscript{24}, however, may provide an insight on the first official Hackathon for the Cultural Heritage Sector organized in Europe\textsuperscript{25}. On the blog post, Milena Popova\textsuperscript{26}, responsible for the promotion of Europeana’s re-use services, writes:

\begin{quote}
At the beginning of April we held the first in the series of hackathons to showcase what cool projects can be done with Europeana collections... We invited an interesting mix of hackers from Europeana partner institutions and freelancers with cultural heritage portfolio and challenged them to try out their ideas for creative reuse of the Europeana content. The only limitation was that all hackathon results will be for non-commercial use only. [188]
\end{quote}

This Hackathon (see Fig. 2\textsuperscript{27}) happened within a bigger context, which was the Europeana project\textsuperscript{28} itself. Europeana is a project developed by the European Commission\textsuperscript{29} to assist and encourage European states to digitize Cultural Heritage. One of the main goals of the project is to provide tools and a free-access services platform for Cultural Heritage so that a wide range of digital content that can be shared and reused for different purposes. The start of the Europeana project can be tracked down to a letter sent by six Heads of State and Government in 2005 urging for the creation of a European virtual library, “aiming to make Europe’s cultural and scientific record accessible for all” [49]. A prototype version of the platform was launched in November 2008, followed by an official release of the platform in February 2009. As a data hub, besides collecting digitized material from cultural institutions across Europe,
and establishing standards for modeling Cultural Heritage metadata, one of the cornerstones of the Europeana project is to stimulate the creative employment of the datasets and APIs [70]. In this sense, Hackathons appear as a fitting alternative to this purpose. In regard to the first Europeana Hackathon in particular, Adrià Mercader, one of the participants, writes on her blog [30], “the team behind Europeana has recently published a preliminary beta version of an API, and the objectives of the Hackathon were both receiving feedback from developers and see what they could came up with after playing with it.” [154] Since then, many Hackathons that deal with Cultural Heritage have been continuously organized, such as the Coding Dürer Hackathon [31], GLAMHack17 - Kulturhackathon [32], Koggethon [33], The Future Museum Challenge [34], GallenKallela hackathon [35], Canadian Museum of History and the Canadian War Museum Hackathon [36], Philadelphia Museum of Art Hackathon [37], The Body: Hackathon [38], and the Science Museum London Digital Lab Hackathon [39] just to name a few.

2.1.1 The Structure of a Hackathon

As the first Hackathon organized by de Raadt, these events do not need to obey a formal or well-defined organizational structure to take place. Having a topic to work on, access to the necessary technology, and people willing to spend time together in order to transform an idea into a prototype are the necessary ingredients to make a Hackathon happen. However, a more structured format has been observed since these events gained in popularity, and has driven interest of companies, institutions, and governments. The Wired Magazine [40] correspondent Steven Leckart, in an article entitled The Hackathon is on: Pitching and programming the next killer app [132], takes a look especially at competitive Hackathons and describes the six main phases that constitute these events. According to Leckart [132],

---

30 http://amercader.net/blog
31 http://codingdurer.de/
32 http://www.openglam.at/#hack
33 http://www.dsm.museum/info/veranstaltungen/koggethon.6408.de.html
34 http://www.europeana-space.eu/hackathons/museums/
35 http://creative-museum.net/329-2/
36 https://www.hackworks.com/museumvx
37 https://philamuseumhackathon.devpost.com/
39 https://www.eventbrite.co.uk/e/science-museum-london-digital-lab-hackathon-tickets-31125271525
40 https://www.wired.com
participants must come up with an idea that can be implemented, form a team, develop a concept, demo a finished prototype, and confront the judges, who will vote for the best product.

<table>
<thead>
<tr>
<th>Collaborative</th>
<th>Competitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction of infrastructures by organizers</td>
<td>Launching</td>
</tr>
<tr>
<td>Understanding of infrastructures and formation of teams</td>
<td>Forming</td>
</tr>
<tr>
<td>Negotiation of team roles</td>
<td>Storming</td>
</tr>
<tr>
<td>Establishment of workflows</td>
<td>Norming</td>
</tr>
<tr>
<td>Construction of the prototype</td>
<td>Performing</td>
</tr>
<tr>
<td>Pitching the project to the jury and/or other participants</td>
<td>Staging</td>
</tr>
<tr>
<td>Voting on the best projects</td>
<td>Selecting</td>
</tr>
</tbody>
</table>

Figure 3: Structures of collaborative and competitive Hackathons

Taking a closer look at group formation and processes specifically, Trainer et al. [229] use the same categories created by Egolf et al.[84] to describe how individuals organize themselves and carry out work during Hackathons. According to Trainer et al. [229] four different phases happen over time. In the Forming phase, individuals need to understand the task and acquire the necessary means to carry it out. Moreover, individuals need also to meet other people, who could be a potential good match for a team. In the Storming phase, individuals need negotiate their role in the team. In the Norming phase, individuals share tasks and organize their working style and process. And finally, during the Performing phase, individuals adhere to the norms previously established and carry out work with minimal emotional interaction. Annika Richtericht [195], while conducting case studies with Hackathons41, could also observe the happening of these phases as described by Egolf et al.[84] and Trainer et al. [229]. According to Richtericht, the three first phases (forming, norming, and storming), however, are “crucially shaped by information exchange on existing skills” [195]. Richtericht [195] additionally added

41 SHD in Eindhoven in August 2014 and Hack4DK in Copenhagen in September 2014.
yet another phase to the group dynamics at Hackathons: *Staging,* which concerns groups having to persuasively pitch their idea to others.

While Trainer et al. [229] and Richtericht [195] focus on the phases related to the interaction within groups of participants, Leckart [132] talks about at least one extra phase that is related to the organizers of the event, and necessary for competitive Hackathons, which is the *prize ceremony.* The authors, however, do not mention another important phase that takes place at the very beginning of either competitive or collaborative Hackathons. The introduction or presentation of the Hackathon is carried out by the organizers, which is described here as *Launching*42. This phase is the defining step for directing the event towards a common goal, usage of a certain technology, or/and setting up rules and values of the event. Based on the contents discussed so far, the figure above (see Fig. 3) provides an overview on the organizational structures of Hackathons. For either collaborative or competitive Hackathons, an introductory phase, carried out by organizers, is presented in order to set the goals and technologies that participants must engage with during the event. It is followed by a second phase aimed at forming teams and negotiating roles of participants. This phase can be subdivided into the stages suggested by Trainer et al. [229] and Richtericht [195], where negotiations related to participants’ skill-sets take place. After that, participants employ their skills and work together in order to materialize their ideas. This is the third phase that is equivalent to the *Performing* phase, previously described. Finally, a forth phase requires participants to demonstrate and, sometimes, discuss with others the results they have accomplished. In addition to the stages described until now, two extra phase are required in competitive Hackathons. The *Selecting* phase is necessary to rank the most relevant projects. This phase can be performed by either a jury or the participants voting in the projects they judge the best. It is not uncommon that the jury and the participants decide on winners utilizing both different criteria for different categories of prizes. The selecting phase happens usually together with a ceremony arranged by organizers in order to distribute prizes to the winners of the competition.

---

42 This phase was observed during the case study described in Section A.2.3.1 and the analysis of the documentation of other Hackathons addressed in this dissertation.
2.1.2 Motivations and Profile of Participants

As described by Briscoe et al. [32], a survey from 2012 [160] conducted by TokBox\textsuperscript{43} with 150 attendees of Startup Weekends, Hack Days, and other similar events from across the United States wanted to know, among other questions, the motivations of participants for attending these events. The study found out that, among developers the main motivations are by far learning (86%) and meeting people (82%), followed by changing the world (38%). Among non-developers, networking (56%), finding partners (47%), and learning how to code (41%) were among the most popular answers. More complete results of the study can be seen below (see Fig. 4).

![Figure 4: Motivations to take part in Hackathons according to [160]](https://tokbox.com/about)

Similar motivations can be observed in individuals who are active members of the open-source community, which conceived not only the first Hackathon [176], but also the set of principles that inspired the Hacker Culture in general (see [46], [222]). As Weber [237] points out, working in such complex projects require certain motivational characteristics from individuals. These are people who are willing to voluntarily donate their time and expertise to work in projects that, in great part, will not provide them with direct monetary reward. Open-source projects appeal to people’s individual inclinations, such as having fun with challenging and interesting programming puzzles, contributing to projects that are ideologically and socially relevant, gaining visibility and reputation within the community, and getting

\textsuperscript{43}TokBox is a subsidiary of the telecommunication company Telefónica S.A. that provides solutions for video, voice and messaging to websites and mobile apps (see https://tokbox.com/about).
to know and profiting from more experienced developers [7]. These motivational factors of contributors of open-source projects are aligned with the findings of Briscoe et al. [32].

In the book *Identity and the Museum Visitor Experience* [78], John Falk proposes five identity types of museum visitors and describes their main motivations to engage with culture. According to Falk [78], museum visitors can be sub-divided into:

- **Explorers** - Interested in learning new things, these individuals have a generic, instead of focused, interest towards the content of the institution. They are driven by curiosity and browse through contents that may grab their attention (see [78, p.190]).

- **Facilitators** - The goal of these individuals is to guide and support the visit and the learning of others. A teacher who provide pupils with a guided-tour, or parents who bring their children to a museum fall in this category (see [78, p.192]).

- **Experience seekers** - These are individuals for whom the main motivation to visit cultural institutions is not much due to their content, but because they are perceived as important landmarks. Tourists personify well individuals in this category (see [78, p.196]).

- **Recharger** - These are individuals who use not only museums, but also aquariums, botanical gardens, national parks as a retreat from their stressful routines. Sometimes, cultural institutions are also seen by individuals of these categories as places that provide spiritual experiences (see [78, p.203]).

- **Professional/hobbyist** - These are individuals for whom their professional careers or/and hobbyist passions are aligned with the work and contents of the institution. Their visit and involvement with the institution is motivated by a content-related objective (see [78, p.199]).

Although not taking into consideration participants of Hackathons directed at Cultural Heritage, Falk’s description of the *professional/hobbyist* provide a good definition of these individuals’ profile. According to Falk, this is “the smallest category of visitor to most institutions, but they are the often disproportionately influential” [78, p.200]. They are people to whom the institutions pay great attention, because of the professional careers, expertise, and knowledge they hold. Their engagement with the institution is focused and goal-oriented, instead of generic, as for e.g. *explorers*.

An online survey conducted from August to October 2017 with 108 individuals (see A.2.1.1) revealed that participants of Hackathons are in some cases professionals working for cultural institutions
and occupying positions in curatorial and research departments of these institutions. From the 108 interviewees, almost half (45% - 49 individuals) stated having taken part in Hackathons one or more times. From this specific group of individuals, when asked about their professions in a multiple response question, 6 interviewees responded they worked as curators. Other professions, such as researchers (32.7% - 16 occurrences out of 49), programmers (32.7% - 16 occurrences out of 49), and managers (20.4% - 10 occurrences out of 49) were among the most popular ones. Only 1 individual declared him or herself as a student.

The professional maturity of participants of Hackathons is also suggested by their age groups (see Fig. 5). From the 49 interviewees that stated having taken part in Hackathons, the age group 35 to 44 years old appears as the largest group with 17 (34.7%) individuals. This group is followed by the age group 45 to 54 years old (32% - 16 individuals), 25 to 34 years old (22.4% - 11 individuals), and 55 or older (10.2% - 5 individuals). None of the interviewees belonged to the age groups 18 to 24 years old and under 17 years old. Individuals belonging from the age group 54 to 25 years old encounter themselves either in the beginning of their professional careers or are already well-established professionals. These results provide additional support to Falk’s findings on the visitor identity type professional/hobbyist.

The term hacker encompasses a variety of different professionals.

The mature age group of participants of Hackathons for the Cultural Sector contradicts the common sense.

Figure 5: Age groups of hackers according to survey A.2.1.1

44 Among the curators, 2 (12%) individuals declared themselves as being also researchers, 1 individual declared him or herself as being also an artist, and finally the 3 individuals left stated working in management positions.
It is surprising to observe that the great majority of Hackathon participants (81.6% - 40 individuals out of 49) describe themselves as non-programmers, what suggest a high degree of multidisciplinarity in Hackathons for the Cultural Heritage Sector. The same multidisciplinarity can be observed in Hackathons for different sectors. The diversity of participants supplies teams and projects with highly multidisciplinary skill-sets. Although there is a hidden consensus that tend to define Hackathons as computer programming events [32], they are nowadays not only exclusively directed to individuals possessing skills relevant to software development, such as programmers, UI, UX, or graphic designers. Depending on the topic of the Hackathon, a multidisciplinary team is formed that must necessarily include individuals, such as journalists - e.g. Hacking Journalism45, health-care professionals - e.g. MIT Hacking Medicine46, scientists - e.g. Science Hack Day47, Cultural Heritage professionals and enthusiasts - e.g. Coding Da Vinci48, or students - e.g. Hack like a Girl49.

2.1.3 Motivation of Organizers

Companies, institutions, and governments engage with Hackathons for many different reasons. The ability to build and test prototypes in just a few days, finding use cases for new technologies, training adopters, promoting technologies to target groups, and recruiting new talents are just a few examples. Most importantly, Hackathons work as a lab for experimenting with technology that is capable of generating innovation. Facebook’s Like Button, Chat, and Timeline, for instance, were conceptualized during internal Hackathons [214]. Hackers must create their prototypes around certain kinds of data, APIs, programming languages, hardware (among other technologies) that are supplied by the organizers. External Hackathons are also arranged by organizations as a way to get access to highly motivated individuals with the attitude and knowledge to find new employment to their resources. Although organizations usually run Hackathons around a well-defined set of technologies and datasets, the conceptualization of the applications is in its great majority up

45 http://hackingjournalism.com/
46 http://hackingmedicine.mit.edu/
47 http://sciencehackday.org/
48 https://codingdavinci.de/
49 http://hacklikeagirl.weebly.com/
to the participants to decide. Technologies and datasets provide a powerful and stable platform on top of which negotiations, design decisions, and intensive development happen as necessary steps in order to materialize an idea.

Figure 6: Participants of the Kultur-Hackathon Coding da Vinci

Taking the Cultural Heritage Sector in particular, Hackathons are organized mainly around datasets. The project Kultur-Hackathon Coding da Vinci\(^50\) (CdV - see Fig. 6\(^51\)) is a good example to illustrate the power of Hackathons as events capable of producing high quality, useful, and innovative outcomes based on datasets. The project is a joint initiative of the Open Knowledge Foundation\(^52\), Wikimedia\(^53\), the Servicestelle Digitalisierung Berlin\(^54\), and the Deutsche Digitale Bibliothek\(^55\), and has run yearly since 2014. In total, from 2014 to 2017, 71 applications were created using the datasets provided by more than 70 institutions, such as the Archäologisches Museum Hamburg, Bayerische Staatsbibliothek, Berlinische Galerie, Bibliothek und Informationsvermittlungsstelle für das Bezirksamt, Botanische Gärten und Botanisches Museum Berlin-Dahlem, Dänisches Nationalmuseum, GRIPS Theater Berlin, Jüdisches Museum Berlin, Universitätsbibliothek der Bauhaus-Universität Weimar, Verein für Computergenealogie just to name a few. Unlike other Hackathons, the CdV represents a good case study,

\(^{50}\)https://codingdavinci.de
\(^{52}\)https://okfn.de
\(^{53}\)https://wikimedia.de
\(^{54}\)https://www.servicestelle-digitalisierung.de/
\(^{55}\)https://www.deutsche-digitale-bibliothek.de/
because it is not focused on one specific dataset or API, but gives
great possibility of choice to the participants, since many institutions
take part at once. The Hackathons organized by Europeana, for
example, happens around only the Europeana API [188]. According
to the project report of 2014:

Das übergeordnete Projektziel lag in der Entwicklung einer
möglichst großen Anzahl von prototypischen Applikationen
mit unmittelbarem Nutzen für Endanwender und
Kultureinrichtungen. Gleichzeitig wurden Kulturinstitutionen
dazu ermutigt, ihre digitalisierten Sammlungsbestände frei
zugänglich und nutzbar zu machen. Kultureinrichtungen und
Teilnehmern sollte eine Plattform geboten werden, um sich
auszutauschen und gemeinsam Ideen für die aktive Nutzung
des digitalen Kulturerbes zu entwickeln. Die vertraglich
definierten Ziele der Veranstalter wurden erreicht und sogar
übertroffen. [102, p.3]

By making their datasets publicly and freely available online,
Cultural Institutions provide a material that hackers use to design
with\textsuperscript{56}. The result of this design process is a series of prototypical
software with immediate benefits for end-users and Cultural
Institutions, as stated by the quotation above. Cultural Institutions
have at once several applications being developed with their content
for almost no cost, besides curating, preparing, and serving their data.
The visitor, and eventually anyone interested in the cultural content
provided by these institutions, is able to download the apps on their
devices or use it online via the Web, if this is the case. Because
they provide interpretive material content on top of their datasets,
the applications provide value to institutions. The data presented in
the datasets are themselves illegible to the public. The applications
transform this raw data material into content that can be digested by
anyone. A few examples of applications developed for the CdV range
from simple mobile apps, games, and visualizations to augmented
and virtual reality experiences. The following paragraphs present a
few examples:

\textsuperscript{56} More on that on Chapters 4 and 5.
One of the applications produced as part of the CdV 2016, the mobile application Zeitblick\(^{57}\) (see Fig. 7), for example, used the dataset\(^{58}\) provided by the Museum für Kunst und Gewerbe\(^{59}\) to return historical portraits based on selfies taken by users. Depending on face expressions, the algorithm recognized and compared the head rotation and emotional features of both the user, and the human figures presented on the portraits. Simple variations on the selfie, such as being serious or smiling, and moving the head up or down caused the application to retrieve different artworks. From the CdV 2015, an interesting example was the mobile application Midiola\(^{60}\) (see Fig. 7). The app used the dataset from the Deutsches Museum\(^{61}\) to enable users to browse through old music rolls, which are storage media used in mechanical musical instruments, and to digitally analyze them in order to play songs. The software also let users to take pictures of physical music rolls, extract their patterns, and play their content based on the pictures taken.

\(^{57}\) https://codingdavinci.de/projects/2016/zeitblick.html
\(^{58}\) https://github.com/MKGHamburg/MKGCollectionOnlineLIDO_XML
\(^{59}\) http://www.mkg-hamburg.de
\(^{60}\) https://codingdavinci.de/projects/2015/midiola.html#project-name
\(^{61}\) http://www.deutsches-museum.de/
An addition example is the IOS application from 2014 *Alt-Berlin*\(^62\) (see Fig. 8\(^63\)) that combined datasets from four different sources to power distinctive functionalities of the application. The Alt-Berlin app used the historical photos of the *Stadtmuseum Berlin*\(^64\) to display to the user, on a contemporary map, the urban landscape from the year 1400 until the 2000s. In addition, the app also allowed overlay of old and new maps of Berlin using the data from the *Senatsverwaltung für Stadtentwicklung und Umwelt Berlin*\(^65\), displayed the changes in location and shape of the medieval district walls using the data from the *Amt für Statistik Berlin-Brandenburg*\(^66\). And finally, it presented before/after pictures to display the Kreuzberg district from the year 1887 to 2007 using *Wikipedia*\(^67\) content. Also from 2014, the *DNB - Data Explorer*\(^68\) (see Fig. 8) used the data from the *Deutsche Nationalbibliothek*\(^69\) to present a series of interactive visualizations based on information about people and places. One of these visualizations, for instance, shows the geographical distribution and population density of certain German cities for a time period that can be set by the user.

\(^62\) https://codingdavinci.de/projects/2014/altberlin.html
\(^63\) Picture source: http://www.sebastianmeier.eu/2014/06/21/deutsche-national-bibliothek-data-explorer/
\(^64\) https://www.stadtmuseum.de/
\(^65\) http://www.stadtentwicklung.berlin.de/
\(^66\) https://www.statistik-berlin-brandenburg.de/
\(^67\) https://www.wikipedia.de/
\(^68\) https://codingdavinci.de/projects/2014/dnb-data-explorer.html#
\(^69\) http://www.dnb.de/
Last but not least, the application Skelex (see Fig. 9), developed during the 2017 edition of the CdV uses the data from the Museum für Naturkunde Berlin to immerse the user into a virtual reality environment where he or she can explore skeletons of serpent heads. The user can virtually grab, rotate, and pull the different bone structures apart. The user can also measure the bones with the help of a virtual ruler. According to the team [112], the application aims at providing an educational experience for museum settings.

The examples above are only a small set of possibilities that can be achieved by implementing Hackathons within the Cultural Heritage Sector. One of the most noticeable uniqueness of Hackathons for Cultural Heritage, if compared with other kinds, is the usage of data repositories, the so called Digital Collections, for constructing shared interpretations about heritage by digitally fabricating Digital Interpretive Artifacts. As previously explained, the hacking process in this kind of Hackathons consists of not only being able to design and program an application, but also having the ability to convert illegible data repositories into interpreted content that can be understood by the audience and served as easy-to-handle interactive media. The next chapter will examine the historical motivations of Cultural Institutions behind their efforts in sharing their data repositories with communities, and how this open attitude relates to the concept of Heritage Interpretation.

---

70 Picture source: https://cdv-skelex.github.io/
71 http://gbif.naturkundemuseum-berlin.de/hackathon/insektenkasten/
72 See definition and discussion in Chapter 5.
73 See definition and discussion in Chapter 3.
In 1957, Freeman Tilden released a seminal work on Heritage Interpretation, called *Interpreting our Heritage* [227], which has hugely impacted how Cultural Institutions relate with their audiences. Between journalism and fiction writing, Tilden’s earlier careers had great influence in his later works on the way interpretation takes place in the Cultural Heritage Sector. Especially on the chapter three of the book, when discussing the differences between information and interpretation, Tilden talks about two journalistic principles present in the beginning of the 20th century that often collided with each other: on the one hand, a journalism dedicated to communicate only facts, and, on the other hand, a journalism where facts are only the background for a compelling story. According to Tilden [227, p.19], the differences between these different journalistic approaches were evident during the reporting by two main journals from New York City on the San Francisco’s earthquake of 1906. While the *Times* insisted in describing only the facts about the earthquake, the *Sun’s* journalist Will Irwin, who was a San Francisco native, was able to write a news story lamenting the destruction of the city that appealed to the readers’ emotions. The readers, according to Tilden:

...saw, felt, and heard - and lamented the loss of something that had instantly become theirs. This was Interpretation: the revelation of the soul of a city. It was based upon fact, but they were not the facts of the earthquake destruction. [227, p.19]

This new kind of journalism was a powerful method for both grasping people’s attention and creating interest for places and times that many had no previous emotional connection to. Tilden realized that the same could be applied to provoke curiosity and care about heritage. With this goal in mind, he proposed six principles for effective Heritage Interpretation [227, p.11-47]:

- “Any interpretation that does not somehow relate what is being displayed or described to something within the personality or experience of the visitor will be sterile.”
• “Information as such is not interpretation. Interpretation is revelation based upon information. But they are entirely different things. However, all interpretation includes information.”

• “Interpretation is an art, which combines many arts, whether the materials presented are scientific, historical or architectural. Any art is in some degree teachable.”

• “The chief aim of interpretation is not instruction, but provocation.”

• “Interpretation should aim to present a whole rather than a part, and must address itself to the whole man rather than any phase.”

• “Interpretation addressed to children (say, up to the age of twelve) should not be a dilution of the presentation to adults, but should follow a fundamentally different approach. To be at its best it will require a separate program.”

Besides his experiences as a journalist, one could argue that the ideas of Tilden towards interpretation were perhaps influenced positively and/or negatively by the theories of cognitive development being formulated during his professional life. In 1950, Jean Piaget published an article that later lead to the release of his seminal book _The Psychology of the Child_ [185], which advocated the importance of the individual’s own experiences in learning; as well as the role of processes such as assimilation and accommodation in acquiring new knowledge based on previous one. On the other hand, the first edition of Skinner’s _Science and Human Behavior_ [211] had been published in 1951, also driving behavioral and learning theories that over-emphasized positive and negative reinforcements for memorizing information; therefore a more desensitized view on the way one’s knowledge acquisition occurs. Tilden pointed out to the affective quality of interpretation, distancing his concept of interpretation from the dry instruction happening in classrooms. However, he kept unchanged the roles of the interpreter, as the one possessing and transmitting knowledge, and the visitor, as one lacking and receiving knowledge. Nevertheless, for Tilden, interpretation appeals to the direct experiences of visitors with sites and objects. He emphasizes the role of the interpreter as someone who is able to make the meaningful connections between audiences and the heritage through storytelling. Heritage Interpretation is “an educational activity which aims to reveal meaning and relationships through the use of original objects, by firsthand experience, and by illustrative media,

1 For Tilden, the concept of instruction was closely related to the transmitting and receiving information, and not necessarily engaging emotionally with it. [227, p.33]
rather than simply to communicate factual information” [227, p.8]. Based on Tilden, other authors have come up with their own definitions of interpretation. In the book Interpretation for the 21st Century [15] released in 1998, Larry Beck and Ted Cable defined interpretation as a “a process that can help people see beyond their capabilities” [15, p.4]. In an attempt to expand Tilden’s vision to the challenges of the Information Age, Beck et al. [15] proposed fifteen principles as guideline for best practices in Heritage Interpretation. According to Beck et al. [15, p.10-11]:

- “To spark an interest, interpreters must relate the subject to the lives of visitors.”
- “The purpose of interpretation goes beyond providing information to reveal deeper meaning and truth.”
- “The interpretive presentation as a work of art should be designed as a story that informs, entertains, and enlightens.”
- “The purpose of the interpretive story is to inspire and to provoke people to broaden their horizons.”
- “Interpretation should present a complete theme or thesis and address the whole person.”
- “Interpretation for children, teenagers, and seniors when these comprise uniform groups should follow fundamentally different approaches.”
- “Every place has a history. Interpreters can bring the past alive to make the present more enjoyable and the future more meaningful.”
- “High technology can reveal the world in exciting new ways. However, incorporating this technology into the interpretive program must be done with foresight and care.”
- “Interpreters must concern themselves with the quantity and quality (selection and accuracy) of information presented. Focused, well-researched interpretation will be more powerful than a longer discourse.”
- “Before applying the arts in interpretation, the interpreter must be familiar with basic communication techniques. Quality interpretation depends on the interpreter’s knowledge and skills, which should be developed continually.”
- “Interpretive writing should address what readers would like to know, with the authority of wisdom and the humility and care that comes with it.”
- “The overall interpretive program must be capable of attracting support – financial, volunteer, political, administrative – whatever support is needed for the program to flourish.”
- “Interpretation should instill in people the ability, and the desire to sense the beauty in their surroundings – to provide spiritual uplift and to encourage resource preservation.”
• “Interpreters can promote optimal experiences through intentional and thoughtful program and facility design.”

• “Passion is the essential ingredient for powerful and effective interpretation – passion for the resource and for those people who come to be inspired by the same.”

Even when discussing about technology employed in the Cultural Heritage Sector, Beck et al. [15] did not deviate from the direction and essence of what Tilden had defined as effective Heritage Interpretation. Technology was seen only as an extra support material that could be helpful for enhancing storytelling. Chapter eight (Eighth Principle: High-Tech Gadgetry), for example, offers good insights on how Cultural Institutions oriented themselves towards the increasing presence of technology in their practice not only in the 90s, but also in the later decades to come. Beck et al. recognized the benefit of technology in enabling the audience’s capacity “to go beyond their senses to see things as we have never seen them before. High technology can reveal startling new perceptions about our world” [15, p.102]. However, the authors could not think of these systems beyond one-directional information transmitters. As we will see, the new models of communications and interaction driven by the transformations in technology at the turn of the century provoked deep structural changes in society as a whole, including the Cultural Heritage Sector. These new technologies changed old structures and introduced new ways to interpret heritage.

3.1 Disrupting Interpretation

3.1.1 Interpretation as Dialog

The vision that Cultural Institutions are the primary authority and the only entities capable of producing and providing unidirectional interpretive content and meaningful experiences was strongly questioned by Nina Simon in her book Participatory Museum [209] released in 2010. Although not basing her analysis solely on the role of technology in Cultural Institutions, but explaining how principles and practices of the Web 2.0 have influenced the Cultural Sector, Nina Simon presented several case studies demonstrating a change in paradigm that has deeply transformed Heritage Interpretation. The main argumentation of Simon is that the maturing of Web
Technologies in the 2000s established channels and tools for anyone to become a potential creator of content. While the Web 1.0 was seen as medium through which static documents were displayed to be consumed, the Web 2.0 invited users to contribute. Wikis, blogs, and social media websites were just a few examples of pioneering models that succeeded under this new paradigm. Institutions, in a constant struggle to keep being relevant, have slowly adopted similar approaches to reengage with their audiences. This new attitude, according to Simon, considers the institution as:

...a platform that connects different users who act as content creators, distributors, consumers, critics, and collaborators. That means the institutions cannot guarantee the consistency of visitor experience. Instead, the institution provides opportunities for diverse visitor co-produced experiences. [209, p.2]

In the fall of 2001, three years later to the release of Beck et al.’s book, an important event radically shifted the paradigm upon which technology was based. The dot-com bubble [142], and the technological achievements that were developed during the years that preceded it, established and consolidated new models for making business, communicating, and organizing that has deeply transformed society in all levels. Companies such as Amazon.com (founded in 1994) and eBay (founded in 1995), which have survived the crash, and projects such as Wikipedia (launched in 2001) were able to take advantage of the architecture of participation [168] of the Web, by setting up their systems as platforms that expanded the role of the user as passive recipient of online content. While eBay opened the possibility for anyone to buy or sell their own paraphernalia online, Amazon enabled buyers to review and rate books, what deeply disrupted the current business model of publishers - now afraid of losing sales caused by negative reviews. Also taking advantage of the collective intelligence of the network, Wikipedia’s model was surprisingly able to produce high-quality content [146, p.21], showing the full potential of open systems when individuals are given the tools for co-creating. In addition, social media platforms, such as

---

2 As pointed out by Simon [209, p.I], a Survey conducted by the National Endowment for the Arts in the United States, released in 2009, revealed that museums, galleries, and performing art institutions had continuously lost public attendance if compared to the previous decades. [239]

3 https://www.amazon.com/

4 https://www.ebay.com/
Facebook, had an enormous impact in societies around the globe. For Cultural Institutions, these platforms represented also a rupture that did not ask for permission to disrupt their practices. As Nanci Proctor points out, “people share their own photos, videos, and links about and to museums around the world through platforms that are not in the museum’s control” [189]. Cultural Institutions felt the changes provoked by this rupture in paradigm in that their authority, role, and relevance were shaken by the audiences’ access of information, and possibility to create and disseminate content. As an active element in the cultural sector equation, the audience drove Cultural Heritage professionals to rethink the role of their institutions in terms of not only educational, but also entertainment aspects. In this sense, John Falk in his book Identity and the Museum Visitor Experience [78] says that:

> We need to understand the museum-going first and foremost as a leisure experience. Thus, to understand museum-going as a leisure experience, we need to understand more about the broader leisure landscape of the twenty-first century [78, p.41].

The ease to which content can be created, shared, and consumed online by members of social media platforms collide with the uniqueness and authority of institutions. In addition to the massive dissemination of copyrighted cultural material that is distributed online - over which institutions have no control of, and original productions made by ordinary people, social media platforms represent an environment where institutions, companies, governments, small groups, and single individuals have the chance to compete among themselves for visibility among audiences. Just to give an example, the content produced by independent video bloggers can achieve far greater reach than traditional TV broadcasters on the same platform\(^5\). The authority, influence, and popularity of content produced on these platforms compete side-by-side, independently of the producer. In this sense, “whether or not museums are actively embracing Flickr, Wikipedia, YouTube, Facebook, Twitter, and the rest, their visitors are.” [189]. Therefore, being active on social media is not an extra added value, but a survival necessity, as digital media become an integral part of anyone’s life.

---

\(^5\) The most popular YouTube channel of a video blogger (PewDiePie by the comedian Felix Arvid Ulf Kjellberg) has 54.1 million subscribers, and generates an income of around 15 million dollars per year [113]. In comparison, the YouTube channel of BBC counts to 1.871.298 subscribers [8], and the Museum of Modern Art in New York has 151.701 subscribers [225].
To sum up, not only Tilden’s, but also Beck and Cable’s principles are still, in great part, influencing guidelines for Heritage Interpretation. However, these principles do not account for the long-term transformations that technology, especially digitization, caused within the cultural field. Digital natives, who have grown accustomed to the empowerment provided by online networked platforms, started to expect from Cultural Institutions new forms of engagement that went beyond an unidirectional communicational strategy that considers the institution as an active transmitter of knowledge, and the public as passive receivers of interpretation; and a hierarchical and almost sacred self-perception\(^6\) of the institution, considering itself as an unquestionable authority. Even cornerstones principles in Tilden’s seminal work is questionable nowadays if one considers the following passage:

\[
\text{The adult visitor who happens to be the auditor or reader of interpretation has no general awe of the interpreter. He takes it for granted that the latter possesses special knowledge that he himself lacks, and he respects both that knowledge and the possessor of it (especially if he is in uniform) to exactly that extent. } [227, p.11]
\]

On the one hand, by considering the visitor as someone who lacks knowledge, and, on the other hand, by using words such as reader and auditor, when talking about individuals whose only expected role is to listen, clearly shows the values and conceptions institutions had not long ago towards their audiences. As a matter of fact, some institutions nowadays take a rather radical approach against the idea of Heritage Interpretation as a unidirectional communication process. Minnie Scott [17], a curator of the Interpretation Department of the Tate Modern (London, UK) explains that:

\[
\text{‘interpretation’ suggests that we tell visitors what artworks mean and what to think of them. The word is associated with translation too, so there’s the implication that art is a foreign language to most people and that some sort of interpreter is needed to make sense of it. All this actually goes against what we’re really trying to do. } [17]
\]

The concept of interpretation for the Tate Modern is more about providing audiences with the space for sharing knowledge and

---

\(^6\) Which is in embedded in the ideology of the White Cube [167].
the tools for enabling their own understanding of cultural content, than providing an authority’s view on the collections. Scott adds that the institution is “increasingly becoming a place for dialogue, not only between visitors and Tate but among fellow visitors who have shared experiences and constructed different interpretations.” [17] More evidence will be given in following sections of this chapter to support the contra-argumentation to why Tilden’s view on the public does not hold under the new paradigm, which Cultural Institutions operate nowadays.

3.1.2 Interpretation as Crowdsourcing

As we have seen, the empowerment of the public granted by the democratization of information and tools for co-creating on the Web disrupted the whole sector and obliged institutions to open their physical and virtual spaces to communities, transforming also Heritage Interpretation in a co-creative process. The public has been invited to take part in the curatorial process and co-create stories and narratives through crowdsourcing. In an article for the Wire Magazine from 2006 [226], Jeff Howe coined the term crowdsourcing, which he defines as “the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call.” [226] Over the Internet, communities were formed not because of physical proximity, but solely because they shared the same interests. This phenomenon, that started organically, has increased its relevance because of successful projects, such as Wikipedia, which overtook the place of traditional encyclopedias, such as the Encyclopedia Britannica7, becoming the number one reference. Johan Oomen and Lora Aroyo propose the following classification for crowdsourcing in regard to the Cultural Heritage projects [173]:

- **Correction and Transcription Tasks** - Inviting users to correct and/or transcribe outputs of digitisation processes.
- **Contextualisation** - Adding contextual knowledge to objects, e.g. by telling stories or writing articles/wiki pages with contextual data.
- **Complementing Collection** - Active pursuit of additional objects to be included in a (Web) exhibit or collection.
- **Classification** - Gathering descriptive metadata related to objects in a collection. Social tagging is a well-known example.

---

7 https://www.britannica.com/
- **Co-curation** - Using inspiration/expertise of non-professional curators to create (Web)exhibits.
- **Crowdfunding** - Collective cooperation of people who pool their money and other resources together to support efforts initiated by others.

Particularly by the end of the 2000s, projects that employed such strategies, especially co-curation, started to pop up in the field. They established partnerships between institutions and communities. In June 2008, for example, the Brooklyn Museum⁸ (New York, USA) launched a 100% community-based photography exhibition called *Click! A Crowd-Curated Exhibition* [33]. The exhibition was inspired by the book of James Surowiecki entitled *The Wisdom of Crowds* [220], which, in a nutshell, argues that decisions coming from groups are often wiser than any single individual. The project was divided into three different consecutive phases. An open call for artists’ contributions, an evaluation procedure for the audience to vote on submissions, and the exhibition itself displaying the printed version of the photographs. From March 1 to March 31, 2008, contributors submitted a total of 389 images that responded to the topic *Changing Faces of Brooklyn*. The photographs were evaluated by the general public (April 1–May 23, 2008) through an online evaluation tool, which offered the possibility to judge the images’ aesthetics, and their relevance in regard to the exhibition’s theme. In total, 3,344 evaluators gave 410,089 feedbacks on the photos. By the end of the process, the top 20% of the pictures were physically displayed during the exhibition that ran from June 27 to August 10, 2008. The project won on the category Best Exhibition of the *Museums and the Web 2009 - Best of the Web Context* [143].

The exhibition called *50/50: Audience and Experts Curate the Paper Collection* [1] organized by the Walker Art Center⁹ (Minneapolis, USA) is another example of crowdsourcing as an strategy for experimenting with a participatory curatorial process. During six weeks (August 1, 2010 - September 11, 2010) around 1900 individuals provided the institution with 231,719 positive or negative feedbacks on 183 images from the museum’s paper collection. The audience used the museum’s website as well as on-site interactive kiosks to register their vote. From the initial set of 183 images, some 100 artworks chosen by the public were displayed side-by-side with other 100 pictures

---

⁸ [https://www.brooklynmuseum.org/](https://www.brooklynmuseum.org/)
⁹ [https://walkerart.org/](https://walkerart.org/)
chosen by the curator Darsie Alexander. This curatorial partnership was an attempt to rethink the relationship between the Walker Art Center and its audience. As Alexander points out, “people always have opinions. This exhibition gives those opinions a little bit of agency so that people feel they actually have a voice” [105]. Alexander, however, seems to have a rather unambitious view on her co-curated exhibition organized within an institution with almost 100 years of existence and an attendance rate of 700,000 visitors per year. As a relevant public space, by researching, interpreting, and exhibiting its collection, the institution exerts a social, political and cultural influence that is capable of gathering attention to certain topics, defining cultural identities, and disputing conflicting narratives [107]. By taking part in such co-curatorial processes, the public starts to influence important cultural processes. Active interpretation, therefore, is also power [163]. Ng et al. argument that “museums are inherently political as every decision made is based on a specific point of view or framework, and they cannot claim to be neutral spaces given their origins in colonialism and imperialism” [163].

A bolder initiative was carried out by The Wing Luke Asian Museum10 (Seattle, USA), a Smithsonian Institution affiliate that dedicates itself to issues related to culture, art and history of Asian Pacific Americans. The museum launched a well-defined and permanent procedure for co-creating exhibitions. The museum decided to radically open itself to the community, and establish open calls for exhibition proposals. Not only the proposals come from the community, but also they are evaluated and selected by a partnership between community advisors and staff members. The main requirement for the proposals to be accepted is that they should be “based on topic, significance, and relevance to the museum’s mission.” [209, p.266]. Once the proposal is approved, a participatory six-stage process is carried out by an Exhibit Team that includes (1) a Core Community Committee composed of ten to fifteen community members, who are the primary decision-makers and responsible for the main tasks involved in the preparation of the exhibition, such as themes, content, form of the exhibit, etc [180]; (2) the Museum Staff, who facilitate work by advising, administrating, and managing the community [180]; and (3) Participating Community Members, who are informally engaged with the project, but nevertheless contribute in the research process, loaning artifacts, outreaching other people in the

10 http://www.wingluke.org/
community, serving as volunteers, and so on. Regarding the stages of the participatory process, they include [180]:

- **Initial Outreach** - The museum staff and community members learn about each other and each other’s dynamics. Knowledge about exhibit design and the museum’s resources are also shared with the community at this stage.

- **Exhibit Development** - A series of meetings to define the vision, main messages, themes, content and form of the exhibit. In addition, all the necessary materials for Exhibit Design are gathered in this phase. The community committee has the priority regarding the decisions made during the meetings.

- **Research and Gathering** - Museum staff and community engage in gathering material, such as photographs and documents, and conducting research in libraries, online databases, historic societies, etc. Interviews and their analysis are also carried out in this phase.

- **Exhibit Design** - Based on the decisions of the community committee, the actual design of the exhibition space is executed. Specific artifacts are selected to create a narrative with objects.

- **Exhibit Opening and Follow-up** - Museum members, the community, and the main stream media are invited to the Exhibit Opening. While the exhibition runs, public and educational programs are carried out. By the end of the exhibition, documentation is created and archived for future study.

### 3.1.3 Interpretation as Digital Fabrication

Long-term transformations brought by technology in the Cultural Sector has consolidated new forms of curatorship and interpretation, namely co-curatorship and co-interpretation. While the approach for creating cultural narratives within the traditional curatorial practice had been based on steadiness, centralized expertise, and hierarchy; by the end of the 2000s, it has become ever-changing, distributed, and horizontal (see [189], [209]). As demonstrated in 3.1.2, technology has produced the material conditions that drove a change of mindset and provided platforms on top of which cultural discourses started to be produced by engaging multiple voices and diverse opinions. In addition, since Cultural Institutions have begun to incorporate Makerspaces (also known as Hackerspaces, Hacklabs, Fabrication Labs or Fablabs) and Hackathons to their repertoire, participation has also meant making. Indeed, one could argue that workshops within Cultural Institutions is not a new idea, and precede the Web. There
were in fact a few pioneering small-scale initiatives that resemble contemporary practices. As Briley Rasmussen explains [191], Victor D’Amico, the founding director of MoMA’s Department of Education from 1937 to 1969, conceptualized several projects for MoMA’s young audiences that combined art together with teaching practices inspired by John Dewey’s child-centered pedagogy - learning by doing, [66] and Jean Piaget’s theories of cognitive development [185]. A report [42] produced by MoMA and published in 1960 describes the project called Children’s Art Carnival that consisted of creating a space in the facilities of the museum where children from 4 to 12 years old were provided with inspirations and art-making supplies. Besides an inspirational space, in which children listened to music, played with modern-art-themed puzzles, and interacted with sculptures, the museum offered also the Studio-Workshop that were children were given with collages and construction materials. In this space:

children worked independently and were assisted by an educator only when they didn’t know how to get started or when they needed help. It was a space designed to foster curiosity, discovery, experimentation, personal expression, and creativity just for children. [191].

Although D’Amico’s projects were limited in scale, reach, and impact11; and a vision of one individual rather than a mutual and broader understanding of participation in the Cultural Heritage Sector, they were the precursors of a more contemporary view of fabrication as method for the interpretation of heritage. The institution was seen as a platform of mediation, which provided assistance instead of instruction. In addition, the children were given construction objects that were used to mediate their interpretation of art, in an open-ended way. D’Amico transformed interpretation into a construction process, and provided objects-to-interpret-with12. With the introduction and popularization of personal computers, this pioneering approach that combined education, creativity, and fabrication as the means for learning gained a consistent theoretical foundation in the 1980s with the work of Seymour Papert, who is

---

11 Since they are confined to the facilities of the museum, therefore not reaching the undefined network of people [226] that characterizes crowdsourcing projects, and, finally, having no impact in the curatorial process of MoMA’s exhibitions.

12 Objects-to-interpret-with is used here as a counterpart of objects-to-think-with, which is a term used in the context of the Constructionist Learning Theory [182] as an epistemological framework where fabricating is both, constructing understanding and constructing artifacts.
considered “the father of the maker movement” [138, p.17], because of his theory of Constructionism. The theory advocates that learning in its most powerful form occurs “by constructing knowledge through the act of making something sharable” [138, p.21].

3.1.3.1 Digital Interpretation

In his groundbreaking book Mindstorms: Children, Computers, and Powerful Ideas [182], Papert reasoned about the possibility of the computer as machines with an enormous potential to “enhance thinking and change patterns of access to knowledge” [182, p.3]. Although Papert used Piaget’s theory of constructivism as a foundation for constructionism13, he stressed the influence of cultural ecosystems and tools for providing scaffolding to enable formal thinking, therefore pushing constructionism towards a Vygotskian14 understanding of cognitive development, which emphasizes culture and social interaction instead. Therefore, for Papert, formal thinking is developed not only because children reach a certain phase in their cognitive development, but also because the culture the children live in exposes them to physical and abstract artifacts that, through experience and social interaction, enable formal concepts to be explored concretely15. According to Papert, Piaget’s theory favors “the kind of learning that happens without deliberate teaching” [182, p.31]. Therefore, Piaget does not propose a curriculum to foster knowledge acquisition. By emphasizing nurture instead of nature, Papert’s ideas are empowering, because knowledge becomes available to anyone if the proper means are given, at earlier or later stages. In this sense, Papert saw especially computation as a powerful enabler, because computation materializes abstract concepts. Thus, Papert suggested that computation could elevate mental processes in more essential and conceptual ways. A great part of Papert’s work went in the direction of implementing ways to create digital artifacts to foster knowledge acquisition. Especial effort went initially into the development of the LOGO Programming Language [182], which was a branch of LISP [151], but optimized to teach math concepts. LOGO

---

13 With strong convergence points between the two theories, such as the idea that new knowledge is built on previous one, and the lack of separation between the learning process and what is being learned. (see [182, p.158])

14 Lev Vygotsky’s Cultural-Historical Theory focus on the role of culture and social interaction as instrumental in the cognitive development. (see [198])

15 With computers, “knowledge that was accessible only through formal processes can now be approached concretely,” [182, p.21]
was the precursor of Visual Programming Languages\textsuperscript{16} (VPLs), such as Scratch \textsuperscript{193}. The turtle (see blue component on Fig. \textsuperscript{17}) was used as a virtual avatar that entangled with the body of the learner and therefore provided shared concepts, such as front, back, left, and right. In addition, it served as a cursor capable of translating typed commands into shapes and visual transformations to be displayed on a screen, among additional possibilities. As Papert described it, "The Turtle serves no other purpose than of being good to program and good to think with." \textsuperscript{182}, p.11

Figure 10: From the module Images of Recursion - Forever Programs

From the beginning, the goal of Programming Languages was to provide helpful representations that would both free professionals from having to deal with binary procedures\textsuperscript{18}, and support individuals to construct models and tackle questions that concerned their own fields of activity. In this sense, there is nothing special to LOGO and Scratch besides being designed to facilitate the understanding of basic programming concepts by "mirroring" a few particular kinds of thought processes, and, at the same time, enabling certain kinds of expression by executing an idea computationally\textsuperscript{19}. In 1965, the development of the Simula Programming Language \textsuperscript{58} configured already a big step taken in the direction of establishing

\textsuperscript{16} According to Noone et al., "a \textit{VPL} is any programming language where users are able to manipulate the underlying code in some graphical fashion rather than the traditional text-based approach." \textsuperscript{165}

\textsuperscript{17} Picture source: https://www.youtube.com/watch?v=1jLNMi5mGqw.

\textsuperscript{18} As in the earlier computers. See also the Altair in chapter 4.

\textsuperscript{19} If, of course, this idea can be somehow computed.
a more human-like symbolic representation layer on top of binary instructions. It provided advanced concepts such as classes [58]. Efforts like Simula were important because they served to hide the complexity of 0s and 1s, and, at the same time, enabled high-level symbolic manipulations. Simula was of the main inspirations for Alan Kay and his team at Xerox PARC towards the goal of consolidating a model that is currently known as Object-Oriented Programming (OOP), which was interestingly inspired by biology. Alan Kay “thought of objects being like biological cells, only able to communicate with messages” [215]. By bringing Programming Languages close to the way the individual understands certain aspects of the world, Programming Languages became also a cognitive tool to think about the world, however in faster and greater ways. In order to understand and emphasize the way Programming Languages, as information-based tool-to-think-with, work in advancing human cognitive capabilities, let us put into perspective Natural Language, which is their primal precursor:

So, say I need to multiply 2,631 by 734. In primary school, children learn a simple procedure for doing this, which works by stepping right to left through the digits of 734, taking each digit and multiplying it with each digit of 2,631, noting the places where they might have to ‘carry’ a number across multiplications, and then adding the results. Moreover, the procedure also has a spatial component, a way of laying out the digits so that numbers will be counted to the right powers (1s, 10s, 100s, and so on). Easy. But now consider being presented with the numbers as Roman numerals and being asked to multiply MMDCXXXI by DCCXXXIV. The numbers are the same, but the representation is quite different, and no rote procedure enables me to transform the problem posed in these terms into a series of operations on individual numerals. Unlike the Indo-Arabic numerals we are used to, the Roman system is not a positional notation; similarly, the appearance of a particular glyph (e.g., X) does not definitively convey a particular number, because other glyphs that appear before it or after it will change its meaning. [73, p.7]

Spoken language itself, as cognitive technology [86], and more specifically formal language, such as the alphabet, notation, nomenclature, as tools of thought [117], are capable of affording certain cognitive capabilities, such as increasing acceleration of information
processing and amplification of human memory\textsuperscript{20}. The quotation above, by Dourish, elucidates yet another influence of the medium, in this case \textit{notation}, in facilitating/impeding or enabling/disabling certain cognitive processes to occur. The Indo-Arabic Number System scales, while the Roman does not. The former affords therefore the brain to think in much greater quantities. Under Marshall McLuhan's perspective, “language does for intelligence what the wheel does for the feet and the body. It enables them to move from thing to thing with greater ease and speed…” \textsuperscript{[152]}. McLuhan saw all technology as extensions of the human organism, but made a special distinction when talking about electric technology. He understood electric technology/media as extensions of our cognitive functions. The computer, as a machine of symbolic computation, a semiotic machine, is perceived here therefore as an extension of the mind, capable of amplifying its capabilities by providing a special language that scaffolds structured thoughts.

As for Programming Languages, their different emphasis on format, stylistic characteristics, and features are suitable for modeling different phenomena and levels of complexity. As Thomas Green and Marian Petre point out \textsuperscript{[98]}, “programming requires mapping between a problem world and a program world. The closer the programming world is to the problem world, the easier the problem-solving ought to be” \textsuperscript{[98]}. For example, \textit{Procedural Programming}\textsuperscript{21} and \textit{OOP} paradigms expect the programmer to think in different ways, either emphasizing a problem to be tackled as a sequence of statements, or as a...

\textsuperscript{20} The idea encapsulated in Linguistic Determinism by Benjamin Lee Whorf, which tell us that the language one speaks, determines the concept that one understands, has already been disproved by many contemporary linguists (see \textsuperscript{[187]}, \textsuperscript{[192]}, \textsuperscript{[86]}). The main criticism is focused on the belief by Whorf that structured thought is made possible by language, and therefore we think because we speak. The often-used example to justify Linguistic Determinism is the debate around tongues lacking e.g. a noun for the English word “time”. For Whorf, the fact that such languages have no word for the concept of time is a sign that the speakers of these languages are unable to understand future or past tenses. By breaking the notion that thought and language are the same thing, Steven Pinker \textsuperscript{[187]} makes a strong distinction between what one thinks internally, and how the message is conveyed by language. He gives examples e.g. of deaf children who although lacking language are able to create one for themselves. Although Pinker’s arguments indicate the far-reaching capabilities of the mind by not having its creative, computational, and representational competences restrained by language, new research on the topic suggests that language has some influence in our thoughts by providing a framework for our attention and memory (see \textsuperscript{[192]}, \textsuperscript{[86]}).

\textsuperscript{21} It is defined as a “programming language in which the user states a specific set of instructions that the computer must perform in a given sequence” \textsuperscript{[221]}. 
network of interdependent objects. In the same way, Programming Languages based on the Imperative Programming Paradigm, as e.g. Java, requires the programmer to think about how something is done, while Programming Languages based on the Declarative Programming Paradigm, as e.g. SQL, requires the programmer to describe what he or she wants done (see [186]).

Figure 11: Scratch Visual Programming Environment

In addition, although presenting problems concerning scalability, VPLs lower the entry-level of difficulty in regard to the skills individuals need to acquire in order to be introduced to programming [165]. Mitchel Resnick et al. [193], when talking about his inspirations for developing the blocks-based grammar of Scratch (see Fig. 12), described how new ideas, together with structures and stories, come into existence when children combine different LEGO bricks together. Therefore, Resnick et al. [186] saw in the brick a powerful metaphor for enabling programming at young ages. As the authors explain:

We wanted the process of programming in Scratch to have a similar feel. The Scratch grammar is based on a collection of graphical 'programming blocks' children snap together to create programs. As with Lego bricks, connectors on the blocks suggest how they should be put together. Children can start by simply tinkering with the bricks, snapping them together in different sequences and combinations to see what happens. There is none of the obscure syntax or punctuation of traditional programming languages. The floor is low and the experience playful. [193]
It is essential to understand programming, one of the building blocks of computation, in order to comprehend Heritage Interpretation under an updated perspective that encompasses not only platforms of participation, but also Digital Fabrication. As McLuhan explains, our relationship with technology does not flow in one direction only. By shaping and using tools, tools themselves also shape us. They impose their own logic, processes, and paradigms on the individual and the society. We adapt language as much as we are adapted by it. In this sense, it is not strange to think that a story about heritage executed through e.g. Scratch would be carried out as a systematic arrangement of programmable blocks. This interpretation would necessarily need to deal with basic algorithmic thinking capacities, such as abstractions and pattern generalizations; systematic processing of information; symbol systems and representations; algorithmic notions of flow of control; structured problem decomposition; iterative, recursive, and parallel thinking; conditional logic; efficiency and performance constraints; and debugging and systematic error detection [99].

3.1.3.2 Digital Interpretive Artifacts

As discussed in section 3.1.3.1, programming languages are cognitive tools for constructing models in the computer, and realizing computable models in the world. These cognitive tools, as well as the outcomes produced with and by it (computer programs), redefine the human condition by enhancing and creating new capabilities. For example, thanks to Machine Learning algorithms and visualization techniques, we are now able to classify large amounts
of data and understand the underlying structures of millions of objects. The Google’s project Curator Table\(^{22}\) is an example of digital technology amplifying the curator’s capability to comprehend the hidden patterns of millions of artworks that are somehow connected among them. Amit Sood, director of Google’s Cultural Institute and Art Project, explained in a TED Talk [212] the idea for conceptualizing this application. According to Sood:

\[
\text{Beyond the pretty picture, beyond the nice visualization, what is the purpose, how is this useful? This next idea comes from discussions with curators that we’ve been having at museums, who, by the way, I’ve fallen in love with, because they dedicate their whole life to try to tell these stories. One of the curators told me, “Amit, what would it be like if you could create a virtual curator’s table where all these six million objects are displayed in a way for us to look at the connections between them?” You can spend a lot of time, trust me, looking at different objects and understanding where they come from. It’s a crazy Matrix experience. [212]}
\]

Computer programs have become then powerful instruments to tell stories (see also Chapter 5). Notwithstanding the enhancements and amplifications, these instruments are constrained not only by computation, but also by the materiality of the platforms that instantiate them (see Section 4.1). And, because of that, they present unique characteristics in regard to their functionality, usability, aesthetics, content, and value that come into play during their development and usage. Clarisse de Souza [213], when proposing a theory of Semiotics Engineering, takes a special look at software as intellectual artifacts that carry messages from designers to users, and act therefore as “the designer’s deputy” [213]. By seeing programs as artifacts of communication, De Souza emphasizes the artificial aspect of computer programs as simulations embedding algorithmic narratives “created by humans” [213], but as part of a reduced communication process, “present in all Computer-Human Interaction” [162], as discussed by Nake. De Souza not only explains that digital artifacts communicate the designer’s messages and intentions by means of computable sign systems, which are “constrained by formal computational factors” [213], but also offers a hybrid perspective on computer systems by merging engineering, as an applied science,

\[22\text{https://artsexperiments.withgoogle.com/curatortable/}\]
to the design practice, as a creative approach to problem-solving. Although De Souza’s reasoning is directed at User Interfaces (UIs), the implications that come from the acknowledgment that computer systems are constrained must understand UIs under a broader context. In this sense, Alan Kay expands the notion of UI when he puts it under a perspective of service. According to Kay, the “‘User interface’ is not just about kinds of inputs, outputs and screen organizations, but about the notion of service — that is, the desired content of the interactions, and the larger goals of the interactions.” [120]

The communication processes involved in computer applications are therefore complex. According to Kay, the UI embeds larger goals that are communicated to the user during interaction. In addition, digital ecosystems, in which digital artifacts are inserted in, must therefore be understood in regard to parallel forces of amplification and reduction of potentialities. Concerning especially the latter, the social media platforms that provoked disruption and, at the same time, opened new participatory possibilities in the Cultural Heritage Sector are now the focus of increasing criticism, because of the way they are designed, how they operate, and what they provoke. Current social network models, such as the one implemented by Facebook, Twitter, Instagram, and so on, are engineered under a habit-forming product design approach that monetizes on the attention/time users spend online. Nir Eyal and Ryan Hoover describe the Hooked Model, which uses insights from the Behavior Model for Persuasive Design developed by the behavioral psychologist Brian J. Fogg, to explain how social media businesses exploit e.g. cravings or negative emotions to drive individuals to engage online to satisfy a desire or compensate discomfort. According to Eyal et al. “feelings of boredom, loneliness, frustration, confusion, and indecisiveness often instigate a slight pain or irritation and prompt an almost instantaneous and often mindless action to quell the negative sensation.” [82, p.68]. Based on the analysis of users behaviors and emotions, social media companies design effortless strategies, tiny actions, such as checking Facebook’s News Feed or getting likes on a post, act as rewards that release dopamine.

---

23 See Section 4.2.
24 https://www.facebook.com/
25 https://twitter.com/
26 https://www.instagram.com/
27 A recent news piece on The Guardian revealed that leaked documents aimed at advertisers describe that Facebook is now able to identify when teenagers feel insecure, worthless and need a confidence boost [136].
in the brain (see [82, p.16-17]). In addition to that, an experiment done in 2012 with Facebook users [127] tested the assumption that positive and negative emotions could be contagious and influence the behavior of individuals online. Kramer et al. [127] showed that the emotional tone of posts could be altered by either increasing or decreasing the quantity of negative and positive updates in the users’ News Feed28. Negative posts led to users producing negative messages, the contrary was observed by increasing the quantity of positive posts. The format of the News Feed is dictated by subtle and powerful computer algorithms, invisible for the eye, but effective for programming the mind.

If understanding, appreciation, and conservation are the cornerstones of Heritage Interpretation [103], and Heritage Interpretation is “an educational activity” [227, p.8], such mindless and narrow strategies that try to influence behaviors and determine outcomes should not define the interactions (services) provided by Digital Interpretive Artifacts (DIAs), since these artifacts need to drive the construction of understandings and provoke reflection. Ideally, such reflection must come from a conversation [206] between the individual and the possibilities of digital technology, dictated mainly by the format of its models and interactions. This conversation comes from a reflection-in-action, which is, according to Donald Schön, “closely tied to the experience of surprise” [206]. Moreover, a DIA must surprise and provoke us [227, p.32], and most importantly be a part of a more complex interpretive process that does not dominate attention, but shares it. For example, in museums, the role of DIAs is to provoke curiosity to the collections. In libraries, they should instigate the interest of reading great books. In cities, they should draw attention to important landmarks. To sum up, DIAs can be defined as computer applications that are used to tell stories about heritage and culture through the (re-)interpretation of cultural assets. DIAs should aim at promoting understanding of heritage through provocation and surprise, instead of instruction. Finally, because of their origin in Cultural Institutions’ contexts, these applications are themselves cultural artifacts that express perspectives and narratives that reinforce or contradict cultural, historical, and/or social discourses.

28 The experiment was done with 689,002 users.
3.2 Constructionist Heritage Interpretation

Constructionist Heritage Interpretation can be defined as the process of co-constructing DIAs. Such short definition must however account for the discussed topics so far. To begin with, one must have an important consideration in mind: Heritage Interpretation should be understood as a special kind of informal educational activity [227, p.8], which has been gradually diversified over the years especially by the increasing digitization of Cultural Institutions (see Sections 3.1.1 and 3.1.2). Through digital platforms of participation, Heritage Interpretation gains even more complexity when it is employed through Digital Fabrication. Storytelling is operationalized by manipulating and constructing representations through information-based tools and externalized as shared digital objects. As in constructionism, the process of producing understandings about the world is done through the development of mental models. In the case of Constructionist Heritage Interpretation, these mental models may relate to the recontextualization of Digital Collections (see Chapter 5) through algorithmic manipulation (see Section 3.1.3). The computational artifact, as an externalization of a mental model about heritage, is the shared object that will instrumentalize and mediate the conversations, negotiations, decisions, and understandings not only in between the individual and the digital material (amplifications and constraints - see Section 3.1.3.2), but also among the different actors involved in the process, namely the institution, other members of the community, and the audience (multiple voices and opinions - see Section 3.1.1). Finally, the challenge and the broader goal in Constructionist Heritage Interpretation must be the same as in Heritage Interpretation, which is to create such an engaging and enriching process so that conservation of heritage is achieved, not before driving understanding and generating appreciation in the participants of this endeavor. Further discussions about Constructionist Heritage Interpretation and its applicability to Hackathons as a method for the interpretation of heritage are presented in Chapter 7.
"Are you building your own computer?", asked Fred Moore in the first Homebrew Computer Club’s action call from March 1975 (see Fig. 13). Located in Menlo Park (California, USA), the Homebrew was, some would say, the most prominent computer club among all others that started to pop up around mid 1970’s (see [159], [137]). According to the historian Jonathan Gottfried [145], these computer clubs, and specially the Homebrew, helped to shape what later on was known as Hackathons. During these gatherings, like-minded individuals met to swap ideas, to build their own computers, and to create some kind of black-magic boxes.

On the bottom of the action call a note handwritten by Moore says: “Hope you can come. There will be other Altair builders there”. The main reason behind the creation of the Homebrew Computer Club was the release of the first microcomputer, the MITS’ Altair 8800 (see Fig. 14), released in also January 1975. Students, engineers, machinists, physicians, hobbyists, among others who had little or no access to computers outside the club, saw in it the opportunity to play with

---

1 MITS stands for Instrumentation Telemetry Systems, a company from Albuquerque, New Mexico, USA. [38, p.107]

---

The Homebrew Computer Club provided the necessary support for hackers to assemble and expand the Altair.
these machines, to expand and enhance them by creating new parts and programs, and to find like minded people to think together about hardware and software-related topics.

Not only the meetings, but especially the technology, the Altair, provided a suitable platform\textsuperscript{2} and an ecosystem\textsuperscript{3} capable of driving creativity among the members of the club, whose innovations strongly impacted the directions of the Computer Industry and the society as a whole. There are a few reasons why the Altair was crucial as a platform capable of driving innovation. The Altair was:

- **Extensible** - The design of the Altair was prepared to be extended mainly thanks to its \textit{S-100 bus board}\textsuperscript{4}. The S-100 bus board came with empty sockets that allowed different cards to be attached to it, besides the one socket containing the microprocessor. The standardized bus meant that individuals and companies could make cards for the Altair in order to extend its capabilities. The possibility of expansion drove the early hobbyists to think about ways to fill up the empty sockets,

---

\textsuperscript{2} See Section \textit{4.1}.
\textsuperscript{3} See Section \textit{4.2}.
\textsuperscript{4} An expansion bus is a communication system that transfers data between components, such as sound cards, video cards, network cards, microprocessor, memory, etc.
and later on various companies were founded around producing computer peripherals, such as Cromemco and Processor Technology Corporation [81].

- **Open-ended** - The Altair was not sold as a final product to be used, but as part of a computer-kit that needed to be assembled [4]. Therefore, the term user, usually applied to describe individuals who interact with Personal Computers today, was not an adequate term to describe the earlier adopters of the Altair. They were makers, and because of that, it was natural the formation of an ecosystem around the technology. The clubs provided mutual support in the understanding and the further development of the technology, which no one knew what exactly to do with. The Altair was an open-ended project that provided a foundation upon which new extensions could be integrated to [137, p.206].

- **Accessible** - Based on the microprocessor Intel 8080, the Altair turned the ownership of a computer into reality for individuals during times when computers were not affordable machines. In addition, the Altair was compact, fitting on the desktop, and its processing power enabled the creation of interesting applications\(^5\). In addition to that, a programming language called Basic and developed by Bill Gates also developed in order to facilitate the creation of programs [137, p.227].

- **Open** - The Altair was shipped with not only a user manual, but also its detail schematics. The access to the schematics [3] enabled a deeper understanding of the architecture of the computer, which facilitated the design of peripheral devices that needed to communicate with it. Besides the creation of hardware extensions to the Altair, the open schematics also lead to variations of the architecture. The Apple I and the Sol-20 computers, for example, were created based on the Altair (see [159], [81]).

In the following sections of this chapter, it will be discussed how these early empowering concepts, such as extensibility, open-endedness, accessibility, and openness have been constantly present though out the computer history, despite attempts by market forces to push for more control and restriction. These concepts have instead proven their strength in driving creativity and innovation not only in more traditional areas of the Computer Science Industry, but also in the Cultural Heritage Sector\(^6\).

---

5 Other microprocessors, such as the Intel 8008, were too slow to be useful.
6 See Section 4.2.1.
4.1 Platforms: Extensibility and Open-Endedness

Extensibility and Open-endedness, in the sense of having no well-defined predetermined limits or boundaries to what one can create with technology, are two of the reasons that explain why the Altair 8800 could be a foundation and a platform for new and revolutionary ideas that transformed the applicability of the computer. The architecture of the Altair, especially concerning its standardized and extensible S-100 bus, not only challenged, but also provoked and inspired hackers to find ways to fill up its empty sockets. Bob Marsh, one of the first members of the Homebrew Computer Club soon realized that the modular, and therefore extensible architecture of the Altair represented a platform that would open up “the beginning of a new era” [137, p.207]. About Marsh, Levy writes:

The flashing LEDs on the Altair were exciting, but he knew that — hackers being hackers — there would be a demand for all sorts of peripheral devices, devices this MITS company obviously could not provide. But someone would have to, because the Altair was the basis for a fantastic system to build new systems, new worlds. [137, p.206]

In platforms, extensibility is achieved through standardization and specification [25]. A stable core [10], with a well-defined set of interfaces and functionalities allow for derivate products to be “efficiently developed and produced” [156]. The set of components that form a platform can be understood as “technological building blocks” [210] that can be reused, rearranged, and innovated upon. In this sense, Baldwin et al. [10], point out that platform definitions identify the reuse or sharing of common elements as key characteristics, and that all platforms are “modularizations of complex systems in which certain components (the platform itself) remain stable, while others (the complements) are encouraged to vary in cross-section or over time” [10]. In the case of the Altair, it is clear the advantages brought by its S-100 bus; a stable core which other peripherals could be attached to. Another important characteristic of the Altair lays in the fact that this computer could only be bought as a kit (see Fig. 14). The standardized parts, or the building blocks, invited hackers to understand how each component had to be arranged together. This knowledge, in addition
to the open schematics [3], enabled hackers to rearrange and reuse the kit and third-party’s components to create new peripherals, which could be interfaced with the S-100 bus. The modular architecture of the Altair afforded extensibility and innovation. The open-ended possibilities provided by the combination of core components in new and innovative ways, led to extensions that the producers of the Altair could never have foreseen. The Altair and the Homebrew Computer Club hugely influenced and shaped individuals and companies that would become the leaders in the computer industries, such as Apple [7] and Microsoft [8]. Based on the Altair, Steve Wozniak presented a prototype of the Apple I to the club [216], Lee Felsenstein demonstrated his alphanumeric video display adapter [133], Harry Garland created innovative image sensor and attached a camera to the Altair [64], creating the very first all digital image capturing system, and Bill Gates started Microsoft developing a programming language to run on it called BASIC [22].

4.1.1 Modularity

A complex system can be managed by dividing it up into smaller pieces and looking at each one separately. When the complexity of one of the elements crosses a certain threshold, that complexity can be isolated by defining a separate abstraction that has a simple interface. The abstraction hides the complexity of the element, the interface indicates how the element interacts with the larger system. [9, p.65]

The excerpt above comprises the main characteristics of modularity, which is a key platform principle, as described by Baldwin and Clark [9]. Concepts such as abstraction, information hiding, and interface were used not only to handle hardware, but also software complexity. Operating Systems (OSs) and bundles [9], for example, are modular approaches that enabled computers to broaden their boundaries. The very early computers presented no separation between software and hardware. Both were closely tied together [39]. The full hardware specification to boot the computer and run the application had to be programmed as a monolithic and procedural block of instructions. Besides being inefficient, this presented a series of compatibility

---

7 https://www.apple.com/
8 https://www.microsoft.com/de-de/
9 Synonym of Apps.
problems always when a new version of a computer was released. The increasing hardware complexity drove engineers to have a modular approach to these systems.

In the early 60s, before the Altair, IBM created the first modular computer, the System/360 [92] (see Fig. 15). This model was a landmark in the computer industry, because it “offered the possibility for other firms to enter and compete in the computer marketplace, on the basis of providing modules that would ‘slot in’ or ‘plug in’ IBM System/360 architecture.” [92] Soon, the industry realized that the modularity principle could also be applied to separate the OS from the hardware, and the applications from the OS. The modular approach to computer systems brought many advantages to the field especially regarding cost, manageability, and evolution of complex systems. According to Baldwin et. al.:

> By promoting the reuse of core components, such partitioning can reduce the cost of variety and innovation at the system level. The whole system does not have to be invented or rebuilt from scratch to generate a new product, accommodate heterogeneous tastes, or respond to change in the external environment. In this fashion, the platform system as a whole becomes evolvable: it can adapt at low cost without losing its identity or continuity design. [10]

---

The very notion of computer languages also represented another decoupling and modular fragmentation between hardware and software, given that in the 1940s engineers "programmed" by manipulating wires and switches\textsuperscript{11}. In addition, the goal of programming languages was also to provide a helpful vocabulary that would both free professionals from having to deal with binary procedures, and support individuals to construct models that could tackle questions concerning their own fields of interest (see Section 3.1.3.1). Computer languages themselves have also become modular. Object-Oriented Programming, for instance, brought software modularity to a higher level, because of concepts such as encapsulation, polymorphism, inheritance, and so on. Code could be broken down into small and independent chunks, and integrated into a complex program with ease. Another example of modularity that concerns hardware and software is the architecture of the Internet. Its reliability and efficiency comes on one hand from a decentralized and distributed network of devices, and on the other hand from protocols capable of coordinating numerous and independent small packages that form the documents, media, and services transmitted on the network; a method called \textit{packet switching}\textsuperscript{12}.

Modularity, as a principle for open-ended innovation, was not only what made the Altair a successful platform, but also many other projects that came afterwards. Hackathons, as events for fast-prototyping, depend on modular technology (APIs\textsuperscript{13}, SDKs\textsuperscript{14}, Data Standards\textsuperscript{15}, etc) to create new ideas and construct applications by combining different building blocks together \cite{210}. Here it is possible to draw a comparison with the also modular LEGO system\textsuperscript{16}, where different sizes, shapes, and colors of its bricks are assembled together to form a variety of toys. Lego provided the modular metaphor used by Resnick to create the Programming Language called Scratch \cite{193}, which reduced complexity of programming in a way so that children were able to be introduced to programming principles and create their own applications (see Section 3.1.3.1).

\textsuperscript{11} See front panel of e.g. System/360 Model 91 on Fig. 15.
\textsuperscript{12} In telecommunications and computing, packet switching is "a mode of data transmission in which a message is broken into a number of parts or packets which are sent independently, over whatever route is optimum for each packet, and reassembled at the destination." \cite{245}
\textsuperscript{13} API stands for Application Programming Interface.
\textsuperscript{14} SDK stands for Software Development Kit.
\textsuperscript{15} See discussion in Chapter 5.
\textsuperscript{16} \small https://www.lego.com/en-au/aboutus

\footnotesize LEGO is a modular system where the bricks can be combined together thanks to standardized interfaces.
Modularity therefore provides an effective strategy to deal with complexity, and enable systems to provide fast adaptation and rapid evolvability. These characteristics allow the great variation of prototypes to be build based on only one or a combination of several technologies. In addition, different system parts have the possibility to develop independently from one another. Because of that, specialized businesses and disciplines might evolve based on modules. As Simon pointed out:

> Technological building blocks (that can be technologies, products, or services) that act as a foundation on top of which an array of firms, organized in a set of interdependent firms (sometimes called an industry ‘ecosystem’), develop a set of inter-related products, technologies and services. [210]

Finally, modularity reduces the entry level upon which individuals and companies innovate, because one does not need to reinvent the whole system from the ground up, but is able to use core components as building blocks and foundations to extend a system in combinatorial ways. Modularity is a very important reason for the appearance of rich ecosystems that have evolved towards greater diversity and decentralization, in that it “creates at least as many options as there are modules” [9, p.236], and it “shifts design options from the ‘center’ to the ‘periphery’ of the system. Hence, modularity decentralizes the options inherent in a design” [9, p.237]. In the next section, I will explore the concept of ecosystem and how it related not only to movements within the Computer Science field, but also to the Cultural Heritage Sector.

### 4.2 Ecosystems: Accessibility and Openness

The definition of digital ecosystems bases itself on biological ecosystems, which are “loosely coupled, domain clustered environment inhabited by species, each proactive and responsive regarding its own benefit while conserving the environment” [26]. Within ecosystems, all living and non-living elements are intertwined, and have the potential to affect other species and the environment as a whole. According to Tim O’Reilly, “each of us depends on thousands, if not millions, of other organisms, each pursuing its own selfish goals, yet somehow weaving a
cooperative web that, for the most part, benefits all” [170]. O’Reilly used this analogy to try to explain, under his point of view, how the Open-Source Movement functioned. He continues by saying that:

*Each developer builds for his own use, and that of his friends, but also makes it easy for collateral benefits to accrue to others he or she doesn’t know. And the open source developer contributes even his failures back into the environment, to enrich the soil from which other innovations can grow* [170].

Digital ecosystems encompass algorithms, technologies, processes, people, and organizations that depend on, and cooperate and compete with each other. The Open-Source Ecosystem, the place of origin of Hackathons, is perhaps one of the most well known examples of a complex digital environment enabled by a collaborative network of developers who collectively push technological progress forward. Its development model is focused especially on both the premise of participants/users being considered as co-developers and free-license agreements, which overcome restrictions regarding copyright infringements, giving anyone permission to build on others work. In addition, the *cooperative web capable of benefiting all* is only possible because of a non-hierarchical networked production processes that are able to cope with highly complex and large amounts of data and code. Besides the motivations of the open-source community already discussed in the section 2.1.2, the distributed and collaborative configuration of the Open-Source Ecosystem raise questions primarily regarding17:

- **Complexity** - How to deal with geographically dispersed communities that do not obey to a centralized control?
- **Coordination** - How to coordinate contributions from different developers working parallel without a hierarchical structure of authority?

Some strategies, such as *Distributed Revision Control Applications*, provide answers to coordination of software development projects in highly complex decentralized and disperse environments, such as the Internet. The currently most popular version control application [233] called Git18 is an open-source project created initially to support another large-scale open-source project, the Linux Operating System (see [41]). In a nutshell, these applications record changes in the

---

17 As suggested by [237].
18 See [https://git-scm.com/](https://git-scm.com/)
software source code by enabling the developer to take snapshots along the development. The snapshot represents a saved state of the source code at a certain time. In the Git terminology, these snapshots are called commits, and a collection of commits is called a repository. A commit can be recalled later on if necessary. A previous working version of an application can be easily restored, if e.g. a later implementation provokes the code to fail compilation. In addition, these version control applications allow the developer to branch the source code. A branch represents a new line of development that deviates from the main one. Branching allows not only time-based versions of the code, but also parallel versions that can co-exist simultaneously and evolve independently. Branching gives great flexibility e.g. for implementing and testing new features before integrating them into the official code. In addition, branching facilitates the simultaneous development of an application by multiple programmers, since each programmer is able to create a parallel version of the application, which is extended in order to hold only one particular feature. This modularization of source code versions represents great benefit to the coordination of complexity during software development. By the end of the development cycle, all features created by different developers can be merged into the main branch, assembling therefore a unique master version.

Figure 16: Merging conflicting code versions with a Git GUI
While Git is a version control tool, GitHub\(^{19}\) is the online service that holds Git repositories. Current figures point to 26 million registered users and over 67 million projects [96]. With GitHub, a complex development ecosystem is created having the Git algorithm as a platform. This service not only provides all the functionalities of Git, such as version control, but also grants universal access to open-source code, as well as data and documents, to anyone with Internet connection. This has a beneficial impact on the quality of an open-source project, because the 26 millions registered users are also potential contributors. By making the code publicly available for others to access, it becomes a social artifact that benefits other developers, who eventually have similar needs, and the evolution of the code itself. Especially in complex projects, that requires interdisciplinary efforts, individuals with different competences are able to enhance and extend software in a way that only one individual is not capable of. This digital ecosystem has proven to be able to harness the collective intelligence of the crowd, and produce highly professional and reliable software\(^{20}\). Moreover, GitHub adds a valuable extra functionality to repositories, which is the possibility to *fork* them. Forking enable anyone with a GitHub account to create a copy of a repository. Therefore, developers are able to gain ownership over the copy\(^{21}\), and experiment with it without changing the original project. According to the GitHub website, “forks are used to either propose changes to someone else’s project or to use someone else’s project as a starting point for your own idea” [83].

4.2.1 Two Ecosystems: Cultural Heritage and Computer Science

Recently, two distinct ecosystems started overlapping. Not only Cultural Heritage Institutions have invited hackers to create new applications with their data (Hackathons), but also institutions themselves have begun experimenting with such technologies. Institutions are using the Git algorithm and the GitHub service to enhance the way they produce and distribute content to their

---

\(^{19}\) https://github.com/

\(^{20}\) In GitHub can be found projects such as Linux (https://github.com/torvalds/linux), BitCoin (https://github.com/bitcoin/bitcoin), d3 (https://github.com/d3/d3), and so on.

\(^{21}\) The possibilities of reuse depends on the License the original developer chooses for his or her project.
audiences. A large number of open-source projects directed to Cultural Heritage are available on GitHub. They range from datasets, to planning documents and Client APIs. Just to name a few, large museums such as the Metropolitan Museum of Art22 (New York, USA), the Tate Modern23 (London, UK), and the Museum für Kunst und Gewerbe24 (Hamburg, DE) have made available datasets with information about their collections available as GitHub repositories. The Andy Warhol Museum25 (Pittsburgh, USA) has even published its digital strategy as a Git repository in order for the files to be easily shared, revised, and reworked. According to the museum:

*We are using GitHub to draft and publish this strategy, which makes a catalog of previous versions accessible. In addition, the source files can be easily shared and repurposed within the cultural sector. We intend to revisit and update this strategy regularly as tactics and technologies evolve.* [83].

Besides datasets and documents shared through repositories, some institutions also invest in developing their own APIs26, Software Libraries/Packages27, Frameworks28, and SDKs29. The Cooper Hewitt, Smithsonian Design Museum30 (New York, USA) has developed a comprehensive set of open-source tools and published them online as public repositories on GitHub. This initiative was one of the results of an intensive three-year renovation during which the museum redefined its mission and role. By the end of 2014, a new digital strategy was launched and an in-house dedicated tech team was formed under the name Cooper Hewitt Labs. According to Pat Knapp, the first task of the lab was “to create a digital infrastructure upon which content and services could be created and shared with the public” [125]. This new digital strategy goes along with the mission of the

22 https://www.metmuseum.org/ - GitHub Repo: https://github.com/metmuseum/openaccess
26 Application Programming Interface - the interface of a library, more specifically a set of callable functions and methods in a library.
27 In modular programming, a Library a set of predefined common functions that can be reused by the programmer when writing software.
28 A group of Computer Libraries.
29 A Software Development Kit is a library or group of libraries that aid programmers in developing code that is made to depend on a particular system.
30 https://www.cooperhewitt.org/ - GitHub Repo: https://github.com/cooperhewitt
museum in “making its collections, knowledge, and resources as accessible and useful as possible” [175]. Among others, the set of open-source tools include e.g. an API so a developer can easily connect his or her application with the museum’s online services\textsuperscript{31}, the typeface of the museum’s brand identity\textsuperscript{32}, a little program to extract color palettes from images called RobotEyes\textsuperscript{33}, and the source code behind the Cooper Hewitt Pen\textsuperscript{34} (see Fig. 17), which is a device created to redesign the visitor experience by that let visitors of the museum to virtually collect artifacts and interact with exhibits by e.g. rendering 3D furniture or finding objects based on drawn patterns.

![Cooper Hewitt Pen](image)

Figure 17: 'Collecting' artifacts with the Cooper Hewitt Pen

A very special characteristic unites the old Altair and the publicly open GitHub repositories. They both have invited individuals to hack. The Altair (as a hardware kit), and GitHub (as a collection of open-source repositories) have disclosed how things were built and provided the necessary \textit{bricks to construct with}. Most importantly, these platforms have been embraced by the community because of their openness. Open projects are able to respond more efficiently to the community needs, because the evolution of these projects are under the direct influence/development of the same community which benefit from the hardware, code, and datasets online. In the next chapter, it will be discussed more in details about how standardization and modularity play out in concern to datasets, which is the most significant material provided by Cultural Heritage Institutions to communities of makers.

\textsuperscript{31} https://github.com/cooperhewitt/go-cooperhewitt-api
\textsuperscript{32} https://github.com/cooperhewitt/cooperhewitt-typeface
\textsuperscript{33} https://github.com/cooperhewitt/py-cooperhewitt-roboteyes-colors
\textsuperscript{34} https://github.com/cooperhewitt/ofxCooperHewittOscPen
Storytelling is considered by most memory institutions as one of the main traditional methods for interpreting heritage. Especially in museums, whose missions aim at not only preserving cultural artifacts, but also carrying on an educational agenda, storytelling is seen as “the real work of museums” [16], because “museums are all about stories!” [85]. In museums, storytelling happens through the objects of a collection, which are triggers and anchors for understanding relevant historical and cultural events [85]. In addition, stories can be used for making sense of the objects themselves, in terms of their purpose and meaning. Stories are therefore the medium that connects the collection and the audience. In fact, stories and their narratives have served human kind during millennia as a powerful medium, which is intrinsically natural to human beings, who “make sense of the world and themselves through narrative.” [16] Narratives provide a point of view that inspires emotional connection and internal dialogs. As Katherine Hayles [106] puts it:

In its emphasis on causality, agency and temporal progression, narrative provides not just specific explanations but frameworks for explanations that allow people to understand and predict how other people and the world around them will act and react. [106]

Although this chapter is not directed at discussing storytelling, but instead the role of data as building blocks of Digital Interpretive Artifacts1 (DIAs), addressing the issues involved in digital/data-driven narratives as a product of the manipulation and recontextualization of Digital Collections is essential to understand the characteristics and affordances of data as a creative material. In Hackathons, the interpretation of heritage happens through the intensive reflexive conversations2 with Digital Collections, which will influence design decisions and the stories that will be told. This Constructionist Heritage Interpretation3, as the production of digital/data-driven narratives

---

1 See Sections 3.1.3.2 and 7.1.1.1.
2 As spoken by Donald Schön [206].
3 See Sections 3.2 and 7.1.1.1.
that give collections meaning, is however important not only to Hackathons, but to any institution wanting to engage educationally with their audiences. As traditional storytelling has had to share its ground with new forms of digital/data-driven storytelling, great emphasis has been placed on interactivity, non-linearity, adaptive content, collaboration, and participation happening through digital media. These new possibilities brought by database-driven technologies in the Cultural Sector has led to deep transformations in e.g. curatorial processes. On the one hand, while principles and practices of the Web 2.0 have disrupted the Cultural Sector; one the other hand, “the rise of Web 2.0 has been associated with a series of new forms of data storage” [73]. Understanding data is therefore necessary to comprehend the transformations that have been rapidly taking place in the cultural sector.

One of the possible questions one could raise regards the possibilities of databases as platforms for storytelling. Lev Manovich [147], for example, confronts the format and characteristics of traditional narratives, as the ones present in novels and movies, to databases that do not organize its elements linearly, but instead “are collections of individual items, where every item has the same significance as any other” [147]. For Manovich, the database is therefore a unique genre of new media that is essentially non-linear. Especially when being employed on the Web to enable platforms of participation, databases provide great fluidity to content. The collection is ever-changing, since adding, deleting, and modifying elements can be performed with ease. For Manovich, the database as a unique genre is therefore the opposite of narratives. In this sense, Manovich compares the contradicting characteristics present in the linearity of stories as the “cause-and-effect trajectory of seemingly unordered items (events)” [147] to algorithms; and the concurrent description of actual set of unordered items (objects) of collections to databases. Hayles [106] however does not see narratives and databases as being defined by opposition and competition. The database opens not only “new opportunities for narratology but also for critical inquiry” [106, p.183]. In a “possibility space” [106], databases confer to narratives combinatorial capabilities “enabling events be interpreted in various ways” [106, p.183]. Databases have therefore a defining and strong impact in the narratives that are generated from them. The production of narratives from collections is indeed what

---

4 See Section 3.1.
5 See Section 3.1.1.
museums have been doing physically for many decades by selecting specific artifacts, and combining them chronologically, thematically, taxonomically in the exhibition space in order to tell stories through them.

In the context of Information Systems for Cultural Heritage, to compare databases to digital collections in the way it has been done so far does not provide the accurate understanding that is necessary here. There are important distinctions between databases, data models, data structures, data values, data formats, and the database software that instantiate them all. For Manovich, databases are seen simultaneously as “a structured collection of data” [147] and the means “for fast search and retrieval by a computer” [147]. Hayles define databases as a more complex "artifact" that is comprised of “technologies of data compilation, storage, transmission and retrieval, which have their own constraints and possibilities deeply affecting how databases are built” [106, p.167].

Databases should be understood as a data repository where the data collected is persistently stored. The software that retrieves, adds, deletes, and updates this collection is the Database Management System6 (see [111, ch.2]). In addition, in case of relational databases, for example, one can notice yet a further distinction between the content comprised by data values7, and the schema, which is a conceptual model that organizes a collection of data fields and structures8. The data repository is then stored obeying a format, which “refers to the particular encoding of information contained within a file.” [75]9. In other words, the content of a data repository, as a collection of encoded and structured data, is organized by a conceptual model/schema that defines the data values that are stored in specific data fields as data structures. However, the data repository is not in any sense dynamic if not run as a computer process. Nevertheless, this static format must afford computation by being machine-readable.

In case of Cultural Heritage, Digital Collections refer to the repositories of cultural data, and not the software that generates

---

6 Relational Database Management System (RDMS), if the software is based on the relational model
7 According to Elings et al. “data values are the information proper stored in the data fields.” [75]
8 According to Elings et al. “data fields are the named units of information, often also referred to as 'elements' or 'categories', which are organized into a record by a data structure.” [75]
9 According to Elings et al. “XML, for example, encodes data fields and their values into a hierarchically structured file, and therefore can be called a data format standard.” [75]
the repository, as e.g. a Collection Management System (CMS). The Canadian Rules for Archival Description\(^\text{10}\) provides an insightful definition for Digital Collections\(^\text{11}\), which is:

\[
\text{An artificial accumulation of documents of any provenance brought together on the basis of some common characteristic, e.g. way of acquisition, subject, language, medium, type of document, name of collector, which may be treated for descriptive purposes as a unit under a common title.}\ [54]
\]

Here an important point needs to be made. Digital Collections have not been traditionally created to support the production of narratives, but instead to serve as an indexed inventory that can be easily searched through. However, there is nowadays a considerable effort in trying to employ these repositories as resources for storytelling. This effort in exploring the intrinsic relationship between algorithms, as counterpart of narratives, and databases, as counterpart of digital collections, in order to create compelling and interactive stories can be observed in the award-winning treasure hunt-like game called Collect & Connect \(^\text{47}\) played on an interactive multi-user table that was part of the exhibition Taku Tamaki – Auckland Stories\(^\text{12}\) in 2015 at the Auckland War Memorial Museum\(^\text{13}\) (Auckland, NZ). The game drove visitors to follow “the footsteps of a curator to uncover collection objects to match a specific theme – pretty much like a real curator at a museum would.” \(^\text{47}\) Presenting a meta-narrative as algorithm - the narrative of the game itself (to collect and connect artifacts) and the new narratives created during the game play (the collected and connected artifacts), the game was also directly served by the Auckland Museum’s Collections Online API\(^\text{14}\) that enabled straight access to records produced with the CMS of the institution. By embedding the digital collection into a gameplay (narrative), the game “provided a unique way of playfully accessing the Museum’s collections data in a very visual and graphic way that was very different from conventional search/browse interfaces.” \(^\text{47}\)

\(^{10}\) http://www.cdncouncilarchives.ca/archdesrules.html

\(^{11}\) Definition listed in a comprehensive review of definitions of collections compiled by Currall et al. \(^\text{54}\)


\(^{13}\) http://www.aucklandmuseum.com/

\(^{14}\) http://api.aucklandmuseum.com/
A few concepts already discussed in previous chapters, such as 4.1 Platforms as a means for Extensibility and Open-Endedness, are necessary to be brought into the spotlight in order to comprehend the interdependences between algorithms as DIAs\textsuperscript{15}, and Digital Collections as one of the platforms for Constructionist Heritage Interpretation\textsuperscript{16}. As previously explained, one of the principles that defines platforms is modularity, which enables extensibility and open-ended possibilities by arranging and rearranging a set of core components. These core components are designed based on standardization and specification so that compatibility is preserved in lesser or higher degrees throughout the system. By allowing uncomplicated combination of different components (building blocks) together, platforms are effective in producing many different variations of outcomes with speed and at low cost. In this sense, databases are nothing else than modularization of information, or, in order words, the further reduction of information into small pieces - data, which, in higher amounts, constitute the fabric/raw material to be given form by computational means. The augmentation of modularization through standardization affords greater interoperability\textsuperscript{17}, which holds the potential of driving the formation of ecosystems. Digital storytelling, in regard to Digital Collections, is therefore susceptible to the same principles as data repositories. Its uniqueness, however, can be found in the effort

\textsuperscript{15}See Sections 3.1.3.2 and 7.1.1.1.
\textsuperscript{16}See Sections 3.2 and 7.1.1.1.
\textsuperscript{17}Interoperability is defined as “The ability of computer systems or software to exchange and make use of information.” [244]
of designing conceptual models that are capable of expressing the particularities of narratives. As Hooland et al. puts it:

*The evolution from an unstructured narrative to a highly structured representation of metadata requires the development of schemas in order to make the metadata interoperable. By slicing up unstructured descriptive narratives into well structured fields, we need to render the meaning of the different fields (also called attributes) explicit by documenting them in a schema.* [111, p.13]

Based on what has been discussed so far, this chapter will analyze the potentials and problems of Data Standards and their Conceptual models, and CMSs as the main tools for representing, standardizing, and producing Digital Collections, which are the raw material for the construction of DIAs in Hackathons for Cultural Heritage. As high-level interpretive layer, Data Standards and CMSs have an fundamental impact in the organization and capacity of representation of data repositories, their content, and consequently their use as an interpretive and creative material.

### 5.1 Data Standards and CMSs

The creation of systems of classification is an ancient human practice. The Babylonians are known for creating the first library of clay tablets organized by a librarian already in 1700 B.C.E. [203, p.256]. In 700 B.C.E, the Royal Library of Nineveh in Mesopotamia arranged its collection in “a classified, numbered order” [203, p.256]. The assemblage of collections is therefore closely associated with the implementation of strategies to cataloguing its items. Thus, representing and managing information in Cultural Institutions have always been a central topic in their practice [34].
Not long ago, card catalogues and file cabinets (see Fig. 19) were traditionally used as a means for organizing institutions’ informational assets. Needless to say that such analog technologies presented several problems regarding e.g. the speed in searching for specific information, the space available on the card for describing an object, the work involved in constantly keeping the order of the cards, the difficulty in modifying the contents of a card, and other issues related to the use of ink & paper. In case the artifact was catalogued primarily according to their creators’ names, meaning that the alphabetically organized letters on the drawers of a cabinet would represent the artist or author’s initials, using the same cards to sort objects e.g. by their chronological order would require adding additional procedures to the already tedious work of the staff, who would need to go through each card individually in order to accomplish such work [150]. However, even presenting difficulties, which were intrinsic to the analog media of the past, the internal work of institutions was significantly facilitated by following specific rules expressed in Data Structure Standards/Schemata. In this sense, by systematically organizing and documenting institutions’ informational assets, artifacts could be easily accessed. These early standards used the alphabet as the main organizing tool.

Figure 19: File cabinets at the Oakdale Workmen’s Institute

Card catalogues and file cabinets were traditionally used for organizing collections.

Alphabetic ordering was one traditional method/schema for organizing inventories.

18 Picture source: https://museum.wales/media/42777/Cardcatalogue.jpg
5.1.1 Data Standards: Powerful yet incomplete

In Europe, the Department of Printed Books at the British Museum (London, UK) released in 1841 the first volume of the Catalogue of Printed Books in the British Museum [27]. The catalogue contained one of the first rigorous set of rules, called Rules for the Compilation of the Catalogue, that “became the foundation for all subsequent catalog rules of the nineteenth and twentieth centuries, and are at the origins of the ISBD of the twenty-first century and of digital cataloging elements such as Dublin Core” [18]. Just to give an example, the rules vary from establishing very objective formatting standards, such as “I. Titles to be written on slips, uniform in size.” [27], and “II. Titles to be arranged alphabetically, according to the English alphabet only... under the surname of the author, whenever it appears in the title, or in any other part of the book” [27]; to rules that reflect social and cultural attitudes of the nineteenth century in Britain, such as “V. Works of Jewish Rabbis, as well as works of Oriental writers in general, to be entered under their first name.” [27] By complying rigorously to this set of rules, the staff of the British Museum produced a sizable ordered list of the institution’s book collection with a neat format as presented in the figure below (see Fig. 20).

Figure 20: Analog entries generated using the British Museum rules

5.1.1.1 Data Standards and the Computer

In America, museums started adopting computers, the so-called mainframes, already in the earlier 60s driven by the initial promise that these systems would eliminate repetitive and time-consuming tasks, such as “sorting records, searching for information, and tabulating

---

19 The catalogue was compiled in more than 400 pages.
results” [150]. However, as Paul Marty explains [150], soon the institutions realized that computation could change not only how museum professionals performed such traditional tasks, but also the reason why information was collected in the first place. In the search for exploring the possibilities of digital technology, and creating new practices in the field, the Museum Computer Network (MCN) was created in 1967. The MCN had as vision the establishment of “shared national and international museum data banks that would not only improve museum management practices, but would provide unprecedented access for researchers.” [158] In the beginning, the MCN saw in mainframes the path for realizing its vision. Based on this technology, the MCN developed the system General Retrieval and Information Processor for Humanities Oriented Studies (GRIPHOS) as an information management system that was shared by fifteen museums from the New York area. The system GRIPHOS “enabled museums to store, retrieve and apply textual markup to their records” [223]. However, with the popularization of personal computers in the 80s and the Internet, the MCN moved away from the idea of a centralized system. On the Web, the focus was instead on Data Standards that would be able to provide interoperability, which can be defined by the capacity of exchanging information across different systems [158], and the maximization of the utilization of datasets to answer more complex and precise queries. However, a universal Data Standard capable of handling and sharing all data of all institutions has been ever since a very difficult problem to tackle. That is because representing and interpreting heritage is not an exact science, but depend on social, historical, psychological, ideological, and many other factors that come into play as in the case of the Rules for the Compilation of the Catalogue, as explained above. In addition, as Paul Marty puts it:

The identification and gathering of this information are driven by the requirements of different museum professionals as they assess what is needed for their own use and for the use of potential museum visitors. These needs often vary from institution to institution and from visitor to visitor. Students in an art history class, for example, come to a museum searching for appropriate examples to use in their papers. Scholars researching a particular topic need to know how many prints by a given artist exist in the museum’s collection or how many paintings deal with a certain subject. Museum curators planning a new exhibit require information about each object’s
Consensus around a universal Data Standard answering the requirements of museum work has been proven challenging.

If, on one hand, a larger number of institutions may be reached by creating a more general shared vocabulary; on the other hand, the lack of specificity of terms may hinder the capacity of institutions in precisely classifying their collections. This is exactly one of the main challenges of Data Standards, including e.g. the standard Dublin Core (DC), which is organized by the Dublin Core Metadata Initiative (DCMI)\(^\text{20}\). The roots of the DCMI can be traced back in 1994 during the 2nd International World Wide Web Conference (Chicago, US), where Stuart Weibel and Eric Miller, both researchers of the Online Computer Library Center (OCLC)\(^\text{21}\) discussed about the difficulties associated in finding resources on the Web, which at that time had around ”only” 500.000 addressable objects (see \([\text{57}]\)). In 1995, a workshop called OCLC/NCSA Metadata Workshop was held in the headquarters of the OCLC (Dublin, US), where “more than 50 people discussed how a core set of semantics for Web-based resources would be extremely useful for categorizing the Web for easier search and retrieval” \([\text{150}]\).

DC was created not only with the Web in view, but also as an alternative to another Data Standard called MAchine-Readable Cataloging (MARC)\(^\text{22}\), which, according to Jeffrey Beall \([\text{13}]\), was problematic because of the “high level of training required for information specialists to create valid MARC records” \([\text{13}]\). In addition to that, MARC was considered to be too library-specific, and the creators of DC wished for a standard with a broader application range that could also be applied to other institutions than only libraries. Thus, the DC’s vocabulary was created to be broad and generic, consisting of only an essential - core - set of descriptive terms \([\text{56}]\). For librarians, however, the lower-definition of DC caused problems regarding ”being far too general and unable to cope with specific needs” \([\text{6}]\). DC was certainly user-friendlier and presented a lower adoption cost in comparison to other Data Standards because of its set of 15 human-readable metadata elements\(^\text{23,24}\) for describing resources. However, although

\(\text{http://dublincore.org/}\)

\(\text{https://www.oclc.org/}\)

\(\text{The following discussion focuses on the most popular version of MARC, which is MARC 21, released in 1996 (see \(\text{https://www.loc.gov/marc/}\)).}\)

\(\text{23 Including Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, and Rights.}\)

\(\text{24 The DCMI has later on released the DCMI Metadata Terms that not only consists of a bigger set of terms, but also is described as a formal ontology (see \([\text{55}]\)).}\)
**MARC** used a complicated set of codes\(^\text{25}\), it was able to express bibliographic data with more precision.

![Diagram](image)

**Figure 21**: Expressing temporal information with the CIDOC-CRM

As **DC** was considered inadequate to fulfill the requirements of some Cultural Heritage Institutions searching for more expressive models that could capture the characteristics and facts around cultural objects, as well as support the work related to administration, preservation, and conservation of heritage, the **International Council of Museums (CIDOC)** started to conceptualize a formal ontology\(^\text{26,27}\) (see Fig. 21.) called **CIDOC-CRM**\(^\text{28}\) (Conceptual Reference Model) that - again - was supposed to fulfill the requirements of institutions. This time however the strategy was to increase the complexity of the model in order to provide a broader and precise description of factual data, and an architecture that was capable of delivering the sufficient interoperability to consolidate information from museums, libraries and archives\(^\text{69}\). Therefore, while **DC** aimed at simplicity in order for standardizing information, the **CIDOC-CRM**’s strategy consisted of

\(^{25}\) See [https://www.loc.gov/marc/bibliographic/bdsummary.html](https://www.loc.gov/marc/bibliographic/bdsummary.html)

\(^{26}\) For a definition of ontologies in Computer Science, see [100] and [91].

\(^{27}\) The decision to build the model as an ontology (Object-Oriented Approach), came from the experience of the **CIDOC** with a previous project that tried to standardize Cultural Heritage information by developing a Relational Data Model (Relational-Oriented Approach). According to Gill [95], the **CIDOC** “produced a relational data model to support the **CIDOC** Information Categories in the mid-1990’s. Although highly valuable as an intellectual exercise, it was too complex to be implemented as working system, which was in large part the motivation for beginning work on the **CIDOC-CRM** in 1996.” [95]

providing the means for describing collections and administrative work based on a more complex model that could be understood by people (human-readable), as well as processed by machines (machine-readable). On the date of its release\textsuperscript{29}, the CIDOC-CRM consisted of 80 classes and 130 properties\textsuperscript{30} conceptualized by an interdisciplinary team of professionals with “a background in museology, history of arts, archaeology, natural history, physics, computer science, philosophy, and others” [69]. The CIDOC-CRM website [110] provides the following definition to the model:

The CIDOC-CRM is intended to promote a shared understanding of cultural heritage information by providing a common and extensible semantic framework that any cultural heritage information can be mapped to. It is intended to be a common language for domain experts and implementers to formulate requirements for information systems and to serve as a guide for good practice of conceptual modelling. In this way, it can provide the “semantic glue” needed to mediate between different sources of cultural heritage information, such as that published by museums, libraries and archives. [110]

Although important Cultural Institutions, such as the British Museum, have indeed based the implementation of their Information Systems on the CIDOC-CRM, demonstrating therefore the relevance of the project, the necessity of creating new Data Standards capable to comply with an increasing number of requirements continues. As one of the biggest digitization initiatives of Cultural Heritage to date with more than 51 million items (including artworks, artifacts, books, videos and sounds from across Europe), the Europeana Project\textsuperscript{31} has developed its own data model, called Europeana Data Model (EDM)\textsuperscript{32}. The EDM was developed as a solution for the lack of granularity\textsuperscript{33} of the previous Data Standard used by the Europeana project called Europeana Semantic Elements (ESE), which was an extension of the Dublin Core Element Set. The problems between the ESE and the EDM were similar to the ones already discussed previously between DC

\textsuperscript{29} September, 2000.
\textsuperscript{30} The current version of the CRM consists of 89 classes and 151 unique properties (see [65]).
\textsuperscript{31} More about the Europeana Project, please see Section 2.1.0.1
\textsuperscript{32} https://pro.europeana.eu/resources/standardization-tools/edm-documentation
\textsuperscript{33} Granularity is the level of depth that a given data model is able to represent information. “Granularity concerns the size of the elements that are the objects of attention and action within the system—and, in turn, with the entities they might represent.” [73]
and CIDOC-CRM. Interestingly, the development of the EDM has not considered the CIDOC-CRM as a foundation for representing heritage. Instead, the EDM was based on another Data Standard called Simple Knowledge Organization System (SKOS)\(^{34,35}\), which is a standard to support the conceptualization of Knowledge Organization Systems (KOS) such as “thesauri, classification schemes, subject heading systems and taxonomies within the framework of the Semantic Web” [116]. In this sense, SKOS is a more general Data Standard, because it focuses on the backbone of organization of knowledge. At the same time, it offers more advanced semantics than ESE. According to Doerr et al.:

> SKOS features a main class to describe concepts. Adapting standards like ISO2788 to a concept-based modelling approach, it coins properties for the labels of these concepts (e.g., skos:prefLabel for the preferred label of a concept, skos:altLabel for the alternative ones), for semantic relationships between these concepts (skos:narrower, skos:broader, skos:related) and for general concept documentation (skos:scopeNote, skos:definition, etc.) [70]

In addition to SKOS, the EDM was also based in a reinterpretation of DC, not only because DC provided a compact vocabulary as discussed previously (facilitating therefore the adoption of the standard by Cultural Institutions), but also because the EDM developers wanted to provide a certain degree of compatibility with the legacy data already created using ESE and the new EDM [70]. Finally, the EDM also incorporated another Data Standard called Friend-of-a-Friend (FOAF)\(^{36}\), which is an ontology used to describe people and people-related terms that can be used in structured data [31]. As a matter of fact, although the undoubted relevance and contribution of the projects discussed in this section (MARC, DC, CIDOC-CRM, ESE, EDM, SKOS, and FOAF), they are only the tip of the iceberg. Jenn Riley compiled an incomplete Glossary of Metadata Standards [196] that lists more than 100 standards, which specifically for museums contains the Art & Architecture Thesaurus (AAT)\(^{37}\), Cataloging Cultural Objects (CCO)\(^{38}\), Categories for the Description of Works of Art, (CDWA)\(^{39}\), MuseumDat\(^{40}\),

---

\(^{34}\) https://www.w3.org/2004/02/skos/
\(^{35}\) More information about SKOS, see Section 6.1.3.1.
\(^{36}\) http://www.foaf-project.org/
\(^{37}\) http://www.aacr2.org/
\(^{38}\) http://www.vrafoundation.org/ccoweb/
\(^{39}\) http://www.getty.edu/research/conducting_research/standards/cdwa/
\(^{40}\) http://www.museumdat.org/index.php?ln=en

---

Ranging from very specific to generic, there are a great number of data standards for Cultural Heritage currently available.
Data Standards and the Representation of Heritage

The number and variety of Data Standards present today is just evidence to the ever-changing nature of culture and its sign systems. Modeling, encoding, and interpreting heritage are not static processes. Therefore, no matter if highly expressive and extensive, or general and limited, Data Standards will always be incomplete. In this sense, most of the current Data Standards are built with extensibility in mind. The CIDOC-CRM has e.g. a number of extension that are created on top of its core in order to provide solutions to various and specific requirements. Just to name a few, some of these extensions are the FRBRoo (a conceptual model for bibliographic information in object-oriented formalism), PRESSoo (a conceptual model for bibliographic information pertaining to serials and other continuing resources), CRMInf (an extension of CIDOC-CRM to support argumentation), CRMsci (a conceptual model for scientific observation), CRMgeo (an extension of the CIDOC-CRM to provide linkage between the standards of the geospatial and the Cultural Heritage community), and so on. These projects show the continuous need to update and create new systems capable of capturing new realities/possibilities. Data Standards, as semiotic systems, can be comprehended as susceptible to the same semiotic processes as other sign systems. Extensibility and revision is understood as semiosis. About this, Charles Sanders Peirce wrote:

The object of representation can be nothing but a representation of which the first representation is the interpretant. But an endless series of representations, each representing the one behind it, may be conceived to have an absolute object at its limit. The meaning of a representation can be nothing but a representation. In fact, it is nothing but the representation itself conceived as stripped of irrelevant clothing. But this

http://www.collectiontrust.org.uk/spectrum
http://www.getty.edu/research/conducting_research/vocabularies/tgn/
http://www.getty.edu/research/conducting_research/vocabularies/ulan/
https://www.ifla.org/publications/node/11240
https://www.ifla.org/node/11415
http://www.cidoc-crm.org/crminf/ModelVersion/version-0.7
http://www.cidoc-crm.org/Resources/crmsci-the-scientific-observation-model-1
clothing never can be completely stripped off; it is only changed for something more diaphanous. So there is an infinite regression here. Finally, the interpretant is nothing but another representation to which the torch of truth is handed along; and as representation, it has its interpretant again. Lo, another infinite series. [184, vol.1, p.339]

The representation of heritage, done also through the production and alternation of conceptual models, is a flexible and infinite semiotic process that produces constant and unlimited (re-)signification. In addition to being subject of the effects of constant revisions and adaptations, sign systems, including the ones used by Data Standards, depend greatly on the intentions that motivate their design. There is therefore no right or wrong, but instead a suitable standard for a given context and time.

5.1.2 CMSs: The producers of Digital Collections

Data Standards are however conceptual models that need to be implemented and instantiated as Information Systems for Cultural Heritage - more specifically CMSs. CMSs provide museums, archives, and libraries with strategies to “collect, store, organize, search and display Cultural Heritage objects or their (metadata) representations in a digital environment” [51, p.339]. The outcome of the work organized by CMSs is the data they collect as a dataset, which is described here as a raw material for the creation of DIAs. Due to issues such as compatibility with legacy systems, current engineering and curatorial practices, and end-user expectations, CMSs tend to favor an inflexible architecture and constrained UI. Because of the central relevance of CMSs in the curatorial and research practices, these factors affect directly not only how the intellectual work inside Cultural Institutions is carried out, but they have a profound impact in the content that is represented and quality of datasets that are generated through these systems.

As stated by Hooland et al., “Collection management systems, archival inventories and library catalogues are all built on top of a RDMS.” [111, p.25] Indeed, the great majority of CMSs - such as The Museum System50, eHive51, Qi Keepthinking52, EMu Collection Management
System\textsuperscript{53}, and so on\textsuperscript{54} are developed on the basis of a relational database model. While relational database models have proven their efficacy for handling the data layer of a variety of applications, they enforce a strong separation between variable content and a fixed data structure [73, p.113] that limits the system’s ability to quickly adapt. Relational databases are not designed to handle frequent changes. Such modifications in the schema require professional assistance, and depend on complex procedures, such as schema migrations\textsuperscript{55}. About the inflexibility of the schema in the relational model to keep up with adaptations, Dourish writes:

\begin{quote}
The significance of the separation between structure and content lies in the radical differences in their malleability. Content is infinitely malleable... That schema itself, however, is much more rigid. While there are mechanisms in most relational databases for columns to be added to or deleted from the database structure, such changes are limited and invalidate data contents, limiting their uses still further. [72, p.91]
\end{quote}

In addition, as explained by Seth Hooland et al., an even more problematic outcome is produced if a new content type needs to be introduced, since “adding an extra table requires the database manager to rethink the entire schema of the database, as adding an extra table might imply a degrading of the normalization process.” [111, p.26]

The technical considerations are however not the only issue. Christiane Weber\textsuperscript{56}, a member of the Research and Education Department of the International Tracing Service\textsuperscript{57} (ITS - Bad Arolsen, Germany), has provided an interesting insight of the close relationship between Information Systems and the content produced as the result of the research practice happening in her institution. According to Weber, in regard to the institutional mandate, an extra text field added to the CMS might mean years of research on thousand of items in the collection. That is because this extra field would need to be implemented as an extra column in a table of a relational database. A conceptual alteration in the representation of the object of research imposes an institutional need that must be fulfilled. Thus, the relational model, in a certain degree, determines

\begin{footnotesize}
\textsuperscript{53} https://emu.axiell.com/
\textsuperscript{54} See Appendix B.
\textsuperscript{55} Schema migration consists of transferring data between computer storage types or file formats.
\textsuperscript{56} Interview (see Section ??).
\textsuperscript{57} https://www.its-arolsen.org/en/
\end{footnotesize}
institutional workflows, because it consistently and systematically applies the alteration to all items defined by the schema. Inserting another content type property by implementing it as a new column on a table would automatically modify all represented items that fall into the respective content type. The extra cells, externalized as e.g. text fields on the UI of the Information System would then need to be filled in. Such high cost, in terms of intellectual work and software development necessary to update the system, refrains institutions to experiment and prototype with digital representations and interpretations, since decisions in this direction need extra consideration. The consequences of what is and what is not represented cannot be ignored. According to Paul Dourish [73]:

*Databases are collections, and frequently what they collect are records and archives of human activity; but databases also have particular shapes, and the shape of the database determines how effectively different kinds of records can be maintained. Objects, activities, or people who cannot easily be represented in a database become, to some extent, invisible to the institutional processes that databases represent. A simple example—conceptually simple, but politically complex—concerns how databases record gender, and the consequences for any case in which a simple, fixed, lifelong male/female binary is insufficient.* [73, p.29]

The format and content of Digital Collections produced by CMSs, together with the visibility and invisibility of things by them represented are defining factors for the creation of DIAs. That is because Digital Collections intermediate the design process. During the interviews and workshops with curators and other cultural institution professionals, it was evident the intertwined relationship between datasets and their affordances as creative resources, as concepts for the design of such applications are bounded to the content of datasets. Throughout the design process, such as the ones happening in Hackathons, the dataset is the raw material that designers will find themselves having a reflexive conversation with [204]. On the one hand, the format of the dataset will impose its constraints and affordances, enforcing that “programmers understand the ways in which digital structures can resist their will, every bit as much as clay, wood, or stone.” [72, p.91] On the other hand, the invisibility

---

58 See Appendix A.
of representations will also affect design decisions. For example, if a CMS does not implement a model for geotagging objects, the dataset produced will consequently lack this information, which will impede the implementation of certain location-based concepts, or require extra work in order to extend the dataset. In the same way, taking the example given above by Dourish, if a dataset does not include gender variations beyond the binary male/female categorization, digital storytelling based on this dataset would render the invisibility of individuals belonging in the intersection of the gender spectrum. The same intrinsic relationship between datasets and their possibilities was found in the results of the survey *Online Repositories for Cultural Heritage*\(^{59}\) when interviewees were asked\(^{60}\) whether they needed to adapt the concept of an application because the data they had available was insufficient to meet the requirements of an initial concept.

![Figure 22: The influence of datasets in the design process](image)

According to the survey, among the 76 interviewees who answered this question (see Fig. 22), 58 individuals (76.32%) said they needed to adapt the concepts of their applications, because the dataset did not meet the necessary requirements for their implementation. Only 18 individuals (23.68%) stated they have not found themselves in such situation. Moreover, the survey also showed the need for new tools capable of providing extra flexibility to the design process in concern to the entities represented in data repositories. The participants were asked\(^{61}\) whether they would like to have a tool that would enable

---

59 See Section A.2.1.1.

60 Question number 6 of Section A.2.1.1.

61 Question number 9 of Section A.2.1.1.
them to easily extend and modify digital representations of existing datasets so that the possibilities of what can be created with them is augmented. Out of 75 individuals who answered this question, the large majority, 62 interviewees (82.67%), stated that they would like to have access to such a tool in order to enable them to easily extend and modify existing data repositories. Only 13 individuals (17.33%) responded that they were satisfied with existing data repositories (see Fig. 23).

![Figure 23: The need for easily extending and modifying datasets](image)

5.1.2.1 **Focus of Representation**

In addition to the lack of flexibility in regard to their schemata that hinders the capacity of the system to adapt, CMSs present problems concerning the lack of a more comprehensive and appropriate knowledge representation that suits the requirements for storytelling, because they are engineered with the focus on management and administrative work [14]. Elana Carpinone [40] gathered a comprehensive list of features commonly found in CMSs. According to Carpinone, these features are [40, p.26-27]:

- **Cataloguing/Object** - "includes fields for basic pertinent object information such as: accession number, catalog number, components, object name, artist or creator, date, culture, dimensions, materials and techniques, school, period, description, condition, current location, value, provenance, exhibition history, source, and picture."

- **Acquisitions** - "includes fields for donor name and contact information, type of acquisition, accession number, date of transferring document, promised gifts, date received and how the object arrived at the museum, any
donor restrictions, seller and purchase price. Typically a CMS allows the museum to use its own established accession numbering system.”

- **Deaccessions** - “file for catalog records for objects removed from the collection. Includes fields for method of disposal, date, and reason for deaccessioning”

- **Thesaurus** - “controls the museum nomenclature for consistent cataloging terms”

- **Loans** - “tracking incoming and outgoing loans. Including catalog record for each object, lender name and contact information, special lender requirements, credit line, shipping information, insurance value, condition, picture, and loan history.”

- **Exhibitions** - “information related to exhibits such as: exhibit name, location, duration, exhibition history, objects in exhibition, lenders, shipping information, insurance, installation notes, budget, and pictures of the exhibit.”

- **Shipments/Transport** - “information about arrangements for transferring objects to different locations or venues. Includes fields for packing notes, crate dimensions, number of crates in shipment, objects in each crate, venue information, dates received and sent, courier, carrier, and customs information.”

- **Condition/Conservation** - “fields to record an object’s condition, date of inspection, inspector, and conservation reports.”

- **Search** - “query methods, language or terms that can be used to search the database according to different entry points.”

- **Reports** - “most CMSs allow for the generation of reports based on the information selected by the user pulled from information in the database either through a built-in report maker or a plug-in for other software such as Crystal Reports.”

- **Copyright/Reproductions** - “includes fields such as copyright restrictions, permissions and copyright owner.”

- **Importing/Exporting data** - “allows for data to be imported and exported in a variety of formats such as XML or DOC.”

- **Barcoding** - “allows the program to create and print barcode labels that include information such as: the accession number, title, and picture. It scans and reads barcodes, track objects and aids inventory control.”

Besides the features mentioned above, CMSs can also provide solutions for organizing and managing bibliographies, insurance, events, contacts, repatriation, tasks, valuation, fund-raising/campaigns, among others. Modeling content is in great part restricted to the factual description of objects\(^{62}\) and their

\(^{62}\) Read above in Cataloguing/Object.
categorization through taxonomies or thesauri\textsuperscript{63}. As found during the survey *Datasets and the Curatorial Process*\textsuperscript{64}, Cultural Institutions are already utilizing their datasets to other purposes than management and administration of assets. Participants were asked\textsuperscript{65} whether the cultural institution they work for employ datasets in different contexts than managing collections (see Fig. 24). The results showed that out of 60 individuals who answered this question, the majority, 42 individuals (70\%), responded positively. According to them\textsuperscript{66}, the employment of datasets where indeed broad, ranging from the production of applications\textsuperscript{67} to the provision of content to third-party services\textsuperscript{68} and events\textsuperscript{69}.

<table>
<thead>
<tr>
<th>Does your institution use its datasets in different contexts than collection management?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Yes</strong></td>
</tr>
<tr>
<td>30.00%</td>
</tr>
</tbody>
</table>

Figure 24: The utilization of digital collections beyond search and retrieval

Especially in the case of off-the-shelf CMSs, apart from description fields, the implementation of conceptual models capable of expressively capturing phenomena out of the scope of collection management is lacking \[14\]. Although the GLAM sector has understandably pushed for standardization, the employment of Digital Collections as the basis for Digital Heritage Interpretation would profit considerably from conceptual models that could express e.g. the personal understanding of a curator on a certain topic that is not described by current standards. Flexibly enriched datasets would provide greater interpretive possibilities that would benefit especially

\[63\] Read above in *Thesaurus*.

\[64\] See Section \textit{A.2.1.2}.

\[65\] Question number 2 of Section \textit{A.2.1.2}.

\[66\] Question number 3 of Section \textit{A.2.1.2}. See Section \textit{A.2.1.2} for detailed information on the findings.

\[67\] Such as virtual exhibitions, educational games, virtual reality apps, etc.

\[68\] Such as Europeana, Wikidata, Wikimedia Commons, OpenStreeMap, ArtUK, etc.

\[69\] Such as hackathons.
museums. When comparing museums with other cultural institutions such as libraries, the staff of libraries does not provide interpretation about the contents of books, but they instead “interpret the information need of the users and consequently pointing them to the right book” [217, p.12]. In regard to archives, the main requirement is to provide factual information that is necessary to reconstruct the original context of the archival material. Although historical interpretations may come from the handling of the historical documents, archives, as institutions, must make sure that the information remains preserved, intact and somehow “unfiltered and free of alteration” [217, p.16]. As for museums, interpretation is a lot more colorful and is understood beyond searching tools and finding aids\(^{70}\). That is because, museums have become a stage for negotiating “societal demands” [97, p.10] and “seek meaning, not just a place to view objects” [53, p.150], in this sense they “interpret and contextualize objects based on the present knowledge and are driven by an educational agenda. Museum objects gain significance and get interpreted by being part of a certain collection and being chosen by a curator” [217, p.16]. Again, the issue of invisibility in regard to the representation of information appears with urgency as part of a more inclusive Heritage Interpretation. However, the need for Information Systems restricted to administrative capabilities concerning libraries and archives is valid as long as these institutions do not engage with educational projects or start to take part in e.g. Hackathons. If archives and libraries do so, the requirements for the information represented through their systems would be equivalent to the ones of museums, since they would engage with the production of meaning by interpreting their collections\(^{71}\).

5.1.2.2 *Form-based User Interfaces*

In regard to their User Interfaces, the great majority of CMSs follow a predominantly form-based design approach to the way information is inserted, retrieved, modified, and visualized. Among their advantages, *Form-Based User Interfaces* require minimum training to be understood and used [45]. That is because they model simple

---

\(^{70}\) According to the Glossary of Archival and Records Terminology, finding aids is a “description of records that gives the repository physical and intellectual control over the materials and that assists users to gain access to and understand the materials.” [243]

\(^{71}\) As an example, although the ITS shares it activities as a tracing service and archive, it has recently invested in developing educational projects. The ITS has also taken part in the 2017 edition of the CdV. For more information about the ITS, see Section 5.1.3.
and well-defined use cases/workflows and enforce them by guiding and restricting the interactions of the user through forms and their input elements\(^72\) \([240]\), such as radio buttons, check-boxes, combo-boxes, drop-down lists, tabs, text fields, text areas, and so on. In contrast, other approaches to interface design that are based on e.g. Object-Oriented User Interfaces (OOUI) \([48]\) do not strongly enforce any workflow, but instead present the user with several simultaneous scenarios that are understood by the unique functionalities (via e.g. context-sensitive menus \([45]\)) and behaviors of each object, and the interactions among themselves\(^73\). For instance, dropping an icon of a document on the trash icon is the metaphor for temporarily discarding files, and “just as in the real world, the user can open the trash can and retrieve objects that have been discarded” \([48, \text{p.6}]\).

\[\text{Figure 25: UI of The Museum System}\]

In case of CMSs highly based on an Form-Based Approach, it would be a mistake however to claim that their forms are devoid of metaphors. The picture above (see Fig. 25) shows the Form-Based User Interface of a popular CMS called The Museum System. On the inferior half of the picture, it is possible to see the tab called Front

\(^72\) Inputs elements can be defined as “atomic UI components designed to receive user feedback upon which all forms are built” \([240, \text{p.81}]\)

\(^73\) Here, it must be observed the different (higher-lesser) degrees of object orientation that can be implemented as interface.
Card selected in orange. The picture demonstrates simply and clearly not only the visual resemblance with old card catalogues (see Fig. 26), but also a equivalent terminology used on the interface.

Although the form serves indeed as metaphor for the card catalogue and therefore provides the staff with continuity and familiarity with old analog practices74 (e.g. filling out cards), a system’s interface predominately based on forms are neither stimulating in concern to creativity nor they afford a more comprehensive understanding on how artifacts, terms, individuals, and other represented concepts relate among themselves. That is because these systems are not designed to support intellectual work in terms of producing interpretations about heritage, but instead to enforce workflows. As pointed out by Coldewey et al., a form-based approach is ideal either to skilled users “need to process well-defined use cases as fast as possible” [45] or untrained or semi-skilled users, who “often need step-by-step assistance while performing their tasks” [45].

The institutional consistency imposed by the relational model together with the enforcement of simple input procedures that do not require special training to be carried out will always have space in Cultural Institutions, because such qualities are required in their practice. However, different approaches that provide more freedom and dynamism for manipulating cultural objects can be beneficial to Digital Heritage Interpretation. The design approach of OOUIs that employ more advanced metaphors and interaction either/both in visual or/and tangible ways (see Fig. 18) provide more direct access to the possibility space described by Hayles75. As already stated, in interfaces displaying higher degree of object-orientedness, meaning that objects map logically and consistently the domain they try to

---

74 Especially concerning senior staff, who have worked in the field before institutions started to digitize their collections and workflows.
75 See introduction to this chapter.
model, the user is introduced with scenarios, which together compose short narratives that are comprehended by the functionalities of the objects that are represented in the metaphor. The **OOUI** becomes a platform that provides combinatorial capabilities that appear during their interactions. In this sense, the objects in such UIs work as building blocks for interpretation that is achieved by rearranging and combining them in various ways. Experimentation is therefore enabled. And, exactly this experimentation is what is not desirable in systems that base their interaction mainly through forms, because their goal is the anticipation of well-defined procedures.

5.1.3 **Case Study: OuSArchiv**

The following case study\[^{76}\] aims at analyzing and understanding how few concepts dealt with above come into play in a real case scenario. The case study takes a closer look at the *OuSArchiv*, which is a custom Java-based **CMS** developed by the software company *Ossenberg Digitalisierung und Software GmbH*\[^{77}\] that works on top of Microsoft SQL Server\[^{78}\], which is a Relational Databased Management System. The *OuSArchiv* was initially developed for and has been continuously adapted to meet the requirements of the International Tracing Service\[^{79}\] (**ITS** or IDS in German: Internationaler Suchdienst - Bad Arolsen, Germany). The case study explores in particular:

- the influences and interconnections between the physical archive, institutional work, data standards, and the **CMS**;
- the design approach in regard to the user interface and its affordances;
- the structure, quality, and integration of the digital collections in the context of the **ITS’** education department.

5.1.3.1 **The ITS and its Physical Archive**

The **ITS** was established after the Second World War as a tracing service in charge of searching for missing victims of the Nazi Regime and clarifying the fate of other missing individuals who could not be found alive. Between the 1950s and the 1990s the service also provided evidence of incarceration in concentration camps and of

---

\[^{76}\] See Section A.2.1.3 for additional information on the case study.

\[^{77}\] [http://www.ousgmbh.de/](http://www.ousgmbh.de/)


forced labor to applicants requesting such information in order to claim compensation or pension from the German state.

Figure 27: ITS’ physical archive

During its existence, the ITS has accumulated 30 million original documents about 17.5 million individuals, and has created a Central Name Index (CNI) containing around 50 million reference cards that were used to help researchers navigate through the collection. The ITS’ archive was structured hierarchically into three main themes that captures the chronological order of the events that occurred during the WWII, as stated by Charles-Claude Biedermann: “Die seither gewählte Archivordnung stellt in einem gewissen Sinn ein Abbild der Chronologie der Ereignisse während und nach dem Krieg dar” [21, p.23]. These categories are ([21, p.27], [179, ch.4]):

- **Main Group 1: Incarceration Documents** - including documents relative to subcategories such as Camps and ghettos, Transports, Prisons, Gestapo-Indexes, Various organizations, and Archival registers.

- **Main Group 2: Wartime Documents** - including documents relative to subcategories such as Implementation of the Allied Orders, Registrations of foreigners, and Post-war Evaluations of Various Organizations.

- **Main Group 3: Post-war Documents** - including documents relative to subcategories such as Evidence of Abode and Emigration, Relief Programs of Various Organizations, and Child-Tracing Service.

Before the digitization of the archives, the ITS had developed its own methods and internal data standards in order to cope with both
the initial social mandate as a tracing service, meaning the need of the
ITS to provide documented evidence of Nazi persecution in response
to requests of relatives of victims; and the specificities of the unique
kinds of documents produced during and after the Nazi regime. Built on top of the gathered documents, the CNI (see Fig. 28) was initially created as a “tracing tool” \[\text{[114]}\] for the ITS to assist the staff to promptly navigate, identify and find files in a stock of 30 million original documents. “For every name appearing on any document, the name index contains an index card showing the document’s exact location” \[\text{[114]}\]. The CNI was later on recognized by UNESCO\(^8\) as a Memory of the World Register\(^8\), which is a programme that lists documentary heritage with “world significance and outstanding universal value” \[\text{[153]}\]. The first version of the CNI was created immediately after the defeat of the Nazi regime by the Allies, who registered the displaced victims due to persecution or forced labor \[\text{[114]}\].

CNI was created as a tracing tool to help the staff to navigate 30 million original documents.

One of the unique characteristic of the CNI as a meta-archive is an alphabetical-phonetical system of classification that was especially developed by the ITS as an ordering principle for the CNI, since a purely alphabetical classification was not suitable. That is because, for a family name, there were numerous spellings that were used across European countries. Therefore, “names are grouped according to their pronunciation and not their spelling” \[\text{[190]}\]. Just as an example,

---

\(^8\) [http://www.unesco.org/]

\(^8\) [http://www.unesco.org/new/en/communication-and-information/memorystic-of-the-world/register/]
“Schwartz, of which there are 35,000 in the central index, is spelt in 42 different ways (Schwarz, Szwarz, Shwars, Svarz, Swartz or Szvarcz, etc.). Again, Weiss, of which there are 46,000, has 33 variations.” [190]

5.1.3.2 The Digital Archive and The OuSArchiv

The ITS begun in 1988 to digitize the CNI and has ever since expanded digitization to other documents of its collection. According to Christian Groh, head of ITS Archive, 90 percent of the physical collection has been processed so far. The initial goals of the digitization process was to not only protect and preserve archival holdings, but also improve efficiency in processing the applications of victims’ relatives. The digitization process consisted mainly in scanning the physical documents, partially transcribing their contents, and digitally indexing them with help of a CMS. One of the requirements of the ITS in regard to the CMS was that its data model should reflect the already existing hierarchical structure of the physical archive, which had supported the ITS’ cataloguing practice and the institutional work already in place. In this regard, Uwe Ossenberg, the owner of the company that has developed the CMS currently used by the ITS, answers when asked about the planning of the information architecture of the OuSArchiv: “Schon bei Beginn der Digitalisierung im ITS und bei der Auswahl des ersten Software-Pakets für die Verwaltung der Digitalisate (das war noch nicht unsere Software) wurde

82 Interview (see Section ??).
83 Interview (see Section ??).
In terms of its UI, it is possible to classify the OuSArchiv as hybrid, because it has characteristics of both Form-Based and OOUIs. As described above, the main goal of the OuSArchiv was to digitally represent as much as possible the archival objects, hierarchical structure, and workflows already established before the digitization process. The Form-Based Approach (see Attribute panel on Fig. 29) was used to implement the input procedures required to represent archival objects’ metadata, such as:

- **attributes** - which are entries used to map the textual information contained on the original documents and “may be the subject of search processes, which are supposed to locate or display all objects with specific characteristics.” [179, p.6-3] Examples of attributes are first and last names, date and place of birth, prisoner’s number, religion, name of camp, and so on;

- **archival descriptions** - which are special entries based on the international archival standard called General International Standard Archival Description84 (ISAD(G)) that “captures comprehensive information relating to groups of documents, chiefly inventorying or cataloguing data” [179, p.6-4];

- **references** - which are entries referencing records of the CNI among each other or records of other stocks. These links follow references types such as Further CNI object to same person, Further inquiry to same person, Reference to archival document, Archival document associated to inquired person, and Referenced document not yet imported (see [179, p.6-22])

---

A partial Object-Oriented Approach was used to represent the archival objects and structure already in place before the digitization of the ITS' collection. On the OuSArchiv's UI, these objects are depicted by the tree diagram displayed on the Navigation Area (as shown on Fig. 29 and Fig. 30), giving a good overview on how the collection is organized. The level of Object-Orientedness of the Navigation Area is however low because it focuses only on the visualization of different archival objects and their hierarchical structure. It does not allow the user to manipulate objects through e.g. dragging-and-dropping. In addition, Context Menus, which are accessed by right-clicking on the icons, offer only options to copy metadata information to the Clipboard. Other features, such as deleting, duplicating, or moving objects are not available, making it not possible to change the structure of the digital archive represented in the system. The lack of easy-to-use options, such as dragging-and-dropping, impose in some sense the existing analog structure over the digital one since . Understandably, as the digital archive was based on the analog collection and changes in the archival structure is not desirable. The OuSArchiv was therefore not designed in a way to enable the easy reorganization of the archival structure, but to enforce the already existing. In addition to that, the system does not allow the creation of different archival objects, but only the kinds that have equivalents in the analog archive. More detailed information about each of the objects and their hierarchical structure can be seen below (Fig. 31):
The managing body of the ITS decided in 2006 to open the archive for the general public. That decision led to the ratification of a formal agreement in 2007 by all the member states in the International Commission\textsuperscript{85} to establish a Research and Education Department (RnED) at the ITS that became active in 2008. As described by Margit

---

85 The International Commission is a supervisory body that was established in 1955 “consisting of representatives of Belgium, France, Israel, Italy, Luxembourg, the Netherlands, the United Kingdom of Great Britain and Northern Ireland, the United States of America and Germany, while the Federal Republic of Germany committed itself to financing this institution.” \cite{179}, p.2-26

---

5.1.3.3 The New Directions of the ITS' Digital Collection

The managing body of the ITS decided in 2006 to open the archive for the general public. That decision led to the ratification of a formal agreement in 2007 by all the member states in the International Commission\textsuperscript{85} to establish a Research and Education Department (RnED) at the ITS that became active in 2008. As described by Margit

---

85 The International Commission is a supervisory body that was established in 1955 “consisting of representatives of Belgium, France, Israel, Italy, Luxembourg, the Netherlands, the United Kingdom of Great Britain and Northern Ireland, the United States of America and Germany, while the Federal Republic of Germany committed itself to financing this institution.” \cite{179}, p.2-26

---

Figure 31: Hierarchy of the archival objects as described in \cite{179}
Vogt\textsuperscript{86}, a current member of the \textit{ITS' RnED}, one of very first difficulties faced by this newly founded department when trying to lay its basis was to obtain “a clear overview of the inconsistent structure and organizational principle of the physical and virtual documents”. As a result of this inquiry, three new educational and research thematic focuses appeared:

- **Incarceration Documents** - concerning concentration camps, ghettos, prisons, labor camps, labor reform camps, etc;
- **Forced Labor Documents** - concerning insurance cards, labor books, civil registration cards, employers’ lists, hospitals’ patients’ lists, etc;
- **Displaced Persons Documents** - concerning applications for IRO (International Refugee Organization) assistance, DP registration cards, emigration lists, DP identity cards, etc.

In regard to the educational mandate in particular, the \textit{ITS} aims primarily at enabling educators and teachers\textsuperscript{87} to work pedagogically with the documents found in the above categories by e.g. inviting teachers and their pupils to visit the institution and learn how Nazi politics impacted people’s lives. By not only analyzing the special characteristics of the documents, but also taking into account the context from which they emerged, students should be able to answer questions such as: who produced this document?, why was this document created?, to which function this document served? Besides this more traditional approach that focuses on the analog collection, the \textit{ITS} plans to expand its research and educational agenda digitally. For instance, the \textit{ITS} is currently developing an Web application to “enhance the understanding of documents held in the \textit{ITS} archive by offering a historical contextualization of the main types of cards, questionnaires and forms in an interactive online guide.”\textsuperscript{[236]} The Web application, called e-Guide is scheduled to be released by middle 2018 and will target international relatives who receive documents from the \textit{ITS}, university students taking part in educational projects at the \textit{ITS}, workshop participants, archive users in Bad Arolsen, and users of the online archive on the Web. As stated by Christiane Weber\textsuperscript{88}, a member of the \textit{RnED}: “We experienced in the past that all those people are aware that every scribble and every abbreviation on the cards can have a meaning. So

\textsuperscript{86} Interview (see Section ??).

\textsuperscript{87} More recently, in the context of the \textit{ITS} educational activities, other kinds of institutions besides schools have become the target group of the \textit{RnED}, as the overall management strategy of the \textit{ITS} is to “make the institution more visible, to raise public awareness of its existence” (Margit Vogt - Interview; see Section ??).

\textsuperscript{88} Interview (see Section ??).
the ITS decided to explain this abbreviations – together with the historical context of how the document was issued and what it was used for – in an interactive e-guide. The digital approach allows us to present the information free of charge on our internet site and it can be accessed from everywhere in the world.”

The CMS used by the ITS however does not accommodate the new content produced by the RnED. For the e-Guide in particular, the information created about the background and contextual information of the cards are not inputted in the OuSArchiv, although they are already produced to in a structured way in order to be presented on a Web application. Instead, the e-Guide will be based on a Content Management System (CntMS) called Typo3, which is a platform for building websites that works on a separated and independent environment, implementing its own data structures. As other projects carried out by the RnED, the e-Guide is not integrated with the OuSArchiv. However, as Weber explains, the ITS’ Information Technology team is currently working on a solution (called Dokumentenerkennung) to automatically tag each document contained in the digital archive with the document descriptions created by the RnED. Dokumentenerkennung is currently on the test phase and its implementation will depend on the results of the tests. No matter whether the solution is indeed implemented, the intention shows the need of a strategies for data consolidation within the institution, and the necessity of further software development to support the merging of educational interpretive content together with the existing archival structure, which (as already explained in Section 5.1.2.1) concerns with capturing only structural and factual information. In any case, if projects do not concern strongly to the archive, but are interpretations build on top of it, isolated solutions are favored in relation to deep structural modifications, since they are costly. As Hooland et al. explains:

In practice, these modifications are often avoided, as there is no time to fundamentally rethink the structure of the database. In this context, people often rely on lightweight and ad hoc solutions, such as creating a standalone spreadsheet. This type of short-term decision causes, over a period of years, tremendous issues with data consistency, as reference data are scattered across different applications. We can therefore conclude that it

89 https://typo3.org/
is not a trivial matter to update and maintain a database, due to
the complexity of modifying the database schema. [111, p.27-28]

Finally, the **RnED** has also started to engage with Hackathons. The **ITS** was one of the participant institutions of the 2017 edition of the CdV. Two of the 15 competing teams were interested in working with the dataset provided by the **ITS**, namely the team responsible for the project *Visualisierung jüdischen Lebens*[^90] and the team responsible for the project *Marbles of Remembrance*[^91]. The **ITS** created a data repository out of the archival unit *Kartei der Reichsvereinigung der Juden*, which contained the metadata and pictures of three different kinds of cards: *Karten der Schülerkartei* (with 11,000 cards), *Karten der Verstorbenen, Emigrations- und Ausländerkartei* (with 21,100 cards), and *Familienverweiskarten* (with 221 cards). However, although the considerable amount of cards represented, the teams working with the **ITS’** data repository faced problems when trying to implement their concepts, because the metadata captured only very few information present in the cards. Barbara Fischer, Curator for cultural partnerships by Wikimedia Germany, describes the insufficiency of the dataset produced by the **ITS’** OuSArchiv and the necessity of hackers to find different complementary data sources to be able to realize their concepts. Emphasizing especially the project *Marbles of Remembrance*, Fischer says:


The extension and integration of the **ITS** data repository with tailor-made datasets produced by the Marbles team, and external

[^90]: https://codingdavinci.de/projects/2017/
[^91]: https://codingdavinci.de/projects/2017/
additional data sources were essential for the success of the Marbles project in the competition, since the extra datasets provided the necessary supplementation for the insufficient content of the ITS dataset in regard to the concept of the application Marbles of Remembrance. These insufficiencies, apparent in the design process, came from the decision of team members to work with the stories of victims of Nazi persecution, instead with the card themselves. This change of focus made evident that more complex stories were impossible to be told based on the original represented data, because they described the content of the cards only - or, in other words, they were not an attempt to represent the individuals on the cards. This kind of invisibility had therefore to be dealt with by further researching on individuals and digitally representing them in such a way that the new data repository could be reused by a DIA. In the next chapter (see Chapter 6), it will be explained in detail not only about the project Marbles of Remembrance as another case study, but also about one of the tools used to enrich the ITS data repository, the platform Artfacts.

5.1.3.4 Conclusion

The ITS case makes it a good example of workflows and archival traditions that are translated into Information Systems, which then reinforce digitally these workflows and traditions by means also of the inner technological structures that constitute these systems. The ITS is an institution built around of its archive. Most of the workflows in place were implemented to deal with thousands of historical documents. The advantage of the OuSArchiv is to enforce procedurally that documents are organized, researched, and processed in specific ways. By means of its Relational Model and User Interface, experienced or inexperienced employees must adapt themselves to the system in order to be able to follow the necessary requirements to fulfill their duties. According to Christian Groh, some employees are responsible for only scanning the physical documents. These individuals, no matter if unaware of the historical and contextual importance of the documents, are nevertheless guided by the procedural rhetoric of the system to input the scanned files.

---

92 Such as the dataset Stolpersteine Berlin: https://datenregister.berlin.de/en/dataset/liste-der-stolpersteine-berlin
93 Interview (see Section ??).
94 Term coined by Ian Bogost to define a form of rhetoric that happens procedurally as a result of interaction [24].
to the database, and contribute therefore to many aspects of the institutional mandate of the ITS without the need of certification in e.g. historical research.

The digital materialization of procedures and organization that happens through the practical instantiation of abstract constructs, such as Software, is a well-know issue discussed by authors such as Paul Leonardi\(^{95}\) [135], Wanda Orlikowski\(^{96}\) [178], and Paul Dourish\(^{97}\) [71]. The Relational Model of the OuSArchiv, which was not designed to handle frequent changes\(^{98}\), and the strong connection between the digital archive to the physicality of thousands of documents organized hierarchically, rigidify in a certain way the institution. In addition, the OuSArchiv, as any other conventional Information System, impose its structural form and procedural demands to the institution, mediating organizational and human relations. As Dourish points out when referring to databases, “the materialities of that model make themselves visible through the ways in which the model drives the institutional action” [71, p 136]. On one hand, such solid rigid structure provides the consistency necessary to enforce that

---

95 Leonardi [135] makes an important observation when he points out to the fact that the concept of materiality is not always directly and intrinsically attached to physicality. By exploring the question “can digital artifacts have materiality?” in the context of organizations, he points out that the materiality can assume other definitions than the one regarding exclusively to physical objects. The author quotes the Oxford English Dictionary as it offers definitions of materiality as being expressed by practical instantiation. Materiality as an expression of practical instantiation refers to an applied abstract or theoretical construct that becomes material when put into practice, which is defined as the process of instantiating. As an example, Leonardi [135] talks about procedures encoded into software to ensure the quality of products by engineers. The software guides engineers to go through the necessary steps in order to make sure a product is completed as specified. By following the procedure, engineers materialize what before was purely abstract. The author writes: “If material is defined as having physical matter, the software is not material. But under the second definition of material, the software clearly helps to instantiate the abstract idea of management.” [135]

96 According to Orlikowski, “it is only when repeatedly drawn on in use that technological properties become constituted by users as particular rules and resources that shape their action” [178]. In this sense, dormant functionalities of digital artifacts exist as encoded constructs in some kind of medium, but are not materialized until they are run by the computer system. Therefore, the materiality of the software comes from its significance, since unimportant features are not instantiated. In addition, the same digital artifact can be used in many distinct ways depending on how users employ it. Therefore, a digital artifact could potentially be materialized in numerous and different ways, depending how the set of functionalities are combined and put into practice.

97 As extensively discussed in this chapter.

98 As explained in Section 5.1.2.
important procedures are carried out, but on the other hand, the system is unable to quickly accommodate the new workflows and content being created by e.g. the ITS’ RnED.

When considering specifically the employment of data repositories for hackathons, the traditional purpose of datasets as catalogues is shifted to the one of a raw digital material from which narratives can be created through the fabrication of DIAs. As discussed previously, the reinterpretation of heritage happens not only based on the information present in the dataset (in terms of quality, expressivity, and content), but also based on the missing information. For example, the dataset *Karten der Schülerkartei* is important, because it enables stories to be told about school children who lived during the Nazi Regime. Otherwise, the lack of such dataset would contribute to the invisibility of such topic. The existence of Data Repositories is therefore an important kind of not only institutional, but also historical and social memory, whose value surpass Finding Aids when used as a creative material.

---

99 Section 5.1.2
The Hackathon, as a method for the interpretation of heritage, requires from Cultural Institutions and their constituents the adoption of principles from different domains that, at first glance, may seem contradictory to the traditional practice of institutions, because of their agile, audacious, and experimental nature. For example, core hacking principles are intrinsically related to the notion of appropriation of tangible and intangible materials, their remix, and resignification\(^1\). In addition, open-source principles require free access and distribution of digital content so that an ecosystem can be created and profit from collective, collaborative, and decentralized contributions\(^2\). Finally, platform principles concerning standardized and modularized infrastructures are vital to the conceptualization, operationalization, and fabrication of these contributions\(^3\). Aware of the benefits such principles can bring to institutions\(^4\), a series of efforts in the Cultural Sector are being employed to articulate and negotiate the loss of authority and control\(^5\) over the intellectual property of Cultural Institutions that these principles push for. Besides projects that make use of third-party platforms of participation to invite the public to co-create\(^6\), public licenses, such as Creative Commons\(^7\), are currently being applied to Digital Collections available online\(^8\) enabling the fast circulation and appropriation of digital materials by granting them with automatic permissions for reuse. This paradigm shift can be also observed by the appearance of other Cultural Institutions that focus especially in issues related to cultural data. The Open Knowledge Foundation\(^9\), an institution built around the idea that Knowledge, Information,

\(^{1}\) See Chapter 2.
\(^{2}\) See Section 4.2.
\(^{3}\) See Section 4.1 and Chapter 5.
\(^{4}\) See Section 2.1.3.
\(^{5}\) See Section 3.1.
\(^{6}\) See Section 3.1.2.
\(^{7}\) https://creativecommons.org/
\(^{8}\) https://creativecommons.org/tag/europeana/
\(^{9}\) https://okfn.org/
and Data should be free of restrictions\textsuperscript{10}, has partnered with the European Commission\textsuperscript{11,12} to co-found the OpenGLAM initiative\textsuperscript{13,14}, whose goal is to help “cultural institutions to open up their content and data through hands-on workshops, documentation and guidance” [177]. In practical terms, the OpenGLAM initiative promotes the awareness in concern to the benefits institutions can obtain if they change their mindset implement licenses and technologies that allow the free distribution and access to Digital Collections via platforms, channels, and services. APIs and online data repositories that are based on open and machine-readable\textsuperscript{15} file formats are a few instances of such technologies, which are at the core of e.g. Hackathons. It is no coincidence that the Open-Knowledge Foundation is one of the responsible bodies for organizing annually the Hackathon Coding da Vinci (CdV).

\subsection*{6.0.1 Two-speed IT Architecture}

These methods, projects, and initiatives that aim at collaborating with communities and finding new creative meanings and usages to Digital Collections are evidence that Cultural Institutions are not monolithic entities. However, as demonstrated in the previous chapter\textsuperscript{16}, inflexible infrastructures are not only a reality, but also a necessity since they enforce straightforward workflows and solid practices. Flexibility and inflexibility, openness and closeness, novelty and tradition however do not need to be dealt in binary terms, but they can instead co-exist. Johan Oomen et al. [218] build on the ideas of Oliver Bossert et al. [29] to propose a Two-speed IT Architecture that is adapted to the Cultural Sector. This Two-speed IT Approach consists of adopting two different digital strategies - slow and fast - in order to preserve the stability of core institutional practices, but

\begin{itemize}
  \item The core values of the foundation are based on the Open Definition, which states that “A piece of data or content is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to give credit to the author and/or making any resulting work available under the same terms as the original work” [177].
  \item https://ec.europa.eu/commission/index_en
  \item As part of the DM\textsuperscript{2}E project (see https://pro.europeana.eu/project/dm2e).
  \item https://openglam.org/
  \item The acronym “GLAM” refers to Galleries, Libraries, Archives, and Museums, but it is applied to indicate memory institutions in general.
  \item “Data in a data format that can be automatically read and processed by a computer, such as CSV, JSON, XML, etc. Machine-readable data must be structured data, such as CSV, JSON, XML, etc.” [174]
  \item See Chapter 5.
\end{itemize}
at the same time innovate by adapting and responding faster to the needs of communities. According to Johan Oomen et al. [218, p.51]:

- the Slow-speed IT Strategy consists of "standardized and off-the-shelf solutions that are used to secure 24/7 service. The solutions are updated following service-level agreements with suppliers. In the heritage domain, good examples are systems for managing storage, cataloguing, play-out and ordering. Given the impact, the frequency of updating applications in the ‘slow’ ecosystem is measured in months or years rather than weeks."

- the Fast-speed IT Strategy consists of "tailor-made solutions that cater to very specific user requirements and are used to experiment with new technologies. Opposed to systems that are ‘core’ (for instance the storage systems), applications developed in the ‘fast’ speed do not have very stringent requirements regarding stability and minimum ‘uptime’(i.e. they are in some cases maintained by developers themselves). For instance: experimental visualizations of datasets, automatic metadata extraction services and online magazines linked to current exhibits."

In technical terms, Oomen et al. [218] include under the definition of a Slow-speed IT strategy Information Systems responsible for storage and cataloguing, which, in the case of Cultural Institutions, are responsibilities delegated to Collection Management Systems (CMSs). On the other hand, the Fast-speed IT Strategy enables experimentation through the development of tailor-made solutions that are built upon e.g. online data repositories and APIs (see [218, p.50]). It is indeed possible to see examples of this Two-speed IT Architecture Approach in the Cultural Heritage Sector in case of institutions that e.g. adopt the guidelines proposed by the OpenGLAM initiative and take part in Hackathons.

Taking the case of the International Tracing Service (ITS) as focal point, it is observed a few characteristics of the Two-speed IT Approach being implemented in different departments of the institution. The Slow-speed IT Approach can be seen in the work of the Archival Description Department that works closely with the software company Ossenberg Digitalisierung und Software GmbH in order to develop and implement consistent methodologies for improving the search and retrieval capabilities of the CMS OuSArchiv, which does not accommodate well structural and frequent changes. This department and the OuSArchiv are fundamental for the execution of the main mandate of the ITS as a tracing service and archive.

\[17\] See Section 5.1.2.
\[18\] See Section 5.1.3.1.
As explained in chapter 5, especially when memory institutions start to engage with educational projects, greater flexibility in comparison to the one provided by the slow-paced infrastructure of CMSs is required, because educational projects need to promptly respond to the demands and communicate with heterogeneous communities, instead of modeling and enforcing internal workflows. In this sense, in parallel to the slow-paced Archival Description Department, the recently created Education and Research Department (RnED) develops its own independent IT projects, such as the e-Guide, which offers extra interpretive content in order to historically contextualize archival documents. As previously explained, the e-Guide has been developed on top of an independent and segregated IT infrastructure - an Open-source Content Management System (CMS) for constructing websites called Typo3. The textual material has also been produced independently by the RnED without the need to conform to e.g. the data model of the OuSArchiv, but instead the one imposed by Typo3. Because Typo3 is too specific for the requirements of building websites, it is unlikely that the data created for the e-Guide can be easily reused on other platforms.

Besides the IT projects such as the e-Guide, a much more radical Fast-speed IT Approach was carried out when the ITS decided to take part in the CdV. Especially in regard to Hackathons for Cultural Heritage, which are based on Digital Collections, fast-prototyping is achieved by modular technology that provides a platform through which the instantiation of new concepts is derived from the arrangement and rearrangement of data elements. Specifically in the case of Digital Collections, platform principles are expressed by the data model that underlies data structures, and data elements (building blocks) are the individual machine-readable data units that conform to the specification of the data model. The ITS has provided the datasets in two different machine-readable file formats, namely EAD-XML and CSV. The projects created during the Hackathon do in some degree conform to the data model of the OuSArchiv, since the datasets used in the projects were produced by the export capabilities of this CMS. In this sense, greater interoperability is

---

19 See Section 5.1.3.3.
20 [https://typo3.org/](https://typo3.org/)
21 [https://www.loc.gov/ead/](https://www.loc.gov/ead/)
23 Interoperability is defined as “The ability of computer systems or software to exchange and make use of information.” [244]
present between the OuSArchiv and the applications developed on top of the datasets generated by it.

![Typo3's Administrative User Interface](image)

**Figure 32: Typo3’s Administrative User Interface**

Both Fast-speed IT Approaches discussed above present advantages and disadvantages. Typo3, for instance, does not provide consistency and interoperability in regard to the core institutional information architecture in place, but it does provide a cheap solution in terms of cost, technical knowledge, and implementation time. Typo3 does not require advanced programming skills for creating websites. It offers instead an UI that enables non-programmers to add predefined and ready-to-use modules (such as pages, views, lists, etc) to the main structure of a website (see Fig. 32 - Content Menu and Navigation Frame). Websites can therefore be built mostly in a visual way. In addition, further modifications structure- and content-wise can be performed with ease, meaning also that no technical skills are needed. However, Typo3 does not serve to general purposes, but it is too specific for building websites. In Hackathons, on the other hand, it is expected that hackers work directly with datasets to built tailor-made applications that are capable of expressing the value of the collections by recontextualizing them. One of the advantages of this approach is the greater interconnectivity between the core institutional information architecture and the applications created on top of it. However, because of the complexity involved in working with datasets, individuals without technical skills will necessarily depend on programmers to conceptualize and implement such applications.
In view of the problems presented here and in the context of a Two-Speed IT Infrastructure aimed at supporting the Cultural Sector, this chapter will propose a platform, called Artfacts, which follows the low-code/no-code principle, which mitigates problems such as the technical background required to work with Digital Collections in order to deploy Digital Interpretive Artifacts (DIAs). As demonstrated, Hackathons are multidisciplinary events that gather individuals from a variety of backgrounds, including curators and researchers of Cultural Institutions who do not have training in programming. Therefore, the Artfacts’ Data Model (ADM) aims at user-friendliness and pragmatism in regard to the employment of datasets as raw material for storytelling. Although its core presents low granularity, the ADM is interoperable, because it is based on widely adopted Data Standards. In addition, the ADM can be easily extended through strategies offered by the platform’s GUI. In concern to its GUI, the Artfacts Platform aims at greater universality, instead of being too specific to a particular application. The cartography of information through hierarchical and rhizomatic networks offers an experimental approach to the data-driven framework of institutions, because it does not impose workflows. The combinatorial capabilities provided by the modular approach of its GUI enables not only greater flexibility in terms of applicability, but also serves as a tool for the communication and analysis of Digital Collections’ contents. The next paragraphs, the characteristics of the platform mentioned above will be discussed in detail.

The low-code/no-code principle, used by e.g. Typo3, is an interesting approach for enabling non-programmers to develop (web) applications.

---

24 Intended for developers, the low-code principle proposes that software development can be accelerated and facilitated through the configuration of modules, instead of purely traditional programming. Intended for “citizen developers” (non-programmers), the no-code principle proposes the development of applications done purely through the configuration of modules and the GUI (see [23], [194]).

25 See Sections 2.1.2, 6.3.2, and A.1.1.1.

26 Granularity is the level of depth that a given data model is able to represent information. “Granularity concerns the size of the elements that are the objects of attention and action within the system—and, in turn, with the entities they might represent.” [73]

27 See discussion in Sections 5.1.2.2, and 5.1.3.

28 Instead of being too specific to a particular application, as in the case of Typo3.
6.1 Conceptualization and Implementation of Artfacts

The following sections present and discuss the conceptualization, implementation, and evaluation of a fast-speed Platform called Artfacts, which was designed within the context of the Two-Speed IT Framework developed by Oliver Bossert et al. [29] and adapted by Oomen et al. [218], where a foundational, stable, and slow infrastructure is complemented by an additional creative, experimental, and Fast-speed IT approach capable of promptly responding to the needs of communities. This platform is an attempt to digitally incorporate some of the principles discussed in this thesis and to mitigate problems concerning specialized knowledge required for manipulating Digital Collections and profiting from their affordances as a creative material. In this sense, through the cartography of information, the platform aims at widening the participation of individuals with no technical background in the development and maintenance process of DIAs, no matter whether within Cultural Institutions or events such as Hackathons for Cultural Heritage. Artfacts intermediates the reinterpretation of cultural datasets and the fabrication of interpretive applications by means of a flexible, general, and interoperable Data Model that is able to adapt to the demands of storytellers, and an open-ended Object-Oriented User Interface (OOUI) that enables analysis and experimentation by rearranging data elements into digital narratives.

6.1.1 Overview: Problems, Requirements, Solution, and Target-Groups

6.1.1.1 Problems and Requirements

Although current CMSs are not suitable for properly supporting the interpretation of heritage [29], these systems are commonly used to produce Digital Collections. Belonging to a well-established and traditional Slow-speed IT Infrastructure in the Cultural Sector, CMSs are designed to administer and organize core institutional assets and support institutional workflows. CMSs focus on tasks, such as the digital description of cultural objects, their systematization and indexing, their search and retrieval within large data repositories,
among others\textsuperscript{30}. Therefore, their central relevance cannot be ignored. As core administrative and organizational tools, CMSs restrict themselves to a pre- and well-defined set of description fields arranged in forms\textsuperscript{31}. Being able to enforce workflows and achieve pre-defined goals are important requirements of these systems. However, the requirements of Information Systems to satisfy to the principles of Heritage Interpretation proposed by Tilden\textsuperscript{32} must explore the range of combinatorial possibilities afforded by data objects, whose arrangement will depend largely on what is relevant given the current social context in which the story will be told. This combinatorial and experimental approach is therefore incompatible with pre-defined outcomes. Such Information Systems must afford open-endedness instead.

The invisibility\textsuperscript{33} rendered by commonly used Data Standards for Cultural Heritage, as discussed in Chapter 5, together with the limitations of CMSs and the Digital Collections they produce hinder the digital interpretation of heritage. As Tilden explains, “information as such is not interpretation” [227, p.18], but “interpretation is an art, which combines many arts” [227, p.18]. Therefore, Information Systems that aim at serving as a tool for Heritage Interpretation need to go beyond over focusing on the interoperability of the data model and the enforcement of workflows\textsuperscript{34}. As storytelling depends largely on the storyteller, no predetermined data model is capable of predicting all elements of a story. Therefore, the system must present the individual with data-driven tools in order to enrich the model in ways that creative and artistic expressions, and visibility to ignored topics are enabled and able to provoke critical thinking on others. In this sense, it is proposed that the technological means for affording Digital Heritage Interpretation is via the maximization of usability, speed and flexibility in which cultural data can be reused, recontextualized, and enriched by storytellers. The maximization of usability, speed, and flexibility can be achieved by:

- **General and Extensible Data Model**: as discussed in the previous chapter (see Section 5.1.1.1), the general strategy used by the Dublin Core (DC) presented both advantages and disadvantages. The DC’s abstract model was designed to be simple and generic, and therefore user-friendlier and of easier adoption if compared to other Data

\begin{footnotesize}
\begin{itemize}
\item See Section 5.1.2.1.
\item See Section 5.1.2.2.
\item See Chapter 3.
\item See Section 5.1.2.
\item See Section 5.1.2.2.
\end{itemize}
\end{footnotesize}
Standards. However, it lacked expressivity and granularity for a more
detail representation of objects. A data model that aims at usability
and flexibility should follow the simple and general approach of the
DC’s abstract model\textsuperscript{35}. But, it must provide also user-friendly
strategies for its expansion when necessary. The objective of the data
model proposed here is not to become a standard in the Cultural
Sector, but provide some degree of compatibility and interoperability
with Data Standards adopted by the industry. In this sense, the
extensions to the proposed data model should obey a pragmatic
approach serving to model digital narratives that are intended to
be consumed through DIAs. Interoperability is therefore in second
plan. This pragmatic approach follows the principles of the Fast-speed
IT Infrastructure proposed by Johan Oomen et al., where the IT
infrastructure is based on “tailor-made solutions that cater to very specific
user requirements and are used to experiment with new technologies.” [218]

• Data-driven and OOUI: the data model described above must be
accompanied by a Human-Data Interface that is able to provide
storytellers with the necessary means for not only the analysis and
recontextualization of data objects, but also the extension of the data
model. The main goal of such a data-driven and object-oriented
approach to the GUI is not to enforce workflows, but provide an
open-ended possibility space where narratives are generated from the
objects and their behaviors\textsuperscript{36}. This OOUI must present a low learning
curve so that the application can be used by individuals with no
technical background.

\textsuperscript{35} As found during our Workshop (see A.1.3.1), the higher the expressivity/granularity
of the data model is, the higher is the level of difficulty of the application.
Consequently, a longer amount of time will be needed to create digital
representation. In addition, as the focus on the platform is to create digital
representations that are good enough for being employed in applications, there
is no need for a highly granular and expressive core (data model). The need for
extensibility is however necessary for the reasons explained in this section.

\textsuperscript{36} As explained in Section 5.1.2.2.
6.1.1.2 Solution

Based on the problems and requirements discussed above, the proposed platform Artfacts does not aim at replacing existing CMSs\textsuperscript{37} (see Collection Management System on Fig. 33). The platform aims instead at providing an extra semantic layer on top of the existing infrastructure that enables institutions to create new or (re-)interpret, reuse, and deploy enriched data repositories (see External Data Sources on Fig. 33 and Fig. 34) that fulfill the requirements necessary for creating data models capable of expressing compelling stories about heritage that can be rendered as DIAs (see Digital Interpretive Artifacts on Fig. 34), such as audio-guides, chatbots, interactive multimedia exhibits, augmented reality apps, and so on.

\textsuperscript{37}As found during our Workshop (see Section A.1.3.3), Cultural Institutions have invested a considerable amount of money on their CMSs. In addition, the staff is already accustomed to using them. Therefore, institutions are not willing to replace these systems easily. In addition to that, CMSs serve another purpose than the proposed platform. They are not a part of the creative process within Cultural Institutions. As stated by one of the workshop participants, the CMS “is a database for collections, and not a database for ideas” (see Section A.1.3.3).
The Artfacts Platform was conceptualized within the boundaries of a Fast-speed IT Infrastructure (see Fig. 33), not being therefore a "core" institutional system, but an add-on that offers a pragmatic approach to the way Digital Collections can be reutilized. Artfacts' main goal is to afford the production of digital storytelling through DIAs. Therefore, Artfacts is not a system to be used to look up information resources, but to enable the rapid creation of new lightweight data models that can reuse and recontextualize the information from external data sources, which are imported into the system as Knowledge Graphs (see Data Storage Format Fig. 34). In addition, Artfacts also presents strategies for quickly converting unstructured narratives into structured representations of stories through its Tagging System, and reutilize existing data objects in new contexts providing new points of view.

The analysis and (re-)interpretation of data is done through the cartography of information as flexible object-oriented data structures called Knowledge Maps (see Fig. 35 and OOUI on Fig. 34). At the heart of Artfacts, Knowledge Maps (see Section 6.1.2) are semantic-rich network diagrams based on the Schema.org's Vocabulary and the Simple Knowledge Organization System (SKOS).

---

38 For more information on Knowledge Graphs, see Section 6.1.3.1.
39 See Section 6.1.3.3.
40 See Section 6.2.
41 For more information, see Section 6.1.3.
Knowledge Maps do not enforce workflows, but instead present the user with a possibility space\textsuperscript{42}, in which scenarios are composed by the objects and their functionalities. Creativity is enabled by the Knowledge Maps’ combinatorial capabilities, which are granted by the freedom and dynamism provided for modeling cultural objects and the relationships among them. Besides serving as tools for the (re-)interpretation of cultural information, Knowledge Maps can also be directly used as interfaces to organize and control the flow of system events and user actions of third-party applications\textsuperscript{43}. This universal Human-Data Interface, which is used to respond to fast and dynamic customization of apps, present a low learning curve requiring therefore no technical skills for their manipulation\textsuperscript{44}.

6.1.1.3 Target-Groups

Finally, Artfacts’ methods may serve institutions’ personnel, communities, and audiences as follows:

- (Curators, Researchers, Educators, Designers)/Hackers - through an intuitive OOUI crafted to support the reinterpretation,

\textsuperscript{42} See introduction to Chapter 5.
\textsuperscript{43} See Section 6.2.
\textsuperscript{44} See Usability Test in Section 6.3.1.
production, and reuse of cultural content by means of its cartography (through layer Object-Oriented User Interface on Fig. 34);

• **Programmers/Hackers** - through a comprehensive API working on top of a consistent and flexible data model, which provides interoperability and extensibility enabling the linkage and management of custom third-party applications (through layer Data Storage Format and External Data Sources on Fig. 34);

• **Audiences of Cultural Institutions** - as indirect target group, through a set of ready-to-use, easily adaptable, and cost-effective DIAs, which are powered by new data assets or/and existing data repositories (through layer Digital Interpretive Artifacts on Fig. 34).

The next section will describe the advantages and justify the use of Knowledge Maps as tools for aiding the analysis, manipulation, and (re-)interpretation of data as digital narratives.

### 6.1.2 The Knowledge Map

In the context of Cultural Institutions, the significance of graph organizers can be summarized as follows: to begin with, graph organizers provide an useful framework for sensemaking in that they can be applied as e.g. methods for the organization and analysis of some aspects of curatorial research and planning, and tools for the production and communication of complex information within internal meetings or public spaces45. Their representational capabilities range from the simple hierarchical classification of information to the complex articulation of ideas through a variety of semantic connections among different topics.

---

45 As found out during our interviews (see ??) and workshops (see A.1.3.2 and A.1.3.3).
In museums, for instance, the interpretation of objects depend not uniquely on the objects themselves, but also on the events, people, actions, places, etc that enabled and surrounded their existence. During the workshop *Expertengespräch Vermittlungskonzept Kogge-Halle* (see Appendix A.1.3.2), Dr. Gregor Rohmann, a historian working in the conceptualization of the didactic mediation concerning the new *Bremer Kogge* exhibition in the *Deutsches Schifffahrtsmuseum* (Bremerhaven, Germany), created a series of network diagrams (see Fig. 36) to support the research and explanation of case studies concerning how different historical actors related among themselves to trade during the *Hanseatic Era* (see Fig. 36), the historical context in which the *Bremer Kogge*, the actual museum artifact, belonged to.

---

46 [https://www.dsm.museum/](https://www.dsm.museum/)
Mapping techniques support the interpretation process by allowing this complexity to be dealt within well-defined boundaries. They offer a unique interpretive layer as they, on top of the collection and the information that contextualizes it, form a constellation of interrelated informational elements, which is used to mediate the comprehension of objects. They invite the viewer to think about how elements relate, engage with stories by following multiple paths, and discover unforeseen connections (see Fig. 37). Such visual representations have the power to explain complexity in a way that could not be represented otherwise [108]. Graph organizers make explicit the binding properties, they provide a bird’s-eye view perspective on the multiple objects of a collection that is digestible to the human eye, and, as a creative tool, provide the necessary support for contextualizing and interpreting collections.

6.1.2.1 Background

Graph organizers: historical perspective

Manuel Lima [140], in this book called The book of trees: visualizing branches of knowledge, analyzes more than eight hundred years of the usage of tree diagrams as tools for organize and visualize information. According to Lima [140, p. 44], Aristotle (384-322 BC), one of the greatest and most influential philosophers who greatly influenced the western thought, proposed a structured system with
general categories that served as a framework to classify all things in the world. The Aristotelian syllogism, which laid the foundations of e.g. First Order Logic, enabled ancient Greek philosophers to execute logical inferences for ontological arguments. This Aristotelian method for reasoning required the categorization of things of the world by the objective assessment of their shared characteristics. According to Robinson et al.:

This model is characterized by vertical and fixed linkages, and binary choices, and by the linking of elements only of the same general nature... A concept must typically fit into one and only one place in a classification scheme, and the hierarchical divisions must be made by a single criterion, and must be mutually exclusive. [197]

Porphyry (234-ca.305 AD), as an attempt to simplify and introduce beginners to the Aristotelian categories, produced a textbook called Isagoge, which contained a hierarchical scheme of classification known as Porphyrian tree (see Fig. 38). According to Lima, the tree diagram “essentially portrays the basis of Aristotle’s proposition in a memorable, easy-to-grasp, treelike visual scheme.” [140]

Figure 38: Porphyrian Tree

[Picture source: https://upload.wikimedia.org/wikipedia/commons/f/f1/Arbor_porphyrrii_%28from_Purchotius%27_Institutiones_philosophicae_I%2C_1730%29.png]
By serving as a device for logical argumentations based on major (on the top) and minor (on the bottom) premises to draw conclusions, the hierarchical form of the tree offered the necessary visual aid for logical and sequential reasoning. Porphyry created a rudimentary tool to support individuals in information processing. Tree diagrams are powerful communication tools to represent a great number of systems of knowledge. In the book called Arbor Scientiae, Ramon Llull (ca.1232-ca.1315) was one of the first to compile the representation of sixteen scientific domains using trees as a metaphor [140, p. 36]. Although Llull’s intention was representational only, his diagrams also served as devices for processing information. As Tom Sales points out:

Llull connected his ‘basic concepts’ with lines, and prescribed that the lines had to be followed to combine the concepts and derive the consequences. This was new. Not now, though; we have a name for the device Llull invented: we call it a graph. [202].

Also Charles Darwin used the tree diagram in this book On the Origin of Species by Means of Natural Selection to represent this evolutionary tree of life. According to Lima, because of their hierarchical characteristics, the tree metaphor is “a foundational system for the organization” [140, p. 47].

In another book called Visual Complexity: mapping patterns of information, Lima [139] argues that with the increasingly complexity of contemporary society, tree diagrams are in many cases not suitable to model and represent phenomena in the world, because of their extreme hierarchical simplicity. Especially in regard to modern science and the characteristics of problems scientists wish to solve, Lima [139] remembers a landmark paper written by the mathematician Warren Weaver, where the author divides modern science into three different periods. According to Weaver [235], there are:

- Problems of Simplicity: this phase ranged from the seventeenth to the nineteenth centuries and is characterized by problems concerned with two-variables, and how one variable affects the other.
- Problems of Disorganized Complexity: this phase occurred in the twentieth century and is characterized by problems concerned with multiple variables. However, the interaction between these variables was considered to be random and therefore not definable.
Problems of Organized Complexity: this is the current phase we live in. Scientists are challenged to consider multiple variables as well as to understand how they depend on and interact with each other.

In the phase of organized complexity, it is primordial to think about systems of knowledge representation that are able to not only model multiple variables, but also depict explicitly interactions (or the relationships) that exist in between them. Lima points out to the rhizomatic model, as proposed by Deleuze and Guattari in the context of theory and research, as a post-modern and post-structuralist thinking reaction to an authoritarian hierarchical model of the tree that does not embrace multiplicity [139, p. 44] and impose determined ways of thinking. The network, as a typology of the rhizomatic model, is a flexible and non-hierarchical structure that allows decentralization and flexibility of representation. As also pointed out by Robinson et al. [197]:

The rhizome concept allows, indeed requires, non-hierarchical linkages, made pragmatically as they are needed, horizontally or across any number of levels, and linking elements of disparate nature when appropriate, crossing categories. [197]

KNOWLEDGE MAPS: CONTEXTUAL PERSPECTIVE

Graph organizers such as Mind Maps, which are defined as tree diagrams, together with Concept Maps and Knowledge Maps, which are defined as network diagrams, are used to structure and make sense of information. The main characteristic used to distinguish them both is how the elements are connected among themselves (see Fig. 39). Once organized into a graph, the set of data elements may gain new meanings that are derived from the contextual reasoning offered by the topology of these organizers. Spatial arrangement and meaningful links in between data can be a powerful semantic strategy for sensemaking, which is defined by Gary Klein et al. as “a motivated, continuous effort to understand connections” [123]. In addition, visual elements such as color, position, shape, and so on [43], when well-encoded, enable higher comprehension of information and, therefore, influence human’s cognitive workload [28]. Graph organizers are therefore important tools to support a wide variety of tasks in the context of either formal [67] or informal educational environments [144].
Knowledge Maps should be understood in the context of other tools for representing, organizing, and making sense of information, namely Mind Maps and Concept Maps. The concept of Mind Maps was originally proposed by Tony Buzan [35] as a cognitive tool in the shape of a hierarchical diagram to support mental tasks related to human cognitive processing of information. Therefore, it is also a method used especially in educational environments. These diagrams are also known as spidergrams or spidergraphs because of their appearance. In most of the cases, these diagrams are focused on one main central topic. Subtopics are branched out from a parent topic. And the information represented by these diagrams is structured into levels that represent higher and lower statuses. These are classic tree diagrams in the way they deal with information. They are hierarchical systems of classification in which entities are separated by shared characteristics. The branches in Mind Maps are unlabeled and non-directional [104]. Mind Maps have been used during decades especially as a lightweight cognitive tool for fast brainstorming, note taking, organizing and communicating processes, solving problems, making decisions, and so on ([230], [80]). Although they present very low semantics, they are nevertheless useful tools for demanding cognitive tasks that are crucial in educational and business environments [230]. Especially in education, Mind Maps have been proven a valid approach in supporting assimilation of concepts. As explained by Dhindsa et al. [67], these diagrams “significantly improved information organization in students’ cognitive structures when these gains were compared to those in a classroom where traditional teaching style was used” [67]. Finally, another characteristic of Mind Maps is the use of visual elements, such as colors, shapes, and pictures in order to improve memorability of the represented content (see [242], [79]).
Offering a more flexible and powerful approach for representing knowledge, Concept Maps are network diagrams where relationships do not need necessarily to conform to a hierarchical topology only. These diagrams were first conceptualized by Joseph Novak [166] with the goal to depict relate complex concepts in order to improve understanding of a bigger topic. As Novak points out, the Concept Map “serves as a kind of template or scaffold to help to organize knowledge and to structure it” [166]. As a methodology, one should start the construction of the map from general concepts down to specific examples. This approach used by Concept Maps for organizing information is called top-down [76]. The relationships (also named as edges, connections, cross-links, or simply links) connecting different concepts (or nodes) are of higher importance in Concept Maps if compared with Mind Maps. As a network diagram, concepts are not arranged only by means of hierarchies. Therefore, it is recommended, although not necessary, that the relationships are attributed values to inform how exactly the concepts relate to each other. According to Novak, cross-links are very important, because they “often represent creative leaps on the part of the knowledge producer” [166]. Jon Kolko [126] points out that:

*The Concept Map itself represents the creators’ mental model of a concept, but it also informs and shapes that mental model during creation, as it allows designers to see both the holistic scale of the concept and also critical details within the concept.*
As it affords action-based understanding at both a gross and fine level, both its creation and its usage become tools for sensemaking [126].

Concept Maps also offer a conceptual framework for constructivist assimilation, promoting therefore meaningful learning, because they are effective tools for linking new with old information [242]. Links in Concept Maps may be directional or non-directional, labeled or unlabeled.

Finally, presenting greater similarity to Concept Maps, Knowledge Maps are visual graph organizers that work as a method for supporting the human cognitive handling of numerous and diverse sets of concepts. Ria Hanewald et al. [104] defines Knowledge Maps as “a graphical display of information in which the importance and relationships between the various elements of knowledge are portrayed in the form of a map” [104]. Angela O’Donnell et al. [171] points out that Knowledge Maps contrast with Mind Maps and Concept Maps because of the use of “a common set of labeled links that connect ideas. Some links are domain specific whereas other links are more broadly used. Links have arrowheads to indicate the direction of the relationship between ideas.” [171] As Concept Maps, Knowledge Maps obey a rhizomatic and open-ended model where their links are directional and labeled. Most importantly, their unique characteristic is that they require a standardized vocabulary, and do not have predefined starting and ending nodes [171].

6.1.3 The Artfacts’ Data Model

6.1.3.1 Model’s Foundation

Artfacts’ Data Model (ADM) was designed with four main considerations in mind. Firstly, the ADM needs to be able to support the implementation of digital versions of Knowledge Maps and preserve their capacity of providing a framework for action-based understanding i.e. helping individuals to organize and make sense of cultural information. Secondly, the ADM needs to be machine-readable and interoperable so that mapping between other Data Standards are possible. Thirdly, the ADM needs to be adaptable to the specificities of storytellers. And finally, the ADM should support production and management of DIAs. In order to comply with these requirements, the data model behind Knowledge Maps was based on the Schema.org’s Data Model, for providing a
basic and general controlled vocabulary, and the Simple Knowledge Organization System (SKOS), for providing the essential linkage structure among different data objects.

SCHEMA.ORG

Schema.org, much like DC, is a collaborative initiative that began in 2011 lead by major search engines, as for instance Bing, Google, and Yahoo “with a mission to create, maintain, and promote schemas for structured data on the Internet, on web pages, in email messages, and beyond” [59]. More specifically, the main goal was to provide a single vocabulary and an integrated schema across different sources on the web in order to improve Search Engine Optimization (SEO) [101]. In addition to helping search engines to better find and categorize online information, structured data present extra benefits: the separation between content and presentation, and automated reasoning are two features worthy of notice. From the beginning, these powerful strategies were at the core of the Semantic Web, an idea that was popularized by the seminal paper of similar name [19] written by Tim Berners-Lee in 2001. In the paper, Berners-Lee advocated the need for transitioning the Web from a purely document-based medium, into a network of interconnected data objects that were to be the basis for many different applications. The Schema.org initiative is one of the instances of the bigger vision of the Semantic Web (as well as other Data Standards), in which online information are able to be processed by machines and used e.g. to deliver contextualized answers to natural language queries, and animate several applications that interpreted the same data repositories in different ways. Google’s Rich Snippets (currently called Rich Results51) was one of the first applications of Schema.org (see Fig. 41). This feature provided the users with additional pieces of information on their search results, such as search result items containing breadcrumbs, reviews, ratings, etc. Another more recent application of the Schema.org also for enriching Google’s search results is the so-called Knowledge Graph. The Knowledge Graph52 is a “graph data model of typed entities with named properties” [101], or, in other words, a network where objects and their relationships, which are described as semantic triples

51 https://developers.google.com/search/docs/guides/search-features
52 https://developers.google.com/knowledge-graph/
(subject/node → verb/property/link → object/node) are first-class citizens. On the contrary of Schema.org, the ADM’s intention is neither to help search engines to try to “comprehend” what web pages contain, nor to be a standard vocabulary on the Web. The contribution of Schema.org to the ADM is instead to provide an understanding and a foundation of things/objects commonly represented on the Web and the relationships among them. In addition, as a Data Standard widely adopted, Schema.org also provides interoperability among many different resources on the Web that use this vocabulary, such as google.com, bing.com, yahoo.com, nytimes.com, guardian.com, imdb.com, linkedin.com, ebay.com, youtube.com, yelp.com, last.fm, myspace.com, just to name a few. In addition to that, as stated by Freire et al., Schema.org poses no obstacles to Cultural Institutions to “deliver metadata in full compliance with Europeana requirements an with the desired semantic quality” [88]. Otherwise, resources that use different vocabularies may be mapped to the Schema.org’s Data

In regard to Artfacts’ architecture, the Knowledge Graph is the model used to process and store information (Fig. 34).

Figure 41: Snippets displaying objects from Google’s Knowledge Graph

Schema.org is a popular Data Standard used to represent many objects and concepts on the Web.
Model (see [148], [74], [155]). Generic and granular, Schema.org’s vocabulary gives the possibility to describe a higher variety of different kinds of objects. By comparison, the DC Metadata Element Set and the DCMI Metadata Terms focus on describing mostly what would be equivalent to one entity type of Schema.org, which is Creative Work. Therefore, DC’s abstract model lacks the granularity necessary to represent a variety of objects that are required to model more complex scenarios. Schema.org’s core vocabulary, on the other hand, is generic and diverse. It encompasses entity types (or classes), such as [90]:

• **schema:Action**: “An action performed by a direct agent and indirect participants upon a direct object. Optionally happens at a location with the help of an inanimate instrument. The execution of the action may produce a result. Specific action sub-type documentation specifies the exact expectation of each argument/role”.

• **schema:CreativeWork**: “The most generic kind of creative work, including books, movies, photographs, software programs, etc”.

• **schema:Event**: “An event happening at a certain time and location, such as a concert, lecture, or festival. Ticketing information may be added via the offers property. Repeated events may be structured as separate Event objects”.

• **schema:Intangible**: “A utility class that serves as the umbrella for a number of ‘intangible’ things such as quantities, structured values, etc”.

• **schema:Organization**: “An organization such as a school, NGO, corporation, club, etc”.

• **schema:Person**: “A person (alive, dead, undead, or fictional)”.

• **schema:Place**: “Entities that have a somewhat fixed, physical extension”.

• **schema:Product**: “Any offered product or service. For example: a pair of shoes; a concert ticket; the rental of a car; a haircut; or an episode of a TV show streamed online”.

**SIMPLE KNOWLEDGE ORGANIZATION SYSTEM**

SKOS is a W3C\(^{54}\) recommended Data Standard Ontology\(^{55}\) primarily used for expressing the basic structure and content of

---

\(^{54}\) Founded and headed by Tim Berners-Lee, the World Wide Web Consortium (W3C - https://www.w3.org/) is one of the main international standards organization for the Web.

\(^{55}\) According to Berners-Lee, ontology in regard to Information Systems, is a “document or file that formally defines the relations among terms. The most typical kind of ontology for the Web has a taxonomy and a set of inference rules.” [19]
concept schemes [157] used in Knowledge Organization Systems (KOS), which include thesauri, classification schemes, subject heading systems, taxonomies and folksonomies [116]. SKOS itself is based on the Resource Description Framework (RDF) and the Resource Description Framework Schema (RDFS). SKOS builds on top a set of RDF properties and RDFS classes “to express the content and structure of a concept scheme as an RDF graph” [157][58]. On top of the framework of the Semantic Web, many vocabularies are structured based on SKOS’ conceptual model. Just to name a few, the UNESCO Thesaurus [59], Library of Congress Subject Headings [60], the UK Public sector vocabularies [61], the Getty Vocabularies [62], and the Europeana Collections [63].

![Figure 42: Hierarchical contextualization of concepts with SKOS](https://www.w3.org/TR/2005/WD-swbp-skos-core-guide-20051102/#secconcept)

Figure above (see Fig. 42[64]) offers an intuition of how SKOS contextualizes two entities, "animals" and "mammals", respectively in relation to each other as broader (more general in meaning) and narrower (more specific in meaning) concepts. As a graph-based model, SKOS therefore provides simple and yet consistent basis for

---

56 RDF is a framework for tagging and describing data in the Web (see [124])
57 RDF Schema provides a vocabulary so that RDF data can express e.g. classes [https://www.w3.org/TR/rdf-schema/]
58 The RDF Graph is a set of triples, which are the underlying structure of any expression in RDF. A RDF Trip consists of two types of nodes, namely subject and object, joined by a relationship, which is called the predicate or property. The RDF Graph is labeled and directed.
59 http://skos.um.es/unescothes/
60 http://id.loc.gov/authorities/subjects.html
61 http://standards.esd.org.uk/
62 http://vocab.getty.edu/
63 https://www.europeana.eu/portal/en
64 Picture source: https://www.w3.org/TR/2005/WD-swbp-skos-core-guide-20051102/#secconcept

131
a the organization of a decentralized Knowledge Graph, which can be comprised of [200]:

- **Concept Class**
  - `skos:Concept`: “An abstract idea or notion; a unit of thought.”

- **Labeling Properties**
  - `skos:prefLabel`: “The preferred lexical label for a resource, in a given language.”
  - `skos:hiddenLabel`: “A lexical label for a resource that should be hidden when generating visual displays of the resource, but should still be accessible to free text search operations.”
  - `skos:altSymbol`: “An alternative symbolic label for a resource.”

- **Documentation Properties**
  - `skos:note`: “A general note, for any purpose.”
    - `skos:definition`: “A statement or formal explanation of the meaning of a concept.”
    - `skos:scopeNote`: “A note that helps to clarify the meaning of a concept.”
    - `skos:example`: “An example of the use of a concept.”
    - `skos:historyNote`: “A note about the past state/use/meaning of a concept.”
    - `skos:editorialNote`: “A note for an editor, translator or maintainer of the vocabulary.”
    - `skos:changeNote`: “A note about a modification to a concept.”

- **Semantic Relations**
  - `skos:semanticRelation`: “A concept related by meaning.”
    - `skos:broader`: “A concept that is more general in meaning. Broader concepts are typically rendered as parents in a concept hierarchy (tree)” (see Fig. 42).65
    - `skos:narrower`: “A concept that is more specific in meaning. Narrower concepts are typically rendered as children in a concept hierarchy (tree)” (see Fig. 42).66
    - `skos:related`: “A concept with which there is an associative semantic relationship.”67

---

65 `skos:broader` is a transitive property. The properties `skos:broader` and `skos:narrower` are each other’s inverse.

66 `skos:narrower` is a transitive property. The properties `skos:narrower` and `skos:broader` are each other’s inverse.

67 `skos:related` is a symmetric property.
• Collections of Concepts
  – skos:Collection: "A meaningful collection of concepts."
    » skos:member: "A member of a collection."
  – skos:OrderedCollection: "An ordered collection of concepts, where both the grouping and the ordering are meaningful."
    » skos:memberList: "An RDF list containing the members of an ordered collection."

• Concept Schemes
  – skos:ConceptScheme: "A set of concepts, optionally including statements about semantic relationships between those concepts."
    » skos:inScheme: "A concept scheme in which the concept is included."

In regard to the ADM, the SKOS’ conceptual model was used to provide a foundation for relating different entities/objects together in a consistent way, and to perform inferences so that the graph can be expanded automatically (see Section 6.1.3.3). The ADM basis itself in only a few items of the SKOS vocabulary, namely the classes skos:Concept and skos:Collection, and the properties skos:member, skos:broader, skos:narrower, and skos:related, in order to structure hierarchies (of relationship types Is-a and Is-part-of) and networks that accept cross-links.

6.1.3.2 Model’s Description

Based on Schema.org and SKOS as foundations, the ADM can be described as such:\textsuperscript{68}:

• skos:Collection
  – E1 artifacts:CollectionNode: A collection of collections representing a statement.
    » E2 artifacts:MainCollectionNode: A collection of collections representing the entities of a statement.
    » E3 artifacts:EntityCollectionNode: A collection of concepts representing the subject or object of a statement.

• skos:Concept
  – E4 artifacts:Node: The most generic type of entity in the context of a skos:Concept.
    » E5 artifacts:Action: Equivalent to schema:Action. An action performed by a direct agent and indirect participants upon a direct object. Optionally happens at a location with the help of an

\textsuperscript{68} For a detailed description of the model, please see Appendix C.
inanimate instrument. The execution of the action may produce a result.

* E6 artfacts: **Artifact**: Equivalent to schema:CreativeWork. The most generic kind of creative work, including artworks, books, movies, photographs, software programs, etc.

* E7 artfacts: **Concept**: Equivalent to schema:Intangible. A utility class that serves as the umbrella for a number of ‘intangible’ things, such as an abstract idea or notion; a unit of thought.

* E8 artfacts: **Event**: Equivalent to schema:Event. An event happening at a certain time and location, such as a concert, lecture, or festival. Repeated events may be structured as separate Event objects.

* E9 artfacts: **Institution**: Equivalent to schema:Organization. An institution such as a museum, library, archive, school, NGO, corporation, club, etc.

* E10 artfacts: **Location**: Equivalent to schema:Place. Entities that have a somewhat fixed, physical extension.

* E11 artfacts: **Person**: Equivalent to schema:Person. A person (alive, dead, undead, or fictional).

* E12 artfacts: **Quality**: Equivalent to schema:QualitativeValue. A value for the characteristic of an entity, e.g. large, medium, small-sized cultural institution.

* E13 artfacts: **Quantity**: Equivalent to schema:Quantity. Quantities such as distance, time, mass, weight, etc. Particular instances of say Mass are entities like ‘3 Kg’ or ‘4 milligrams’.

* E14 artfacts: **PropertyNode**: Attaches additional property that offers structured values to an entity.

  · E15 artfacts: **RelationshipValue**: Attaches a text value to a property of an entity.

  · E16 artfacts: **NodeClass**: The classification of an entity.

  · E17 artfacts: **ExtraValue**: Attaches a text value to a property of an entity.

  · E18 artfacts: **Boolean**: Attaches a boolean value to an entity (True or False).

  · E19 artfacts: **Unit**: Attaches to an entity a structured value indicating the quantity, unit of measurement, and business function of goods included in a bundle offer.

  · E20 artfacts: **URI**: Attaches a Data type:URI to an entity.

  · E21 artfacts: **GPS**: Attaches the geographic coordinates of a place or event to an entity.

  · E22 artfacts: **Date**: Attaches to an entity a date value in ISO 8601 date format.
· E23 artifacts: **Medium**: Attaches to an entity a media object, such as an image, video, or audio object embedded in a web page or a downloadable dataset i.e. DataDownload.

· E24 artifacts: **Hook**: Attaches to an entity an agent for controlling a device or application.

· E25 artifacts: **WebAddress**: Attaches a Data type:URL to an entity.

### 6.1.3.3 Model’s Implementation

**Artfacts’ Graphical User Interface**

![Figure 43: Components of the Artfacts’ GUI](image)

The Artfacts’ GUI was designed with a clear requirement in mind in regard to the user experience and interaction. The Knowledge Maps should be prioritized. In this sense, the many different components of the interface that are responsible for operationalizing the creation and management of Knowledge Maps (such as panels, tabs, and menus) had to be draggable and retractable in order to maximize the space available for visualizing the network diagrams. For instance, it is possible to see the different workspace modes when all components are collapsed (as shown on Fig. 35) in comparison to the same components expanded (as shown on Fig. 43). On the figure below (see Fig. 43), the following components are displayed:
• **Content Panel:** This is one of the most important components of the Artfacts’ User Interface. This panel is used to create nodes (KMNodes, KGNodes) and links (KMLinks, KGLink) that constitute respectively Knowledge Maps and Knowledge Graphs. Besides providing an text field where labels/titles can be added to nodes, this panel presents other components, such as:

  - **Statement Area:** This component is a text area where the user can paste or type a statement. In Artfacts’, a statement is defined as a set of strings that can contain only one subject, and may contain one or numerous objects.

  - **Core Tagging Vocabulary:** This component, presented as two button groups, enables the tagging of highlighted strings in the Statement Area into nine different entity types. They are artfacts:Action, artfacts:Artifact, artfacts:Concept, artfacts:Event, artfacts:Institution, artfacts:Location, artfacts:Person, artfacts:Quality, and artfacts:Quantity.

  - **Entity Area:** This component provides a list of entities, and the option to set these entities as either subject or object(s). Moreover, additional metadata can attached to entities and their relationships. The following properties are allowed: artfacts:RelationshipValue, artfacts:NodeClass, artfacts:ExtraValue, artfacts:Boolean, artfacts:Unit, artfacts:URI, artfacts:GPS, artfacts:Date, artfacts:Medium, artfacts:Hook, and artfacts:WebAddress.

• **List Panel:** This component presents the user with a list view of all Knowledge Map Nodes (KMNodes) contained in an Artfacts’ project. The list can be sorted alphabetically by label or entity type, and chronologically by the date when the node was created or modified. This panel also allows nodes to be selected, and activated or deactivated on the visualization of the network diagram.

• **Search & Filter Tab:** This component enables the user to performed more complex queries on the Knowledge Map. It is possible e.g. to disable all nodes belonging to a specific entity type, or a group of entity types. It is also possible to enable and disable the nodes of a Knowledge Map based on a text search on statements, and then save the query as contexts. Several contexts can be saved and activated presenting therefore different cuts of a Knowledge Map.

---

69 In the context of the Artfacts’ Platform, Knowledge Graphs are comprised of nodes (KGNode) and links (KGLink). As stated in Section 6.1.1.2, the Knowledge Graph is the format in which information is stored and computed. Therefore, KGNodes and KGLinks will be used to process and create the Knowledge Maps (see Tagging System), which are also constituted of nodes and links described here as Knowledge Map Nodes (KMNodes) and Knowledge Map Links (KMLinks). The Knowledge Map is an interface element that is used in the context of HCI.
• Context Menu: This component is activated by right-clicking on a node. The menu offers functionalities, such as deleting the selected node, controlling the visibility of the selected node or a group of nodes connected to it, producing data visualizations, and retrieving other data objects from the Knowledge Graph that are directly related to the selected node.

**TAGGING SYSTEM**

Figure 44: Research questions displayed on the walls of the exhibition

The production of content, including interpretive materials, take place within Cultural Institutions also as part of the process of research, which is equivalent in many aspects to the research conducted in academia. This strong research aspect of institutions was evident not only in the already discussed case study presented in Chapter 5 (section 5.1.3), but also during the interview with curators conducted at the beginning of this research project (see Appendix ??) and the workshop *Expertengespräch Vermittlungskonzept Kogge-Halle* (see Appendix A.1.3.2). Particularly in the case of the workshop, which discussed the exhibition design of the *Kogge-Halle* in the *Deutsches Schiffahrtsmuseum*, the strong connection with the research done in the institution was to become present through e.g. research questions displayed in the exhibition space. According to the workshop participants, as a self-designated *Forschungsmuseum*.

---

70 See also Appendix A.2.1.3.
71 [https://www.dsm.museum/](https://www.dsm.museum/)
(research museum), the exhibition design should invite the visitor to put him- or herself in the shoes of a researcher. The audience should be presented with research questions and try to answer them by experimenting with the exhibits. In the case of the German Maritime Museum, research is an important mediator that is used beyond the research department of the institution. The scientific method also mediates how the institution organizes its interpretive program. Therefore, a tool that is supposed to be integrated within the established infrastructures and support institutional workflows needs to provide also strategies that are compatible with research processes.

![Figure 45: Coding with Atlas.ti](image)

Taking the research aspect of Cultural Institutions into account, the Artfacts Platform provides the user with the option to construct Knowledge Maps by tagging entities from statements\(^ {72}\) using its Core Tagging Vocabulary. The conceptualization of this feature drew inspiration from Qualitative Data Analysis Systems (QDASs - Atlas.ti\(^ {73}\),

---

72 A statement is defined as a set of strings that can contain only one subject \(\text{artfacts:hasCollectionWithSubject}(x, y)\) - see Fig. 48), and may contain one or numerous objects \(\text{artfacts:hasCollectionWithObject}(x, y)\) - see Fig. 48).

73 [https://atlasti.com/de/](https://atlasti.com/de/)
MAXQDA\textsuperscript{74}, NVivo\textsuperscript{75}, QDAMiner\textsuperscript{76}, etc) that use Grounded Theory\textsuperscript{77} as method for the analysis of collected media resources, such as text, audio, and video. As in Grounded Theory, these systems employ Coding as a fundamental analytic process, and the Code as its core unit. According to Johnny Saldaña, “a code in qualitative inquiry is most often a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of language-based or visual data.” \textsuperscript{201} In QDAS, a portion of these language-based or visual data is defined as indicators\textsuperscript{78} (see [109]). On the figure below (see Fig. 45), it is possible to see a text excerpt (indicator), highlighted in blue, being coded by the keyword Storytelling, which is being indicated on the side panel on the right-hand side of the User Interface of the software Atlas.ti.

The Coding System employed by QDASs is however different from the Tagging System employed by Artfacts\textsuperscript{79}. That is because Artfacts’ Tagging System is more granular and aims ideally at the extraction of triples from statements. In order to be successful, the tagging process must necessarily extract at least one entity that is assigned as subject (artfacts:hasCollectionWithSubject(x, y) - see Fig. 48) of the statement. However, a statement may also contain one or several objects (artfacts:hasCollectionWithObject(x, y) - see Fig. 48), which are defined as such as soon as the subject of the statement is identified. All the entities are converted into nodes, and relationships are also defined between the subject and object(s). Labels can be also assigned to the relationships by adding an extra property to them.

\begin{itemize}
  \item \textsuperscript{74} https://www.maxqda.com/
  \item \textsuperscript{75} http://www.qsrinternational.com/nvivo/nvivo-products
  \item \textsuperscript{76} https://provalisresearch.com/products/qualitative-data-analysis-software/
  \item \textsuperscript{77} Grounded Theory is a research method that is used to develop a theory by detecting patterns in data. In other words, “Grounded Theory is the generation of theories from data.” \textsuperscript{234}
  \item \textsuperscript{78} These portions of relevant information can be also defined as quotations in the case of QDASs, such as Atlas.ti.)
  \item \textsuperscript{79} For insights in how the Artfacts Platform supports research, please see Section 6.3.1.3.
\end{itemize}
For example, the Knowledge Map shown on the figure above (see Fig. 46) partially depicts how the guided tours for the chatbot *Marbles of Remembrance* (see Section 6.3.2.1) was modeled. A tour was defined as having several acts, which were associated to a particular address in Berlin, Germany. Each act should start with an introduction, and be accompanied by messages associated to an avatar, in this case *Maayan Freier*, who would be responsible entity for interacting with the user. A part of this Knowledge Map was extracted from the following statement: "Act 1 starts at *Auguststraße 11-13* and it contains *Introduction, MSG1 - Follow the Map, MSG2 - Present Mom’s Picture, MSG3 - About 2. jüdische Volksschule*." The keywords in bold represent entities that can also be seen on the Knowledge Map (see Fig. 46). On the figure below (see Fig. 47), it is possible to see the tagged statement on the component *Content Panel* of the Artfacts’ User Interface.
This statement can be formally described as follows, the KGNode artifacts:MainCollectionNode(id:“mco2”) is linked to the following KGNodes:

- artifacts:EntityCollectionNode(id:“mco2ec01”, label:“Act 1”, type:“Event”)
- artifacts:EntityCollectionNode(id:“mco2ec02”, label:“Auguststraße 11-13”, type:“Location”)
- artifacts:EntityCollectionNode(id:“mco2ec03”, label:“Introduction”, type:“Artifact”)
- artifacts:EntityCollectionNode(id:“mco2ec04”, label:“MSG1 - Follow the Map”, type:“Artifact”)
- artifacts:EntityCollectionNode(id:“mco2ec05”, label:“MSG2 - Present Mom’s Picture”, type:“Artifact”)
- artifacts:EntityCollectionNode(id:“mco2ec06”, label:“MSG3 - About 2. jüdische Volksschule”, type:“Artifact”)

The KGNode id:“mco2ec01” is connected to the KGNode id:“mco2” by the KGLink described below and used to represent the subject of a statement:

- artifacts:hasCollectionWithSubject(id:“mco2”, id:“mco2ec01”)

The other KGNodes (id:“mco2ec02”, id:“mco2ec03”, id:“mco2ec04”, id:“mco2ec05”, id:“mco2ec06”) are connected to the KGNode id:“mco2” by the KGLinks described below and used to represent the objects of a statement:
Based on this Knowledge Graph, Artfacts creates automatically the KGNodes representing the following entities:

- artifacts:Event(label: "Act 1")
- artifacts:Location(label: "Auguststraße 11-13")
- artifacts:Artifact(label: "Introduction")
- artifacts:Artifact(label: "MSG 1 - Follow the Map")
- artifacts:Artifact(label: "MSG 2 - Present Mom's Picture")
- artifacts:Artifact(label: "MSG 3 - About 2. jüdische Volksschule")

Then, the KGLinks below are established between artifacts:EntityCollectionNodes and artifacts:Nodes:

- artifacts:hasMainEntity(id: "mco2ec01", label: "Act 1")
- artifacts:hasMainEntity(id: "mco2ec02", label: "Auguststraße 11-13")
- artifacts:hasMainEntity(id: "mco2ec03", label: "Introduction")
- artifacts:hasMainEntity(id: "mco2ec04", label: "MSG 1 - Follow the Map")
- artifacts:hasMainEntity(id: "mco2ec05", label: "MSG 2 - Present Mom's Picture")
- artifacts:hasMainEntity(id: "mco2ec06", label: "MSG 3 - About 2. jüdische Volksschule")

Finally, based on the following rules:

\[
\text{hasCollectionWithSubject}(x, y) \land \text{hasMainEntity}(y, z) \Rightarrow \text{hasSubjectEntity}(x, y) \\
\text{hasCollectionWithObject}(x, y) \land \text{hasMainEntity}(y, z) \Rightarrow \text{hasObjectEntity}(x, y) \\
\text{hasSubjectEntity}(x, y) \land \text{hasObjectEntity}(x, z) \Rightarrow \text{relatedDirected}(y, z)
\]
Rhizomatic relationships (KGLinks also defined as cross-links) are established between artfacts:Nodes, namely artfacts:relatedDirected (as shown in Fig. 49). The artfacts:Nodes (also defined as KMNodes) and inferred relationships are the elements that are possible to be visualized as a Knowledge Map.

Besides rhizomatic relationships, Artfacts is also capable of modeling hierarchical relationships for classes, defined as relationships of the type Is-a, and for schemes, defined as relationships of the type Is-part-of. Based on the definitions of relationships provided on the figure above (see Fig. 48), an intuition for both hierarchical models is provides in figures 50 and 51 respectively.
Figure 49: Rhizomatic relationships
Figure 50: Hierarchical structure for classes
Figure 51: Hierarchical structure for schemes
6.2 Use Case Scenarios

The following use case scenarios were conceptualized based on the feedbacks and insights obtained during the workshops, usability tests, case studies, and interviews with experts from different Cultural Institutions. The goal of the scenarios is to provide an intuition of the practical use of the Artfacts Platform with institutional contexts.

6.2.1 Scenario 1 - Exhibition Conceptualization Support

![Architectural Plan and Collection Gallery](image)

Figure 52: Architectural Plan and Collection Gallery

Ms. Müller works as a curator for a medium sized museum in Hamburg. The museum is planning a new temporary exhibition about posters produced at the turn of the 19th century. The exhibition is set to take place at two big halls that can hold up to 54 items altogether. Ms. Müller would like to have a quick overview of her curatorial possibilities. To begin with, she wants to know what her museum has in its collection. Therefore, she uses Artfacts to quickly query for posters produced between 1890 and 1910. The result shows 125 items. Ms. Müller needs now to answer four questions: (1.) what is the best way to tell a story with the objects, (2.) how these objects must to be arranged in the museum to best tell this story, (3.) what objects must be discarded, (4.) and what objects must be externally

---

80 See Appendix A.
acquired in order to fill up the space and support the narrative of the exhibition.

Ms. Müller creates a new project and imports the 125 items into the workspace of the application (Fig. 52). The additional metadata connected with the items such as concepts, geographic locations, and artists are also loaded into the project. In the attempt to answer the first and second questions, Ms. Müller uses the platform to place the objects on the architectural plan of the museum. Although she has an evenly amount of artworks per artist, after producing an Area Diagram (Fig. 53) on the artists’ countries of origin, she realized that most of the posters were produced by Austrian and French artists. The unbalanced representativeness of artworks from different European locations makes her discard the idea of presenting objects by country of origin. Instead, she want to explore another alternative. She thinks that exhibiting artworks by themes is a more interesting approach for the visitor. In order to get more insights on how the artworks relate among themselves, she produced a Sankey diagram between artworks and the topics related to them (Fig. 53). However, the diagram shows that among all the objects only 17 of them have shared topics. After using the application to generate a Picture Gallery of the objects, she determines that the database of the museum certainly lacks valuable information about the interpretation of objects. Ms. Müller requests extra curatorial research from the research department. She creates new relationships
among the artworks, and asks the research team to validate the links and produce new content.

![Figure 54: Timeline Report](image)

The research team has immediate access to Ms. Müller's project. They find the Timeline diagram (Fig. 54) that can be produced based on artists and artworks to be especially helpful to situate their research. As research progresses, they annotate new information and create new relationships among objects. By reviewing the work of the research team, she decides that the exhibition can explore how gender was differently portrayed at the turn of the century, especially in regard to contemporary values. From the 125 artworks the museum possesses, Ms. Müller decides to present only 36, since they either talk about gender or have male, female, and other gender representations on them. Finally, Ms. Müller uses Artfacts to search for other artworks her museum could borrow in order to complete the exhibition. She knows that a partner museum in Berlin has a comprehensive archive of not only homemade, but also international posters. Ms. Müller accesses the partner museum’s Knowledge Base (available online as an open repository) with Artfacts. This time, she focuses specifically on gender-related posters. After selecting a few artworks, Ms. Müller arranges them in the exhibition space. She uses her tablet to walk around the exhibition space in order to validate if her exhibition planning suits well the visitor experience. She is glad she can use Artfacts to accelerate and manage the complexities of the exhibition conceptualization.
6.2.2 Scenario 2 - iBeacon-based Audioguide

As a next phase of the exhibition design, Ms. Müller needs to think about what kinds of interpretive support materials should the museum use, and how they need to be implemented within the exhibition space in order to facilitate the comprehension of the exhibition content by visitors. By browsing through possibilities based on the experience of other museums she knows, a solution implemented by the Metropolitan Museum of Art\(^81\) (New York, USA) catches Ms. Müller’s attention in particular. iBeacon-based audioguides used a low-cost location-aware service based on Bluetooth Low Energy (BLE) technology to provide visitors with contextual information on artworks\(^82\). This technology enables targeted content to be delivered precisely as the visitors approach artworks inside the museum. Ms. Müller then talks to the director and the IT department of her museum in order to propose the development of a iBeacon-based audioguide that is able to reuse the information of the Artfacts’ project. Ms. Müller carefully marks the most important artworks she would like the tours to talk about on the architectural plan of her Artfacts’ project, and uses it for presenting her idea to the museum director and IT department. An important

\(^81\) [https://www.metmuseum.org/](https://www.metmuseum.org/)

\(^82\) [https://www.metmuseum.org/blogs/digital-underground/2015/beacons](https://www.metmuseum.org/blogs/digital-underground/2015/beacons)
advantage of the Artfacts platform is the ability to not only visualize how the audio content is arranged on the architectural plan, but also easily manage the guided-tours by further editing her Artfacts’ project. The development of the audioguide application is approved. The application will cover 30 artworks, which means that 30 iBeacon tags are ordered. The attachment of the audio recordings together with the IDs of the iBeacons to the 30 artworks is performed with ease by attaching two PropertyNodes (see E14: Section 6.1.3.2 and Appendix C), namely Hook (see E24: Section 6.1.3.2 and Appendix C) and Medium (see E23: Section 6.1.3.2 and Appendix C), to each artworks object. Once the narrations and the iBeacon IDs are attached to the artworks, Ms. Müller also shares the Artfacts’ project with the company responsible for installing the iBeacons next to the Artworks in the exhibition space. The construction company is able to follow the information on the architectural plan of Artfacts’ project, in order to know precisely where each one of the iBeacons needs to be placed. Once the installation is done, visitors can use the audioguide application by either acquiring a device at the museum entrance, or downloading the application to their mobile phones.

6.2.3 Scenario 3 - Hackathon

The intuition concerning the use case scenario 3 can be abstracted from the description of the case study described on section 6.3.2.

6.3 Evaluation

6.3.1 Usability Test: Artfacts’ Vocabulary and the Tagging System

A usability test was applied in May 2016 in order to evaluate the first version of the platform mainly in regard to one of its requirements: In the context of its OOUI, the application had to support individuals with no technical background (namely curators, researchers, educators, etc) in the collection of information and its cartography. The method that Artfacts employed for dealing with

83 See Section A.1.3.4.
84 All these individuals are part of the group that comprehends hackers (see Section 2.1.2).
85 See Section 6.1.1.1.
this matter, Entity Tagging\textsuperscript{86}, had to present a low learning curve during the process of tagging important entities in Statements and converting the extracted metadata into Knowledge Maps. The test therefore focused on the manual and auto-tag functionalities. In the first version of Artfacts\textsuperscript{87}, the tagging system partially implemented the Schema.org’s Data Model for generating simplified Knowledge Maps based on co-occurrence\textsuperscript{88}. Relationships were established in between KMNodes depending on shared/co-occurrence entities. The graph visualization was generated using the Force Layout\textsuperscript{89} of the library D3.js\textsuperscript{90}. KMNodes were pulled towards each other depending on the existence of relationships in between them. With this strategy, it was possible to identify different topics that should most likely be understood together, because they share strong correlations. And, on the contrary, it was also possible to identify topics that do not closely relate, because they are shown far apart.

6.3.1.1 Method

The test was applied to a mixed-gender group of ten participants\textsuperscript{91}, which included research assistants and students. The participants were presented with four different textual sources of similar topic, level of difficulty, and length. Then, participants were asked to create two different maps by utilizing the information of two sources each. A Knowledge Map had to be generated by tagging key entities of the first two texts, and a Mind Map had to be generated by manually diagramming key concepts of the other two last texts. The participants were able to freely decide how they wanted to construct the maps. However, participants were asked to use respectively the first version of the Artfacts Platform, and a popular Mind Map web application called Bubbl.us\textsuperscript{92}, where diagrams are created by manually drawing lines between node elements, instead of tagging. The Mind Map application was used to provide a standpoint so that the contrasts between the two different approaches could be easily compared by participants. Before the creation of each map, it was

---

\textsuperscript{86} See Section 6.1.3.3.

\textsuperscript{87} See Section A.1.2.

\textsuperscript{88} According to Jérôme Kunegis, “co-occurrence networks represent the simultaneous appearance of items. Co-occurrence networks are unipartite and unweighted.” [129]

\textsuperscript{89} The Force Layout is applied to a node-and-link graph and it “combines physical simulation and iterative constraint relaxation for stable graph layout.” [30]

\textsuperscript{90} https://d3js.org/

\textsuperscript{91} Amount considered as acceptable according to Nielsen [164].

\textsuperscript{92} https://bubbl.us/
provided instructions on how to use each one of the applications. The participants were required to know only a few key functionalities necessary to create the maps. Time was given for text comprehension and map construction, with the same amount of time distributed between the Knowledge Map and Mind Map phases (See Fig. 56). The test was entirely performed on a desktop computer. Furthermore, all the participants were recorded and were asked to think out aloud while creating the maps.

![Flowchart representing the usability test workflow](image)

Figure 56: Flowchart representing the usability test workflow

Besides observing the participants during their interaction with both applications, it was also given to them two different questionnaires to be answered after the interaction with the software. An open-ended questionnaire was given in order to know their impressions on the efficacy of Knowledge Maps in comparison with Mind Maps in regard to constructing, modeling, and expressing the information contained in the texts. The following questions were asked:
1. In your opinion, which one of the map approaches would facilitate/aid your research/study better? Why?

2. In your opinion, which one of the maps is able to better represent the information contained in the texts? Why?

3. Did the process of making a Knowledge Map helped you to better understand or notice some information that was not obvious before?

4. Did the process of making a Mind Map helped you to better understand or notice some information that was not obvious before?

5. Please, explain aloud your personal mental strategy to organize the information into a Knowledge Map.

6. Please, explain aloud your personal mental strategy to organize the information into a Mind Map.

The other questionnaire contained the standardized questions of the System Usability Scale (SUS) (see [119, ch.21]). The SUS test is a tool for measuring the usability of software applications. It gives software engineers and designers a classification mechanism to measure effectiveness (how well are the users’ objectives achieved), efficiency (how much effort and resources should be user spend for achieving these objectives), and satisfaction (how satisfactory was the experience). The main measure to understand the results of SUS is based on the average results obtained by applying this test over the years, in this sense a “SUS score above a 68 would be considered above average and anything below 68 is below average” [2]. The SUS average is shown as the gray horizontal line crossing the vertical bars of the Figure 57. The SUS questionnaire contained the following questions, based on which the participants had to give a scale from strongly disagree (1) to strongly agree (5) (see [2]):

1. I think that I would like to use this system frequently.

2. I found the system unnecessarily complex.

3. I thought the system was easy to use.

4. I think that I would need the support of a technical person to be able to use this system.

5. I found the various functions in this system were well integrated.

6. I thought there was too much inconsistency in this system.

7. I would imagine that most people would learn to use this system very quickly.

8. I found the system very cumbersome to use.

9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

6.3.1.2 Results

The evaluation of Knowledge Maps created with the aid of the manual and auto-tag functionalities was positive. Nine in ten participants perceived Knowledge Maps as better suited for supporting them during their research or study. Among the reasons, participants claimed that the tagging system could help them in keeping track of Statements systematically, because it automatically connected different Statements based on their shared key entities. The standard vocabulary composed of eight\(^{93}\) different tagging categories was seen as advantageous, because it provided consistency among tagged entities. Participants were able to quickly identify the different entity types that composed the map and how they were connected. In addition, the highlighting of keyword on the Content Panel (see Fig. 47) assisted the fast identification of key concepts within the text excerpts. The auto-tag functionality was seen as relevant in binding new Statements with pre-existing ones, and therefore re-contextualizing new and old information. Overall, participants declared that using the auto-tag functionality to establish linkage between Statements was done effortlessly and faster if compared with Bubbl.us, on which relationships needed to be drawn.

When asked which of the two applications could better support the representation of the information contained in the texts, six of the participants marked Artfacts, three participants marked Bubbl.us, and one participant was not sure about which application to choose. Some of the participants who preferred the mind mapping tool found that they had more control over the final look of the map, because the tool allowed them to position the node/box on a specific area of the screen, and because the relationships could be manually drawn between nodes/boxes\(^{94}\). Finally, participants had to independently grade both map approaches on how effective they were in allowing them to notice meaningful connections between information that was not obvious during reading. Most of participants (seven out of ten) claimed that the Knowledge Map approach was effective in enabling

\(^{93}\) The entity type Action was added later on in order to improve the system for controlling system events.

\(^{94}\) In the first version of Artfacts, the user had no control concerning the positioning of KMNodes on the screen. Based on the result of this test, this capability was added in the second version.
them to perceive connections among information. The other three participants claimed that this approach was very effective. In regard to Mind Maps, most of the participants said that this approach was reasonably effective, while one declared Mind Maps were very effective, and one found this approach not effective. According to participants, the reason why they found Knowledge Maps effective in perceiving connections among KMNodes is due to the automatic detection and highlight of entities through the auto-tag functionality and the establishment of relationships in between nodes, which is also done automatically.

![Figure 57: Results of the SUS test](image)

The results of the SUS test\(^5\) in regard to the usability of the tagging strategy for the creation of Knowledge Maps were also positive, with an average score of 71.75 (represented by the black horizontal line of Fig. 57 - 3.75 points above the average of 68 points - see [11]). Since the Artfacts Platform required participants to learn some of its key functionalities for the construction of Knowledge Maps, it was expected some problems in regard to the manipulation of the interface. However, as observed and demonstrated by the SUS test, the interface elements and system workflow necessary for the creation of Knowledge Maps through the manual and auto-tag features was easily understood and performed by participants without problems. Taking the average of questions individually for all participants, the test showed us that participants were confident using the system, as suggested by the average SUS score of 75.0 on question 9 of the SUS questionnaire (“I felt confident using the system”)

\(^5\) See Section A.1.3.4 for the complete punctuation table.
and found it easy to use, as suggested by the SUS score of 85.0 on question 3 (“I thought the system was easy to use” [119, ch.21]).

6.3.1.3 Discussion and Conclusion

As stated above, the great majority of the interviewed individuals affirmed that Artfacts was more suitable for supporting the collection of information and its cartography than the mind mapping tool. These affirmations are backed by the following observations pointed out by interviewees:

- By storing a set of metadata information of the textual sources, Artfacts was seen as a valuable tool for managing and keeping track of content in a systematic way;

- The pre-defined vocabulary offered by Artfacts for categorizing information provided a unified framework that afforded consistency in how the information of different sources was structured, since all entities have to follow into the standard and generic categories independently of the textual source;

- Reviewing information already stored through the platform was facilitated not only by the visual graph representation, but also by the highlighted keywords on the Statement Area of the Content Panel (see Fig. 43). This point was especially important for the accurate identification of entities within the text excerpts. The highlights also obeyed a pre-defined schema of colors to help the user to identify different categories of entities.

One of the main observations noticed during the interviews was that Artfacts offered a multi-step and straightforward method for creating Knowledge Maps through its tagging strategy, which enabled users to work directly with text excerpts extracted from the textual sources in four steps, which required the participants to: (1) divide the textual source into clear Statements/KMNodes, (2) identify important keywords that captured the meaning of these KMNodes, (3) tag the keywords according to a pre-defined vocabulary, (4) use the auto-tag functionality for receiving recommendations from the system about shared entities. Artfacts’ Knowledge Maps were created as a result of the described four-step procedure, which produced the necessary metadata to automatically establish cross-references among KMNodes stored in the system. In this sense, Artfacts was able to connect information the user did not anticipate. On the contrary, Bubbl.us appeared as more mentally overwhelming for participants as it offered no method for working directly with the textual sources. Participants were required to come up with their
own mental strategies to reshape the information into a tree diagram before being able to manually drawing the map. As identified through our observations and questionnaires, the mental strategy for creating Mind Maps by the majority of participants was to understand the texts, identify central and less important ideas, and, through a mental exercise, translate them into branches and sub-branches.

Both applications were able to assist participants in the synthesis of information. However, Artfacts seemed to have provided participants with a four-step method that is better suited to cope with greater amounts of information, because it offered more than diagram drawing functionalities. However, the over-simplicity of Mind Maps can be advantageous depending on the desired final objective of the representation, once it limits itself to depict the category-subcategory logic or a macro-to-micro hierarchical ordering of topics. This limitation is however an impediment when more ambitious representations are required, since with Mind Maps, one needs to adapt the information to the predictable shape of the tree. The rhizomatic form of Knowledge Maps, on the other hand, which is created as a result of the tagged content, cannot be predicted. The relationships in Knowledge Maps are arbitrary and can assume many different kinds of connections.

It was clear from the beginning that it was necessary to increase the semantics of the ADM in order to provide greater interoperability with external data repositories, and to be able to utilize Artfacts’ projects to organize and control the flow of system events and user actions of third party applications (another requirement). For that matter, it was incorporated later on to the ADM another set of types from the Schema.org’s core vocabulary. In addition, the ADM was remodeled as an extension of SKOS. The challenge was to increase the expressivity/granularity of the model while preserving the good usability of the system and a low learning curve. In this sense, this usability test was important, because it indicated the tagging system as a easy-to-use strategy to manage the complexity of the rhizomatic cartography of information.
6.3.2 Case Study: The Hackathon Coding da Vinci

This case study was organized in order to gather insights on the limitations and effectiveness of Artfacts for serving as platform that provides solutions for developing and managing DIAs. In this regard, the 2017 edition of the CdV offered an optimal setting for assessing the platform. The CdV has not only obtained the support from more than 70 institutions (see Section 2.1.3 for more information about the motivation of organizers in regard to the CdV), but also been able to attract different individuals from a variety of technical, artistic, and cultural backgrounds who are interested in pushing the boundaries of what can be create with Digital Collection (see Section 2.1.2 for more information about the motivations and profiles of participants in regard to the CdV). Concerning the research goals, besides trying to gather insights on ideal use cases supported by the platform, the case study also focused on analyzing how well the Artfacts Platform was able to cope with the demands of an interdisciplinary team within a creative working environment and under a strict deadline. Such environments, which are based on experimentation, expect DIAs’ prototypes to respond quickly to changes in concept, structure, and content. The platform needs therefore to accommodate fast-changing requirements during intensive iterative development cycles (see Fast-speed IT Approach in Section 6.0.1). In this sense, in an ideal scenario, all team members should retain a certain degree of autonomy to try out different design possibilities. That can be afforded by user-friendly and flexible methods for modeling data structures that can be fast deployed and integrated in DIAs. As part of the study, the chatbot Marbles of Remembrance (Murmeln der Erinnerung) was conceptualized and implemented. Not only details in regard to the technical development are discussed, but also observations that concern the employment and deployment of storytelling with Artfacts.

6.3.2.1 The Chatbot Marbles of Remembrance

conceptualization

DIAs are applications of a particular kind, because they bring together different ecosystems and practices namely Cultural Heritage

96 This case study should be understood as a part of the case study presented in Chapter 5 (see Section 5.1.3).
and the interpretation of collections, and Computer Science and software development. Adding these practices and ecosystems to the context of Hackathons, which are experimental and driven by a hands-on attitude, just increases the complexity of the picture. In this regard, it is advocated that a design perspective, instead of an engineering one, is better suited to understand the particularities of the creative process that takes place in such events. During Hackathons for Cultural Heritage, DIs are created from the interpretation, modification, and instantiation of Digital Collections, which are the creative material hackers use to design/hack with (see Chapters 3 and 5).

In particular to the project Marbles of Remembrance (Marbles; see [62]), the concept depended on subjective and objective factors. Among them: (a.) the variety and affordances of datasets, (b.) the particular skills of team members, (c.) their individual motivations, (d.) a certain degree of affinity in regard to the personalities of participants, (e.) the history of past Hackathons in terms of already implemented projects, and (f.) the relevance of the location (e.g. city of Berlin) where the event takes place. It is unclear if these factors obey a certain kind of order (based on e.g. importance), as teams and concepts seem to be formed in an almost chaotic way assuming they were not previously predefined. As for Marbles, the concept of developing a chatbot app appeared spontaneously during the Forming Phase (see Section 2.1.1), which happened just after organizers presented goals, technologies, and datasets. According to Trainer et al. during this phase “individuals attempt to identify the nature of the task and what information is required for it. To get to know one another they exchange personal information” [229]. In chronological terms, the overview provided by 19 representatives from different CIs presenting 31 datasets (see a.) appeared therefore to have a certain impact on the origin of the concept.

It was observed that, during discussions among participants, a certain kind of natural selection (considering b.,c.,d.,e.,f.) takes place concerning the formation of teams. An exchange of different concepts starts to emerge. Participants try to convince others about their ideas, adapt parts of the concept to accommodate different interests, or carry on with it individually. However, teams that e.g. cannot engage programmers are likely to disperse. In the case of the project Marbles, all team members were interested in working with
a subset (Karten der Schülerkartei\(^{97}\)) comprised of almost 11 thousand registration cards of Jewish children that suffered persecution during the Nazi Regime. The dataset was provided by the ITS (see Section 5.1.3). The idea of a chatbot seemed to be suitable to animate the dataset, because of the emotional nature of the topic. A chatbot app would provide the means for engaging users in first-person narratives, in which the children would tell their stories by taking the user to a tour around Berlin, and showing physical landmarks and personal memories that were important in their lives. Implementing storytelling through a chatbot appeared as a strong candidate for a compelling concept especially because the Marbles team did not want to repeat the formats of past CdV projects, as e.g. rendering data visualizations from datasets. Finally, since the ITS dataset did not contain any information regarding the children’ addresses, the team decided to include an additional dataset. The Stolpersteine Berlin\(^ {98}\) dataset completed the necessary pool of information needed to create location-based narrations.

**IMPLEMENTATION**

---

Chatbots offer an interesting strategy to engage young audiences with first-person narratives.

---

\(^{97}\) Karten der Schülerkartei is a subset of Kartei der Reichsvereinigung der Juden - [https://codingdavinci.de/daten/](https://codingdavinci.de/daten/)

The chatbot presented the following main functionalities: natural language queries (image on the right - Fig. 58), guided tours (image on the middle – Fig. 58), and notifications. To begin with, users were able to ask open questions about the persons registered in both datasets *Karten der Schülerkartei* and *Stolpersteine Berlin*. Depending on the metadata available, the chatbot was able to respond whether the person was one of the persecuted children, show his or her registration card, birth and death dates, and whether the person was able to survive and emigrate to another country. Furthermore, since the records of both datasets were crossed, the chatbot could show the person’s address on the map as well as information about his or her relatives. Secondly, the chatbot provided guided-tours following the narratives of particular children. These stories are the most compelling part of the concept, because the user has the chance to get a deeper impression on how the Jewish children lived, how they saw the world, and visit the places that were part of their routine during the Second World War. The guided tours are location-based, meaning that the sub-parts of the stories are connected to certain landmarks in the city. Finally, because the Marbles team used the Telegram API\textsuperscript{99}, the user was able to share its live location\textsuperscript{100} with the chatbot. Depending on the current GPS coordinates of the user, the chatbot app would notify him or her Stolpersteine within a certain radius.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{architecture.png}
\caption{Architecture of the chatbot}
\end{figure}

In regard to its architecture, besides the Artfacts Platform, the chatbot used Telegram (Server API and the Telegraf Client API\textsuperscript{101}).

\begin{itemize}
\item \textsuperscript{99} \url{https://core.telegram.org/api}
\item \textsuperscript{100} \url{https://core.telegram.org/bots/api#sendlocation}
\item \textsuperscript{101} \url{https://github.com/telegraf/telegraf}
\end{itemize}
for handling the communication with the user. The app identified the type of message\(^{102}\) and redirected it to the proper component for it to be further processed (Message Analyzer – Fig. 59). Natural language queries were sent to an external webservice called DialogFlow\(^{103}\) so that entities\(^{104}\) and intents\(^{105}\) could be identified, and communicated back to the chatbot app. The system compared then the intent received from DialogFlow with a dictionary of methods. Thus, depending on the entities and intent, the chatbot would know whether it was necessary to request further information from the Artfacts Platform, or the response could be handled by the chatbot itself. In case the component Message Analyzer identified the input from Telegram as a guided-tour request, the chatbot app would redirect the message to the component Story Module, which was in charge to accessing and translating Knowledge Maps contained in Artfacts’ projects into the guided-tour narratives.

![Content Panel with expanded Entity Property](image)

**Figure 60: Content Panel with expanded Entity Property**

On the Artfacts Platform, the Knowledge Maps employed to model the stories were composed by different categories of KMNodes that were defined by adding extra metadata to the subject of the statements using `artfacts:Hook`. In this sense, a schema was defined obeying the following hierarchical structure: for a KMNode of type `artfacts:Hook("Story")`, there was one KMNode that assigned an actor to the story of type `artfacts:Hook("Actor")`, and several other KMNodes of type `artfacts:Hook("StoryAct")`. Each

\(^{102}\) As e.g. natural language query, GPS coordinates, or guided-tours requests.

\(^{103}\) [https://dialogflow.com/](https://dialogflow.com/)

\(^{104}\) Entities are strings in the text that represent categories of objects, such as names, dates, locations, etc (see \[68\]).

\(^{105}\) Intents are mappings that connect a group of entities to a certain system action (see \[68\]).
KMNode of type artifacts:Hook("StoryAct") was related to a KMNode of type artifacts:Hook("Location") that contained, among others, GPS metadata defined by artifacts:GPS. Moreover, KMNodes of type artifacts:Hook("StoryAct") could also contain several KMNodes of type artifacts:Hook("Narrative") and artifacts:Hook("Speech"), which were responsible for storing information relative to the messages used for the exchange of messages with the user. KMNodes of type artifacts:Hook("Speech") could also be connected to KMNodes of type artifacts:Hook("Medium"), used to reference audio, video, and picture files. Finally, KMNodes of type artifacts:Hook("Narrative") and artifacts:Hook("Speech") could be connected to KMNodes of type artifacts:Hook("UserAction"), which had as functions either pausing the story and waiting for the user to press a button, or watching for a confirmation that the user has reached a certain landmark. On the picture above, it is possible to see the property artifacts:Hook being assigned to the subject of the entity "Act 1" (See also Fig. 46 for a partial representation of a Knowledge Map used to model a tour).

6.3.2.2 Observations

The team was composed by a programmer and three other members with either little or no experience in programming. While the programmer was in charge of inquiring into software libraries, designing the architecture, and coding the chatbot, the tasks of other team members – defined here as content producers - consisted of
understanding and refining the datasets, and producing extra content by researching on specific children selected to be the subjects of the guided-tours. By discussing with the content producers in order to understand how the structure of the narratives could be represented, the programmer defined the schema presented above that was used for modeling the stories and controlling the user experience of their tours. Once this schema was explained to and agreed by the content producers, modifying the narrative was done easily on a no-code basis. The autonomy provided by Artfacts to the content produces was crucial for creating and concurrently improving the guided-tours, because they were able to try out the tours by themselves on the sites, and immediately modify their structure and content if they realized that e.g. a landmark should be presented first, or a message was too long to be read while on a busy street. Thus, using Artfacts raised the overall quality and speed of the development process, and contributed potentially to the success of the project.

6.3.2.3 Conclusion

This case study intended to identify ideal use cases and analyze how the platform would cope with the development of DIAs in competitive and creative environments of Hackathons. Thus, the study was carried out during the 2017 edition of the CdV. The results suggest that the affordances of the platform in enabling the reuse of existing datasets from Cultural Institutions, and their quick recontextualization by Knowledge Maps was effective for creating compelling narratives. In addition, the ability of content producers to iterate faster and try out a greater quantity of different design possibilities during the development process, highly likely contributed to the success of the project, which was honored with a prize at the Hackathon finals. The following press releases can be found about the project:

- *Die Spur der Daten*. By Claudia Reinhard in *Frankfurter Allgemeine*¹⁰⁶;
- *Smartphone-Chat mit Zeitzeugen*. By Nadine Emmerich in *Evangelischer Pressdienst*¹⁰⁷;

---

¹⁰⁷ [https://www.epd.de/landesdienst/landesdienst-ost/schwerpunktartikel/smartphone-chat-mit-zeitzeugen](https://www.epd.de/landesdienst/landesdienst-ost/schwerpunktartikel/smartphone-chat-mit-zeitzeugen)
Finally, it is important to point out some limitations of Artfacts. Although the Tagging System offered a better solution for the semi-automatic cartography of information, if compared with other tools based on manual layout as a strategy for designing maps, the Artfacts’ GUI presented scalability problems that can be described in two points. To begin with, the library used for rendering the visualizations, namely D3.js, is based on SVG, which is used by Web browsers to decode and render vector graphics. This technology is not hardware accelerated, and because of that presents performance issues if too many SVG objects are displayed simultaneously. During the CdV, hackers had some performance issues in handling Artfacts’ projects containing Knowledge Maps with more than 140 KGNodes. A possible solution for this issue might be in exchanging D3.js for a library capable of utilizing the hardware acceleration capabilities of the computer’s graphic card. Lastly, besides this technical issue, Knowledge Maps with too many nodes and relationships are cognitively overwhelming; especially when depicting a rhizomatic structure. Using lists to navigate through the maps and filters to deactivate nodes and relationships have helped the mitigation of cognitive overload. However, further research on new strategies to reduce cognitive overload is required, such as the further modularization of Knowledge Map, the interconnection of Artfacts’ projects, and the use of recommender systems to provide assistance in creating and navigating maps.

Further conclusions about the integration of the Artfacts Platform within the context of Hackathons are discussed in Chapter 7.

---

108 Currently available under https://www.podcast.de/episode/368848440/
112 https://d3js.org/
CONCLUSION

This thesis is situated within a new materialistic paradigm \[52\] of increased abstraction, intangibility, and digitality that also pervades the domain of Cultural Heritage. This new materialistic approach sees especially intangible materiality as something that is not self-contained, but it manifests its existence through the transformations it provokes in the world. The Actor-Network Theory \[130\] offers interesting insights in this regard. Human and non-human actors are not defined by the boundaries of their physicality, but by how they affect a network comprised of numerous interdependent entities. Especially in the case of the digital, its materiality is indivisible from the processes evoked for, through, and by them. The digital material is therefore itself an agent of (human) transformation, either we like it or not.

This paradigm shift can be perceived in the Cultural Sector by the constant expansion and revision of the term Cultural Heritage. Nowadays, Cultural Heritage is used to refer not only to traditional works of art, but also a range of intangible objects and traditions that hold social and historical significance\[1.\] The domain of Cultural Heritage, and consequently the domain of Cultural Institutions built to protect and preserve cultural memory, has become increasingly abstract. One cannot disregard the contributions of our digitally mediated world in the revisions of the term heritage. The digital is nowadays ubiquitous and is an integral part of institutionalized culture. On the one hand, digital technologies, including computer programs, have themselves become heritage\[2.\]. On the other hand, these technologies have provided the basis for supporting workflows within Cultural Institutions, especially the so-called GLAMs, and originated methods to understand and preserve the institutions’ collections.

Within networks/ecosystems of interdependent human and non-human agents, and within networks of interdependent networks, sets of data, which in the context of this thesis are understood as

---

1 See Chapter 1.
2 See Chapter 1.
Digital Collections\(^3\), provide the material conditions for a range of transformations that are non-deterministic, but nevertheless potential. The Hackathon for Cultural Heritage is one of the manifestations of these potentialities afforded by Digital Collections\(^4\). However, as we will see, the digital infrastructure and the processes they organize do not provide the complete set of agencies necessary for the realization of these events, because, as previously said, they are contained in the intersection of various ecosystems\(^5\). Based on the contributions of this work, the following paragraphs will discuss the dynamics and interactions between digital materials, tools, and outcomes with concepts, which are unique to the Cultural Sector, such as Heritage Interpretation. In addition, this chapter will contextualize and discuss how these materials, tools, and outcomes are operationalized in Hackathons so that shared interpretations about heritage are *constructed*. Last but not least, the chapter will discuss about the benefits and limitations of the Artfacts Platform within the context of institutional infrastructures and affiliated and non-affiliated communities.

### 7.1 Contributions

#### 7.1.1 Hackathons for Cultural Heritage as a Constructionist Method for Heritage Interpretation

This thesis proposes events known as *Hackathons for Cultural Heritage* as a constructionist method for the interpretation of heritage. Hackathons were originated from both the evolution of and revolution caused by digital computers. Originally, they have been implemented as a collaborative method for solving computer-related problems or conceptualizing new possibilities based on specific infrastructures\(^6\). Only later on, when Cultural Institutions had undergone intensive digitization, Hackathons started to be part of their repertoire. Because of the special nature of Cultural Institutions, Hackathons for Cultural Heritage cannot be understood in the same way as their counterparts happening in a purely engineering domain. Problem-solving and conceptualization through

\(^3\) For definition, see introduction to Chapter 5.
\(^4\) If certain principles are observed (see Chapter 4 and 5).
\(^5\) See Section 4.2.
\(^6\) See Chapter 2.
collaborative programming are entangled with the significance of the content matter they intend to deal with: the institutions’ collections\textsuperscript{7}. Not only, but especially in museums, the preservation of cultural artifacts is accompanied by educational mandates, which aim at supporting their audiences to understand and, ultimately, appreciate heritage. The cornerstone of educational departments in cultural institutions is a practice known as Heritage Interpretation, which aims at establishing meaningful and emotional connections between the public and heritage through storytelling\textsuperscript{8}. Heritage Interpretation has already a long tradition in Cultural Institutions. It has been used to e.g. contextualize artifacts in regard to their social and historical backgrounds, and draw parallels between the past, present, and even future through the objects of the collection\textsuperscript{9}. Although data-driven technologies have significantly changed the format in which Heritage Interpretation occurs, it has not however altered its essence. Collections have been digitized, but the need for making sense, appreciating, and preserving heritage remains.

7.1.1.1 Digital Interpretive Artifacts & Constructionist Heritage Interpretation

The Hackathon for Cultural Heritage is therefore an additional method for the interpretation of heritage that happens through the algorithmic recontextualization of digitally represented objects of collections as narratives, which are externalized as shared computer applications, and, for the most part, open source\textsuperscript{10}. These computer applications are defined here as Digital Interpretive Artifacts\textsuperscript{11} (DIAs), because they comprise digital interpretations of heritage. Interpreting and making sense of heritage happens not only through the utilization of DIAs by the audience of an institution, but as a process of digital fabrication and co-construction\textsuperscript{12}. In this sense, the thesis proposes constructionism as a theoretical foundation to understand how sense making occurs in these events by their participants\textsuperscript{13}. The term Constructionist Heritage Interpretation is introduced here in order to emphasize the means through which this Heritage Interpretation arises on top of infrastructures and principles that

\textsuperscript{7} See Chapters 3 and 5.
\textsuperscript{8} See Chapters 3 and 5.
\textsuperscript{9} See Appendix A.1.3.2.
\textsuperscript{10} See Section 4.2.
\textsuperscript{11} See Section 3.1.3.2.
\textsuperscript{12} See Section 3.1.3.
\textsuperscript{13} See Section 3.2.
are unique to the Cultural Sector. In constructionism, the process of constructing understandings about the world depends on the development of mental models, which is the product of personal appropriation of knowledge through situated *microworlds*\(^\text{14}\) that provide concrete contact with materials and supporting structures. In the case of Constructionist Heritage Interpretation, the development of mental models is situated in the intersection between institutional (i.e. GLAMs) and hacking (i.e. Hackathons) environments. It is also situated in the shared institutional infrastructures (i.e. format) that are used to preserve and enable the management of collections (i.e. content).

Constructionist Heritage Interpretation cannot be understood as a simple programming task. When constructing narratives, the storyteller/hacker/designer is required to step outside him or herself and think how to communicate the content to someone else in a way so that it can be comprehended and enjoyed. In addition, stories come to be as the result of *reflexive conversations*\(^\text{206}\) with not only oneself, team members, and staff of institutions, but also the dynamic media used to produce and convey the digital narrative. Constant conceptualization and reconceptualization, programming and debugging, planning and replanning, evaluating and reevaluating, reflecting and redefining happen towards an general idea that evolves by overcoming constraints and building up on successes\(^\text{205}\), instead of a task with a predefined goal. Hackathons are open-ended and experimental in nature. The construction of *DIAs*, the designers deputy\(^\text{15}\), expects therefore the construction and combination of different kinds of knowledge: knowledge of programming languages, knowledge of software engineering and user experience design, knowledge of information representation and databases, knowledge of communication and interactive storytelling, and knowledge of heritage, just to name a few.

---

14 Papert describes *microworlds* as a “subset of reality or a constructed reality whose structure matches that of a given cognitive mechanism so as to provide an environment where the latter can operate effectively. The concept leads to the project of inventing microworlds so structured as to allow a human learner to exercise particular powerful ideas or intellectual skills”\(^\text{50}\).

15 See Section 3.1.3.2.
7.1.1.2 The Hackathon’s Framework

Building on the contributions of Egolf et al. [84], Trainer et al. [229], and Richtericht [195], this thesis adds two extra phases to the framework of hackathons, namely launching and selecting. The launching phase\(^{16}\) consists of organizers introducing the infrastructures hackers will base their work on. The selecting phase\(^{17}\) consists of the jury or/and participants voting for the best

\(^{16}\) See Section 2.1.1.
\(^{17}\) Only in case of competitive Hackathons.
projects in the competition. As an introductory encounter with the infrastructures to be used, the launching phase may have a defining impact in the conceptualization of DIAs, that is because first decisions will be influenced by not only the characteristics of the infrastructure (as e.g. content and format of Digital Collections), but also how they will be presented by organizers to the participants. The subsequent phases (see Fig. 62), namely forming, storming, norming, performing, and staging will require participants to understand the content as well as the affordances and constraints of Digital Collections, evaluate strengths and weaknesses in regard to the participant’s skills so that a team can be organized, discuss the particular motivations of each team member and negotiate team roles, establish workflows so that collaboration can be enabled, and finally co-constructing DIAs.

7.1.1.3 The Hackathon’s Infrastructure

The Hackathon is a collaborative and open-ended method that must be enabled by equally collaborative and open-ended infrastructures. As explained in Chapter 5, a centralized mainframe solution for managing content in museums, materialized as the system GRIPHOS, was soon replaced by a distributed infrastructure comprised of several interconnected micro-computers. This distributed and interoperable infrastructure was made possible by the standardization of information representation. Different systems were able to communicate and process each other’s information provided that standards were adopted. Interoperability by standardization, although presenting some issues, is a cornerstone and powerful principle that remains until today, and it is used for interconnecting a variety of different kinds of Cultural Institutions. Its consolidation has enabled the establishment of ecosystems and initiatives, such as the Europeana project and the Open Knowledge Foundation. The Europeana project, an important platform for the aggregation and distribution of cultural data, is behind one of the first, if not the very first, Hackathon for Cultural Heritage in Europe. And the Open Knowledge Foundation, advocating the benefits of widespread utilization of open standards and data, organizes a series of Hackathons around the world.
including being one of the institutional entities responsible for the Hackathon Coding da Vinci\textsuperscript{23}, one of the cases addressed in this thesis. Thus, data-driven platforms and the processes that take place on top of and because of them are closely intertwined.

On the one hand, standardization\textsuperscript{24} has provided the material conditions for interoperable, distributed, and collaborative systems\textsuperscript{25}. On the other hand, modularity\textsuperscript{26}, a key principle of software platforms and standardized systems, has afforded extensible and open-ended infrastructures, that, when made accessible and open, enables the creation of vibrant ecosystems\textsuperscript{27}. In platforms, standardization provides a stable core upon which several kinds of building blocks can be arranged and rearranged in diverse ways. Communities are formed not only to construct modular extensions on top of platforms, but also around the open possibilities that the combination of different ready-made modules enable. New constructions are extensions of previous constructions. In this sense, Digital Collections are defined here also as platforms for Heritage Interpretation, because they present the same characteristics as platforms, and enable similar underlying possibilities\textsuperscript{28}. The modularity embedded in structured data establishes the possibility space\textsuperscript{29} that is afforded by the combinatorial capability of Digital Collections. Modular data elements are arranged, rearranged, and enriched to produce algorithmic narratives. Hackathons happen exactly within this possibility space. An open, accessible, modular, and extensible infrastructure is therefore essential for enabling Hackathons.

7.1.2 The Artfacts Platform

The Artfacts Platform (see Artfacts within Fast-speed IT Infrastructure on Fig. 62) was conceptualized and implemented within the context of the Two-speed IT Infrastructure\textsuperscript{30} approach and on top of principles, such as standardization and modularity. The platform is intended to provide open-ended, user-friendly, and

\textsuperscript{23} See Sections 2.1.3 and 6.3.2.
\textsuperscript{24} See Section 4.1.
\textsuperscript{25} See Chapter 5.
\textsuperscript{26} See Section 4.1.1.
\textsuperscript{27} See Section 4.2.
\textsuperscript{28} See Chapter 5.
\textsuperscript{29} See Chapter 5.
\textsuperscript{30} See Section 6.0.1.
interoperable interfaces so that the (re-)interpretation of Digital Collections can be supported and facilitated. The conducted empirical research indicated that the term hacker comprises individuals from a variety of professional backgrounds. Not only software developers take part in Hackathons for Cultural Heritage, but also artists, designers and professionals who work for Cultural Institutions, such as curators, researchers, and managers. The number of participants who engage with Hackathons for Cultural Heritage and have no programming background is therefore considerable. In addition to that, although it was identified a new kind of curator, who experiments with programming and engage themselves with the conceptualization and development of application as part of their professional activity, this multidisciplinary background is still far from characterizing curatorship. Therefore, tools that are intended to be integrated within the established infrastructures and support institutional work need to take the lack of programming skills into account.

Part of the Slow-speed IT Infrastructure of institutions (see Fig. 62), Collection Management Systems, although responsible for the creation of Digital Collections, are designed to administer and organize core institutional assets and enforce institutional procedures. The open-endedness necessary to afford a constructionist interpretation of heritage cannot therefore be accomplished through these systems. Interpreting Digital Collections through digital instruments means emphasizing certain data elements in detriment of others, enriching digital representations so that visibility to certain contents are given, and organizing and rendering these representation in a way so that stories can be told about heritage. Therefore, an Information System that aims at facilitating and supporting Heritage Interpretation should provide the storyteller with scenarios that can be composed from the interaction of multiple objects, instead of pre-defined procedures that must be accomplished.

In order to support the (re-)interpretation of Digital Collections by individuals with no technical background, the Artfacts Platform

---

31 See Section 6.1.
32 See Sections 2.1.2, 6.3.2, and A.2.1.1.
33 See Appendix A.2.1.1.
34 See Section 5.1.2.
35 See Section 5.1.2.1 and 5.1.2.2.
36 See Section 5.1.3.
employs graph organizers, understood here as Knowledge Maps\textsuperscript{37}, as the main elements of its GUI. Based on a user-friendly, generic, and extensible data model, Knowledge Maps do not enforce the accomplishment of pre-defined workflows, but instead enable narratives to be created from the combination of multiple objects and their relationships through Artfacts OOUI\textsuperscript{38}. In addition, via the Artfacts API\textsuperscript{39}, the structured and object-oriented approach of Knowledge Maps can be employed as interfaces for organizing and controlling the flow of system events and user actions\textsuperscript{40} in DIAs. By, on the one hand, enabling the enrichment of Digital Collections, and, on the other hand, providing leverage on the design and content of DIAs, the Artfacts Platform contributes especially to visibility\textsuperscript{41} of particular points of view that otherwise would depend on highly specialized knowledge to be accomplished. The project Marbles of Remembrance provides a argument in regard to this issue, because content producers\textsuperscript{42} were able to perform parallel research on specific victims, and, without the need of programming skills, were able to convert this research into structured data that enriched the content contained in the dataset Karten der Schülerkartei\textsuperscript{43}. Therefore, as Cultural Institutions are important social agents, widening the participation and possibilities of expression through data-driven means is then relevant for augmenting the visibility of important topics.

### 7.2 Limitations and Future Work

In regard to proposing Hackathons as a constructionist method for Heritage Interpretation, it is important to point out that this thesis does not advocate Hackathons as a better method in comparison to others, especially the ones that are not done through digital means\textsuperscript{44}. In fact, Hackathons appeal to a small number of individuals and are one of the many interpretive practices that only a few institutions around the world adopt. However, for those individuals who are

\textsuperscript{37} See Section 6.1.2.

\textsuperscript{38} See Sections 5.1.2.2 and 6.1.3.3.

\textsuperscript{39} See Section 6.1.1.2.

\textsuperscript{40} See Section 6.3.2.

\textsuperscript{41} See Section 5.1.2.

\textsuperscript{42} See Section 6.1.2.

\textsuperscript{43} See Section 6.3.2.

\textsuperscript{44} Such as guided tours and other interpretive activities (see Section 3.1.3)
accustomed and inclined to enjoy digital technology, this method might indeed be an effective interpretive approach.

As discussed, Hackathons offer some unique ways to comprehend heritage through digital means. However, the fast approach employed by this method in order to drive participants to come up with their own interpretations about heritage might appear flawed for some. That is because, in comparison, Cultural Institutions are rigorous, methodical, and slow in regard to the content and interpretations they produce. Hackathons, on the other hand, are highly experimental. One should however take into account that Hackathons operate in a different logic, which is more similar to the way Wikipedia and Open-source projects work, where quality comes from quantity [219]. For instance, several prototypes (interpretations) can be built on top of one dataset (see Fig. 62). In this sense, erroneous interpretations might occur as the result of these events, but the outcomes are refined during the selection process, which is supervised by the institutions themselves. In addition to that, as part of a constructionist process that is not bind to pre-defined outcomes, misinterpretations present also opportunities for learning. Due to time and resource constraints, this project was not able to produce evidence in regard to the quality of mental models (interpretations), externalized as Knowledge Maps and the DIAs produced in these events. During the CdV case study45, it was observed however that the participants of the Marbles of Remembrance team did indeed develop appreciation for the work carried out by the ITS and the victims represented in the dataset Karten der Schülerkartei. A more focused and larger study would need to be organized in order to produce evidence in regard to whether Hackathons are able to produce appreciation for heritage.

In concern to the Artfacts Platform, besides some technical limitations already discussed in Sec. 6.3.2.3, it is important to point out that the platform does not provide a solution that completely eliminates the need for programming. Although this is in fact not the objective of the platform, to provide greater leverage to non-programmers would be beneficial to widening the participation in regard to the way heritage is represented and digitally interpreted. Because of the flexible data model and GUI, the platform is already able to adapt to different concepts and implementations of interpretive applications, and offer indeed some leverage to non-programmers. However, further development is

45 See Section 6.3.2.
necessary to augment these capabilities. Finally, Knowledge Maps are proposed here as accessible scaffoldings to deal with information and institutional work. The use case scenarios presented in Sec. 6.2 provide an intuition of how the tool could be integrated as part of the institutional work carried out within institutions. A long-term study would be however required to verify their efficacy and how Knowledge Maps would suit personal working styles.
EMPIRICAL RESEARCH OVERVIEW

Phase 1: Curators
1. Requirements/Analysis
   - 1. Interview (Curators)
2. Implementation
   - 1. Prototype Development (Lisa Platform)
3. Validation
   - 1. Expert Reviews
   - 2. Focus Group Workshops
   - 3. Usability Test

Phase 2: Hackers
1. Implementation
   - 1. Prototype Improvement (Artifacts)
2. Validation
   - 1. Surveys
   - 2. Case Study OuSArchiv
3. Validation
   - 1. Case Study Hackathon
A.1 Phase 1

The phase 1 of this empirical study focused on understanding the curatorial settings in which Heritage Interpretation and content production occur in Cultural Institutions. The importance of that lies in the fact that Cultural Institutions are the entities responsible for providing the resources (raw material) that hackers use to create with. Therefore, it is essential to understand how the production of interpretive materials takes place inside institutions and how different groups deal with it. Special attention was also given to the role played by Information Systems. In this sense, this phase focused on acquiring a broader overview of the landscape of information systems that are direct and indirect utilized by the staff of institutions according to the literature and interviews (see Appendix B).

A.1.1 Requirement Analysis

A.1.1.1 Interview with Curators

OVERVIEW

- **Background**: Cultural Institutions play a fundamental role as informational resource provides for hacker communities.
- **Goal**: To understand how content is produced within institutions.
- **Duration**: This series of interviews took place between February and October 2014.
- **Method**: Semi-structured interviews and qualitative content analysis.
- **Number of Interviewees**: 4.

QUESTIONS

1. Is production of content performed in this institution?
2. How is your role in content production?
3. Can you give some examples of dense content material produced.
4. What tools do you use to aid you at your work?
5. Do you store all data produced during the research process (including your notes)?
6. Would you have an idea of the ratio between data analyzed vs. information that is officially used?
7. Do you have any strategy for reusing the data analyzed and the content produced that were not used (tagged content, stored in different piles, etc.)?

8. How is the information retrieved?

9. When you work on a particular project that requires you to perform research together with others, how is the content produced managed?

10. How do you evaluate (from excellent to poor) both the individual and collective research performed at the institution? Why?

11. Could you list five improvements you would change in the research process done in the institution?

12. Have you or the institution ever invited visitors or ordinary individuals to participate in the research process? If so, how was the experience?

13. Have you or the institution ever used content created by visitors or ordinary individuals in any of the museum media?

14. Does the institution have any channels for communicating with the public? If so, what are they?

SUMMARY OF RELEVANT FINDINGS

• Research and production of content is part of the daily work of cultural institutions. They range from material concerning public programs (debates, residencies, exhibitions, performances, open meetings and courses) to publications and media productions (social media posts, reports, books, papers, pamphlets, websites etc). The work of a curator is vast and diverse, and cannot be limited to the classical view that these individuals concern themselves with the conceptualization of exhibitions only.

• Curators produced large amounts of data and research information in order to create the curated product. Great part of this material is kept as analog private notes. Meetings and discussions may be supported by post-it notes and brainstorming techniques, such as mind mapping.

• None of the interviewed individuals used specialized curatorial software during the research, analysis, and content production process. The main digital solutions utilized included the Web for researching on topics, e-mail for collaborating internally and externally with third parties, Microsoft Excel and Word for organizing workflows, collaborations and producing content. Finally, Zotero was used for managing bibliographies. State owned institutions, as the case of the Städtische Galerie Bremen, had restrictions to utilize the Web and acquire new software even if freeware, because of government regulations and agreements with companies, such
as Microsoft. Collection Management Systems were important as institutional repositories of cultural assets and were fully integrated in digitization initiatives. These softwares were used to store official/curated content. They played little role during the research and content production processes, besides search and retrieval of digitized content. In the e.g. Kunsthalle Bremen, where the digitization process has not been completed yet, the curators also needed to utilize card catalogues to consult the collection.

- The interviewed curators were not skilled with computers, but were able to use their respective Collection Management Systems, or take part in software development projects that require expertise knowledge, and manage well the Microsoft Office Package.

- Collection Management Systems are either purchased off-the-shelf or tailor-made by third parties. As institutional tools, they support search and retrieval of items in the digital collection, but do not tackle specific research and content production problems. According to some participants, these systems also do not help to improve communication among different departments (e.g. curatorial and educational depts.), and productivity during research (besides facilitating digital access to the collection) so that curators can cope with e.g. tight deadlines and pressure from other department for publicize content.

- The institution Associação Cultural Videobrasil was involved in a collaborative and experimental open access project called Plataforma:VB (http://plataforma.videobrasil.org.br/) that required the indirect participation of many artists to provide unique content about themselves and their works. This exploratory tool was conceptualized to offer insights on how different topics, artworks, and artists interconnected. These insights were obtained based on cartography of tagged content exclusively produced by the institution and aimed at contextualizing different entities by displaying them as a network diagram. Users however were unable to modify the content on the website of the project.

- Based on the analysis of the interviews, it was identified that curators, although not having advanced computer skills, mastered the software required for accomplishing their curatorial tasks. This indicates that the introduction of new software is possible, if it is capable of adding value to the curatorial work. There was a lack of software alternatives to support curators during the content production phase. The development of software that would enable curators to experiment with the data gathered during their own research could be beneficial to the reuse of the stored information not only in later projects, but also by colleagues working on similar topics. Such software could have direct impact in the curatorial practice of
institutions and the production of support material that concern the institutions’ audiences.

A.1.2 Implementation

A.1.2.1 Development of Prototype

Figure 63: First version of the prototype

OVERVIEW

- **Background:** Based on the literature review, interviews, requirement analysis, and research on information systems used directly and indirectly by cultural institutions (see Appendix B), a prototype was developed with the objective of implementing a data-driven curatorial workflow, which would provide an extra infrastructure to support research, experiment with interconnections, and produce lightweight datasets based on the curators own knowledge.

- **Goal:** The goal of the prototype was therefore to provide more freedom for alternative and experimental interpretations that are separated/independent from the official resources/infrastructure, but nevertheless interconnected. Such (prototype) model may benefit communities directly or indirectly by providing hackers with alternative datasets/interpretations produced by curators, accessibility to the dataset through an API, and the possibility of hackers utilizing themselves the platform as a means to facilitate the conceptualization and implementation of data-driven interpretive applications. In addition, the development of the working model in an early stage of the research project had as goal to implement an iterative design process so that the collection of data, the evaluation
and validation of requirements could be used to reflect on the concepts dealt in the thesis. Finally, the prototype served to facilitate communication with target-audiences.

- **Duration**: The development of the first version of the prototype took place between December 2014 and April 2015.
- **Method**: Iterative Design Approach.
- **Reference**: More information on the first version of the prototype, please see [61].

A.1.3 **Validation**

A.1.3.1 **Reviews from Experts**

**OVERVIEW**

- **Background**: The interviews were arranged after the implementation of the first version of the platform was concluded.
- **Goal**: To review the requirements, pose questions, and gather feedback mainly on usability aspects of the platform.
- **Duration**: This series of interviews took place between May and December 2015.
- **Method**: Semi-structured interviews and qualitative content analysis.
- **Number of Interviewees**: 8.

**QUESTIONS**

1. Please, provide feedback on the following functionalities of the platform: Toolbar Button, Context Menu Options, Tagging System, Knowledge Map, and Graph Search.
2. How do you evaluate the look and feel of the User Interface?
3. What are the positive and negative aspects you see the platform is able to provide within a curatorial/research setting?
4. Do you have any extra recommendation?

**SUMMARY OF RELEVANT FINDINGS**

- **Positive Feedback**:
  - Because the tool is integrated in the browser as a plugin, it is provided easy access to the web (abundant resources, fast information retrieval).
  - The tool forces the tagging of entities. Therefore, it helps the individuals to think about important key concepts in the text.
– The tool supports reasoning and expression of complex concepts/ideas (graph framework provides the necessary foundation for storing relational information).

– The tool promotes self-re-exploration (if the topic is already represented by the system), because sub-topics (interconnections) naturally appear as the more information is inserted on the Knowledge Maps.

– Initial contact with the Information Design Module seems to be very intuitive and easy to follow.

– The tool seems to be flexible enough to allow the development of personal working styles regarding its functionalities and processes.

– The discovery of sub-topics during the usage of the tool can be very fruitful for museum research, because there might appear interesting themes that can be explored within an exhibition.

– The tool promotes sense making from large amounts of data.

– Applying different filters help to reduce the Knowledge Maps in a way they become readable.

– The categories used by the system seem to be enough for modeling information.

– The application saves time in a research project because it allows the definition of actors and relationships, and helps to visualize them in a way to drive conclusions that would not be possible without it.

– In order to reason about large data sets, reduction of information by selection is necessary. Therefore, machine generated statements based on the Knowledge Maps helps the inference process, because they can be seen as a point of view on the topic.

– The application is useful in the exploration phase of a topic.

• Negative Feedback:

– Lack of current hierarchical strategies for organizing information. Parallel to purely rhizomatic approaches, hierarchical structures are necessary when creating narratives, applications, and taxonomies for visitors. The horizontality the system enforces might not be the right strategy for museum research, as hierarchy is used to construct museum narratives.

– Graph search needs a considerable amount of information in the repository to retrieve meaningful related information.

– Annotate all the data is problematic. Pre-processing (Automation) of data is necessary for better clarity of the Knowledge Maps.
- Transition from the analog space (in the sense that some researchers produce analog material) to the digital space can be time consuming and demanding.

- Uncertain whether the vocabulary, regarding the entities and the relationships, are compatible with the researcher’s personal working style. More time with the tool is needed. The tool might restrict and impose certain procedures of research that might not be desired by the researcher. The researcher needs to get fluent in the tool to develop working styles with the tool.

- Difficult to understand the goal and the functionalities of the system. More explanation is necessary. Using the tool for collecting sources is an affordance the system provides, but is not explicitly explained. Necessary to invest time in order to take advantage of the full capabilities of the tool. (High degree of motivation). Individuals do not immediately understand/recognize the value of mapping information as a reasoning strategy. Learning curve - The tests do not allow for a deeper understanding and usage of the tool. Therefore tests might generate inaccurate results.

- The user interface presents some technical problems regarding usability and bugs.

- The Knowledge Maps created cannot be reused in other platforms, such as website of the institution.

- The tool is not connected with external data repositories, and therefore important cultural data cannot be reused.

- Co-occurrence networks are one, but not the only way information can be represented.

**Recommendations:**

- The graph search needs to take into account nodes that are connected not only through strong relationships, but also weak ones. Paths that connect not direct connected entities is also helpful to read information stored on the maps.

- The Knowledge Maps need to be enriched with more semantics. Otherwise, it is not possible to extract meaning from them, reducing therefore their usefulness.

- Options for custom definition of relationships might be useful to model information in a more accurate way.

- Hierarchical strategies for organizing data might is useful especially in museum contexts.

- More supporting material is needed. Supporting material should be in the format of scenarios displaying common problems the application helps to solve.
The system should provide a summary on a certain topic or map. This is useful to aid sense-making of the data contained on the archives.

A.1.3.2 Workshop 1

OVERVIEW

- **Background**: The cog exhibition is unique, because it makes evident the role of the original object in the orientation and the role of the museum. For the DSM, the Bremen Cog is on the spotlight because of its historical relevance. Therefore, the Bremen Cog receives a great deal of attention, which is practically translated into museum resources as e.g. exhibition space and time spent on research. Recently, the museum is ongoing renovation. The "Kogge-Halle" is seen as first step and key for implementing the new DSM.

- **Goal**: To obtain insights into the interpretation process happening in the German Maritime Museum (DSM - Bremerhaven, Germany) and feedback on the first version of the prototype.

- **Date, Duration and Location**: 23.03.2016, 4 hours (20 min. for presenting and discussing about the first version of the prototype), Leibniz- Gemeinschaft Berlin.

- **Method**: Focus Group (Expertengespräch Vermittlungskonzept Kogge-Halle).

- **Structure and Content**:
  - I. Welcome and Introduction
  - II. Principles of the semi-permanent exhibition
  - III. Principles of the hands-on models
  - IV. Principles of digital media
  - V. Principles of accessibility and family friendliness
  - VI. Visitor research

- **Number of Participants**: 14.

- **Additional Information**: The "Kogge-Halle" was inaugurated in March 2017 (see Fig. 44).

SUMMARY OF RELEVANT FINDINGS

- **General**:
  - The main exhibition of the German Maritime Museum is the Bremer-Kogge (Bremen Cog).
  - One of the concerns of the exhibition is to try to explore topics about the finding, recovery and conservation of the Bremen Cog.

1 Picture extracted from: https://www.youtube.com/watch?v=eIS7fyU1DNy
- The exhibition tries to answer questions, such as: How was the ship constructed? What were the materials used? Who was the crew of the Bremen Cog? How was life on-board? How was the Bremen Cog navigated? What was the socio-historical context in which the Bremen Cog was used (laws, associations, trade, etc.)?

- **Research-oriented Exhibition:**
  - Ms. Weber sees the emphasis on the science showcase as an opportunity to explicitly emphasize the strengths of the DSM as a research museum.
  - The visitor should be able to interact and experiment with “research” (science as mediation).
  - Besides the questions described above, the exhibition also explores questions related to the current research activity carried out at the museum that deals with the topics: materiality, interests, perception. How does a ship come into being (physically and in the imagination)?
  - The new exhibition concept provides researchers a direct view and access to the exhibition area. The goal is to make the connection with science evident to the visitor and provide researchers higher accessibility to the original objects (abolition of the boundaries between restoration and exhibition).
  - The museum wants to give the visitor the impression that the visitor the part of the restoration process.
  - The visitor must be given a very precise explanation of how and why the exhibition and research activity are integrated.
  - The museum wants the visitor to formulate questions that could also be posed in the subject of shipping in the 21st century.
  - The exhibition design must engage the visitor with research questions and give them the chance to participate.
  - The researcher will take on a new role. Frequent announcements of current status of research will be displayed in the exhibition space. The goal is to value the research work and the researcher.
  - The exhibition encourages the conversation between visitors and researchers.

- **Semi-permanent Exhibition:**
  - The goal of the semi-permanent exhibition is to increase flexibility by having part of the content changing frequently.
  - The changes should relate to the new developments of the research being carried out at the museum. "Fresh” research data should be quickly communicated in the exhibition. So that the exhibition space is transformed also in a platform for the sharing of scientific data.
• **Contextualization:**
  - Addressing current topics and making parallel with historical event may bring value to the exhibition. Possibility for new research questions and participation of the visitor.
  - Comparison between the Hanseatic League with the EU (Hanseatic League = EU of the Middle Ages). According to Mr Puhle, one could see to what extent the Hanseatic League is one of the foundations of today’s Europe and where structural similarities exist.
  - Parallels between piracy in the late Middle Ages and today.

• **Hands-on Models and Digital Media:**
  - Hands-on models of the Bremen Cog can touched and rotated. Tangible objects are seen as attractive for young visitors.
  - New technology, such augmented reality glasses are considered. The definition of a concept is still unclear.
  - Location-aware audio-guides promote fast guidance and spatial orientation by means of guiding objects around the exhibition space. The guides also give visitors the possibility to choose among different guided-tours (e.g. guided tours to hands-on objects for the visually impaired).
  - Short interviews with researchers should be included as part of the content of the exhibition (video appearances and audio-guide content).

• **Review on the First Prototype:**
  - The first version of the prototype was seen as suitable as a research tool. However, the participants cannot envision how the tool could be used within the exhibition.
  - The Knowledge Maps created by the application are seen as too complex to be understood by visitors. The information displayed in the exhibition space should be reduced.

**CONCLUSIONS**

• In the case of museums dealing with historical objects, it is evident the connection between research and the exhibition. The new management of the DSM wanted to stress even more this characteristic - The German Maritime Museum as a “research museum”. In order to do so, the interpretation program of the DSM aimed at placing science as a mediator for the interpretation of its collection. The visitor is invited to put him- or herself in the shoes of a researcher. He or she is presented with research questions and try to answer them by experimenting with the exhibits. The challenge with this approach is reducing the complexities of the scientific method to a level that is understandable, interesting and entertaining for
visitors to engage with. It is clear however that in the DSM, not only the research, but also the educational and interpretive programs gravitate around the original objects, especially the Bremen Cog. The still unanswered research questions make this object a source for interpretation (not only scientific, but also curatorial). In the case of the DSM, a tool that is capable to address heritage interpretation should therefore also support scientific research, and facilitate the reduction of complexity necessary to make content suitable to the ordinary visitor.

A.1.3.3 Workshop 2

OVERVIEW

• Background: The workshop was arranged after the implementation of the conclusion of the first version of the platform.

• Goal: To review the requirements, pose questions, and gather feedback mainly on usability aspects of the platform.

• Date, Duration and Location: 27.06.2016, 3:30 hours, MZH Universität Bremen,

• Method: Focus Group.

• Structure and Content:
  – I. Introduction (30 min.) - Mapping: What it is and why you should use it
  – II. Practical Part (60 min.) - Hands on Drawing
  – III. Demo (30 min.) - The Artfacts Platform
  – IV. Practical Part (60 min.): Creating Knowledge Maps
  – V. Discussion (30 min.)

• Number of Participants: 7.

SUMMARY OF RELEVANT FINDINGS

• Although the participants (curators) produce content to be used in platforms, such as Europeana, they do not understand the technology.

• Confirming the results obtained in the first round of interviews, curators are able to manage the CMS, but they do not use the CMS for producing curatorial content. Instead, the Web and books are the main tools for supporting content production.

• The CMS is therefore not a part of the creative process. As stated by one of the participants, the CMS “is a database for collections, and not a database for ideas”.

191
• In regard to Europeana, the Museen Böttcherstraße was the only one collaborating to the project. The curators were not involved with the technical part. The museum has only provided the dataset. The Europeana personnel were responsible for inserting it into the Europeana database.

• Data repositories are created as an export capability of the CMS and usually managed by IT personnel.

• Curators would engage in creating data repositories if the application is really easy to use and understand.

• In regard to the usability of the data model of the application, it is necessary to use the minimum knowledge expressivity/ granularity as possible (the more expressivity/ granularity, the more difficult to the user).

• While it is possible to easily export data out from the CMS, it is difficult to import information back to the CMS in a suitable way. According to one of the participants, if you export data as a spreadsheet, it is difficult to import the spreadsheet back into the system. In addition, the integration with third party applications is also problematic. New information that is created with other apps can only be inserted into the system by copying and pasting.

• The CMS used by the Übersee-Museum (TMS) is not capable of reusing the data of different sources (such as DBPedia) in order to provide more or contextual information on museum objects stored into the system. According to the director of the museum, that would be a useful feature.

• Cultural Institutions have invested a considerable amount of money on their commercial CMSs. Therefore, they are not willing to replace them for systems that do not offer the maintenance and support that traditional companies do. Companies also provide support in converting old data formats into the ones compatible with their software.

• The Übersee-Museum has used mind mapping for exhibition concepts. However, they do not use software for this matter, but draw the maps on paper. Interns are the ones in charge to digitizing the outcomes of the brainstorming sessions.

• Integrating Knowledge Maps into a more comprehensive workflow is difficult for participants to envision. But, connecting different museum objects together with a text recently produced in order to provide insights is seen as a good idea.

• A tool cannot augment complexity. On the contrary, a tool to serve interpretation and the visitor should reduce complexity.
• Museums are searching for solutions that can bring together the analog world of exhibitions and the digital world through digital installations in the museum or using the mobile devices of the visitors.

• In exhibition planning one must switch from a 2D thinking to a 3D thinking, by providing experiences where the visitor can look at objects in different perspectives.

A.1.3.4 Usability Test

OVERVIEW

• Background: The software needed to be accessed in regard to one of its main requirements, which consisted of supporting individuals with no technical background in the collection of information and its cartography.

• Goal: To evaluate the usefulness, effectiveness, satisfaction, and learnability of the Tagging System.

• Date: Throughout May 2016.

• Method: Semi-structured interview, video and sound recording (think out aloud), and System Usability Scale (SUS).

• Number of Participants: 10.

• Additional Information: See Item 6.3.1 for a detailed description.
### SUS - Detailed Results

**System Usability Scale (SUS)**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Statement 1</th>
<th>Statement 2</th>
<th>Statement 3</th>
<th>Statement 4</th>
<th>Statement 5</th>
<th>Statement 6</th>
<th>Statement 7</th>
<th>Statement 8</th>
<th>Statement 9</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>User 1</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>85.0</td>
</tr>
<tr>
<td>User 2</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>85.0</td>
</tr>
<tr>
<td>User 3</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>80.0</td>
</tr>
<tr>
<td>User 4</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>75.0</td>
</tr>
<tr>
<td>User 5</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>85.0</td>
</tr>
<tr>
<td>User 6</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>88.0</td>
</tr>
<tr>
<td>User 7</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>82.0</td>
</tr>
<tr>
<td>User 8</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>75.0</td>
</tr>
<tr>
<td>User 9</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>78.0</td>
</tr>
<tr>
<td>User 10</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>49.0</td>
</tr>
</tbody>
</table>

**Average** 75.75

SUS score above 68 would be considered above average and anything below 48 is below average.

To calculate the SUS score, first sum the score contributions from each item. Each item's score contribution will range from 0 to 4. For items 1,3,5,7, and 9 the score contribution is the scale position minus 1. For items 2,4,6, and 10, the contribution is 5 minus the scale position. Multiply the sum of the scores by 2.5 to obtain the overall value of SUS.
**Phase 2**

The phase 2 of this empirical study focused on getting a clear picture of the "hacker" and understanding the communities in which these individuals are part of. The goal of this phase was also to compare the results previously obtained through the qualitative research done during the first phase with the results obtained via quantitative methods in order to draw more accurate conclusions. In addition to that, two case studies were carried out. The first had as goal to acquiring insights in how CMSs organize the workflows in cultural
institutions. And the second and final case study had as objective to analyze how the second version of the prototype was handled in an actual hackathon setting, and what its contributions and disadvantages really were.

A.2.1 Analysis

A.2.1.1 Online Survey A: Online Repositories for Cultural Heritage

Overview

- **Background & Objective:** Not only quantitative validation for the results gathered by the past qualitative methods were necessary to be obtained, but also it was unclear how communities perceived digital collections in regard to their employment in the development of applications. Therefore, this survey was conducted in order to evaluate how well data repositories (digital collections) released by cultural institutions were responding to the communities’ needs.

- **Duration:** Conducted from August 2017 to October 2017.

- **Number of Interviewees:** 108.

- **Method:** The online survey was conducted through the website SurveyMonkey (https://www.surveymonkey.com/) and utilize the channels described below, because of their representativeness in the cultural heritage sector among communities that work with digital collections:
  - Mailing List of the EuropeanaTech Community: europeana-tech@list.ecompass.nl
  - Mailing List of the Open Knowledge Foundation Community: okfn-en@lists.okfn.org
  - Mailing List of the UK Education and Research Communities: cultural-heritage-datasets@jiscmail.ac.uk

General Results and Preliminary Analysis

Please, see content on the next page.
Question 1:

What is your gender?

<table>
<thead>
<tr>
<th>Answer Choices</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>60.19%</td>
</tr>
<tr>
<td>Male</td>
<td>39.81%</td>
</tr>
</tbody>
</table>

Answered 108
Skipped 0

Question 2:

What is your age?

<table>
<thead>
<tr>
<th>Answer Choices</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>under 17 years</td>
<td>0.00%</td>
</tr>
<tr>
<td>18 to 24</td>
<td>0.00%</td>
</tr>
<tr>
<td>25 to 34</td>
<td>17.59%</td>
</tr>
<tr>
<td>35 to 44</td>
<td>29.63%</td>
</tr>
<tr>
<td>45 to 54</td>
<td>34.26%</td>
</tr>
<tr>
<td>55 or older</td>
<td>18.52%</td>
</tr>
</tbody>
</table>

Answered 108
Skipped 0
**Question 3:**

**What is your profession?**

<table>
<thead>
<tr>
<th>Answer Choices</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curator</td>
<td>18.52%</td>
</tr>
<tr>
<td>Researcher</td>
<td>34.26%</td>
</tr>
<tr>
<td>Artist</td>
<td>5.56%</td>
</tr>
<tr>
<td>Programmer</td>
<td>21.30%</td>
</tr>
<tr>
<td>Graphic Designer</td>
<td>2.78%</td>
</tr>
<tr>
<td>Retired</td>
<td>0.93%</td>
</tr>
<tr>
<td>Student</td>
<td>1.85%</td>
</tr>
<tr>
<td>Other (please specify)</td>
<td>40.74%</td>
</tr>
<tr>
<td><strong>Answered</strong></td>
<td><strong>108</strong></td>
</tr>
<tr>
<td><strong>Skipped</strong></td>
<td><strong>0</strong></td>
</tr>
</tbody>
</table>

- **Curator:** 18.52% (20 responses)
- **Researcher:** 34.26% (37 responses)
- **Artist:** 5.56% (6 responses)
- **Programmer:** 21.30% (23 responses)
- **Graphic Designer:** 2.78% (3 responses)
- **Retired:** 0.93% (1 response)
- **Student:** 1.85% (2 responses)
- **Other (please specify):** 40.74% (44 responses)

**Respondents (other please specify):**

- Librarian
- Editor
- Technology systems and projects manager
- IT
- Project manager digital heritage
- Coordinator projects
- Consultant
- Information scientist
- Librarian
- Project Coordinator
- Project manager
- Freelance, digital
- Application manager
- Repository manager
- Librarian
- Project Manager
- Librarian
- Digital Assets Manager
- Librarian
- Conservator
- Librarian
- Librarian
- Entrepreneur
- Librarian
- Manager
- Standards manager
27 Data manager
28 Psychologist
29 Project Manager
30 Web CMS platform leader
31 manager
32 PR specialist
33 Software Developer
34 Designer,inker
35 Librarian/Data Analyst
36 IT architect
37 Digital Humanities Manager
38 Data Manager
39 Manager of Maker Hub, Creative Cultural Consultant
40 Online Communications
41 Bit of a troublemaker
42 Consultant
43 Interaction Designer
44 Communication officer

**Question 4:**

**How many times have you taken part in hackathons before?**

<table>
<thead>
<tr>
<th>Answer Choices</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>51.96%</td>
</tr>
<tr>
<td>1</td>
<td>8.82%</td>
</tr>
<tr>
<td>2</td>
<td>12.75%</td>
</tr>
<tr>
<td>3</td>
<td>10.78%</td>
</tr>
<tr>
<td>4</td>
<td>1.96%</td>
</tr>
<tr>
<td>5 or more</td>
<td>13.73%</td>
</tr>
</tbody>
</table>

Answered 102
Skipped 6

![Bar chart showing the frequency of hackathon participation](chart.png)
**Question 5:**

Have you worked directly with online data repositories for the cultural heritage field?

<table>
<thead>
<tr>
<th>Answer Choices</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>88</td>
</tr>
<tr>
<td>No</td>
<td>14</td>
</tr>
<tr>
<td>Answered</td>
<td>102</td>
</tr>
<tr>
<td>Skipped</td>
<td>6</td>
</tr>
</tbody>
</table>

**Question 6:**

Have you ever needed to adapt the concept of an app, because the dataset you had available was unsuitable for producing it?

<table>
<thead>
<tr>
<th>Answer Choices</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes, I had to adapt my concept to the data or I had to use other sources</td>
<td>58</td>
</tr>
<tr>
<td>No, The repository(ies) contained all data I needed for me to develop my</td>
<td>18</td>
</tr>
<tr>
<td>Answered</td>
<td>76</td>
</tr>
<tr>
<td>Skipped</td>
<td>32</td>
</tr>
</tbody>
</table>

---

200
Could you please name a few the repositories you have worked with?

<table>
<thead>
<tr>
<th>Repository Name</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKÖ Collection Online (<a href="https://github.com/MK%C3%96Hamburg/MK%C3%96CollectionOnline/LDO/XML">https://github.com/MKÖHamburg/MKÖCollectionOnline/LDO/XML</a>)</td>
<td>13</td>
</tr>
<tr>
<td>This question doesn’t apply to me because I haven’t worked with data repositories before</td>
<td>0</td>
</tr>
<tr>
<td>Deutsche Digitale Bibliothek (<a href="https://github.com/Deutsche-Digitale-Bibliothek">https://github.com/Deutsche-Digitale-Bibliothek</a>)</td>
<td>21</td>
</tr>
<tr>
<td>The Metropolitan Museum of Art Open Access (<a href="https://github.com/metmuseum/openaccess/bikinimaster/MetObjects.csv">https://github.com/metmuseum/openaccess/bikinimaster/MetObjects.csv</a>)</td>
<td>13</td>
</tr>
<tr>
<td>Others</td>
<td>43</td>
</tr>
</tbody>
</table>

Please, specify the name of the institution(s) that offered the data repository(ies)

- Answered
- Skipped

**Responses**

1. dlrs-hh.urkoe.de
2. Stadtmuseum Berli, Europäana, ODB, Ethnologisches Museum, Naturkundemuseum, GLAM
3. GLAM
4. National Library of Israel
5. Cenes, Memoria de Madrid
6. normacranes.es, www.itapaia.es
7. gaifca.bnf.fr
8. Europian
9. National Museum of Ireland
10. Europian, Biblioteca nacional de España, biblioteca Bélica
11. CARR, Documentos y Archivos de Aragón (http://www.dataragon.es)
12. various
13. Github in general
15. Wikipedia Commons, using the MediaWiki API
16. europeana
17. Wikimedia Commons, Wikidata, Wikipedia
18. British Museum etc.
20. Cannot remember
21. europian
22. Wikidata, Europian
23. Europian
25. JSTOR
26. File:Barcel JosepHimself11 Llandudno01 png 151 (pg 3)
27. or another from site: Austrian archives
29. oai data
30. Europian
31. Museum Digital
32. Archivo FENP heritage
| National Gallery of Denmark, Rijksmuseum (The Netherlands), Museum für Naturkunde Berlin, Botanischer Garten & Botanisches Museum Berlin-Dahlem, Sound & Vision (The Netherlands), YLE (national Finnish broadcaster) plus archives such as Internet Archive, the Free Music Archive, Pêler Creative Commons collections, Public Domain Review collections and Freesound etc. |
| The Henry Ford (not a repository, but was given a partial data set when I asked if they would be willing to share their data with me.) |
| 35 Academic and research institutions | National Library of the Netherlands, Netherlands Institute for Art History, Dutch National Archives |
| 37 OPenn (http://openn.library.upenn.edu) | Data feeds from Dorset County Museum, University of Exeter, and other non-public sources. |
| Europeana, National Maritime Museum (UK), Flickr, Wikidata, Imperial War Museum, Science Museum (UK) | |
| 42 1FP compliant resources from Libraries worldwide | Science Museum London (just opened their api so I ran Science Museum Hack for them and we’re running another one in early 2018) |
| 46 SMK, Skagens Museum | }
Question 8:

In your opinion, the way the data is structured (LIDO-XML, DC-XML, LD-JSON, etc) is... 

<table>
<thead>
<tr>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complicated: I needed considerable time to understand the structure and parse the data fast and with no further complications, nor simple. Because...</td>
</tr>
<tr>
<td>Simple: I could understand the structure and parse the data fast and with no further complications, nor simple. Because...</td>
</tr>
<tr>
<td>Neither complicated, nor simple. Because...</td>
</tr>
</tbody>
</table>

Answered: 72, Skipped: 36

The level of difficulty in working with datasets depends on the technical knowledge of the user, the use case in which the data is being applied to, previous familiarity with the metadata structure, and depth/level in which the dataset is being used. For some individuals, the use of standards facilitates because they are well documented. For others, standards are complicated.
Question 9:

Would you like to have a tool to allow you to easily extend and modify existing data repositories?

<table>
<thead>
<tr>
<th>Answer Choices</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes. Enriching the dataset would give me more possibilities in what I can do</td>
<td>82.67%</td>
</tr>
<tr>
<td>No. I’m satisfied with the existing repositories!</td>
<td>17.33%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>75</strong></td>
</tr>
</tbody>
</table>

Skipped: 33

COMPARATIVE RESULTS AND PRELIMINARY ANALYSIS

Question 1:

What is your gender?

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1: Curator</td>
<td>57.14%</td>
<td>42.86%</td>
<td>14</td>
</tr>
<tr>
<td>Q2: Researcher</td>
<td>31.25%</td>
<td>68.75%</td>
<td>32</td>
</tr>
<tr>
<td>Q3: Programmer</td>
<td>4.55%</td>
<td>95.45%</td>
<td>22</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>31.03%</td>
<td>68.97%</td>
<td>58</td>
</tr>
</tbody>
</table>

Answered: 58
Skipped: 0

What is your gender?

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1: Curator</td>
<td>57.14%</td>
<td>42.86%</td>
</tr>
<tr>
<td>Q2: Researcher</td>
<td>31.25%</td>
<td>68.75%</td>
</tr>
<tr>
<td>Q3: Programmer</td>
<td>4.55%</td>
<td>95.45%</td>
</tr>
</tbody>
</table>

Compared analysis and filtered by question #5 (Value: YES)
**What is your age?**

<table>
<thead>
<tr>
<th></th>
<th>under 17 years old</th>
<th>18 to 24</th>
<th>25 to 34</th>
<th>35 to 44</th>
<th>45 to 54</th>
<th>55 or older</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q3: Clinician</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>14.29%</td>
<td>2.00%</td>
<td>42.86%</td>
<td>6</td>
</tr>
<tr>
<td>Q3: Researcher</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>18.75%</td>
<td>6.00%</td>
<td>31.25%</td>
<td>10</td>
</tr>
<tr>
<td>Q3: Programmer</td>
<td>0.00%</td>
<td>0.00%</td>
<td>27.27%</td>
<td>46.91%</td>
<td>9.09%</td>
<td>22.72%</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>18.97%</td>
<td>31.03%</td>
<td>31.03%</td>
<td>11</td>
</tr>
</tbody>
</table>

**Answered: 58**  
**Skipped: 0**

**What is your age?**

- **Q3: Clinician**
  - **Under 17 years old:** 0%
  - **18 to 24:** 4.29%
  - **25 to 34:** 13.25%
  - **35 to 44:** 22.38%
  - **45 to 54:** 31.30%
  - **55 or older:** 28.30%

- **Q3: Researcher**
  - **Under 17 years old:** 0%
  - **18 to 24:** 13.25%
  - **25 to 34:** 31.25%
  - **35 to 44:** 18.75%
  - **45 to 54:** 18.75%
  - **55 or older:** 31.03%

- **Q3: Programmer**
  - **Under 17 years old:** 0%
  - **18 to 24:** 27.27%
  - **25 to 34:** 40.91%
  - **35 to 44:** 22.38%
  - **45 to 54:** 9.09%
Question 3:
### How many times have you taken part in hackathons before?

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 or more</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q3: Curator</td>
<td>71.43%</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14.29%</td>
<td>2</td>
</tr>
<tr>
<td>Q3: Researcher</td>
<td>53.13%</td>
<td>17</td>
<td>6</td>
<td>21</td>
<td>7</td>
<td>0.00%</td>
<td>2</td>
</tr>
<tr>
<td>Q3: Programmer</td>
<td>31.82%</td>
<td>7</td>
<td>9</td>
<td>18</td>
<td>1</td>
<td>0.00%</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>50.00%</td>
<td>29</td>
<td>5.17%</td>
<td>16.52%</td>
<td>5.62%</td>
<td>0.00%</td>
<td>12</td>
</tr>
</tbody>
</table>

Answered: 58
Skipped: 0
Have you ever needed to adapt the concept of an app, because the dataset you had available was unsuitable for producing it?

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>This question doesn't apply to me because I haven't worked</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q3: Curator</td>
<td>69.23%</td>
<td>9</td>
<td>30.77%</td>
<td>4</td>
</tr>
<tr>
<td>Q3: Researcher</td>
<td>86.96%</td>
<td>20</td>
<td>13.04%</td>
<td>3</td>
</tr>
<tr>
<td>Q3: Programmer</td>
<td>83.33%</td>
<td>15</td>
<td>16.67%</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>82.22%</td>
<td>37</td>
<td>17.78%</td>
<td>8</td>
</tr>
</tbody>
</table>

Answered: 48, Skipped: 13

Yes, I had to adapt my concept to the data or I had to use other sources to complement the missing data.

No, the repository(ies) contained all data I needed for me to develop my initial concept.

This question doesn't apply to me because I haven't worked with data repository(ies) before.
In your opinion, the way the data is structured (LIDO-XML, DC-XML, LD-JSON, etc) is...

<table>
<thead>
<tr>
<th></th>
<th>Complicated: I needed considerable time to understand the structure and parse the data.</th>
<th>Simple: I could understand the structure and parse the data fast and without any problem.</th>
<th>Neither complicated, nor simple: Because...</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q3: Curator</td>
<td>45.40%</td>
<td>18.16%</td>
<td>36.29%</td>
<td>11</td>
</tr>
<tr>
<td>Q3: Researcher</td>
<td>38.13%</td>
<td>26.09%</td>
<td>34.78%</td>
<td>23</td>
</tr>
<tr>
<td>Q3: Programmer</td>
<td>68.67%</td>
<td>16.67%</td>
<td>18.70%</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>51.16%</td>
<td>15.80%</td>
<td>30.03%</td>
<td>43</td>
</tr>
</tbody>
</table>

In your opinion, the way the data is structured (LIDO-XML, DC-XML, LD-JSON, etc) is...

- **Q3: Curator**
  - Complicated: 45.40%
  - Simple: 18.16%
  - Neither complicated, nor simple: 36.29%

- **Q3: Researcher**
  - Complicated: 38.13%
  - Simple: 26.09%
  - Neither complicated, nor simple: 34.78%

- **Q3: Programmer**
  - Complicated: 68.67%
  - Simple: 16.67%
  - Neither complicated, nor simple: 18.70%

- **Total**
  - Answered: 43
  - Skipped: 15
Would you like to have a tool to allow you to easily extend and modify existing data repositories?

<table>
<thead>
<tr>
<th></th>
<th>Yes, Enriching the dataset would give me more possibilities in what I can create with data!</th>
<th>No, I'm satisfied with the existing repositories!</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2: Curator</td>
<td>11</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>Q3: Researcher</td>
<td>20</td>
<td>3</td>
<td>23</td>
</tr>
<tr>
<td>Q3: Programmer</td>
<td>13</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>37</td>
<td>8</td>
<td>45</td>
</tr>
<tr>
<td>Answered</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skipped</td>
<td>13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Would you like to have a tool to allow you to easily extend and modify existing data repositories?

- Q3: Curator
  - Yes, Enriching the dataset would give me more possibilities in what I can create with data: 84.62%
  - No, I'm satisfied with the existing repositories: 15.38%

- Q3: Researcher
  - Yes, Enriching the dataset would give me more possibilities in what I can create with data: 86.96%
  - No, I'm satisfied with the existing repositories: 13.04%

- Q3: Programmer
  - Yes, Enriching the dataset would give me more possibilities in what I can create with data: 72.22%
  - No, I'm satisfied with the existing repositories: 27.78%
CONCLUSIONS

• Just a very small percentage of interviewees declared themselves as being retired (0.93%) or students (1.85%). The great majority of individuals declared themselves as professionals. Among professionals (101 individuals), almost half (48.5% - 49 individuals) have taken part in hackathons 1 or more times. Among the individuals who declared themselves as curators (18 individuals), 6 individuals (33.3%) stated taking part in hackathons 1 or more times. However, a much higher number of individuals who declared themselves as curators (77.8% - 14 individuals) stated having worked with online data repositories for cultural heritage, which may accommodate task performed by hackers during hackathons for cultural heritage (see Survey 3 for overview of tasks performed by hackers). Taking into account the channels used to gather results, it was already expected a higher engagement with datasets by interviewees. However, the results of the study indicate the existence of a possible new type of curators who are themselves hackers or engage with digital collections. Based on the results, it is possible to suggest that other cultural heritage professionals are themselves hackers.

• Comparing three different categories of professionals (namely curators - 14 individuals, researchers - 32 individuals, and programmers - 22 individuals), and filtering the questionnaire only by positive feedbacks (value = YES) on question number 5, provide interesting insights in regard to the characteristics of these groups. For example, there is a clear male dominance of individuals who declared themselves are programmers (95.45% - 21 individuals). However, a more balanced number appears in the group of individuals who declared themselves as curators (57.14% - 8 individuals), which is composed by a slight female majority. Programmers are also the youngest group in contrast with curators. This opposition between programmers and curators is also seen in question number 4. The number of programmers who engage in hackathons is higher than curators.

• All the 3 groups analyzed affirmed that the dataset could influence the concept of an application (curators: 69.23%, programmers: 83.33%, and researchers: 86.96%). This result clearly shows that the properties of the dataset play an important role in design decisions.

• Programmers are the individuals who most found the structure of standards complicated to be understood and manipulated (66.7% - 12 individuals). On the contrary, only researchers (39.13% - 9 individuals) found working with standards complicated. The most likely explanation for that might be due to the professional activities of researchers in the field of cultural heritage and data, which focuses on the conceptualization and implementation of data standards. Programmers, on the other hand, have a more pragmatic approach
to datasets, using them to the development of applications, instead of concerning themselves with the right data representation.

- All groups affirmed that they would like to have a tool for enriching existing data repositories. Surprisingly, programmers belonged to the group of individuals with less individuals who answered positively to this question (72.22% - 13 individuals), although they have indicated they had more problems to work with datasets. Researchers, on the other hand, belonged to the group most positive to the idea (86.96% - 20 individuals).
A.2.1.2 Online Survey B: Datasets and the Curatorial Process

OVERVIEW

- **Background & Objective:** The qualitative research done during the first phase of this empirical research focused on understanding how the interpretation of heritage occurred in cultural institutions in more general terms. The second phase of the empirical study focused especially on questions about interpretation related to hacker communities. In this sense, understanding how the curatorial process relates to the production of digital collections (the hacker’s raw material) was an important topic to be addressed. The level of engagement with datasets of participants should be taken into account, because of the channels utilized for collecting the data. The goal of the survey was to understand how the curatorial process relates to the production of digital collections.

- **Duration:** Conducted from September to December 2017.

- **Number of Interviewees:** 60.

- **Method:** The online survey was conducted through the website SurveyMonkey (https://www.surveymonkey.com/) and utilize the following channels, because of their representativeness in the cultural heritage sector:
  
  - Mailing List of the EuropeanaTech Community: europeana-tech@list.ecompass.nl
  
  - Mailing List of the Open Knowledge Foundation Community: okfn-en@lists.okfn.org
  
  - Mailing List of the UK Education and Research Communities: cultural-heritage-datasets@jiscmail.ac.uk

GENERAL RESULTS AND PRELIMINARY ANALYSIS

Please, see content on the next page.
Question 1:

What is your job title?

<table>
<thead>
<tr>
<th>Answer Choices</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curator</td>
<td>16.67%</td>
</tr>
<tr>
<td>Collection manager</td>
<td>13.33%</td>
</tr>
<tr>
<td>Exhibition designer</td>
<td>1.67%</td>
</tr>
<tr>
<td>Museum director</td>
<td>0.00%</td>
</tr>
<tr>
<td>Researcher</td>
<td>33.33%</td>
</tr>
<tr>
<td>Data specialist</td>
<td>26.67%</td>
</tr>
<tr>
<td>Other (please specify)</td>
<td>33.33%</td>
</tr>
</tbody>
</table>

Answered: 60
Skipped: 0

Showing 6 words and phrases:
- Librarian
- Project Manager
- Archivist
- Director
- Executive
- Metadata
Does your institution use the datasets of its digital collections in different contexts than collection management?

<table>
<thead>
<tr>
<th>Answer Choices</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>70.00%</td>
</tr>
<tr>
<td>No</td>
<td>30.00%</td>
</tr>
<tr>
<td>Answered</td>
<td></td>
</tr>
<tr>
<td>Skipped</td>
<td></td>
</tr>
</tbody>
</table>

Does your institution use its datasets in different contexts than collection management?
Question 3:

If you answered “Yes” to the question above, could you please explain how datasets are used in your institution?

<table>
<thead>
<tr>
<th>Respondents</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Foremost the datasets are made available on different platforms for others to use. We have all our content on Europeana, alla our high res images on Wikimedia Commons, all datasets for download on GitHub. We have our 3D models on Sketchfab and our digitized texts on Internet Archive. We ourselves use our datasets sometimes in digital and physical exhibitions but we could be better at this.</td>
</tr>
<tr>
<td>2</td>
<td>Predominantly for scientific research. This research was in the past mostly taxonomic / floristic in nature, but more and more it is being used for biogeographical and ecological research. Hot topics are climate change and invasive alien species. Occasionally the collections are used to study ethnobotanical or historical studies. Data linkage to related resources is foreseen to be an important aspect in the future, but this is very new to us at this point.</td>
</tr>
<tr>
<td>3</td>
<td>4 datasets for research purposes (Digital Humanities) Datasets are made available under cc 0 and content/images under cc by-sa under wikidata, wikimedia commons, OpenStreetMap</td>
</tr>
<tr>
<td>4</td>
<td>NA Datasets, dealing with collection and studies of work of art, are created (studies), displayed on a website and on the verge of being reusable (triplestore)</td>
</tr>
<tr>
<td>5</td>
<td>It really depends what you mean by a dataset. For example we have a ‘collection’ of data points that describes the townlands of the west of Ireland which we use in a mapping application. Also, with partners we are creating our own datasets from primary sources, with a view to creating applications and novel visualisations. However, we have other datasets which we just curate by describing and allowing download as tar file.</td>
</tr>
<tr>
<td>6</td>
<td>9 Agregation project with specific datasets and mir projects Museum object management</td>
</tr>
<tr>
<td>7</td>
<td>Publication of object information</td>
</tr>
<tr>
<td>8</td>
<td>11 We use different datasets for insurance themes: statisticstics, reports, articles, For delivery of our online collections records Delivered to third parties like Europeana and ArtUK Unofficially, delivered through an API to a limited number of end users Used currently on a very small scale for visualisations, insight etc used in exhibitions, social media etc</td>
</tr>
<tr>
<td>9</td>
<td>12 Open data sets released for public reuse</td>
</tr>
<tr>
<td>10</td>
<td>Visualisations</td>
</tr>
<tr>
<td>11</td>
<td>13 Presentations, wrapping different groups of objects etc We have used cultural datasets as the main content to be used for a Social Hackathon aimed at developing digital solutions for the valorisation of our cultural heritage</td>
</tr>
<tr>
<td>12</td>
<td>- datasets are used as vouchers in biodiversity research via GBIF - a freely available database allows to use sound recordings for different purposes (science, education etc.)</td>
</tr>
<tr>
<td>13</td>
<td>16 Objektataoge, Forschung, Depot-Management, Monitoring</td>
</tr>
</tbody>
</table>
In general there is a strong difference between the Landesarchiv Baden-Württemberg as a traditional archive and our users. Hence the question is nearly not answerable (means a for us possible answer is missing). But in some cases we use the datasets we preserve for the appraisal of further information and documents which are offered to the Landesarchiv.

Next to collection management we also use datasets to register monuments and subsidise the upkeep of them, among others use.

20 MuseumPlus
21 Publication, Experimentation projects
22 not
Research, exchange of information for attributions, planning exhibition proposals
23 research, innovation
24 To support research
25 Released as open data sets
26 Web publication and research.
28 To research proposes
the bibliographic records are held in a kind of European Union Catalogue of early printed books. Accompanying authority records are stored separately. In the CERL Thesaurus
29 As research resources, as research material
1.-Documentation process
2.-Vocabulary datasets
31 3.-Networking projects/digital libraries
32 Lots of individual researchers' datasets, a few different museum databases
33 Research, restoration, exhibitions and online exhibitions, cultural portals
34 Archival description, use, research, writing
35 Linking data
36 for create virtual museum and Vr applications
37 exhibition design
38 in our digital library - research environment
39 We use datasets to match institutions with content that they’re likely to want to exhibit
40 Datasets are available online through several APIs
41 They are published online in several cultural heritage portals.
42 They are used as learning material, shared on other platforms (Europeana, Wikipedia, flickr, SoundCloud etc) and undergo “big data” research.
43 Online, in exhibitions, for catalogues, as (a) research content (tool), to compare collections between institutions, for loans between institutions
44 L.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Distribution</th>
<th>Applications</th>
<th>Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>API, Online Platforms (Europeana, Wikidata, Wikimedia Commons, OpenStreetMap, ArtUK, Portals, Social Media)</td>
<td>Virtual Exhibitions, Educational Games, Visualizations, Collections Management, Online Collection, Monitoring, VR.</td>
<td>Digital Humanities, Experimentation, Innovation.</td>
<td></td>
</tr>
<tr>
<td>Events</td>
<td>Exhibitions, Hackathons.</td>
<td>Value</td>
<td>Valorization of Cultural Heritage</td>
</tr>
<tr>
<td>Networking</td>
<td>Loans</td>
<td>Production</td>
<td>Writing, Exhibition Design</td>
</tr>
</tbody>
</table>

For Curators, research is the most shared term.
Question 4:
Question 5:

In your opinion, what is/should be the influence of your institution’s data assets in the curatorial process?

<table>
<thead>
<tr>
<th>Answered</th>
<th>49</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skipped</td>
<td>11</td>
</tr>
</tbody>
</table>

Responses:

- Make content available in the best form. Guide in the process of curating this content for exhibitions online or in the museum.
- The data about our museum objects are more visible, better usable by the experts and general public.
- The data collection or capture process should be complete integrated into the workflow of collection management. This would be beneficial for the management and preservation of the collection, at the same time ensuring that value information does not get lost.
- Getting better knowledge on the collections.
- A research institution produces data but there is still a very uneven process in making these data accessible. Personally (with my research unit and research team) I am an activist of open knowledge (even more than open science - to support and facilitate all sorts of coproductions and a strong synergy with already existing online collaborative projects)
- NA
- Our data and research unit are considered as reference data/studies in our domain. We are also working on web semantic models
- To a large extent they drive the curatorial process - this is likely due to the immaturity of our processes which at this early stage are still reactive.
- Helping managing and disseminating, with less restriction.
- At this moment we are starting to make a data set of sound and audiovisual contents. The curatorial process is not part of our work at this moment.
- They are central
- Very good and specialised.
- Very high. I am starting to work much more closely with curators to identify data issues, gaps, best practice, opportunities; also looking to work with research community especially in digital humanities to use datasets for large scale research
- Much higher
- Exploit the truth
- Crucial
- Quite high
- NA
- Data assets are an important tool for curation
- Support services are needed
- Information is crucial.
- The question goes along with a positivist assumption. It’s the archivist who develops the curatorial process. He has to take into account the specific qualities of the data assets. But it’s his choice between different options, all based in the specific qualities of the objects.
- Thanks to our data base other institutions / curators could see and research in our collection before planning an exhibition
- It should be a means to an end instead of being the end
- Defining the purpose of the curation
- Should support curatorial selection
- Informative, archival, for research purposes, as a tool to gain information and to add new information when new research process is in act
- “--"
not clear what you mean
Considerable influence.
They should be embedded in the process.
Analysis existing collection - identify gaps/ strengths
Opportunities to foster digital scholarship
Inventory control, research, source for publications.
It should be significantly more
Not sure I understand the question
In my institution (an University) each researcher curates his/her own data
assets.
1.- Better evaluation of data
2.- Sharing curatorial information with other colleague
3.- Identification of curatorial process
If databases had authoritative information that was kept current they would
be the wellspring of everything we do.
The influence should be high because you can only manage a collection that
is properly documented, preferably in digital format. But many curators don’t
understand the importance of documentation and therefore can’t appreciate
it.
New fields of research
It can be much larger.
Improving project management
We have created a search engine for curators to find works in private hands
for exhibitions. As a result, unexpected connections are being made
internationally where offline alternatives wouldn’t have had the same
serendipity.
They should be expanded to help with exhibitions, cataloguing, etc. But this
depends on the availability (currently limited) of useful data from other
institutions on the web (via Europeana etc.)
Presenting a new model based on institutions’ data
There should be a more consistent/realistic approach towards data (quality,
provability) within any curatorial process.
The curatorial could benefit from both cataloguing data and media assets in
order to prepare exhibitions for instance.
They should be the foundation of the entire process - and more needs to be
done to make these assets accessible.

Analysis
Distribution (Collection & Content) Improving Visibility, Improving Accessibility (synergy in the field, curatorial
  information, discovery of private artworks)
Curatorship Guidance (exhibition planning and design, definition of purpose of curation, support
of curatorial selection), Research (analysis, identification of gaps and strengths,
digital scholarship), Management.
Support Production Capturing and Creation of Digital Content, (Semantic) Models.

For curators, the influence of dataset should be in the process as a whole, exhibition planning and research.
Question 6:

In your opinion, what is/should be the role of the curator in the production of datasets?

**Answered** 51

**Skipped** 9

<table>
<thead>
<tr>
<th>Respondents</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>If the curator has specific knowledge about the objects, this should be included in the dataset.</td>
</tr>
<tr>
<td>2</td>
<td>Refine metadata.</td>
</tr>
<tr>
<td></td>
<td>The curator should provide as accurate as possible data about the museum objects, he can create virtual exhibitions on many themes, or virtual tours...</td>
</tr>
<tr>
<td>3</td>
<td>Firstly, the curator should optimise the workflow to ensure that data is not lost, at the same time seeing that data capture is not a significant time drain of the curation. This is a dynamic operation that should be regularly evaluated. Secondly, the curator should be aware of external existing datasets which may benefit the collection. These may then be incorporated into the present database or linked to if possible. Thirdly, making sure that universal identifiers are allocated to specimens for correct citation, to ensure linkage with other external data repositories.</td>
</tr>
<tr>
<td></td>
<td>- request for production of new datasets</td>
</tr>
<tr>
<td></td>
<td>- participate to the production process</td>
</tr>
<tr>
<td>4</td>
<td>manage the datasets</td>
</tr>
<tr>
<td></td>
<td>At least he/she should be aware of the implications of licenses, metadata, coproductions.</td>
</tr>
<tr>
<td>5</td>
<td>select the relevant information (both in the collection and in the model);</td>
</tr>
<tr>
<td>6</td>
<td>ensure the accuracy and quality of data; consider the use of data</td>
</tr>
<tr>
<td>7</td>
<td>That depends on institutional goals and strategy</td>
</tr>
<tr>
<td>8</td>
<td>To guarantee Quality and specificity</td>
</tr>
<tr>
<td>9</td>
<td>Selection and designing new services and products of information.</td>
</tr>
<tr>
<td>10</td>
<td>Produce, control and keep up to date</td>
</tr>
<tr>
<td>11</td>
<td>Analyser and with complete vision for the selection</td>
</tr>
<tr>
<td></td>
<td>Quality control, including using expertise to guide vocabularies, digital curation, and also interpretation of end results</td>
</tr>
<tr>
<td>12</td>
<td>Curator role focus on reviewing and understanding collection via datasets</td>
</tr>
<tr>
<td>13</td>
<td>strengths and weaknesses of collection</td>
</tr>
<tr>
<td></td>
<td>The curator should supervise projects and looking for funds and international network</td>
</tr>
<tr>
<td></td>
<td>co-working with data specialist</td>
</tr>
<tr>
<td></td>
<td>To describe in what science perspectives the data is described, ie a quality work</td>
</tr>
<tr>
<td>17</td>
<td>N/A</td>
</tr>
<tr>
<td>18</td>
<td>Curator should supervise the production of datasets and by himself create datasets.</td>
</tr>
<tr>
<td></td>
<td>An intermediary curatorial role but also an orientation role together with facilitating communication</td>
</tr>
<tr>
<td>19</td>
<td>projekt, forschungs-, auftrags-, arbeitsabhängig</td>
</tr>
<tr>
<td></td>
<td>The curator should look for the integrity (the physical completeness) and the authenticity (the conformity to the original intellectual content) of the objects. He should impede the physical loss (deletion) and the logical loss (objects being detectable) of the objects.</td>
</tr>
<tr>
<td>21</td>
<td>As the curator knows the collection he knows the needs of the database</td>
</tr>
<tr>
<td>22</td>
<td>The curator should take care of quality, meaning and purpose of the datasets</td>
</tr>
<tr>
<td></td>
<td>The curator should be involved in the creation and management of the datasets</td>
</tr>
<tr>
<td>24</td>
<td>Personal experience and knowledge of the collection and the purpose of the curation</td>
</tr>
<tr>
<td></td>
<td>His/her role is and should be unimportant; the curator should produce meaning not datasets</td>
</tr>
<tr>
<td></td>
<td>Supervisor of the research, organiser of the formation of groups and subgroups of datasets in case of specific research or exhibition ideas/proposals, adding information with new data</td>
</tr>
</tbody>
</table>
I don't know, maybe datasets management.

advise on structure based on the FAIR principles

Bigger role.

Not sure about the production as opposed to the ongoing management. But training to ensure that the dataset is well structured and presented will make curation much more straightforward.

Assist in identification of suitable datasets

Some quality control role

Varies widely, but they should be involved, if the datasets are to be of any use to them.

the curator must be aware of the difficulties and process of the production of datasets to understand better how to use them

they should have a role in the cataloguing process, in how the data is made available to researchers, and they should be consulted about integration/linking of various data sets

A very important role as the curator will be the responsible for a dataset after its creation phase.

1. Checking quality of data and use of metadata

2. Take part of building metadata frame

3. Proposing Project about linking data

The only way that the databases will be authoritative is if the curator is a dataset developer etc.

Curators should produce datasets according to documentation standards using controlled vocabulary. They should be responsible both for the scientific aspect and the accuracy of data.

4. Moderator

The most important role

I think that the curator must be able to interpret the data, not to produce.

Scientific only

Curators will need to use datasets to create new links that are unexpected.

Ensure schema is correct from a research point of view

Define research-oriented / outreach data projects on top of dataset.

It should be central as the curator would renew its content.

A curator could be involved in defining the data architecture that actually describes the collection.

Managing expertise of out-of-the-institution experts

A curator could be an advisor formulating requirements from his own perspective.

The curator is sourcing, acquiring, creating, editing and sharing these assets to assist scholarly research.

Analysis

Management


Production

Expert Knowledge (metadata). Create Interpretive Content (virtual exhibitions, tours, themes). Conceptualize Services and Products, Meaning, Interpretation. Obs.: 2 interviewees said that curators should not be involved with dataset production, but provide meaning.

Awareness

Existing Datasets, Applications and Services, Benefits of Datasets in concern to Curatorial Work, Licenses, Strengths and Weaknesses of Collections, Recognize the Importance of Digitization, Difficulties and Process, Utilization of Datasets, Documentation Standards, His or Her Responsibility.

Obs.: For curators, their main task should be supervising and producing datasets.
Question 7:

Do you see some yet unexplored potential to the use data in the field of Cultural Heritage?

<table>
<thead>
<tr>
<th>Answered</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skipped</td>
<td>10</td>
</tr>
</tbody>
</table>

Respondents | Responses |
---|---|
Big datasets and connected datasets (LOD) is still much unexplored in the field. Here we have the potential of giving our visitors (online and in the museum) the bigger picture in several different ways with the help of rich content (images, 3D, films) and good metadata (LOD) together. They can be used more for educational purposes, for presenting heritage in interesting / amusing ways - attracting young public to the museums, or providing various interactive educational games... We have a number of historical collections in our herbaria. In order to make use of this data, the data first needs to be captured. We have just completed a mass digitisation project that captured 1.2 M specimens - this will be invaluable for doing this type of research. As this has just happened, the potential has not yet been explored.
big data, data mining, IA technics are getting more influence in the Cultural Field, but they are still under used
There are countries where there is no complete database of cultural heritage or the existing information is not yet accessible. Making the dataset is a first necessity and it has a direct impact on conservation and valorisation of this heritage. Also data related to cultural heritage combined with data related to risks (and all the necessary geographical data) is very important to plan interventions and to coordinate interventions in case of calamities.
Yes
Absolutely - unexplored potential abounds for the use of data in everything!
Good luck with your investigations
No
Yes, the sound and audiovisual documents are unexplored.
visualisation has to be improved
no, thanks
Where to start?! So many opportunities. From the smallest projects to enhance records, presentation, discoverability, access, to very large projects using data at scale. For a few very basic examples, see my blog at http://www.catchingtherain.com/|
Yes
More data being released
Stronger understanding of linked data
Don’t know
Yes using data to make visits personalized
Yes
Of course, open cultural data are the most appropriate to experiment new models of interactivity / for the development of new cultural meanings and values?
Institutions have a lot of data but do not use them for support services
Datenbank-Kompatibilität, Datenaustausch,
I can only answer this question in a philosophical way. All historical sciences are working with the assumption that they could find some new questions to the historical material. Hence, a basic principle of archival science is that archived materials are open for other questions than during their production phase, see, Theodore Schellenberg, Modern archives, Chicago 2003 (1956), p. 28-32.
Yes
Yes, there is no togetherness in the data. Building datasets that can be used by municipalities and governmen alike would benefit the insight
Could be interesting in the socialization of the museum’s heritage
Medical applications, especially psychology direction
Yes, for example the metadata of artifacts could lead to a better comparative research
curating archival exhibitions on specific datasets (e.g.: how data meaning and procedures developed and how information has been added, changed, re-written, errors that occurred in the process, new attributions, etc.)

28 Yes, I do. Spatial and close range data for Cultural Heritage monitoring
29 yes - re-use of data for innovation, education, etc.
30 No
31 Definitely.
Yes

32 No different to the unexplored potential of our collections - for researchers
My field is more natural science, but I think more could be done to allow and encourage schools’ access to scientific research done in institutions like museums.
34 yes there is still a lot of work to be done on connecting data (either by linking or by integrating datasets, whichever is more practical). For that to be successful we probably need to develop shared vocabularies and / or mapping of terms. Finally we need to look very carefully at how we facilitate end users that want to make use of our data (what sort of interface do we offer, what sort of data can they reuse, how do we explain to them what they may and may not expect from our datasets)

36 No, data should be preserved and curated to enable reuse.
1.-Restoration projects

37 2.-Keeping risk
Absolutely: curators are not in the drivers’ seats & need to be for data to reach its potential.
Data should be checked and reworked for quality. Often quality is poor and standards are neglected.
Yes. E. g. networking in the history. Efficiency. Larger audience. europeana, open data base and artificial intelligence for automate content aggregation.
Sure. Not everybody uses linked data and another tools for own data. There are a lot of work to do.
39 prospect Interpretation potential
Yes. Look at what Vastari is doing, what Google Cultural Institute / Labs are doing, and the advent of Creative AI - there is much more that can be done
Depends what you mean by ‘unexplored potential’. Use of data for comparison is an established practice in ancient art (our collection) since the 18th century. The use of online data for this purpose is, however, still almost non-existent. What is needed are a) proper tools b)useable ontologies for finding comparanda c) enough online data.
42 not yet
43 Yes of course, working on it
45 Yes.
Yes - the more data available, the more the potential. It concerns me that data assets are only a small percentage of a CH institution’s entire collection though. CHs now need to prepare a one hundred year plan to make all their material digitally available.

Analysis
Integration & Distribution
- Creation of Vocabularies, Linkage, Mapping, Governmental Reuse, Open Data, Data Standards, Automated Content Aggregation.
Quality
- Discoverability, Accessibility, Metadata Quality, Employment of Data Standards.
Usage
- Big Data, Data Mining, Creative AI, Data-driven Curatorship, Socialization, Research (object comparison, finding new questions), Education, Visualizations, Presonalized Visits, Open Data, Models for Interactiveness, Services.
Interpretation
- Insights on the Collection’s Big Picture for Audiences, Presenting Heritage in an Amusing Way, Development of New Curatorial Meanings and Values.
Preservation
- Monitoring (spatial data, geographical data, close range data), Preservation Plan (Intervention and Coordination), Risk Keeping, Restoration.
COMPARATIVE RESULTS AND PRELIMINARY ANALYSIS

Question 1:
Does your institution use its datasets in different contexts than collection management?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1: Curator</td>
<td>70.00%</td>
<td>30.00%</td>
<td>10</td>
</tr>
<tr>
<td>Q1: Researcher</td>
<td>65.00%</td>
<td>35.00%</td>
<td>20</td>
</tr>
<tr>
<td>Q1: Data specialist</td>
<td>81.25%</td>
<td>18.75%</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>71.79%</td>
<td>28.21%</td>
<td>39</td>
</tr>
</tbody>
</table>

Answered: 39  Skipped: 0
How would you rate the degree of impact data-driven projects have in the curatorial process of your institution?

<table>
<thead>
<tr>
<th>Question</th>
<th>No impact</th>
<th>Little impact</th>
<th>Moderate impact</th>
<th>Considerable impact</th>
<th>Total</th>
<th>Weighted Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1: Curator</td>
<td>10.00%</td>
<td>10.00%</td>
<td>50.00%</td>
<td>30.00%</td>
<td>3</td>
<td>25.64%</td>
</tr>
<tr>
<td>Q1: Researcher</td>
<td>30.00%</td>
<td>20.00%</td>
<td>20.00%</td>
<td>20.00%</td>
<td>4</td>
<td>51.29%</td>
</tr>
<tr>
<td>Q1: Data specialist</td>
<td>6.25%</td>
<td>31.25%</td>
<td>31.25%</td>
<td>31.25%</td>
<td>5</td>
<td>41.03%</td>
</tr>
<tr>
<td>Total</td>
<td>17.95%</td>
<td>23.08%</td>
<td>33.33%</td>
<td>25.64%</td>
<td>10</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

Answered: 39
Skipped: 0
SUMMARY OF RELEVANT FINDINGS

- Most of participants (70%) replied their institution utilized digital collections in more in other contexts than collection management. The uses of digital collections were divided as followed:
  - Distribution: Online Platforms (Europeana, Wikidata, Wikimedia Commons, OpenStreeMap, ArtUK, Portals, Social Media), APIs.
  - Applications: Virtual Exhibitions, Educational Games, Visualizations, Online Collection, Monitoring (preservation), Virtual Reality.
  - Research: Digital Humanities, Experimentation, Innovation.
  - Events: Exhibitions, Hackathons.
  - Value: Valorization of Cultural Heritage.
  - Networking: Connecting different artifacts, Managing Loans.
  - Production: Writing, Exhibition Design.

- The majority of participants (60%) stated that data-driven projects (such the ones listed above) have moderate to considerable impact in the curatorial process of their institutions. 40% of participants stated that the impact was either none or little. The impacts were described as follows:
  - Distribution (Collection & Content): Improving visibility, Improving Accessibility (synergy in the field, curatorial information, discovery of private artworks).
  - Curatorship Support: Guidance (exhibition planning and design, definition of purpose of curation, support of curatorial selection), Research (analysis, identification of gaps and strengths, digital scholarship), Management.
  - Production: Capturing and Creation of Digital Content, (semantic) Models.

- Participants were asked about the role of curators in the production of datasets. The responses can be classified as follows:
  - Management: Quality Control Assurance (accuracy, selection, refinement, searchability, authenticity, good representation, updated dataset), Coordination & Optimization of Workflow (request the production of new datasets, supervise the integration with external datasets, prevent data loss, facilitate communication, coordinate collaboration with specialists), Search for Funds, Supervise Research, Establish an International Network, Find Right Purpose of Digital Collections, Propose Projects, Assure Scientific Standards, Define Data Architecture, Formulate Requirements.
– Production: Expert Knowledge (metadata), Create Interpretive Content (virtual exhibitions, tours, themes), Conceptualize Services and Products, Meaning, Interpretation. Obs.: 2 Interviewees said that curators should not be involved with dataset production, but provide meaning.


• Based on the answers of participants, the yet unexplored possibilities in regard of the use of data in the Culture Heritage Sector were classified as such:
  
  – Integration & Distribution: Creation of Vocabularies, Linkage, Mapping, Governmental Reuse, Open Data, Data Standards, Automated Content Aggregation.
  
  – Quality: Discoverability, Accessibility, Metadata Quality, Employment of Data Standards.
  
  – Usage: Big Data, Data Mining, Creative AI, Data-driven Curatorship, Socialization, Research (object comparison, finding new questions), Education, Visualizations, Personalized Visits, Open Data, Models for Interactiveness, Services.
  
  
  – Preservation: Monitoring (spatial data, geographical data, close range data), Preservation Plan (Intervention and Coordination), Risk Keeping, Restoration.

• The comparative analysis showed that curators (80%) belong to the group that recognizes the impact of data-driven project in the curatorial process as moderate or considerable. On the contrary, researchers (60%) belong to the group who stated the impact of data-driven projects in the curatorial process was either none or little. Interestingly, using datasets for research was stated by a few curators as an important aspect of curatorial work. This group sees the value of digital collections as supporting the curatorial work as a whole (research and exhibition planning).

A.2.1.3 Case Study: OuSArchiv

OVERVIEW

• Background & Objective: The literature review helped to understand some of the main problems of CMSs especially in regard to their limitations for supporting the interpretation of heritage. However,
some aspects of the entangled relationship between collection management systems, the institutional work they support, and the users they serve were still unclear. The case study was conducted to provide a closer look into the institutional context in which a CMS is inserted, and how this system supports and determines workflows inside different departments of a cultural institution.

• Date: Conducted from October 2017 to February 2018.
• Methods: Interviews, Software and Documentation Analysis.
• Number of Participants: 6.
• Additional Information: See Item 5.1.3 for a detailed description.

A.2.2 Implementation

A.2.2.1 Improvement of Prototype

• Background: Based on the feedback provided by participants of the workshops, interviews, and usability test, improvements on the prototype were implemented. The data model used was modified and expanded to better fit the requirements of institutions (especially in regard to both design possibilities and compatibility with existing digital collections). The user interface also suffered major modifications based on the results of the usability test (overall increase of the workspace area, draggable and resizable panels, user-defined positioning of nodes, etc). Finally, the API of the prototype was improved so that third-party applications could easily be integrated. This last modification was especially implemented to adapt the concept of the software for the requirements of hackathons.

• Goal: While conceptualization and development of the first version of the prototype targeted the research work done in institutions, the goal of the second version of the prototype was to support the conceptualization and development of applications during hackathons.

• Date: The development of the second version of the prototype took place between January and December 2017.
• Methods: Iterative Design Approach.
• Additional Information: See Chapter 6 for a detailed description.

A.2.3 Validation

A.2.3.1 Case Study: Hackathon Coding da Vinci
• Background: The hackathon "Coding da Vinci" presented an ideal setting to verify the modifications implemented in the second version of the prototype.

• Goal: To test and validate the second version of the prototype in real case scenario.

• Duration, Location: Conducted from October to December 2017. Opening at Hochschule fuer Technik und Wirtschaft (HTW) Berlin. Closing at Jüdisches Museum Berlin.

• Methods: Direct participation and observations, open interviews, software and document analysis.

• Number of Event Participants: 19 representatives from different cultural institutions presented 31 datasets for 120 participants.

• Number of Group Participants: 4 (Leonardo de Araújo, Nina Hentschel, Nicole Mayorga, Adrienn Kovács).

• Jury: Karin Glasemann (Digital Coordinator, Nationalmuseum Sweden), Michael Büchner (Technische Koordination, Deutsche Digitale Bibliothek), Helene Hahn (Projektleiterin und Coding da Vinci Mitbegründerin, Open Knowledge Foundation Deutschland), Thorsten Koch (Mathematiker, Entwickler, Leiter der Servicestelle Digitalisierung Berlin), Kurt Jansson (Stellvertretender Vorsitzender Wikimedia Deutschland e.V., Wikimedia Deutschland).

• Additional Information: See Section 6.3.2 for a detailed description of the case study. See also 2.1.1, 2.1.2, A.2.1.1 (online survey contains also the mailing list of Coding da Vinci)
INFORMATION SYSTEMS

The following analysis offers insights into the characteristics of information systems used directly or indirectly by cultural institutions. Please, see content on the next page.
**Taxonomy**

The taxonomy on the left provides an insight on the hierarchical categorization scheme for analyzing information systems in regard to how content is dealt with. The applied taxonomy is described on the table below and based on the following types of information systems:

1. Resource Management
   - a. Cultural Heritage
   - b. Bibliography
2. Node-Link Diagram Tools
   - a. Mind Map
   - b. Network
3. Code-Based Tools
4. Data Visualization Tools
5. Model Construction and Implementation Tools
6. Assistants
<table>
<thead>
<tr>
<th>Tool</th>
<th>Content Analysis</th>
<th>Content Visualization</th>
<th>Content IO</th>
<th>Content Structure</th>
<th>Content Production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify7</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Table/Spreadsheet/Form</td>
<td>Text</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Tree/Hierarchical</td>
<td>- Serialization Format</td>
<td>- Artifact Mgmt.</td>
<td>- Document, Invoice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Planar: Geospatial</td>
<td>- Tagging</td>
<td>- Specimen</td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Temporal</td>
<td>- Metadata Extraction</td>
<td>- Vocabulary</td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- GIS</td>
<td>- Relational Data Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Temporal Sequence</td>
<td>- Result</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Query-Related Operation</td>
</tr>
<tr>
<td>Arctos</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Table/Spreadsheet/Form</td>
<td>Text</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Tree/Hierarchical</td>
<td>- Serialization Format</td>
<td>- Artifact Mgmt.</td>
<td>- Document, Invoice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Planar: Geospatial</td>
<td>- Tagging</td>
<td>- Specimen</td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Temporal</td>
<td>- Metadata Extraction</td>
<td>- Vocabulary</td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- GIS</td>
<td>- Relational Data Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Result</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Query-Related Operation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td>The Museum System</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Table/Spreadsheet/Form</td>
<td>Text</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Tree/Hierarchical</td>
<td>- Serialization Format</td>
<td>- Artifact Mgmt.</td>
<td>- Document, Invoice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Planar: Geospatial</td>
<td>- Tagging</td>
<td>- Specimen</td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Temporal</td>
<td>- Metadata Extraction</td>
<td>- Vocabulary</td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- GIS</td>
<td>- Relational Data Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Result</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Query-Related Operation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td>eLive</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Table/Spreadsheet/Form</td>
<td>Text</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Tree/Hierarchical</td>
<td>- Serialization Format</td>
<td>- Artifact Mgmt.</td>
<td>- Document, Invoice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Planar: Geospatial</td>
<td>- Tagging</td>
<td>- Specimen</td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Temporal</td>
<td>- Metadata Extraction</td>
<td>- Vocabulary</td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- GIS</td>
<td>- Relational Data Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Result</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Query-Related Operation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td>Qi Keepthinking</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Table/Spreadsheet/Form</td>
<td>Text</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Tree/Hierarchical</td>
<td>- Serialization Format</td>
<td>- Artifact Mgmt.</td>
<td>- Document, Invoice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Planar: Geospatial</td>
<td>- Tagging</td>
<td>- Specimen</td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Temporal</td>
<td>- Metadata Extraction</td>
<td>- Vocabulary</td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- GIS</td>
<td>- Relational Data Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Result</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Query-Related Operation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td>EMu - Collection</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Table/Spreadsheet/Form</td>
<td>Text</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Tree/Hierarchical</td>
<td>- Serialization Format</td>
<td>- Artifact Mgmt.</td>
<td>- Document, Invoice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Planar: Geospatial</td>
<td>- Tagging</td>
<td>- Specimen</td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Temporal</td>
<td>- Metadata Extraction</td>
<td>- Vocabulary</td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- GIS</td>
<td>- Relational Data Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Result</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Query-Related Operation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td>Adlib Museum</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Table/Spreadsheet/Form</td>
<td>Text</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Tree/Hierarchical</td>
<td>- Serialization Format</td>
<td>- Artifact Mgmt.</td>
<td>- Document, Invoice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Planar: Geospatial</td>
<td>- Tagging</td>
<td>- Specimen</td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Temporal</td>
<td>- Metadata Extraction</td>
<td>- Vocabulary</td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- GIS</td>
<td>- Relational Data Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Result</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Query-Related Operation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td>Calm</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Table/Spreadsheet/Form</td>
<td>Text</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Tree/Hierarchical</td>
<td>- Serialization Format</td>
<td>- Artifact Mgmt.</td>
<td>- Document, Invoice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Planar: Geospatial</td>
<td>- Tagging</td>
<td>- Specimen</td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Temporal</td>
<td>- Metadata Extraction</td>
<td>- Vocabulary</td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- GIS</td>
<td>- Relational Data Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Result</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Query-Related Operation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Dataset</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Vocabulary: Taxonomy</td>
</tr>
<tr>
<td>Tool</td>
<td>Type</td>
<td>Nonsequential</td>
<td>Qualitative</td>
<td>Dataset</td>
<td>Resource Management</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------------</td>
<td>-------------</td>
<td>---------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Mimsy X6</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Linear</td>
<td>Tree/Hierarchical</td>
<td>Dataset</td>
</tr>
<tr>
<td>Karma</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Tree/Hierarchical</td>
<td>Planar</td>
<td>Geospatial</td>
</tr>
<tr>
<td>Mendeley</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Linear</td>
<td>Tree/Hierarchical</td>
<td>Dataset</td>
</tr>
<tr>
<td>Papers</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Linear</td>
<td>Tree/Hierarchical</td>
<td>Dataset</td>
</tr>
<tr>
<td>Zotero</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Linear</td>
<td>Tree/Hierarchical</td>
<td>Dataset</td>
</tr>
<tr>
<td>Citavi</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Linear</td>
<td>Tree/Hierarchical</td>
<td>Dataset</td>
</tr>
<tr>
<td>Bubbl.us</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Tree/Hierarchical</td>
<td>Dataset</td>
<td>Manual Content CRUD</td>
</tr>
<tr>
<td>Docear</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Tree/Hierarchical</td>
<td>Dataset</td>
<td>Manual Content CRUD</td>
</tr>
<tr>
<td>MindManager</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Tree/Hierarchical</td>
<td>Sequential</td>
<td>Presentation</td>
</tr>
<tr>
<td>MindMeister</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Tree/Hierarchical</td>
<td>Sequential</td>
<td>Presentation</td>
</tr>
<tr>
<td>Xmind</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Tree/Hierarchical</td>
<td>Temporal</td>
<td>Sequential</td>
</tr>
<tr>
<td>Gephi</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Network: Node-Link</td>
<td>Planar: Geospatial, Infographics</td>
<td>Multidimensional</td>
</tr>
<tr>
<td>Tool</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Graph</td>
<td>Process</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>---------------</td>
<td>----------</td>
<td>----------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Graph Commons</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Graph</td>
<td>Process</td>
</tr>
<tr>
<td>GraphViz</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Graph</td>
<td>Text</td>
</tr>
<tr>
<td>Atlas.ti</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Resource Mgmt.</td>
<td>Text</td>
</tr>
<tr>
<td>MAXQDA</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Resource Mgmt.</td>
<td>Text</td>
</tr>
<tr>
<td>NVivo</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Resource Mgmt.</td>
<td>Text</td>
</tr>
<tr>
<td>QDAMiner</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Resource Mgmt.</td>
<td>Text</td>
</tr>
<tr>
<td>Info gram</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Table/Spreadsheet/Form</td>
<td>Media</td>
</tr>
<tr>
<td>RAW</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Table/Spreadsheet/Form</td>
<td>Media</td>
</tr>
<tr>
<td>Tableau</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Table/Spreadsheet/Form</td>
<td>Media</td>
</tr>
<tr>
<td>Tagged</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Table/Spreadsheet/Form</td>
<td>Media</td>
</tr>
<tr>
<td>Protégé</td>
<td>Qualitative</td>
<td>Nonsequential</td>
<td>Dataset</td>
<td>Vocabulary</td>
<td>Model</td>
</tr>
<tr>
<td>Visual Paradigm</td>
<td>IRIS.AI</td>
<td>DBPedia Spotlight</td>
<td>Open Calais</td>
<td>Zemanta Plugin</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>--------</td>
<td>-------------------</td>
<td>-------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td><strong>Qualitative</strong></td>
<td>• Nonsequential - Tree/Hierarchical - Network: Notation Diagram</td>
<td>• Qualitative</td>
<td>• Qualitative - Linear</td>
<td>• Qualitative - Linear</td>
<td></td>
</tr>
<tr>
<td><strong>Programming Language/Script</strong></td>
<td>• Tagging - Metadata Extraction - Entity Extraction</td>
<td>• Tagging - Entity Extraction</td>
<td>• Tagging - Entity Extraction</td>
<td>• Tagging - Entity Extraction</td>
<td></td>
</tr>
<tr>
<td><strong>Modeling Notation</strong></td>
<td>• Graph - Semantic Network: Node-Based (Knowledge Graph)</td>
<td>• Graph - Semantic Network: Node-Based (Knowledge Graph)</td>
<td>• Graph - Semantic Network: Node-Based (Knowledge Graph)</td>
<td>• Graph - Semantic Network: Node-Based (Knowledge Graph)</td>
<td></td>
</tr>
<tr>
<td><strong>Text</strong></td>
<td>• Text - Unstructured: Annotation</td>
<td>• Text - Structured: Markup-Language</td>
<td>• Text - Structured: Markup-Language</td>
<td>• Text - Structured: Markup-Language</td>
<td></td>
</tr>
<tr>
<td><strong>Result</strong></td>
<td>• Result - Query-Related Operation</td>
<td>• Result - NLP-Related Operation</td>
<td>• Result - NLP-Related Operation</td>
<td>• Result - NLP-Related Operation</td>
<td></td>
</tr>
<tr>
<td><strong>Dataset</strong></td>
<td>• Dataset - Vocabulary: Ontology</td>
<td>• Dataset - Vocabulary: Ontology</td>
<td>• Dataset - Vocabulary: Ontology</td>
<td>• Dataset - Vocabulary: Ontology</td>
<td></td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td>• Model - Notation - Program</td>
<td>• Model - Notation - Program</td>
<td>• Model - Notation - Program</td>
<td>• Model - Notation - Program</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX - ARTFACTS DATA MODEL

**E1 artifacts:CollectionNode**
- Subclass of
  - skos:Collection
- Superclass of
  - artifacts:MainCollectionNode
  - artifacts:EntityCollectionNode
- Scope note
  - A collection of collections representing a statement.
- In First Order Logic
  - artifacts:CollectionNode(x) ⊃ artifacts:Collection(x)

**E2 artifacts:MainCollectionNode**
- Subclass of
  - artifacts:CollectionNode
- Scope note
  - A collection of collections representing the entities of a statement.
- In First Order Logic
  - artifacts:MainCollectionNode(x) ⊃ artifacts:CollectionNode(x)
- Properties
  - hasCollectionWithSubject [artifacts:EntityCollectionNode]
  - hasCollectionWithObject [artifacts:EntityCollectionNode]

**E3 artifacts:EntityCollectionNode**
- Subclass of
  - artifacts:CollectionNode
- Scope note
  - A collection of concepts representing the subject or object of a statement.
- In First Order Logic
  - artifacts:EntityCollectionNode(x) ⊃ artifacts:CollectionNode(x)
**Properties**


**E4 artfacts:Node**

- **Subclass of**
  - skos:Concept
- **Superclass of**
  - artfacts:Action
  - artfacts:Artifact
  - artfacts:Concept
  - artfacts:Event
  - artfacts:Institution
  - artfacts:Location
  - artfacts:Person
  - artfacts:Quality
  - artfacts:Quantity
  - artfacts:PropertyNode

- **Scope note**
  - The most generic type of entity in the context of a skos:Concept.

- **In First Order Logic**
  - artfacts:Node(x) ⊃ skos:Concept(x)

**E5 artfacts:Action**

- **Subclass of**
  - artfacts:Node
- **Equivalent to**
  - schema:Action

- **Scope note**
An action performed by a direct agent and indirect participants upon a direct object. Optionally happens at a location with the help of an inanimate instrument. The execution of the action may produce a result.

- In First Order Logic
  - artifacts:Action(x) ⊃ artifacts:Node(x)
  - artifacts:Action(x) ≡ schema:Action(x)

- Properties

E6 artifacts:Artifact

- Subclass of
  - artifacts:Node

- Equivalent to
  - schema:CreativeWork

- Scope note
  - The most generic kind of creative work, including artworks, books, movies, photographs, software programs, etc.

- In First Order Logic
  - artifacts:Artifact(x) ⊃ artifacts:Node(x)
  - artifacts:Artifact(x) ≡ schema:CreativeWork(x)

- Properties

E7 artifacts:Concept

- Subclass of
  - artifacts:Node

- Equivalent to
  - schema:Intangible
• **Scope note**
  - A utility class that serves as the umbrella for a number of 'intangible' things, such as an abstract idea or notion; a unit of thought.

• **In First Order Logic**
  - `artfacts:Concept(x) ⊃ artfacts:Node(x)`
  - `artfacts:Concept(x) ≡ schema:Intangible(x)`

• **Properties**

E8 **artfacts:Event**

• **Subclass of**
  - `artfacts:Node`

• **Equivalent to**
  - `schema:Event`

• **Scope note**
  - An event happening at a certain time and location, such as a concert, lecture, or festival. Repeated events may be structured as separate Event objects.

• **In First Order Logic**
  - `artfacts:Event(x) ⊃ artfacts:Node(x)`
  - `artfacts:Event(x) ≡ schema:Event(x)`

• **Properties**

E9 **artfacts:Institution**

• **Subclass of**
  - `artfacts:Node`

• **Equivalent to**
  - `schema:Organization`
• Scope note
  – An institution such as a museum, library, archive, school, NGO, corporation, club, etc.

• In First Order Logic
  – `artfacts:Institution(x) ⊃ artfacts:Node(x)`
  – `artfacts:Institution(x) ≡ schema:Organization(x)`

• Properties

E10 `artfacts:Location`

• Subclass of
  – `artfacts:Node`

• Equivalent to
  – `schema:Place`

• Scope note
  – Entities that have a somewhat fixed, physical extension.

• In First Order Logic
  – `artfacts:Location(x) ⊃ artfacts:Node(x)`
  – `artfacts:Location(x) ≡ schema:Place(x)`

• Properties

E11 `artfacts:Person`

• Subclass of
  – `artfacts:Node`

• Equivalent to
  – `schema:Person`

• Scope note
  – A person (alive, dead, undead, or fictional).
• In First Order Logic
  - `artfacts:Person(x) ⊃ artfacts:Node(x)`
  - `artfacts:Person(x) ≡ schema:Person(x)`

• Properties

E12 `artfacts:Quality`

• Subclass of
  - `artfacts:Node`

• Equivalent to
  - `schema:QualitativeValue`

• Scope note
  - A value for the characteristic of an entity, e.g. large, medium, small-sized cultural institution.

• In First Order Logic
  - `artfacts:Quality(x) ⊃ artfacts:Node(x)`
  - `artfacts:Quality(x) ≡ schema:QualitativeValue(x)`

• Properties

E13 `artfacts:Quantity`

• Subclass of
  - `artfacts:Node`

• Equivalent to
  - `schema:Quantity`

• Scope note
  - Quantities such as distance, time, mass, weight, etc. Particular instances of say Mass are entities like '3 Kg' or '4 milligrams'.

• In First Order Logic
- artifacts:Quantity(x) ⊃ artifacts:Node(x)
- artifacts:Quantity(x) ⊃ schema:Quantity(x)

• Properties

E14 artifacts:PropertyNode

• Subclass of
  - artifacts:Node

• Superclass of
  - artifacts:RelationshipValue
  - artifacts:NodeClass
  - artifacts:ExtraValue
  - artifacts:Boolean
  - artifacts:Unit
  - artifacts:URI
  - artifacts:GPS
  - artifacts:Date
  - artifacts:Medium
  - artifacts:Hook
  - artifacts:WebAddress

• Scope note
  - Attaches additional property that offers structured values to an entity.

• In First Order Logic
  - artifacts:PropertyValue(x) ⊃ artifacts:Node(x)

E15 artifacts:RelationshipValue

• Subclass of
  - artifacts:PropertyNode

• Equivalent to
  - schema:PropertyValueSpecification

• Scope note
  - Attaches a text value to a property of an entity.
• In First Order Logic
  - `artfacts:RelationshipValue(x) ⊃ artfacts:PropertyNode(x)`
  - `artfacts:RelationshipValue(x) ≡ schema:PropertyValueSpecification(x)`

• Properties

E16 artfacts:NodeClass

• Subclass of
  - artfacts:PropertyNode

• Scope note
  - The classification of an entity.

• In First Order Logic
  - `artfacts:NodeClass(x) ⊃ artfacts:PropertyNode(x)`

• Properties

E17 artfacts:ExtraValue

• Subclass of
  - artfacts:PropertyNode

• Scope note
  - Attaches a text value to a property of an entity.

• In First Order Logic
  - `artfacts:ExtraValue(x) ⊃ artfacts:PropertyNode(x)`

• Properties
E18 artfacts:Boolean

- Subclass of
  - artfacts:PropertyNode

- Scope note
  - Attaches a boolean value to an entity (True or False).

- In First Order Logic
  - artfacts:Boolean(x) ⊃ artfacts:PropertyNode(x)

- Properties

E19 artfacts:Unit

- Subclass of
  - artfacts:PropertyNode

- Equivalent to
  - schema:TypeAndQuantityNode

- Scope note
  - Attaches to an entity a structured value indicating the quantity, unit of measurement, and business function of goods included in a bundle offer.

- In First Order Logic
  - artfacts:Unit(x) ⊃ artfacts:PropertyNode(x)
  - artfacts:Unit(x) ≡ schema:TypeAndQuantityNode(x)

- Properties

E21 artfacts:GPS

- Subclass of
  - artfacts:PropertyNode

- Equivalent to
  - schema:GeoCoordinates
• **Scope note**
  - Attaches the geographic coordinates of a place or event to an entity.

• **In First Order Logic**
  - `artfacts:GPS(x) ⊃ artfacts:PropertyNode(x)`
  - `artfacts:GPS(x) ≡ schema:GeoCoordinates(x)`

• **Properties**

---

E22 **artfacts:Date**

• **Subclass of**
  - `artfacts:PropertyNode`

• **Scope note**
  - Attaches to an entity a date value in ISO 8601 date format.

• **In First Order Logic**
  - `artfacts:Date(x) ⊃ artfacts:PropertyNode(x)`

• **Properties**

---

E23 **artfacts:Medium**

• **Subclass of**
  - `artfacts:PropertyNode`

• **Equivalent to**
  - `schema:MediaObject`

• **Scope note**
  - Attaches to an entity a media object, such as an image, video, or audio object embedded in a web page or a downloadable dataset i.e. DataDownload.

• **In First Order Logic**
- artifacts:Medium(x) ⊃ artifacts:PropertyNode(x)
- artifacts:Medium(x) ≡ schema:MediaObject(x)

**Properties**
- pointsTo [artifacts:Action ∨ artifacts:Artifact
  ∨ artifacts:Concept ∨ artifacts:Event ∨
  artifacts:Institution ∨ artifacts:Location ∨
  artifacts:Person ∨ artifacts:Quality ∨ artifacts:Quantity
  ∨ artifacts:PropertyNode]

### E24 artifacts:Hook

- **Subclass of**
  - artifacts:PropertyNode
- **Equivalent to**
  - schema:ControlAction
- **Scope note**
  - Attaches to an entity an agent for controlling a device or application.
- **In First Order Logic**
  - artifacts:Hook(x) ⊃ artifacts:PropertyNode(x)
  - artifacts:Hook(x) ≡ schema:ControlAction(x)
- **Properties**
  - pointsTo [artifacts:Action ∨ artifacts:Artifact
    ∨ artifacts:Concept ∨ artifacts:Event ∨
    artifacts:Institution ∨ artifacts:Location ∨
    artifacts:Person ∨ artifacts:Quality ∨ artifacts:Quantity
    ∨ artifacts:PropertyNode]

### E25 artifacts:WebAddress

- **Subclass of**
  - artifacts:PropertyNode
- **Scope note**
  - Attaches a Data type:URL to an entity
- **In First Order Logic**
  - artifacts:WebAddress(x) ⊃ artifacts:PropertyNode(x)
- **Properties**
  - pointsTo [artifacts:Action ∨ artifacts:Artifact
    ∨ artifacts:Concept ∨ artifacts:Event ∨
    artifacts:Institution ∨ artifacts:Location ∨
    artifacts:Person ∨ artifacts:Quality ∨ artifacts:Quantity
    ∨ artifacts:PropertyNode]


[61] Leonardo De Araújo, Tossawat Mokdara, Heidi Schelhowe, and Michael Lund. “Beyond browsing and searching: Design and development of a platform for supporting curatorial research and content creation.” In: Chicago, IL, USA, Apr. 2015.


262


[113] Nathan McAlone Insider Business. *These are the 18 most popular YouTube stars in the world — and some are making millions.* URL: http://www.businessinsider.de/most-popular-youtuber-stars-salaries-2017 (visited on 01/04/2018).


[116] *Introduction to SKOS - SKOS Simple Knowledge Organization System.* URL: https://www.w3.org/2004/02/skos/intro (visited on 02/22/2018).


[121] *King’s College London - About Digital Humanities.* URL: https://www.kcl.ac.uk/artshums/depts/ddh/about/index.aspx (visited on 07/30/2018).


[205] Donald A. Schön. Educating the reflective practitioner: Toward a new design for teaching and learning in the professions. Educating the reflective practitioner: Toward a new design for teaching


[224] TEDx Talks. Creating art with data: Aaron Koblin at TEDxAmazonia. url: https://www.youtube.com/watch?v=cAsEeY-0E1U (visited on 07/24/2018).


274


