

Let's talk about standards – A write-up of a discussion on metadata standardization in the Digital Humanities

Marian Clemens Manz, Julien A. Raemy, Béatrice Gauvin, Vera Chiquet – 2022



Credits: Vera Chiquet, Standards... Roma, 2021.

Standardized schemas that can be used to describe, deliver, and annotate all different kinds of digital and physical objects to make data usage possible across projects and institutions has been and is part of a vast number of research projects. Examples include the development of the [Dublin Core](#) metadata schema in the mid-90s, some of the World Wide Web (W3C) standards like [Resource Description Framework \(RDF\)](#) and the [Web Annotation Data Model](#) -, the [International Image Interoperability Framework \(IIIF\)](#) application programming interfaces (APIs), as well as [TEI/XML](#) for scholarly editing purposes. Research projects have shown that there seems to be a need for using and producing metadata that is standardized, especially for reusability and long-term conservation purposes.

However, is this truly the case? Is it always to the advantage of researchers to rely on standards that sometimes cannot cover all issues?

In the context of the first DHCH event (Digital Visual Media and Metadata) at the Istituto Svizzero in Rome (ISR), the participants – professors, Ph.D.students, postdocs, researchers, and curators – were asked to take part in a battle round to discuss the pro and contra aspects of using standardized metadata schemas in the field of the Digital Humanities.

One standard to unite all – Pro Standardization

Standards connect professionals and institutions all over the world by allowing practitioners to have a common language. It enables people from different communities, countries, and backgrounds to normalize and understand input as well as processes. Standardized metadata offers an easier workflow using and working with data, information, and knowledge within and between groups, projects, and institutions. Sharing open standards means sharing elaboration that may lead to better and more resilient ideas than custom recipes made by a small team.

In the case of data interoperability, a standard can break silos and enable broader use of cultural heritage collections across institutional borders while creating the possibility to put them in a new context. One can argue that standards are not necessarily helping with "accessibility" if there are no proper tools to assist them. However, standardization paves the way for scalable tools. By doing so, it helps to draw people's attention to shared challenges among different communities. They may collaborate or compete with each other. Still, in any case, working collectively towards more extensive digital infrastructures (like repositories) needs standardization to deliver scalable and maintainable products or services. Standards also allow smaller institutions and non-profit organizations (NGOs) to follow community practices and easily reuse external data given the standards are open (which is, for instance, the case for IIIF). If no standard is used, only big companies have time and money to adapt to every case. Researchers work on their individual projects within limited timelines as well as specific institutional and technological frameworks. They find themselves under pressure to deliver their results fast, but there is rarely a real practical, executable plan for what happens to their results after the project ends. Simple data archiving does not meet the needs of these cases, since what they produce is valuable in dynamic as well as in static form. What projects need are digital preservation plans and not simply archiving. Thus, it is easier for new or upcoming archives and collections to adopt tested methodologies and defined standards to be able to publish and work with the data right away. This way, researchers can use the often limited financial and time resources on actual research work.

Freedom in individual specificities – Contra Standardization

The domain of computer science brought standardization, and the humanities seem to have to adapt to these standardization concepts to fit the knowledge world affected by computer science. However, the complexity of the arts and humanities cannot limit itself to standard concepts, and the variety of fields might lose their richness; projects could lose focus of their specific questions when adhering to standards. Research in the humanities is meant to complicate our view of the world

and the things within it, and standardization counteracts this, sometimes in a harmful way. Categorization should be part of a research question. How one perceives their field of work is not something settled but rather should be reworked constantly. By using standards and thereby trying to stabilize the world, what room is there for further innovation? The larger the standard gets, the lower the chance is that it is flexible and furthermore covering the essence of a described entity. By setting explicit rules, standardization tackles complex questions of various domains that are not set already. The uncertainties are getting less, and fields are losing their specifications.

Are some babel towers already being built with metadata standards, that are not reflecting on their ruins? What about after standardization? What if these systems change, fall apart? Who finances the upkeep? Standardization often become black holes for financial and human resources. What seems to happen, in a way, is that researchers, scholars and institutions are all trying out solutions, building and paving the way, but are at the same time unable to create the infrastructural support for making these solutions sustainable. Big tech players will have witnessed these struggles from their privileged position and descend upon them with the full stack, the whole package as a service solution. They will, in the end, appropriate and adapt the research of others according to their needs.

To only work with a set standard would be to participate in a political enterprise to control the past and the present. Putting concepts, categories, and even their relative organization into hierarchies is not a settled question. The very definition of such (as a shared framework) is already a domain that belongs to the research itself. There is a danger of perpetuating universalisms that exclude differences, for example, existing metadata standards that are cultural-biased. In other words, a more pluralistic approach is required, perhaps using open standards and at the same time recognizing that openness is also a potential problem.

Big thanks for the great support and basis of this discussion blog to all participants of the #dhch21 thanks to which this discussion has become a lively debate.