



# The HyperRecord Project

## Sharing Information Resources between Museums, Institutes, Archives

A common project of Musei Capitolini, Bibliotheca Hertziana, Theological Library Maribor

### The HyperRecord System

The HyperRecord system is a framework for the interconnectivity of museums: it connects different museum's knowledge domains and provides access to their information resources over the Web, defining common formats and protocols for storage, identification and retrieval over HTTP – for humans using the web pages and machines using the REST interface.

### Museum Information Resources

A museum relies on three different kinds of information resources for its day-to-day work: (1) text documents; (2) data, i.e. Information mostly organized in flat or tabular format; and (3) museum records, which are a mix of text and data. Normally only those information which are of the data type – and therefore easy to integrate in a traditional database – are preserved during digitization. Instead, the HyperRecord system aims to capture and preserve all information resources alike, including structured text material, in order to preserve a museum's entire information context.

### Records ...

Digital museum records must retain (at least) all the qualities and information of the former paper records. Therefore, not only knowledge bits like inventory number and weight, height etc. have to be digitized but also all information in regard to prior assessment, different restorations, varying collocations in different epochs – in short, the entire history of the object through all periods. Hence a new museum record schema has been developed, with sections grouping all information related to distinct periods while allowing infinite subgroups for different parts.

### ... Data ...

Other forms of information sources are inventory tables, acquisition lists, access lists, etc.: In other words, flat data which can date back to centuries ago in some cases. Digitization structures for such information usually have to be invented ad hoc, because they are always very specific. On developing such a structure, it should always be fully self-explanatory.

### ... and Text Documents

Records and data are further accompanied by additional documentation in the form of letters, restoration reports, inventory descriptions, archival documents etc. Such documentation can also be scientific articles and even entire books, – in short all kind of text documents which add context to the museum items or the history of the museum.

### Digitization Formats

The basic digitization format of the various types of information resources is plain XML, which is able to store both document and data contents, and is an optimum solution for exchange across different platforms and for long-term storage. The following formats are currently in use:

#### TEI, DocBook

For the text documents, existing schemata like TEI and DocBook are used, for which a multitude of tools are available – even for export from OpenOffice, for example.

#### CollectioML

For the museum records, a new schema has been designed using RelaxNG as schema language. Its modular structure allows the extension of its scope across various domains without compromising its validity. It covers everything which so far has been dealt with by traditional data input systems: putting this functionality out in a schema written in a standardized language makes systemic modifications much easier. It can also easily be shared or even exposed as a publicly accessible net resource. As the above, it is Free Software under the GPL: museums, private and public collections can use it freely to create their own digital records.

#### XMP

We use XMP for the description of metadata for all media formats. It proves very handy for the rapidly growing collections of still images, allowing us to manage tens of thousands of images with just a few clicks.

### Referencing System

The different information resources can be connected in a way that precludes any manual intervention. Instead of hard-wired links, Uniform Resource Names will mark the individual items across the different document types: the actual link to a resource is no more than the result of a query to the repository. These URNs are made up of a museum identification code and the inventory number.

### Synchronoptic Views

Since all information is collected with regard to its relevant date, synchronoptic views become possible for the entire collection and its history: we are thus able to see the contents, the history and the context of a collection in a specific slice of time. This includes the museum objects' collocations, their restoration status, etc. since the records themselves are organized in distinct chronological periods.

### Federated Network

Just as references are used to link up resources inside a domain, for use inside the museum LAN, they can also be made to link up to external repositories, whenever they are accessible over the web. This allows to build a system where queries are forwarded across the network from the initial repository, which has only a reference, to another one which has actual ownership of the resource.

### RESTful Querying and Access

Different domains can communicate and exchange information by simply serving the XML information resources in the requested form over a plain HTTP connection, sending all necessary information during the call (the REST mechanism). This method provides much more than simple metadata harvesting over the net: in fact, the real data can be queried, accessed and aggregated to larger entities in real time.

### W3C, ISO, OASIS Standards

All technologies used are W3C, ISO or OASIS standards, such as: XML as source, exchange, and archival format all at the same time; Unicode for the encoding; XSL-T/FO for the transformations into other media; RelaxNG as schema language; XQuery as query language and for addressing and aggregating resources which can reside even in outside domains using the REST API over plain HTTP; etc.

### Free and Open

The use of the GNU General Public License for the schema of museum records makes sure that all documents created can be freely exchanged and shared – now and in the future. Only open and non-proprietary standards can guarantee future compatibility between different platforms and systems, provide freedom of use, unlimited exchange and sharing and unmatched availability of tools. The full availability of the system without licensing costs and without fear of being trapped in proprietary technologies is a prerequisite for a collaboration between larger and smaller heritage organizations.

### Contact

For further information, contact dr. Claudio Parisi Presicce at [c.parisi.presicce@comune.roma.it](mailto:c.parisi.presicce@comune.roma.it) or dr. Klaus E. Werner at [kewerner@hyperrecord.net](mailto:kewerner@hyperrecord.net) or just visit our web sites: <http://hyperrecord.net> and <http://museicapitolini.net>