

The Digital Ark: From Taxonomy to Ontology in 17th-century Collections of Curiosities

Brent Nelson

brent.nelson@usask.ca

University of Saskatchewan

When the famous seventeenth-century gardener John Tradescant named his home, with its collection of rarities and curiosities, "the Ark," he was expressing his desire to compile a microcosm of a wide world of variety beyond common experience. Such collections represented the sum of early modern European experience of the world at a time of rapid scientific and geographical expansion and reflected fundamental epistemological shifts in attitudes toward curiosity, wonder, and credulity on the cusp of the modern age. The rapidly expanding world of exploration, colonization, and commerce in the seventeenth century proliferated with strange and bizarre creatures and artifacts that challenged the traditional limits of knowledge. To meet the need for a complete and accessible record of early modern collections, "The Digital "Ark"" will accumulate a database of artifacts and natural specimens as represented by documentary records of early modern collections (inventories, diaries, correspondence, etc.), contemporary drawings and engravings, as well as digital images and curatorial records of extant remnants of these collections. It will be an extensive record of all known collections of rarities and curiosities in England and Scotland from 1580-1700 for which documentary evidence survives, comprising up to 10,000 specimens and artifacts. This information, both textual and visual, will be delivered in an open-access Web-based virtual museum that will collect and display artifacts and natural specimens drawing from a fully searchable database that will record and classify these items and their descriptions in some two dozen fields of information.

This poster will briefly introduce the project and then focus on the challenges this data poses for

a computational process that involves naming data types and defining relationships between them. The principle challenge comes from two unique aspects of the project:

1. The need to accommodate in the user interface a wide range of source genres in a rationalized and consistent form, while representing the distinct epistemological modes of these diverse forms of representation;
2. The need to respect and reflect the way the data was viewed and understood in this age of transition between humanistic and empirical ways of knowing as we interpret the data set and design a database structure to encode, store, and represent this data in all of its complex relationships.

The poster will have four sections:

1. An introduction providing a brief paragraph on the cultural background illustrated with a bulleted set of statistics and a 17th-century engraving of a typical cabinet of curiosities.
2. A chart depicting the diverse data types that provide the content of the digital ark, including: letters; travel accounts; diaries, inventories and catalogues; discursive prose; poetry; contemporary engravings; drawings and paintings; modern photographs of extant objects; and secondary scholarly sources) along with the characteristics of these genres that complicates the process of defining data structures. Page facsimiles will illustrate these data types.
3. A chart depicting the differences between a taxonomic and an ontological view of data. In brief, the taxonomic approach involves entry into a new body of data and the naming and categorizing process that occurs as one interprets and makes sense of this new data, while the ontological approach involves fixing categories and properties in a determined order of being. The seventeenth century represented a significant shift from ontology, where the nature of existence was received and commonly understood by all, to an age of taxonomy, where the new and strange demanded an open-ended reconsideration of the world of existence and a continual configuration of knowledge. In the computer age, we are experiencing a similar tension in the desire to explore and discover

relationships between data, while at the same time thinking of data representation in terms of ontologies. This chart will represent this tension both in the context of the epistemology of the early modern collections and the context of computer processes that might be employed to represent them.

4. The conclusion will outline two steps that will be taken to address these needs:
 - i. The use of qualitative tagging as a means to interrogating the source documents to find what is there, before determining the tag set that will inform the final data structure, that is, to infer a taxonomy rather than simply impose an assumed ontology.
 - ii. The use of a combination of the TEI structure to represent text-based sources with modified object-based ontologies to represent the objects as depicted in these textual sources and also as depicted in graphical sources, both contemporary engravings and drawings, and modern photographs of extant objects.

Acknowledgments

This project is funded by the Social Sciences and Humanities Council of Canada.

References

- Bradley, John** (2005). 'Documents and Data: Modelling Materials for Humanities Research in XML and Relational Databases'. *Literary and Linguistic Computing: Journal of the Association for Literary and Linguistic Computing*. **20.1**: 133-51.
- Corns, Thomas N.** (2000). 'The Early Modern Search Engine: Indices, Title Pages, Marginalia and Contents'. 'The Renaissance Computer: Knowledge Technology in the First Age of Print'. Neil Rhodes, Jonathan Sawday (eds.). London, England: Routledge.
- Daston, Lorraine** (2007). *Objectivity*. Peter Galison (ed.). New York: Zone Books.
- Daston, Lorraine** (1989). 'The Museum: Its Classical Etymology and Renaissance Genealogy'. *Journal of the History of Collections*. **1.1**: 59-78.
- Gilbert, Neal Ward** (1960). *Renaissance Concepts of Method*. New York: Columbia University Press.
- Grindle, Nick** (2005). 'No Other Sign Or Note than the very Order': Francis Willughby, John Ray and the Importance of Collecting Pictures'. *Journal of the History of Collections*. **17**: 15-22.
- Harmon, Margaret** (1975). *Stretching Man's Mind: A History of Data Processing*. New York: Mason/Charter.
- Leith, Philip** (1991). 'Postmedieval Information Processing and Contemporary Computer Science'. *Media, Consciousness, and Culture: Explorations of Walter Ong's Thought*. Bruce E. Gronbeck, Thomas J. Farrell, Paul A. Soukup (eds.). Newbury Park, CA: Sage Publications, pp. 160-176.
- Liu, Yin, Jeff Smith** (2008). 'A Relational Database Model for Text Encoding'. *Rept. Digital Studies/Le champ numérique*. **0.11**.
- MacGregor, Arthur** (c2007). *Curiosity and Enlightenment: Collectors and Collections from the Sixteenth to the Nineteenth Century*. New Haven, CT: Yale University Press.
- Marcus, Leah** (2000). 'The Silence of the Archive and the Noise of Cyberspace'. *The Renaissance Computer: Knowledge Technology in the First Age of Print*. Neil Rhodes, Jonathan Sawday (eds.). London, England: Routledge, pp. 18-28.
- Obrst, Leo, Howard Liu** (c2003). 'Knowledge Representation, Ontological Engineering, and Topic Maps'. *XML Topic Maps: Creating and using Topic Maps for the Web*. Sam Hunting (ed.). Boston: Addison-Wesley, pp. 103-149.
- Sowa, John F.** (2000). *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Pacific Grove: Brooks/Cole.
- Stafford, Barbara Maria** (1996). *Good Looking: Essays on the Virtue of Images*. Cambridge, Ma.: MIT Press.