



Book Review: Advanced Digital Preservation

Edited by David Giaretta, Springer,

2011, 532 pp., € 99.95, ISBN 978-3-642-16808-6

Inge Angevaare

Coordinator, Netherlands Coalition for Digital Preservation,

inge.angevaare@kb.nl

As discussions during LIBER's 40th Annual Conference demonstrated, the research library is evolving from a collection of books to a service centre that caters to a variety of information needs in the university community to which it belongs. One option is for libraries to support researchers in managing and preserving their digital research data. If your institution is contemplating developing such services or is already engaged in such activities, this book is a must-have.

In this book, Giaretta and his co-authors bring together the entire body of knowledge about digital preservation that the partly EU-funded CASPAR project has delivered. And more.

Although the title of the book may not immediately reveal this quality, the book contains much *basic* information about what digital preservation is and how and why the basic functional and information models for digital preservation, the OAIS standard, came into being.¹ Any library manager with digital collections should read these chapters, because if you are contemplating doing digital preservation, or even if you are outsourcing the actual preservation activities, it is essential that you know enough about the essence of digital preservation to be able to organise your own part effectively. In other words, you should have knowledge about the wide variety of threats to digital objects, both of a technical and an organisational nature, and the importance of clearly defining your policies: what do you collect and preserve and for whom (the designated community). Mastering basic digital preservation terminology will also enable librarians to cooperate with others in the chain of preservation (producers of digital objects, archives, consumers).

Despite all this basic information, the book is titled *Advanced Digital Preservation*. That is because Giaretta argues that the well-known preservation strategies of the digital preservation literature (migration and emulation)² are useful for dealing with simple objects such as PDF's, .tiffs or .jpegs, and play a role in establishing authenticity, but they are not up to the task of preserving complex digital objects such as research data (e.g. data bases, software), especially those that cannot be rendered on a screen and those that are dynamic rather than static. And that is important information for (prospective) data librarians.

The CASPAR project developed its own methodology for dealing with these complex objects, revolving around the Representation Information Network — meaning that every bit of information that is required for the designated community to be able to use the *data object* must accompany the data object in the form of metadata and be stored with it in the archive. As technologies change over time, the amount of metadata required will increase, so the method calls for much flexibility (for example: the representation information of a current digital object may refer to XML and the computer will know what XML is, but there may come a time when XML is no longer used).

But this is already getting quite technical. At this point Giaretta himself would insert an unmistakable warning sign in his text, and tell readers without a technical background to skip this particular chapter. Such informal pointers, the fluid writing style, frequent summaries and an abundance of cross references add to the book's readability despite the technical subject matter.

And the book's flaws? I must admit I am not a technical digital preservation engineer, and thus I cannot judge CASPAR's solutions to digital preservation challenges. Perhaps no-one can, at this point in time, as so much is still unknown about digital preservation practice. But I did notice that very little attention is paid in the book to that other European R&D project on digital preservation, the Planets project and the ensuing Open Planets Foundation (OPF) which was established to continue developing the digital preservation tools developed by Planets. No doubt Giaretta would argue that the tools developed by Planets are not suited for research data, thus the book title *Advanced Digital Preservation*. But the reader and research librarian would be served by an accompanying OPF manual to digital preservation, or, even better, by a joint effort. The present APARSEN project, which, like the CASPAR project, is being led by David Giaretta, promises to give the community an

overview of *all* available tools and practices and an assessment of where they work and where they do not. I am looking forward to what APARSEN will deliver in the next three years.

Notes

¹ An introduction to the OAIS model by the UK Digital Curation Centre (DCC) at <http://www.dcc.ac.uk/resources/curation-reference-manual/chapters-production/using-oais-reference-model-curation>.

² Over time, the software/hardware combinations with which a digital object was created, will not be available anymore. The mainstream strategies to deal with this problem are: migration, whereby the digital object is changed so it can be played on newer software/hardware platforms; and emulation, whereby the digital object is not changed, but software is built to emulate the old software/hardware combination on a newer platform. Or, as Giaretta would describe it in his book: the content of digital objects (rows of zeros and ones) is unfamiliar to people, and we need technology to make it familiar, now and in the future.