An approach to software preservation

Brian Matthews[(1)], Arif Shaon[(1)], Juan Bicarregui[(1)], Catherine Jones [(1)], Jim Woodcock [(2)]

[(1)] e-Science Centre, Science and Technology Facilities Council

STFC Rutherford Appleton Laboratory, Chilton, Didcot, OXON OX11 0QX

EMail: brian.matthews@stfc.ac.uk

[(2)] Department of Computer Science, University of York

Heslington, York, YO10 5DD, UK.

ABSTRACT

Preservation of software components is a key aspect of preservation of data, as processing and analysis software frequently needs to be preserved to maintain the usability of data. However, only a small part of the research which has been carried out to date on the preservation of digital objects has looked specifically at the preservation of software. This is because the preservation of software has been seen as a less urgent problem than the preservation of other digital objects, and also the complexity of software artefacts makes the problem of preserving them a daunting one. Further, the preservation of software is frequently seen as a secondary activity and one with limited usefulness.

Software preservation is thus a relatively underexplored topic of research and there is little practical experience in the field of software preservation as such. The results reported in this paper arose from a UK JISC sponsored study into the significant properties of software for preservation[1], and subsequently in a JISC development project into methods and tools for software preservation[2]. Given the relative immaturity of the field, the studies became an exploration of the notion of software preservation as much as identifying the significant properties.

In this paper, we discuss some of the motivations and approaches taken to preserve software. We also discuss some framework concepts of what it means to preserve software, in particular a notion of *adequacy* of preservation, an aspect of the OAIS concept of authenticity which tests the future performance of software against specified preservation properties. In addition, we analyse the relationship of this software preservation framework with the different components of the OAIS information model in terms of their applicability to the retrieval, reconstruction and replay of software on a future technological platform. We then identify within the framework, a number of additional properties of software against which the adequacy of its behaviour, and hence its preservation may be measured in future. We go on to discuss the application of the software preservation framework in the context of a use case involving the British Atmospheric Data Centre. This includes evaluating the overall efficiency of the framework against a number of BADC software, specifically in terms of its relevance (to the software that it is applied to) and sufficiency (of the information recorded) for long-term preservation of software, considered within the context of the BADC's approach to accommodating changes in the technological environment to ensure effective long-term software maintenance and re-use.

---

[1] Joint Information Systems Committee (JISC) study into the *Significant Properties of Software* (2007).

[2] Joint Information Systems Committee (JISC) sponsored project *Tools and Guidelines for Preserving and Accessing Software Research Outputs* (2007-09)**.**