

## Unlocking the secrets of electronic records

*Adrian Brown, UK National Archives*

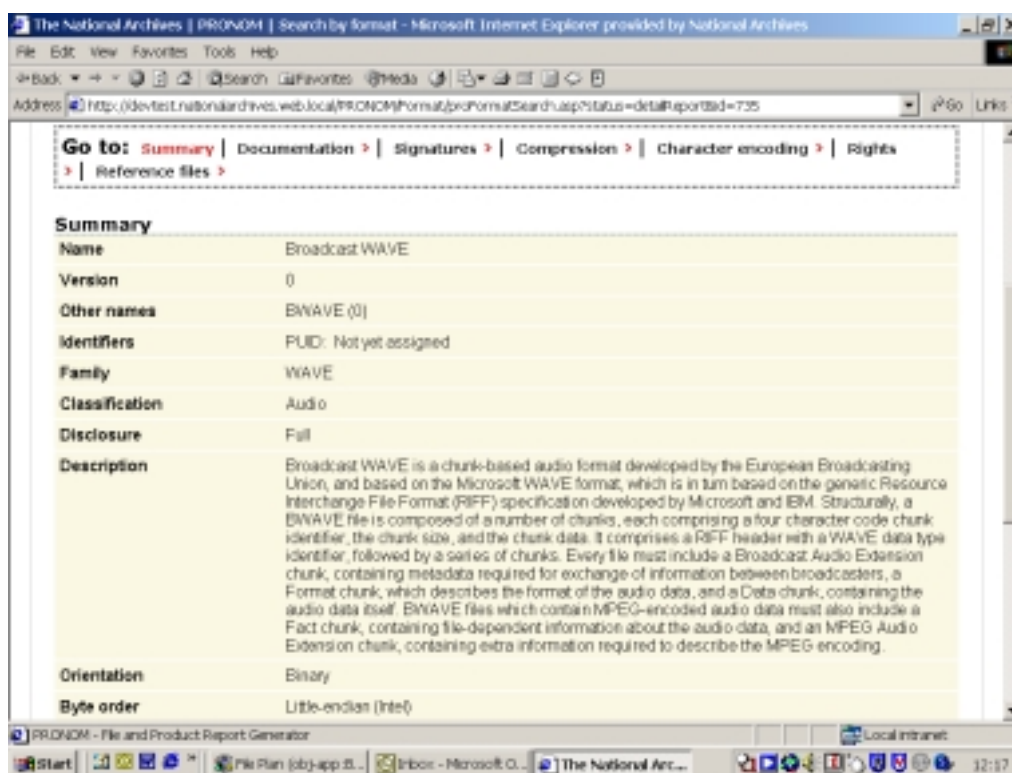
Electronic records pose many challenges for archivists, but these principally arise from a single underlying issue: access to a digital object is entirely dependent on technology. A file in a given format requires software to decode and display it; that software in turn requires a specific combination of hardware, operating systems, and other software to run. Equally, the storage media on which the file is stored requires its own combination of hardware, software and operating system in order to be accessed. Understanding this complex network of technical dependencies therefore lies at the heart of any archiving programme for electronic records.

Furthermore, the technology on which electronic records are so utterly dependent is constantly evolving: existing technologies are redeveloped in new versions, or become obsolete, and entirely new technologies emerge to replace them. This can happen at a very rapid rate, with new versions of software products being released on an annual basis. The challenge for the archivist is therefore not only to understand the nature of these technical dependencies, but also to continually monitor changes in those dependencies which threaten the continued accessibility of electronic records. This monitoring process is known as 'technology watch'.

In order to meet these challenges, the UK National Archives (TNA) has developed a technical registry system called PRONOM. At the heart of PRONOM is a database containing detailed information about the various types of technical component which may be required to access or process an electronic record. These include file formats, software products, operating systems, hardware

components, and storage media. PRONOM has been developed as a resource to support both TNA's own digital archive, and for the wider international community. For this reason, it has been made freely available on the web at ([www.nationalarchives.gov.uk/pronom/](http://www.nationalarchives.gov.uk/pronom/)) where users can search its content, generate reports, and even submit new information for inclusion.

The continuing development of PRONOM is a major project for TNA; the first version was released in 2002, and work is already underway on version 4, which is due to be released in June 2005. PRONOM 4 will build on the existing system to add further levels of detail to the underlying database, and also to develop the first in a series of freestanding tools which will use information contained in the database to provide additional functionality for archivists. The tool currently under development will provide an automatic file format identification service, using binary 'signatures' stored in PRONOM. As signatures are added or updated in the database, the tool will be able to be updated with the latest information. This tool will be used by TNA as part of the accession process for newly-transferred electronic records, enabling every file to be characterised in an automated manner. Such automation is essential when processing large volumes of records – it is simply not feasible to undertake manual processing and cataloguing for millions of files.



The tool will also be made freely-available to download from the TNA website, and will be both platform-independent and easily integrated with other systems, enabling other digital archives to use it within their own systems. The potential for integration will be demonstrated through a JISC<sup>1</sup>-funded project with Southampton University, which is building the tool into a new ingest module for the Eprints digital repository system, used by over 130 archives worldwide.

Looking further into the future, TNA has initiated a major programme called Seamless Flow, which will integrate and automate processes for managing electronic records throughout their lifecycle, from creation, appraisal, selection and transfer from government departments to preservation and dissemination by TNA. The further development of PRONOM will be an integral part of this programme: planned enhancements over the next two to three years include automated technology watch, preservation planning, and migration control services.

In order to maximise the value of PRONOM, TNA welcomes the input and

participation of all interested parties. This does not require major resources: a very practical way to contribute is by submitting new information for inclusion in the database. In particular, TNA is keen to collaborate with any having specialist knowledge of particular types of electronic record, such as scientific data formats and software tools. Information can easily be submitted using an online form, and all sources are fully acknowledged.

Further information about PRONOM, including the latest news on new developments, is available from the PRONOM website at [www.nationalarchives.gov.uk/pronom/](http://www.nationalarchives.gov.uk/pronom/)

*Adrian Brown, UK National Archives*

<sup>1</sup> Joint Information Services Committee (see [www.jisc.ac.uk](http://www.jisc.ac.uk))