BRITISH LIBRARY

The large-scale archival storage of digital objects

The use of digital time stamping in the British Library Digital Object Management Programme

Sean Martin
Sean.Martin@bl.uk

DOM Programme Mission and Vision

Our mission is to enable the United Kingdom to preserve and use its digital output forever

Our vision is to create a management system for digital objects that will:

- store and preserve any type of digital material in perpetuity
- provide access to this material to users with appropriate permissions
- ensure that the material is easy to find
- ensure that users can view the material with contemporary applications
- ensure that users can, where possible, experience material with the original look-and-feel

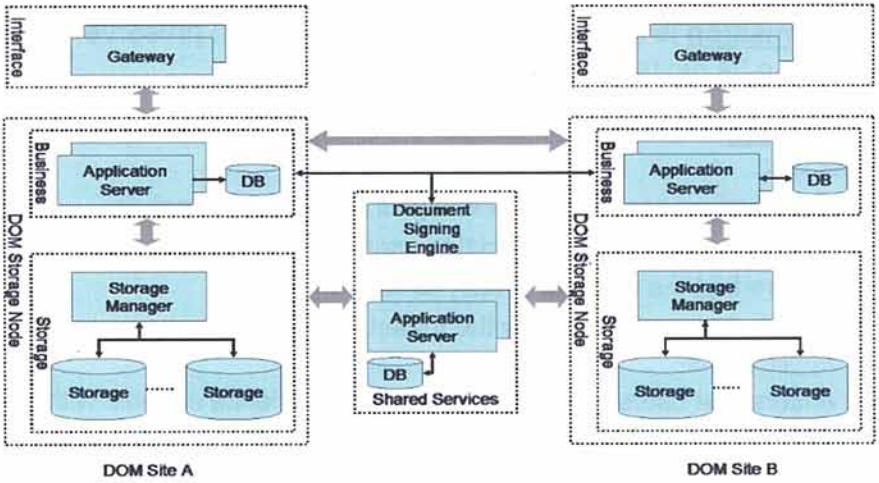
DOM Programme Content Drivers

- Legal deposit legislation for non-print material: royal assent in October 2003 but still needs secondary legislation to bring it into force
- Existing voluntary deposit scheme operational since 2000
- Digitised versions of BL material from early '90s onwards
- New digitisation initiatives: newspapers, sound, etc
- Electronic journals
- Sound Archive's 15TB of material per year (with 50 year collection)
- Web archiving
- Cartography and datasets
- &c &c

Integrity and authenticity of objects within DOM

- Integrity: "*Basis for detecting corruption in the store*"
  - System monitors continuously the object store to detect object corruption
  - Based on using a Secure Hash Algorithm (SHA-1 extended to SHA-512)
  - It would then initiate object recovery
- Authenticity: *"Basis for assurance that a re-presented object is identical to original"*
  - There is a stringent need since a digital object lacks supporting circumstantial evidence for authenticity, such as composition of paper, ink, evidence of change
  - Based on the use of cryptographic digital signing techniques
  - Each object is signed and time-stamped when it is ingested
  - The signature is verified when required
  - The signing mechanism is "tightly" controlled
- Integrity and Authenticity can be determined locally within the architecture

DOM Long-term Storage



The BL stamps content as it is received

Example uses of BL date stamping

Rhone-Poulenc. Source: 'Reports of Patent, Design, and Trade Mark Cases', 1996 no. 4 p125

A European patent in French covered a paste for the preparation of pharmaceuticals. When translated into English, the phrase "35-80%" was mis-copied as "35-50%". Although this clerical error was spotted and corrected, the incorrect English version had been published and therefore the more limited patent was held to be valid. **Rhone-Poulenc appealed. Part of the evidence was the dated BL copy of the correction: a lawyer argued that "on this basis, the date of publication was therefore the date on which the translation was made available to the public in the British Library."** But there were all sorts of arguments about dates and legal details: the company was refused permission to correct the official translation, although the Patent Office was allowed to issue an erratum slip - classic fudge!

Viziball Ltd. Source: 'Reports of Patent, Design, and Trade Mark Cases', 1988 no. 11 p213

A dispute about whether 1 of 2 co-patentees owned more of the patent than the other was taken to appeal. The patent covered squash balls with a reflective lining. An earlier US patent covered the same ground. The appeal judge said, inter alia, **"It is clear from the ... US patent no. 4042236, made available to the public on the shelves of the Science Reference [now British] Library on 6 September 1977, that the use of [this] material on game balls** ... was known before the priority date of the present application." Appeal dismissed.

Three distinct phases for providing assurance

Ingest	Preservation	Access
• *Short term activity*	• Very long term activity	• *Short term activity*
• *Content is transferred from source*	• Content is held within preservation store	• *Content is transferred to destination*
• *Need for a trust model linked to content source*	• No need for a trust model that links two or more parties	• *Need for a trust model linked to content destination*
• *PKI can help provide assurance for receipt of authentic content*	• The BL is a trusted long term repository for multiple content types	• *PKI can help provide assurance for delivery of authentic content*
	• It has no need for assurance provided by short terms organisations such as the current root Certificate Authorities	
	• The Preservation Store will thus use a self signed BL root Certification Authority and an time source derived from the UK NPL	