

Archiving the BBC Website

Why archive the website?

- BBC objectives
 - different from national libraries and institutions
 - role of Information & Archives
- BBC Online
 - benefits for production
 - research/reuse
 - legal and BBC Charter responsibilities
 - historical
 - commercial exploitation
- BBC intranet: Gateway
 - Core Records: business critical information

How to archive the BBC website

- convince New Media production that it **needs** archiving
 - the ephemerality of the internet
- production processes
 - non-standard so difficult to map to an adequate data model for supporting content and archive management systems
 - range of authoring tools
 - fast-changing production methods

How to archive the BBC website

- technology
 - where from and how to capture published content
 - taking regular snapshots would not meet legal requirements, not provide reusable multimedia content nor prove cost effective
 - challenges of capturing constantly 'republished' content and of distinguishing URLs reused over time
 - plug-ins and players required to support stored content
 - 'open technology' to integrate with future production and archive systems

How to archive the BBC website

- storage and access
 - content version control
 - hierarchical storage management
 - tight controls restricting access for editorial and/or legal reasons
- emulation and migration
 - wide and ever-widening range of file formats
 - archiving supporting players
 - maintaining software/hardware environment
 - legally-admissible evidence
 - accurate historical record

How to archive the BBC website

- legal responsibilities
 - strict control over permissions and access to stored material
 - public access to archived content
- I&A to offer expertise to 'media manage' the production process
 - no equivalent of TV and Radio's Broadcast Assistant role or programme reporting procedures
 - constant change in production technologies and techniques
 - transferable information/data management expertise within I&A

How to archive the BBC website

- archiving technologies
 - piloting archive technology: automatic indexing and cataloguing; speech and image recognition
- metadata standards and file-naming conventions
 - possible adoption of Dublin Core within SMEF data model
 - librarians as 'metators'
 - cataloguing for non-mediated retrieval directly by users

How to archive the BBC website

- selection policies and retention schedules
 - diversity and volume of content including material broadcast on other platforms and captured by TV/Radio archives
 - with improved system access aiming to retain 100% of output
 - dictated by technological constraints and longterm costs of storage
 - select according to requested URLs
 - digital object/content management
 - outline joint responsibilities for all departments: archives, production, technology

From BBC Online to BBCi

- BBCi: interactivity on any platform
- New Media Archive Policy
 - Digital Object/Content Management Policy
- New Media Archive System
 - initially for capture and management of BBCi:Online
 - Statement of Requirements/Investment Proposal
 - available software designed for other purposes eg. management of still images in desktop publishing
 - integration with multi-platform authoring tools and end-to-end digital production projects
 - complications with timing, funding and support outside of I&A

Modernising the Archive

- less archiving, more **media (asset) management**
- less archive systems, more **content production and management systems**
- less post-production cataloguing, more **pre-production metadata creation**
- less BBC website, more **digital data management across all New Media platforms**