



Citation for published version:

Pennock, M & Kelly, B 2006, 'Archiving Web Site Resources: A Records Management View' Paper presented at WWW2006, 15th International World Wide Web Conference, Edinburgh, 23/05/06 - 26/05/06, .

Publication date:
2006

[Link to publication](#)

University of Bath

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Archiving Web Site Resources: A Records Management View

Overview

Resources published on the Internet increasingly hold transactional, evidential and record-keeping value, but current Web management and archiving practices are often insufficient to ensure the admissibility and authenticity of the records over time.

This paper proposes the use of records management principles so that Web Site resources with enduring value as records can be properly managed and identified.

Existing applications

Most existing applications utilise either a harvesting approach, or a (legal) deposit approach.

Internet Archive

UK Web Archiving Consortium

Wayback Machine

International Internet Preservation Consortium

Neither approach was developed to meet institutional or organisational requirements for managing and archiving their own Web Site records.

An alternative approach is therefore required for organisational management of Web Site records.

Archiving Web Site Resources: A Records Management View

Bad Practice -

Scenarios

Failure to apply records management principles to Web-based resources and retention can have serious implications:

Incorrect data posted on the Internet and rectified without internal or external acknowledgement can result in financial penalties, cast doubt on organisational accountability, and damage public perceptions.

Resources published on the Web fall under the Data Protection and Freedom of Information Acts; inappropriate data or inability to completely respond to request in good time results in penalties.

Legal Discovery notices are received but later investigation reveals that intranet records were not submitted as their existence was unknown to records managers and penalties are issued.

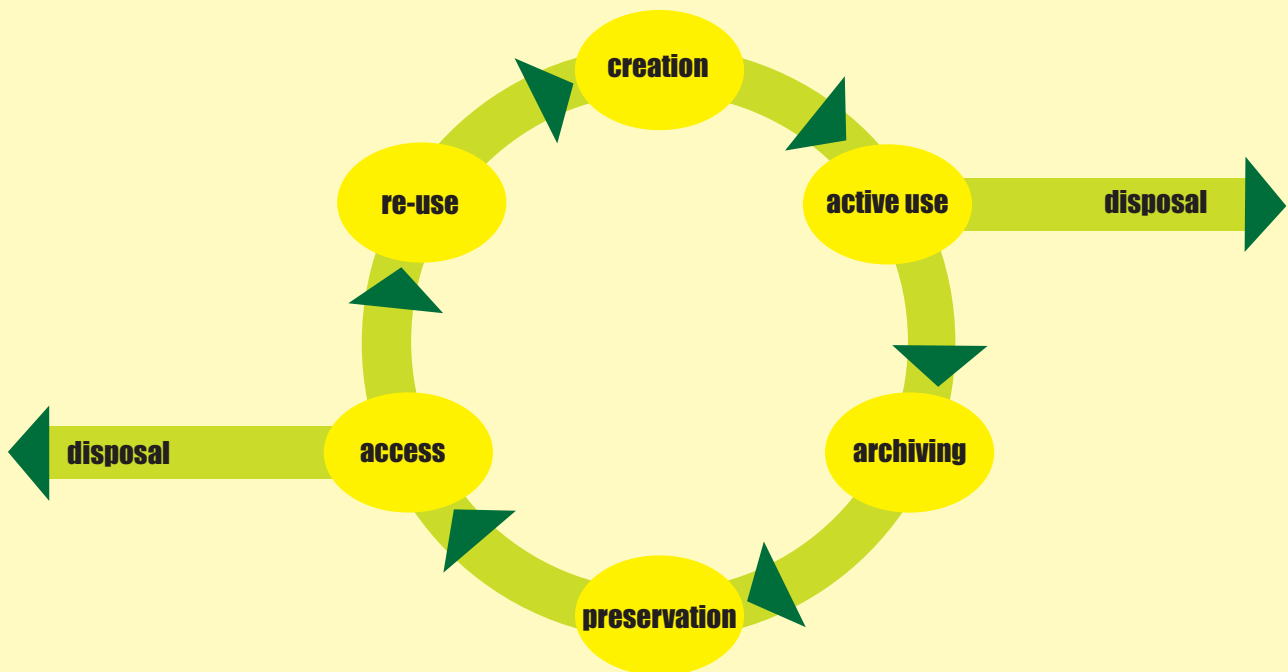
Records Management:

- Limits liability
- Allows accountability
- Contributes towards future archiving and preservation activities
- Ensures that records published on the Web Site and subsequently archived are identifiable, reliable, authentic and re-usable.

Archiving Web Site Resources: A Records Management View

Life-cycle model

The life-cycle model is an important concept for records managers and is the starting point for creating a records management programme. Different stakeholders have different responsibilities at different stages of the model.



Control must be exercised throughout the life cycle of the records if good practices opposite are to be achieved.

Archiving Web Site Resources: A Records Management View

Good Practice

Records management and archival practices should be applied to both Internet and intranet records. Essential good practices include:

- records must be well-formed
- not all data needs to be retained
- valuable records should be explicitly identified and classified according to organisational retention schedules
- valuable records should be archived with appropriate metadata
- non-records should not be archived
- record retention periods must be defined in accordance with organisation and legal requirements
- changes to published records must be recorded in metadata
- web archiving activities must be reinforced by policy
- archived records must be protected from unwarranted alterations
- archived records must be preserved for as long as necessary

Further Information

The paper is published in the WWW2006 Conference Proceedings and is also available online at <http://www.dcc.ac.uk/research/publications/>
For further information please contact Maureen Pennock at M.Pennock@ukoln.ac.uk

