

Better Practice Checklist - 7. Archiving Web Resources

November 2011

Introduction

Australian Government departments and agencies use a range of new technologies for information provision, service delivery and administration.

A key role of the Australian Government Information Management Office (AGIMO) is to identify and promote better practice. This checklist has been created to help agencies to create, capture, manage and retain records of web-based activity for as long as the records are required.

This checklist outlines a number of issues for consideration when determining and implementing appropriate strategies for making and keeping records of online resources. Not all items in the checklist are mandatory, but they have been provided to help agencies to develop a strategy specific to their individual online recordkeeping needs.

This checklist is intended to be a guide to staff responsible for websites, including content authors, website administrators, recordkeeping practitioners and IT staff such as network managers or data administrators. Managers may also find the checklist useful in dealing with contractors, or where these functions are otherwise outsourced. This checklist focuses on non-technical issues.


It should be noted that the checklist is not intended to be comprehensive. Rather, it highlights key issues for agencies. The checklist is iterative and draws on the expertise and experience of practitioners. The subject matter and issues are reviewed and updated to reflect developments.

Acknowledgments

Originally published by the National Office for the Information Economy (NOIE) in 2001 (Version 1), and updated in 2002, (Version 2), this checklist was revised in 2011 with assistance from Australian Government agencies. In particular, AGIMO would like to thank the National Archives of Australia.

Why archive online resources?

Under s. 24 of the Archives Act (1983), Australian Government agencies must retain, destroy or transfer their records to the National Archives of Australia in accordance with records authorities issued by the National Archives. In particular, agencies must comply with the National Archives' policy statement [Archiving Websites: Advice and Policy Statement](#) [ 58 KB].

For practical advice on how to comply with this policy and how to implement checklist items, agencies can use the National Archives' [Archiving Web Resources: Guidelines for Keeping Records of Web-Based Activity in the Commonwealth Government](#) [ 200 KB].

There is no generic solution for creating and maintaining records of web-based activity. The most appropriate records management strategy is likely to use a combination of approaches, depending on the outcome of an analysis of an agency's particular circumstances.

Summary of Checkpoints

- Assign and document records management responsibilities
- Conduct a risk assessment of web activity
- Assess whether you have sufficient records management capability to retain the archived website for as long as you are required to do so
- Ensure that you can create and maintain adequate metadata, including recordkeeping metadata
- Decide on object-driven or event-driven strategies (or a mixture of the two) for ensuring that records of the agency's web-based activity are created and captured
- Consult the National Archives on how often you need to transfer archived websites to their custody
- Ensure that staff are aware of their records management responsibilities to make and keep full and accurate records of web-based activity
- Ensure that the National Library of Australia is informed of government publications that are only available online
- Implement policies and procedures for managing hardware and software dependency and obsolescence
- Consider issues associated with ensuring records of web-based activity remain accessible over time
- Wherever possible, use open standards rather than proprietary standards and platforms
- Conduct regular reviews and monitor developments
- Ensure records of web-based activity that have been identified for transfer to the National Archives of Australia are transferred as soon as practicable

Checkpoints


Assign and document records management responsibilities

A major component of an agency's internal accountability process should be assigning and documenting responsibilities for web-based records management. Three or four different groups of staff may have responsibility for making and keeping records of web-based activity and resources, including content authors, website administrators, records management practitioners and IT staff (such as network managers or data administrators).

For example, responsibility for putting individual records of web activity into a formal recordkeeping system could be assigned to the content author or to recordkeeping staff, rather than to the site administrator.

Developing relevant policies, procedures and standards and ensuring they are followed will be crucial to successfully archiving web-based records.

Conduct a risk assessment of web activity

An agency should be able to quantify the liability or litigation risk regarding the contents of its website, and the need to prove or disprove that particular content was on the site at a given date and time. This applies to both past and present versions of the site. Agencies should provide a disclaimer on their sites. You can find guidance on disclaimers in the [Web Guide website](#) [] or from the Australian Government Solicitor.


Your risk assessment should cover:

- public visibility
- the purposes for which the web resources are used
- the nature of the material on the site
- the audience of the web resources
- how frequently information on the site is changed and updated.

Websites that change frequently but irregularly are likely to pose the most risk to an agency, because such changes are the most difficult to track.


Further guidance on risk management and the conduct of a risk assessment can be obtained from Australian Standard AS 31000:2009, *Risk Management*.


Assess whether you have sufficient records management capability to retain the archived website for as long as you are required to do so

Records of web-based activity need to remain authentic and useable for as long as they are needed as outlined in a records authority. This means that archived web records need to be stored in a secure environment and meet minimum records management requirements. For more information on minimum requirements for systems that keep digital records please consult ISO 16175 - [Principles and Functional Requirements for Records in Electronic Office Environments: use of the standard in the Australian Government](#) [].

☐ Ensure that you can create and maintain adequate metadata, including recordkeeping metadata

Metadata is information attached to web-based resources which allows the records to be found, managed, controlled and understood.

Australian Government agencies' website metadata must comply with the elements outlined in Australian Standard AS 5044-2002, *AGLS Metadata Element Set*. You can find guidance on AGLS 5044 in the [Australian Government Implementation Manual for AGLS Metadata](#) [].

The [Australian Government Recordkeeping Metadata Standard](#) [] specifies the additional metadata you should use to describe the context, management, use, preservation and disposal of information resources, including records of web resources and activity.

You can find further details about applying metadata to website content in [Better Practice Checklist – 6. Use of Metadata for Web Resources](#).

☐ Decide on object-driven or event-driven strategies (or a mixture of the two) for ensuring that records of the agency's web-based activity are created and captured

The type and complexity of web resources being managed will influence the strategy for capture.

- **Object-driven approach.** This approach manages the 'objects' that constitute a website - for example, collections of HTML documents, and/or the various objects assembled to create an HTML document when a user queries the site. This approach could involve taking periodic snapshots of collections of web resources, and tracking changes to the site and logging transaction details.
- **Event-driven approach.** This approach captures 'events' or transactions that occur between the website and the user. This approach may be appropriate for websites that:
 - are dynamic
 - are database-driven
 - rely on stored user profiles, and/or
 - mainly provide transaction services .


This approach involves capturing and storing (with relevant metadata attached):

- the date and time of the event
- the user's IP or domain address of the user
- the user profile
- the query or other action performed
- the resource served to the user.

☐ Consult the National Archives on how often you need to transfer archived websites to their custody

Websites, and website content, need to be retained in accordance with classes in the Australian Functional Disposal Authority (AFDA) and AFDA Express, the legal instruments that set out requirements for agencies to keep, destroy or transfer Australian Government records.


Those records of web-based activity that qualify for transfer to the National Archives of Australia need to be identified and transferred in a timely manner.

For assistance with identifying records of web-based activity that are to be transferred to the National Archives of Australia, please contact the Archives through the [Agency Service Centre](#) [].

Ensure that staff are aware of their records management responsibilities to make and keep full and accurate records of web-based activity

All agency staff are responsible for good records management. You should raise general awareness of these responsibilities, especially the need to create and capture records of web-based activities, and provide the necessary training.

Ensure that the National Library of Australia is informed of government publications that are only available online

The National Library of Australia has established the [PANDORA](#) [] archive to manage online publications and to provide long-term access to them.

Australian Government publications that are available only online are a high priority for this archive, and the NLA should be notified of their existence.

PANDORA archives online publications such as reports, manuals, newsletters, discussion papers and other public accountability documents, as well as any commercial publications. The NLA will negotiate with agencies about access to online commercial publications. The NLA does not usually archive working/exposure drafts, internal agency records, press releases, speeches and public notices. If in doubt, it is best to check with the NLA.

Depositing online content with the NLA does not replace the agency requirement to deposit web records identified as national archives with the National Archives of Australia.

Implement policies and procedures for managing hardware and software dependency and obsolescence

Hardware and software quickly become outdated, and electronic material can become unreadable. Agencies, especially those taking website snapshots for online or offline storage, need to consider a range of hardware and software questions - for example, the estimated physical and/or commercial life of the format in which a website snapshot and related metadata are stored; and the long-term availability of the hardware and operating system needed to read records stored on different media.

Consider issues associated with ensuring records of web-based activity remain accessible over time

These may include:

- managing records by regularly checking and refreshing media
- planning for obsolescence by making sure records can be copied, reformatted or migrated
- using widely supported standards
- implementing measures to protect records against deliberate or accidental alteration
- using persistent identifiers
- ensuring environmental control and monitoring.

Wherever possible, use open standards rather than proprietary standards and platforms

When designing and building websites, you should aim to use software tools and applications that are readily available and fully supported. Factors to consider include maintaining records in widely accepted and technology-neutral storage and data interchange formats (such as XML) and avoiding the use of non-standard HTML language tag extensions.

Conduct regular reviews and monitor developments

Within six months of adopting a strategy for website records management you should review how it is working. The review should look at the effectiveness of the approach, measure performance, identify any changes or corrections required, minimise risk through procedural failure and develop a regular review program.

Because website technology is changing dramatically and rapidly, it is important to monitor web development technologies, emerging trends and business applications, as they will continue to affect strategies and techniques for recordkeeping and archiving websites.

Ensure records of web-based activity that have been identified for transfer to the National Archives of Australia are transferred as soon as practicable

Records of web-based activity which are required to be retained as a national archive should be transferred to the National Archives of Australia as soon as possible after business use has ceased. Storing digital records for the long term is a vast undertaking as records need to remain readable and useable in order to maintain value. By transferring web records that have been identified as national archives under an appropriate disposal authority, agencies may save resources as the responsibility to maintain and appropriately store the records has been removed.

Further information

About the Better practice checklists

Comments about this checklist and suggestions for additional checkpoints can be directed to:

Better Practice
Australian Government Information Management Office
Email: better.practice@finance.gov.au

[Other Better Practice Checklists](#)

Contact for information on this page: [AGIMO Better Practice Team](#)