ELECTRONIC RECORDS MANAGEMENT
AND
ARCHIVES MANAGEMENT POLICY

Guidelines on Electronic Records Management:
Managing Electronic Records
in the Web Environment

NATIONAL ARCHIVE OF MALAYSIA

ARKIB NEGARA MALAYSIA

Projek Pemeliharaan Rekod Elektronik Sektor Awam
(e-SPARK)

<u>Project Documentation</u>

Guidelines on Electronic Records Management:
Managing Electronic Records
in the Web Environment

Managing Electronic Records in the Web Environment

Preface

This Guideline was produced as a result of the e-Spark initiative. Sponsored by the Arkib Negara Malaysia and involving public offices and agencies from across the Government of Malaysia, the purpose of this initiative was to develop policies, standards and practices, technical specifications and training plans to enable the Government of Malaysia to manage records in electronic form. Also included was a strategic plan reflecting the roles and responsibilities of public offices and various central and lead agencies. The Arkib Negara Malaysia, within its legislative mandate to facilitate the management of records in any physical form and to acquire, preserve and make available those of archival value, is the lead public office responsible for facilitating the government-wide management of electronic records. In this capacity and in cooperation with other central agencies and government public offices it is responsible for issuing standards and guidance to government public offices on the management of electronic records.

Managing Electronic Records in the Web Environment is one of a series of guidelines that have been developed to help government public offices and agencies manage electronic records. This guideline should be used in conjunction with the general guideline, *Guidelines on Electronic Records Management* (also available from the Arkib Negara Malaysia). Companion guides are: *Managing Electronic Records in the Structured Environment* and *Managing Electronic Records in the Unstructured Environment.*

These guidelines should also be used in conjunction with *Electronic Records and the Akta Arkib Negara 2003* (available from the Arkib Negara Malaysia). This publication supports the implementation of the Akta Arkib Negara 2003 and the requirement by government departments not to dispose of their records without the approval of the National Archivist and to transfer records assessed as having archival value to the control of the Arkib Negara Malaysia.

For additional information, please contact:

Arkib Negara Malaysia,
Jalan Duta,
50568 Kuala Lumpur
Tel. 603-62010688
Fax. 603-62015679
Web Site: http://arkib.gov.my

Table of Content

Managing Electronic Records in the Web Environment

1. Purpose

The purpose of this section is to provide specific guidance on the management of electronic records generated in the 'web' environment. This section will be of particular interest to web content owners and developers, web masters, program managers, technical specialists, Electronic Service Delivery managers, and registry staff. This guide should be used in conjunction with **Guidelines on Electronic Records Management** (available from the Arkib Negara Malaysia).

2. The Web Environment

The web environment is an environment where work processes are generally associated with the 'publication' and 'communication' of information (though this is changing rapidly with the advent of e-Government initiatives) and the role of the web master is dominant.

This is a rapidly evolving environment. It is the world of 'web content' in which, in the earliest stages of web site evolution, organizations find themselves 'publishing' content onto the web (ergo the issues in this environment tend to be derived from the world of communications, publishing, marketing and library services). But in this era of e-Government, they are also finding themselves managing information that has emerged from defined work processes such as those connected with the development of policy (e.g. the preparation of various drafts of a consultation document placed on the web site or the handling of enquiries placed via the e-mail facility featured on most web sites – similar to 'correspondence management'). Pursuant to the agenda established for many e-government initiatives, many organizations are evolving even further by turning their web sites into gateways or portals in order to support on-line transaction processing (e.g. e-filing of tax returns).

In the early stages of web site evolution, however, the issues are expressed as content management issues (e.g. authenticity, reliability, integrity, security, etc.) and the solutions tend to be derived from the publishing/communications world. In later stages both the issues and the solutions may be more closely aligned with the worlds of records management, data management and applications systems development. Over the longer term, the 'web' environment will reflect the convergence of multiple business processes, multiple disciplines and multiple (increasingly integrated) solutions.

There are two types of web sites within the web environment in terms of target users. The first is the internet-based web site that supports the provision of information and services to the **private organizations and members of the public** while the second is the intranet-based web site that supports the provision of information and services to **government officers.** The guidance in this Section applies to both internet-based and intranet-based web sites.

A *web site* is a collection of electronic resources:

- That is made available in a particular domain of the Internet, for the communication of information and/or the conduct of business transactions.
- That share a common domain name, normally belonging to a single or defined group of organizations and having as their virtual location (or Uniform Resource Identifier) a hierarchical (or other) relationship with the main domain content (often referred to as the 'home page'); and providing a body of interlinked information resources that is navigable using browser technology'.

The following types of information are commonly found in both internet-based and intranet-based web sites:

- **E-mail messages** transmitted to or received from private organization and members of the public via the Internet and messages transmitted to or received from government officers via the internet/intranet.
- **Transaction-based information** generated in support of business process on the internet/intranet. In this case, the users are interacting with an underlying application system with an interface on the internet/intranet.
- **Static and dynamic postings** available on the web sites as a result of a defined programme, service or business activity.
- **Discussions** transmitted through the bulletin boards, discussion forum, list servers or chat rooms set up on the internet/intranet.
- **Web pages** themselves that collectively comprise the web site supported on the internet/intranet.

Typically, website content might include a combination of:

- Static pages
- Compound pages formed by displaying together content from a variety of sources
- Dynamic pages formed according to the user's expressed preferences from a variety of sources
- Active server pages (for the entry of simple information to interact with databases that are an integral part of the site such as a search facility or a telephone directory)
- Web forms capturing information for processing in separate database applications
- Graphics, audio files, and video clips
- Linked documents such as publications

3. Challenges to Managing Electronic Records in the Web Environment

As with other government records, public offices are required to identify, capture, preserve and dispose of the information of record status created or received on the web sites in public office recordkeeping systems to meet their business and accountability needs.

The challenges presented by web sites from an electronic record keeping perspective are as follows:

Immature technology

- The main driver for the technology to date has been mainly one-way communication; underlying management mechanisms for more involved use, such as e-business has only recently been introduced.
- The technology is rapidly evolving, resulting in unstable technological standards.

Content management

- Some websites have very rapidly changing content as full advantage is taken of the 'immediate publication' potential of the web. Quality control of the content is not always properly considered before posting.
- The facilities for publishing to websites directly from standard office software can encourage uncontrolled publishing.
- Version control can be haphazard.
- There is a perception that 'content management' solutions actually take care of records management considerations. In general, they provide document management facilities with little or no records management functionality.

Records capture

- Capturing the content of websites in other formats, electronic or otherwise, will usually result in loss of functionality or present a degraded reproduction of the site, especially over time.
- Compound documents, drawing together in the user's browser a selection of resources from different electronic objects, pose particular challenges from a record keeping perspective.
- Web site content represents the first or one of the earliest types of record that cannot be accurately represented by a hard copy print out. Printing to paper might be considered for the representation of a few single-version static pages with very rudimentary linkages, but may be inadequate for more complex web content.
- Website resources that have the status of records may have been produced and retained in an unmanaged environment, often with little input from records managers.

These challenges can lead to the following outcomes:

- The loss of operationally critical information
- Increased administrative costs associated with re-creating existing information
- Non-compliance with relevant legislation and policy
- Declining levels of client satisfaction
- The untimely loss of corporate memory

4. Electronic Records Management in the Web Environment[1]

With the increasing use of web sites for information dissemination and delivery of programmes and services, public offices should recognize the possible impacts as follows in the context of recordkeeping:

- The possibility that information available on web sites does not have an official copy filed in the recordkeeping system.
- The possibility that the government is exposed to political, legal or financial risks if public offices are not aware of or do not control properly what information has been available on their web sites and when.
- The possibility that apart from the information content, public offices do not capture and maintain the web pages or the entire web site and thus cannot meet their business and accountability requirements for preserving records completeness and contextual information with the records content as the record was made available.

To minimize the above impacts on the government, public offices should take the following steps in managing web records:

- Ensure that all government officers are made aware of the risks and benefits associated with the use of websites and understand and comply with legislative and policy requirements governing the management of records.
- Establish an accountability framework for website management that:
  - Includes website creation policies and practices that specify who has the authority to create websites and under what conditions the legal or public affairs specialists should be consulted.
  - Establishes a coordinating mechanism to ensure that web administrators, content managers, records officers, librarians and publication managers coordinate their efforts.
  - Incorporates website management functions and activities within the overall record keeping and publication management functions.
  - Develops policies, standards and practices for managing websites and defines roles and responsibilities for activities such as the following:

---

[1] Much of the guidance in this section was derived from guidance developed for the Government of Canada by the National Archives of Canada (2002) as well as, *Archiving Web Resources: Guidelines for Keeping Records of Web-based Activity in the Commonwealth Government*, produced by the National Archives of Australia, 2001

- o Posting control. Public offices should ensure that only authorized individuals are permitted to post information on public office websites.
- o Certification. Public offices should ensure that those who maintain accountability for the information content—that is, content managers—have carried out their record-keeping and publication management responsibilities before releasing a document for posting on the web.
    - For static documents, this includes capturing the source document in the corporate record-keeping system, as well as updating the associated metadata with information relevant to the posting process and, where warranted, identifying and depositing published material in the library or with the Arkib Negara Malaysia.
    - For dynamic documents, this entails ensuring that the record-keeping system has captured significant changes, capturing relevant metadata regarding the posting process and in, cooperation with the librarian, identifying and capturing significant changes to published material in the public office library.
    - For business or service transactions, this entails capturing those transactions that the public office has defined as business activities.

- Conduct accountability exposure analysis and determine the level of business and accountable risks
- Establish control procedures to ensure that all material posted on the Web is systematically assessed for timeliness, relevance and quality, both at the time of posting and periodically thereafter.
- Define criteria, procedures and responsibilities for managing different versions of the same document. This includes defining the conditions under which the public office should maintain various versions of a document.
- Ensure that the public office has established appropriate retention periods and that it disposes of web content in accordance with records retention schedules.
- Conduct accountability exposure analysis and determine the level of business and accountability risks.
- Identify categories of information that should be kept as records for business and/or accountability needs.
- Develop and implement policy and procedures to capture an official copy of that information together with relevant contextual information in the recordkeeping system to ensure the integrity of records and facilitate information access.
- Assess business and accountability requirements and coordinate with the Arkib Negara Malaysia to draw up records retention and disposal schedule for the records (see *Electronic Records and the Akta Arkib Negara 2003*, available from the Arkib Negara Malaysia).
- Develop preservation and migration strategies to ensure that the records can be maintained for as long as required.
- Dispose of the web records according to records disposal schedules approved by the Arkib Negara Malaysia (see *Electronic Records and the Akta Arkib Negara 2003*, available from the Arkib Negara Malaysia).

5. Accountability Exposure Analysis

Public offices should analyze the accountability exposure[2] of their web sites so that the acceptable level of business risks and suitable course of recordkeeping actions can be determined (see Appendix 1).

It is the responsibility of the officers who are involved in the administration of the web site to conduct the accountability exposure analysis. As and when required, the officers concerned may need to seek advice from the legal advisors and senior management. The analysis will help public offices determine the level of accountability exposure (i.e. business risks) of the web site and the adequacy of existing recordkeeping policy and procedures.

While accountability needs vary from public office to public office, they should recognize that the key aspects of accountability are the ability to keep evidence of past decisions and to reconstruct what occurred in the past. In the analysis process, public offices should consider the following factors within the context of the current state of recordkeeping system:

5.1 Level of public scrutiny

As a general rule, the higher the level of public scrutiny of a public office, the greater the risk of it being held responsible for the content of its web site. Public offices should recognize that the nature of their responsibilities and the extent of their dealing with clients (i.e. government officers, members of public or private organizations) will determine their level of public scrutiny (and criticism). Public offices should also review their level of public scrutiny before new programs, services or business are introduced to users.

5.2 Purpose of the web site

Public offices may employ the web site for diverse business purposes, such as dissemination of information, communication with clients, promotion and delivery of programs and services, policy consultation, conducting business transactions, and any combination of these. Each of these purposes will pose different level of business risks. Therefore, public offices should assess the risk level of their web site with reference to the purpose(s) of the web site and their level of public scrutiny.

5.3 Complexity of the web site

From a recordkeeping point of view, the level of interactivity (i.e. interactions with underlying databases) and existence of dynamic postings determine the complexity of the web site. The more interactivity and dynamic postings on the web site, the more efforts should be made to ensure that relevant web content is captured and maintained in the recordkeeping system.

---

[2] The concept and the use of "accountability exposure analysis" in the context of management of web were first defined by Charles R. McClure and J. Timonthy Spreh. See *Analysis and Development of Model Quality Guidelines for Electronic Records Management on State and Federal Websites* (January 1998; available at http://www.istweb.syr.edu/~mcclure/nhprc/nhprc_title.html.

5.4 Frequency of content change

Subject to the nature of the information posted on the web, public offices will change the content at regular intervals or on a needed basis to make sure that the information is accurate and up-to-date. For the purposes of organizational accountability and site maintainability, it is important that records relating to web content replacement and updating should be made and kept. Web sites with content that changes irregularly and frequently are of higher risk level because it is difficult to trace and capture the content before and after the changes. In this process, public offices should determine the frequency and content changes and the nature and depth of the content being changed.

The result of the above analysis provides the basis to determine the strengths and weakness of the current recordkeeping system. Suitable recordkeeping strategies and course of actions can be developed according to the identified risk level.

For assistance in conducting the accountability exposure analysis and business risk assessment, public offices should contact the Arkib Negara Malaysia.

6. Identifying and Managing Records in the Web Environment

Before taking any recordkeeping actions, public offices should clearly identify which information on the web constitutes official records. In this process, public offices may find the following questions useful:
- Is the information a unique occurrence (i.e. the content is not held elsewhere)?
- Is the information important to the public office in terms of business and accountability needs?
- Is the web version of the information in its own right important to the public office in terms of business and accountability needs (even though the same content has been kept in the recordkeeping system)

If the answer to either question (b) or (c) is positive, the information on the web should be considered official records and be maintained in the recordkeeping system.

6.1 Management of static and dynamic postings

Web site postings may take the form of a static document (i.e. "static postings" [3]) posted by the public office and may be created/populated dynamically from a database or a series of electronic files behind the web site providing a real-time response to an interactive event, such as a query or request (i.e. "dynamic postings" [4]).

---

[3] Examples of static postings include organization chart, description of programmes and services, media release, published books and articles, consultation papers.
[4] The content of dynamic postings is created "on-the-fly" in response of an interactive event. In most cases, this is done by using active server pages (ASP) or common gateway interface (CGI). Examples of dynamic postings include searching results of a web-enabled catalogues or an underlying database in real time, web site statistics (such as visitors counter) and personalization settings.

Compared with the dynamic postings, the management of the static postings should be straightforward. The steps involved are:

- Identify and file documents about the policy and procedures for the preparation and posting process of the content, including initiation, deliberation, approval and content posting.
- File an official copy of the static posting with relevant contextual information in the recordkeeping system before posting the content on the web site.
- Maintain an historical log of the postings in the recordkeeping system for the web site that is classified as medium or higher level of business and accountability risks. The historical logs may include the following information:
    - o Title or name of the posting;
    - o File reference number of the posting;
    - o Location path of the posting on the web site;
    - o Version number of the postings;
    - o Name, designation and contact details (such as e-mail address and phone number) of the author;
    - o Date and time of the initial posting;
    - o Date and time of the significant and last modifications (if applicable);
    - o Date and time of the replacement or withdrawal; and
    - o Disposal arrangement of the posting after replacement or withdrawal.

- Take snapshots of the web site formed of static information at regular intervals or when the content is changed, especially in the case of web sites of high business and accountability risks.

As the source material of the dynamic postings is held elsewhere and not an integral part of the web page on which the result of an interactive and real-time event is displayed, "stabling" dynamic information and capturing it in a recordkeeping system could be a challenging task.

From a recordkeeping standpoint, public offices should determine whether the event is to produce records before the given interactive and real time event happens on the web site. If records exist or will be created, public offices should develop the proper procedures to capture the information into the recordkeeping system.

For the content that is highly interactive (or dynamic), public offices should capture the "events" (the transactions between the web site and the users) rather than the content that was posted on the web site at the time of visiting or making the query or request. In this case, details of the "events" that should be captured include:

- date and time of the event
- internet protocol (IP) or domain address of the user
- user profile (if any)
- query or other action performed
- details of the source information (file names and location path of the underlying databases or electronic files)

As far as the web site of high business and accountability risks is concerned, public offices should consider taking snapshots of the web site at particular points in time and maintain complete audit logs to ensure that the actual user experience can be reconstructed as far as possible.

Public offices should note that no solution that can reconstruct exactly the content of dynamic postings at all points in time is available at the moment. Therefore, where business and accountability risks are particular high, public offices should exercise caution when making this type of content available on the web site.

6.2 Management of discussions

Before setting up any discussion thread on a specific topic in a bulletin board, discussion forum, list servers or chat rooms in the course of business activities, public offices should assign a reference number (or an identifier) to the topic according to the records classification scheme. Related discussions can then be identified and captured in the recordkeeping system.

In addition to the content, public offices should capture the relevant contextual information concerning the discussions given below in the recordkeeping system as far as practicable:
▪ Topic of the discussion
▪ Name, designation and contact details (including organization, e-mail address, phone number, etc.) of the author
▪ File reference number
▪ Date and time of the posting

6.3 Management of web pages

Capturing an entire web site or particular web pages should take the form of snapshots of the web site or the relevant web pages that possess the information content required for keeping. To help determine the scope of information to be captured (i.e. whether the entire web site or particular web pages) and frequency of taking the snapshot, public offices should determine the level of business and accountability risks they face in maintaining the web sites (see Appendix 1).

Public offices should only capture the snapshots of the web sites that are classified as having a high level of business and accountability risk. The frequency of taking the snapshots also depends on the risk level. Public offices should ensure that the snapshots are transferred to the recordkeeping system as soon as they have been produced. A comprehensive index of web content over time is also recommended for web sites determined to have high-risk level.

Public offices should retain the snapshots of some types of web sites for substantial historical significance. Examples of these web sites include:

- The first web site of a public office
- The web site that documents how a public office conducted its first electronic commerce (i.e. "e-commerce")
- The web site after major modification in terms of its design, structure or mode of information presentation

For advice and assistance in appraising the historical value of the web sites, public offices should contact the Arkib Negara Malaysia.

Electronic forms of the snapshots are always the recommended format to maintain the snapshots as it preserves the hyperlinks, animations, audio and video clips and the "look and feel" of the web site. A number of commercial software packages are designed to capture snapshots of user-defined level of details (including hyperlinks) of web sites. For advice in selecting appropriate software package for this purpose, public offices should contact the Arkib Negara Malaysia.

6.4 Management of Records of Web Site Development and Maintenance

Public offices should recognize that the development and maintenance of a web site is a business process and, as such, they are accountable for the records of the transactions associated with this process. As a web site evolves, policy, procedures and other documentation relating to its design, maintenance and information posting process should be documented and maintained in the recordkeeping system.

To ensure that the above records can be captured into the recordkeeping system and linked properly to other related records, public offices should check if they are covered by the public office records classification scheme. If necessary, public offices may be required to add new subject term(s) to or expand the scope of existing subject term(s) in the records classification scheme. For assistance in maintaining the records classification scheme, public offices should contact the Arkib Negara Malaysia.

Appendix 1

Levels of Accountability Exposure and Recordkeeping Response

I.   Web Site of Low Level of Accountability Exposure

Characteristics	Recordkeeping Response
● Public office has a single web site containing only copies of official publications. ● Public office controls who may post to web site. ● Experience shows that public office publications are not controversial. ● Publications have never been the subject of litigation and there is little or no legal risk. ● Publications generate no unfavourable press reaction.	● Public office keeps adequate records documenting its information systems, accountability framework, standard operating procedures for its web site, etc. ● Public office has well-established procedures for recordkeeping of official publications, independent of its web site, and these procedures capture publications into recordkeeping systems in both paper and electronic formats.
	Records Management Analysis:
	Public office conducts periodic review of web site to determine whether accountability exposure increases.

II. Web Site of Medium Level of Accountability Exposure

Characteristics	Recordkeeping Response
● Public office web sites grow in number and complexity. ● Span-of-control problem grows; many offices are allowed posting to web sites. ● Quality-control problem grows; e.g., several web sites overlap in content but present conflicting information. ● Types of web site uses multiply; both purpose-prepared and interactive real time materials occur on web sites. ● Original materials, not elsewhere captured in recordkeeping systems, sometimes appear on web sites. ● Legal views have not been sought about potential legal liability of web sites. ● Substantial variability in management controls over whether record-quality materials are transferred into recordkeeping systems.	Public office is keeping basic web site records as specified under the Low Level of Accountability Exposure. **Records Management Analysis:** Public office must take some additional measures to ensure ability to reconstruct past web site contents. **Proposed Recordkeeping Response:** ● In addition to above measures, Public office also established an historical log describing contents of web sites. ● Historical log itself is an official record; it is transferred to the public office recordkeeping system and brought under records retention and disposal schedule.

III.   Web Site of High Level of Accountability Exposure

Characteristics	Recordkeeping Response
- Public office web sites grow numerous and highly complex, with sub-sites and sub-sub-sites emerging. - Contents of web sites include ephemeral 'bulletin board' postings, official publications, original materials not captured elsewhere in recordkeeping systems, official hearings and other interactively in real time. - Many different administrative arrangements arise for web sites. - Public office operates under intense public scrutiny; its publications are controversial. - Legal advices suggested that web site carries substantial liability for public office. - Litigation against public office is frequent and acrimonious. - Press coverage and usage of materials covered in web sites is frequent; public interest groups watch public office closely and also often visit web sites. - Legislative interest in public office is active and vigilant. - At the same time, public demands increasing electronic access to public office information holdings and expansion of web sites offerings. This condition makes web sites practically essential to public office mission.	Public office keeps basic web site records as specified under Low Level of Accountability Exposure and also historical log or similar response as specified under Moderate Level. **Records Management Analysis:** Public office must take additional measures to ensure precise reconstruction of exact copy of past web site contents. **Proposed Recordkeeping Response:** - Public office takes periodic 'snapshots' (electronic copies) of entire web sites. Public office should be capable of reproducing entire site contents exactly as they appeared. - Public office maintains comprehensive index of web site contents over time. - Periodicity of snapshots (hourly, daily, weekly, etc.) is frequently reviewed. - Snapshots are official records, transferred to public office recordkeeping system and brought under records retention and disposal schedules.