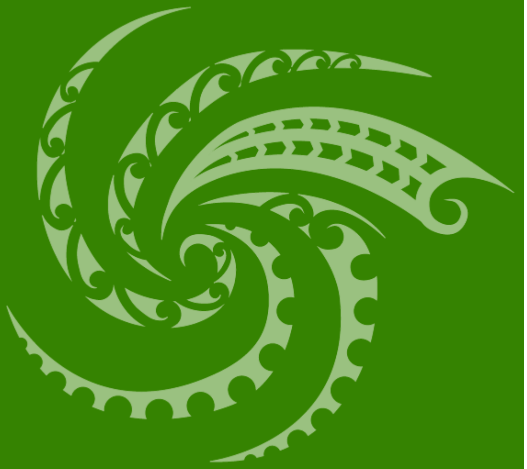

Managing websites as records

Best practice guidance

April 2023

Te Rua Mahara o te Kāwanatanga
ARCHIVES
NEW ZEALAND
Part of the Department of Internal Affairs

New Zealand Government

Document details

Document Identifier: 23/G23

Version	Date	Description	Revision due
0.1	Nov 2022	Development draft	
0.2	Dec 2022	Final draft	
1.0	Apr 2023	Final for approval and publication.	Apr 2026

Contact for enquiries

Government Recordkeeping Directorate
Archives New Zealand
Phone: +64 4 499 5595
Email: rkadvice@dia.govt.nz

Licence

Crown copyright ©. This copyright work is licensed under the Creative Commons Attribution 3.0 New Zealand licence. In essence, you are free to copy, distribute and adapt the work, as long as you attribute the work to Archives New Zealand, Department of Internal Affairs and abide by the other licence terms. To view a copy of this licence, visit http://creativecommons.org/licenses/by/3.0/nz/.

1 Introduction 4
1.1 1.1 Purpose 4
1.2 1.2 Scope 5
1.3 1.3 Key terms 5
2 2 What is a web record? 6
2.1 Is content on the website re-published from another source? 6
2.2 Do web records exist on third party web sites and applications? 6
2.3 Information about web projects and the management of websites are records 7
3 3 Creating and maintaining web records 7
3.1 3.1 Systems for managing web records 7
3.2 3.2 Risks to web records 8
3.3 3.3 Deciding what web records to keep 9
3.4 3.4 Access to web records 9
3.5 3.5 Digital continuity 9
4 4 Information and records management for websites 11
4.1 4.1 Strategy and policy 11
4.2 4.2 Sentencing and disposal 12
4.3 4.3 Using metadata for description and context 13
4.4 4.4 Implementing a migration strategy 13

1 Introduction

Websites are a collection of related web pages, images, videos, or other digital assets, written in mark-up language such as HTML or XML, accessed through an IP address or domain name. New Zealand public offices and local authorities (public sector organisations) use websites to fulfil a range of functions from delivery of basic information to online transactions.

Web content and website activity on public sector organisations' websites are public or local authority records under the Public Records Act 2005 (the Act). Web content is currently considered as a publication and the website as the digital publishing platform. Where websites perform functions, provide services or contain information that is not replicated offline, they are the primary sources of an organisation's business activity. The Act requires public sector organisations to create and maintain full and accurate records of their affairs in accordance with normal prudent business practice. Web content and website activity form part of these web records.

1.1 Purpose

This document is a best practice guide for public sector organisations on managing websites as records in line with their organisation-wide information and records management strategies, policies and systems.

To ensure this guide is applicable to many different processes and systems, its focus is on informing decision-making rather than delivering a procedure. The information given is not a prescription for implementation. It aims to provide direction and present options for organisations to consider when applying information and records management policies to websites. The guide will assist planning for the management of websites, including existing and legacy web records.

The guide is intended for:

- records, knowledge and information management professionals and their support staff
- information technology professionals
- communications professionals
- web managers, developers and other web professionals
- all other staff with responsibility for the creation and management of websites.

1.2 Scope

This guide applies to information and records on a public sector organisation's website as well information and records about the creation, maintenance and functioning of that website.

It is applicable to all websites including third party sites, intranets and secured extranets created and maintained by public sector organisations. It includes information and records in any format, and regardless of whether access is restricted or public.

It may be easier to manage web records where they exist on websites that conform to the New Zealand Government Web Standards.

Discontinued or decommissioned websites and non-current web records need to be managed until they can be destroyed or archived under an approved disposal authority.

For information on web archiving methods and processes, see the factsheet *Web archiving* (23/F31).

1.3 Key terms

The following key terms relate specifically to the web and web records and are used throughout this guide:

- **Content Management System (CMS)**

A system or software package that allows authoring and publishing of web pages or content.

- **Deep Web**

Websites or web pages that cannot be discovered bypublic search engines and are often not accessible by the general public (for example, password protected sites, or sites restricted to users on specified networks). Publicly available web pages that are created dynamically (for example, as the result of a specific search) are also classed as deep web, as these pages exist briefly, on demand, and thus do not exist when a search engine spider indexes the site. A spider is a very simple form of an automated browser. When a spider visits a web page it first reads the text on the page and then tries to follow links from that page off to other web pages and websites.

- **Dynamic Website**

A website that uses databases and/or logical programming to deliver pages, look and feel or content based on variables such as date, user, or randomly selected elements.

- **Extranet**

A private, secured website that shares part of an organisation's information or operations with suppliers, vendors, partners, customers or other businesses.

- **Intranet**

A private website used to distribute information within an organisation.

- **Static Website**

A website where pages are 'served' with the same information to every user. Nothing on the page changes unless the page is edited. The website does not use a database and does not use scripting to perform background logic related to user, page content, layout or structure.

- **Website**

A coherent collection of one or more related web pages that together provide common use or functionality. It includes static web pages, dynamically generated web pages, and web applications.

2 What web records are

A web record may be any information in whole or part that appears on a website and provides evidence of business activity. Websites include 'deep web' content such as information found on intranets, extranets, and secured websites. It also includes records in informal web applications such as wikis, blogs, forums, and shared workspaces.

In some situations, an entire website may be managed as a record, for example, a website devoted to a single project. However, in most instances, a web record will be an element of a website, for example, a page, content item, image, document, submitted information, or log entry recording an action taken. The web record may have multiple pages, items, or log entries, and will have separate metadata about changes made over time.

2.1 Deciding how to manage web records

Organisations must make a clear decision on how to manage entire websites and parts of a website. When making this decision, remember to consider republished content and web records on third-party websites and applications. Organisations should also consider information about web projects and the management of websites.

2.1.1 Website content republished from another source

Some content on websites is likely to have been re-published from another source, with the original managed as a record outside of the website. For example, annual reports and strategic plans may be captured and managed in an organisation's corporate recordkeeping system.

In situations like these, the document on the website is a duplicate which may not need to be kept unless there is good reason, for example, where the additional functionality and context the website gives changes the interpretation of the information. Organisations still need to consider keeping change logs to document when the information was available to the public.

2.1.2 Web records on third-party websites and applications

Web records located on third-party sites need to be managed as public records and be accessible to the organisation and the public, for example, through an Official Information Act 1982 or Local Government Information and Meetings Act 1987 request. The creation and management of these records should be informed by information and records management policies and procedures and included in any contracts or agreements.

Examples of third-party websites include:

- online project management websites such as Basecamp and social networking sites such as LinkedIn, Facebook and Twitter
- sites operated by other organisations such as Google Analytics, Wikipedia, extranets, shared workspaces, survey tools and third-party forums
- file sharing sites such as Google Docs, Issuu and Peer-to-Peer.

2.1.3 Web projects and the management of websites

Information about web projects and the management of websites are records. This may include:

- website commissioning information such as project documentation, email correspondence, strategies, and reports
- website management information such as site statistics, site reports, user testing results, and communications about website operations including hosting and support are also records
- system documentation recording the functionality and ongoing maintenance requirements for the website that ensures its management can be understood over time.

3 Creating and maintaining web records

The creation and maintenance of web sites and web records should form part of organisation-wide information and records management processes and support good business practice.

Minimum requirements for the creation and maintenance of web records are described in the mandatory *Information and records management standard* (16/S1).

Important factors to consider in choosing an approach to creating and maintaining web records are:

- the volume and complexity of the web records
- what the records are about, and the risk associated with the records
- the ability of existing or proposed systems to manage web records
- existing information and records management policies and procedures
- existing appraisal and disposal tools such as our General Disposal Authorities and other disposal authorities.

When planning the implementation of systems to manage websites, procedures and training for all staff should also be developed to inform the creation and maintenance of web records.

3.1 Capturing web records

Capturing or maintaining the context of a web record might require reproduction of the look and feel of the site, page or element. Capturing the look and feel of a website includes recording its graphical user interface or software package, including the colours, shapes, layout and typeface (the look) as well as the behaviour of dynamic elements such as buttons, boxes, and menus (the feel).

3.2 Systems for managing web records

Approaches to choosing a system to manage web records include:

- using a web publishing system such as a content management system (CMS), as a recordkeeping system. A CMS allows for the publishing of web content and can auto-capture metadata about that content giving the information context
- using an electronic information and records management system (EDRMS), to manage content published on the web. This may require other control mechanisms to ensure there is a record of when content is published. It may also require other methods to capture the look and feel of the website
- integrating a web publishing system with an information and records management system so web records management tasks are shared between systems. This will depend on the functionality of both systems.

The majority of websites today are dynamic and data-driven, published and managed by a CMS. While all references in this guide relate to web publishing using a CMS, the techniques described also apply to static websites. This is because creating and publishing web pages is in essence the same as that performed in an automated fashion by a CMS.

3.3 Risks to web records

As web records can change frequently, there is the potential for loss and inaccuracy. Web records may also be short-lived in their published form, and the process of publication may differ from other formats in the scrutiny received before being published. Organisations must be aware of the risks posed by failing to capture changes made to web records.

The level of risk will depend on the type of information being created and published on the web, and the nature and frequency of the changes. Organisations should undertake a risk assessment to inform decisions on how often changes made to a web site must be captured or documented.

Assessing potential risks to web records and finding ways to mitigate these through implementing systems, procedures and training will benefit organisations through improved access to and confidence in the quality of information. Good management of web records can also mitigate business risk.

These risks can include:

- failure to meet legislative and regulatory requirements
- poor decisions made on inaccurate or incomplete information (including by members of the public)
- information not available or accessible for accountability, legal challenge or evidential purposes and loss of information that has cultural heritage value
- static web sites may be undermanaged, not updated technologically, and vulnerable.

3.4 Deciding what web records to keep

Appraisal of web records ensures consistency of high-level decision making about the management and disposal of public records and supports good records management. This will be determined by the way web records are created and maintained, the values inherent in their content and context, and the functions and activities they document.

The appraisal of web records requires consideration of all elements of a website, including:

- pages
- content items
- files
- background processes
- look and feel
- automated or manual transactions
- evidence of business activity in creating and maintaining the website.

The appraisal process generally does not consider the format of the record as this does not usually have a bearing on its value. However, some web records may be defined by format. In this case, the appraisal process should identify where the format of the record provides specific information and records management requirements. For example, information that is part of the organisational record as a static or ‘fixed’ document, but which is also on the website with interactive functions, such as allowing comments for consultation feedback, may need to be kept as a record because of the additional functionality provided by the website.

3.5 Access to web records

All information and records created by public sector organisations, including web records, should meet the access requirements of relevant legislation. Key legislation that informs access to information includes, but is not limited to, the Act, the Official Information Act, the Local Government Official Information and Meetings Act, and the Privacy Act 2020.

When establishing access requirements, public sector organisations need to also consider web records that are not freely accessible on public websites such as information on intranets and extranets. If an organisation chooses to maintain and archive web records in a separate system from other organisational records such as a CMS, they will need to ensure access obligations are met.

3.6 Digital continuity for web records

Digital information must be proactively managed and cared for to ensure that it is accessible for as long as it is needed, for both current and ongoing business purpose and as archives in the future if appropriate.

Ongoing and managed processes need to be in place when maintaining web records over time to ensure digital continuity. This includes managing the risks associated with unauthorised access to records as well as events which may damage or destroy them.

Key principles to consider for ensuring continuity of digital information are:

- Web records are controlled so that they are easily identified and retrieved, without damage or loss. This could be achieved by using CMS rollback functionality.
- Web records are well described so both the technical aspects and content of the records are documented ensuring they remain accessible over time.
- Disaster management programmes are established and maintained to minimise risks. For web records, IT disaster recovery plans and tests are directly applicable, and websites should be tested to ensure a site can be recovered. Recovery includes the successful retrieval of all background data required for preservation and access. The purpose of disaster recovery is to mitigate the risk of loss events, and for data recovery. It is not meant to be an ongoing information and records management strategy or as a replacement for information and records management processes.
- Web records are secure against theft, vandalism, misuse, or inadvertent release. Information and records are secured against unauthorised alteration, deletion or disposal. Web records should have read, write, and delete permissions applied according to the organisation's information and records management and privacy policies.
- Web records are stored in systems which will ensure they are managed for as long as they are required. This means systems are run on supported platforms, hardware is maintained, and tests are carried out to check the information is still accessible.

Ensuring continuity of web records is particularly important when websites are migrated between different CMS products. Planning and testing for continued access to information and records in these situations should be part of the migration.

For information on web archiving methods and processes, see the factsheet *Web archiving* (23/F31).

4 Information and records management for websites

4.1 Strategy and policies

Developing an organisational approach to web records involves managing them as part of an information and records management strategy. An information and records management strategy is a combination of people, policies, procedures, methods, technology, institutional culture, data and knowledge. An organisation develops such a strategy to assess needs, implement practices, manage change and put software and technology support in place.

To support an organisation's objectives and main functions, they should consider web records when developing information and records management policies.

- State the principles of managing web records, and identify the policies used by the organisation to manage those records.
- Assign responsibility for the capture or creation and management of web records and communicate this to all staff.

- Put in place monitoring and review processes to ensure the policy stays current and continues to support business needs.

Consider the creation and maintenance of web records when developing information and records management policies to support the organisation's objectives and main functions.

- State the principles of managing web records, and identify the policies used by the organisation to manage those records.
- Assign responsibility for the capture or creation and management of web records and communicate this to all staff.
- Put in place monitoring and review processes to ensure the policies stay current and continue to support business needs.

Content that is removed from a website should be managed in accordance with the organisation's information and records management policies and processes, including standards and disposal authorities issued by the Chief Archivist under the Act:

- Ensure content to be disposed of is covered by the organisation's disposal authority
- No web records may be disposed of unless a disposal authority is in place
- Duplicate information on websites may be disposed of using *General Disposal Authority 7* (GDA7)
- *General Disposal Authority 6* (GDA6) may also be appropriate to use for disposal of some web records.

Document any specific systems, processes and tools used to create, maintain and dispose of web records in relevant information and records management procedures. All staff responsible for undertaking these activities should receive training on these policies and procedures.

4.2 Sentencing and disposal

It's good business practice to sentence information and records (i.e., determine their final authorised disposal requirements) as close as possible to the point of creation. This way they can be managed according to their value and retained for no longer than is necessary. This is particularly important with digital information such as web records, where adherence to retention periods can avoid extra migration costs or retention of unnecessary data volume.

A CMS may be used to automate part of the sentencing and disposal process by utilising the content scheduling function to delete a web record according to the assigned disposal action. When organisations are considering using a CMS for web records management, thought should be given to the disposal functionality required.

Not all CMS products will have scheduling functionality included. In some systems existing functionality may need to be modified to meet requirements. In an EDRMS, disposal can be linked to the classification structure, and the disposal action automatically applied to a web record when it is captured in the system.

Any disposal of web records must be documented. For further guidance on authorised disposal, see our website.

4.3 Using metadata for description and context

Web records, irrespective of the system in which they are managed, must have metadata. Applying metadata to information ensures that the information has meaning, can be found, can be relied upon to be what it appears to be, and can be moved safely from one system to another.

Much of the metadata associated with web records will be generated by a CMS and should be captured as part of record creation and the ongoing management process. Consider what metadata is needed to adequately describe web records when a new CMS is being implemented or there is a major upgrade. Web records managed in other electronic systems should also have metadata associated with them.

Minimum requirements for metadata are described in the mandatory *Information and records management standard* (16/S1). For further information and guidance on metadata, see our website.

4.4 Implementing a migration strategy for web records

Migration means moving a web record from its current system (for example, an existing CMS) to a new one, or as part of a significant version upgrade of a system (often called a technology refresh). Successful migration strategies enable web records to be maintained over time when they are moved from legacy systems and ensures that any archived web records remain available in the current system.

Migration strategies are most useful when using a CMS as a records management system, as they maintain the required metadata and ensure access. Successful migration involves consideration of the following:

- Migrated web records must include all required metadata and any associated information (such as versions or other transactions) that completes the record.
- Any links or pointers to a web record must remain intact or provide information about how to access the record. For example, create persistent links by updating page URLs so that there is automatic redirection from the old location to the new location.
- Information that may change with a new system such as the look and feel or context of the web record, should be captured as part of the record.

For information on web archiving methods and processes, see the factsheet *Web archiving* (23/F31).