

Information Technology Policy

Management of Web Records

ITP Number ITP-INFRM004	Effective Date September 21, 2007
Category Information	Supersedes None
Contact RA-ITCentral@pa.gov	Scheduled Review June 2022

1. Purpose

This Information Technology Policy (ITP) establishes guidance regarding the retention and disposition of web resources that meet the criteria of a record. This is in accordance with Management Directive 210.5 *The Commonwealth of Pennsylvania State Records Management Program*.

2. Scope

This ITP applies to all offices, departments, boards, commissions and councils under the Governor’s jurisdiction (hereinafter referred to as “agencies”). Agencies not under the Governor’s jurisdiction are strongly encouraged to follow this ITP.

Third-party vendors, licensors, contractors, or suppliers shall meet the policy requirements of this ITP that are applicable to the products and services provided to the Commonwealth.

3. Policy

The Commonwealth’s presence on the Internet is an important method for providing information and services to citizens and the business community. The Internet is a venue for the transaction of business and the delivery of services, both public and private. Record of this activity is necessary for accountability in cyberspace as it is in the more traditional material world.

While planning for a website, analysis of the content of the website needs to take place to determine which elements are records and the associated retention periods that have been established for each series of records. Web records may include content generated externally by customers, suppliers, or partners, who may submit or exchange information through an agency website. Web records may be generated internally and include documents such as policy statements, contracts, or other information that agency customers or partners rely upon to conduct business. Such internal records may include publications that are Web-only versions and are to adhere to Management Directive 205.23 *Submission of Commonwealth Publications to the State Library for Collection and Distribution to Other Designated Libraries*. Whether generated externally or internally, the contents of websites are to be analyzed and scheduled. Web content management systems are being employed to manage aspects of Web content control such as versioning, security, and workflows, and they should also be implemented to manage the retention of records.

When planning a website in order to meet business needs, an agency needs to make sure its website is trustworthy. Trustworthy websites involve numerous qualities. A trustworthy website is one that is reliable and in which the content can be trusted as full, accurate,

complete, and unaltered; it is authentic and has been created by the entity for which it is identified; and it is usable - its content can be easily located and retrieved.

An untrustworthy website may lead to risks. These risks include the inability to document or validate transactions, the inability to reconstruct views of Web content that were created dynamically, and the loss of versions of policies, procedures or other content that were only posted to the Web with no other copies retained.

Other ITPs and policies may be referenced that impact the construction and maintenance of a website and include guidance on accessibility, search engines, file name standards, hardware and software standards, privacy concerns, and Right-to Know policies.

All Web content, including, but not limited to, forms, policy, charts, graphics, and announcements, shall be inventoried and analyzed to define records series, and maintained and disposed of in accordance with approved records retention and disposition schedules.

The life cycle of a Web record shall be determined prior to or during creation and managed according to established policies and procedures, reference ITP-INFRM001 *Life Cycle of Records*.

These policies shall be based upon the business needs of the agency to either retain content or destroy content. Destroying content that is no longer useful reduces costs, confusion, and contradictions. Likewise, maintaining content systematically and with appropriate metadata reduces the cost of locating information when it is needed and allows the use of more automation in managing records and further reduces costs.

No Web applications or content shall be developed without taking into consideration the life cycle, retention, and disposition of the resulting records. If Web content is generated via a Web application, ITP INFRM005 *System Design Review of Electronic Systems* applies.

Retention and Scheduling of Web Records:

Some Web content may already be governed by existing records retention and disposition schedules. For example, Web records may be covered under general records retention and disposition schedules and, in particular, the General Administrative Schedule. Some Web records may be duplicate copies of records maintained in another format or system and, thus, already scheduled on existing agency-specific records retention and disposition schedules.

Some Web content may only be partially governed by existing schedules or may be entirely new records. For example, databases that operate with a Web interface shall be scheduled as part of the process of scheduling the records within the database, mentioning that they are Web interactive. In some cases, the Web interface or data entry forms related to such a database may be new records and may be scheduled separately if they govern the content of the database. In particular, versions of a Web form used to input data into a database shall be retained as long as the data is retained as part of the documentation of how that data was generated.

When determining retention periods for website-related records, as with other records, the agency needs to assess how long the information is needed to satisfy business needs and provide government accountability. If specific Web content is available in places other than the website, consider whether the existence of the information in other records affects the retention needs for the Web records. If the information is unique to the website, the Web version is, by definition, the official record copy and needs to be scheduled.

The options for analyzing the addition of a website to a records retention and disposition schedule depends on associated record management risk factors. An agency may want to breakdown the website; however, this is not recommended because most websites have a variety of content types and each is to be evaluated separately for possible records management risks. A website may be analyzed as a whole, minus one or two portions that may have different characteristics, such as an e-government application. An agency may also need to break down the website by function for evaluation and scheduling purposes. This could result in a large number of independent records series. Whatever specific process is used, the entire website is to ultimately be inventoried with records series defined and scheduled.

As an example, some Web sites may have very dynamic content that may be a record (i.e., Web pages that reflect the status of various systems). If the information is used as the basis for decisions, then a log shall be created and maintained as the record. If the information is similar to a status report, then it shall be kept only as long as it is useful. All sorts of transitory announcements are entered or displayed on websites (i.e., jobs announcements and conferences) and, in most cases, it is important to keep a copy of both the announcement and the date the copy was posted and removed. Such records are to be scheduled. In many cases, the Web announcement is not the official copy, but it may still be important to retain the knowledge of when it was posted and removed from the Web site for a period of time.

The inventory process shall be coordinated with the agency Web content managers and the agency records coordinator, as well as the website manager/Web master for implementation.

Records Management Requirements for Websites:

- All agency websites shall be inventoried to define records series and establish records retention and disposition schedules for all records on the websites.
- When records on a website duplicate records stored in other formats or systems, a clear determination shall be made regarding which copy is the duplicate and which is the official record copy that is subject to the records retention and disposition schedule.
- The duplicate- and information-only Web pages are to be destroyed when they are no longer needed as indicated on the General Administration Records Retention and Disposition Schedule.
- As Web content management systems are implemented, they shall be integrated with records management systems, if possible, or have records management requirements specified in the content management system as records are created or ingested. For example, expiration dates for documents shall correspond to retention periods (unless the document on the website is a duplicate as indicated in the previous paragraph).
- When record series of official record copy containing long-term and archival records exist on a Web site, special arrangements shall be made to preserve them. They may be maintained in a separate system or format or moved to such a system after their active life on the website. If maintained on the website itself, security and access rights as well as backup provisions are to be more rigorous. For Web publications, if a permanent copy is kept on the agency servers, the State Library shall accept a persistent URL (i.e., a permanent link to the document) instead of the document itself according to the policy set forth in Management Directive 205.23 *Submission of Commonwealth Publications to the*

State Library for Distribution to Other Designated Libraries. The link shall be permanently maintained with redirects to its current location from the old address.

- Version control is especially important with policies, contracts, and other documents, available via the Web, which are updated or revised. This is critical for any document in which visitors to the Web site may base business decisions or actions. The date and time that specific records are posted and removed may be needed for business or administrative needs. Revised or rescinded policies are to be retained according to the Records Retention and Disposition Schedule(s). Policies will generally be archived and subject to permanent retention. It is recommended that agencies remove revised or rescinded policies from public-facing Web sites in order to prevent those policies from being mistaken with current policies. Agencies are to make provisions to keep copies of each version of policies along with the dates that the policies were issued, revised, and/or rescinded utilizing a configuration management system, document management system, agency file shares, or hard copy. Likewise, Web-based forms or other documents related to the entry of information into a database shall be maintained as long as the content generated by the forms is scheduled to be retained. For Web publications, all Web-only document versions are to be forwarded to the State Library of Pennsylvania as set forth in Management Directive 205.23 *Submission of Commonwealth Publications to the State Library for Distribution to Other Designated Libraries*.

4. Responsibilities

- 4.1 Agencies** shall comply with the requirements as outlined in this ITP.
- 4.2 Office of Administration, Office of Information Technology** shall comply with the requirements as outlined in this ITP.
- 4.3 Third-party vendors, licensors, contractors, or suppliers** shall comply with the requirements as outlined in this ITP.

5. Related ITPs/Other References

Definitions of associated terms of this policy are published on the Office of Administration's public portal: <http://www.oa.pa.gov/Policies/Pages/Glossary.aspx>

Commonwealth policies, including Executive Orders, Management Directives, and IT Policies are published on the Office of Administration's public portal:

<http://www.oa.pa.gov/Policies/Pages/default.aspx>

- Management Directive 205.23 *Submission of Commonwealth Publications to the State Library for Collection and Distribution*
- Management Directive 205.34 Amended *Commonwealth of Pennsylvania Information Technology Acceptable Use Policy*
- Management Directive 210.5 *The Commonwealth of Pennsylvania State Records Management Program*
- ITP-ACC001 *Information Technology Digital Accessibility Policy*
- ITP-INFRM005 *System Design Review of Electronic Systems*

6. Authority

Executive Order 2016-06 *Enterprise Information Technology Governance*

7. Publication Version Control

It is the [Authorized User](#)'s responsibility to ensure they have the latest version of this publication, which appears on <https://itcentral.pa.gov> for Commonwealth personnel and on the Office of Administration public portal: <http://www.oa.pa.gov/Policies/Pages/default.aspx>. Questions regarding this publication are to be directed to RA-ITCentral@pa.gov.

8. Exemption from This Policy

In the event an agency chooses to seek an exemption from the guidance within this ITP, a request for a policy waiver is to be submitted via the enterprise IT policy waiver process. Refer to [ITP-BUS004 IT Policy Waiver Review Process](#) for guidance.

This chart contains a history of this publication's revisions. Redline documents detail the revisions and are available to CWOPA users only.

Version	Date	Purpose of Revision	Redline Link
Original	09/21/2007	Base Document	
Revision	11/18/2010	ITP Refresh	
Revision	06/22/2021	<ul style="list-style-type: none">ITP RefreshAdded to current ITP templateAdded third-party vendors to Scope and Responsibilities SectionsUpdated Related ITPs and Exemption Section	Revised IT Policy Redline <06/22/2021>