



Version 1.7, 19 avril 2024

Archives Web Suisse

Sites web représentatifs sur la Suisse

Une collection commune de bibliothèques cantonales suisses, de bibliothèques et archives spécialisées et de la Bibliothèque nationale suisse

Glossaire

Modifications dans le document

Version	Date	Remarque
1.0	22.02.2006	Création
1.1	01.04.2008	Actualisation
1.2	01.05.2009	Actualisation
1.3	15.01.2011	Actualisation
1.4	05.04.2012	Actualisation
1.5	12.12.2012	Actualisation
1.6	05.02.2016	Actualisation
1.7	19.04.2024	Actualisation

1	Table des matières	
1	Table des matières	2
2	Introduction	3
3	Glossaire	3

2 Introduction

Le glossaire ci-après contient le vocabulaire professionnel et les abréviations utilisés dans les notices des Archives Web Suisse.

3 Glossaire

AACR	Anglo-American Cataloguing Rules
Access	Processus dans le modèle OAIS: mise à disposition des données d'archives pour l'utilisation.
Administration	Processus dans le modèle OAIS: surveillance des processus de travail dans le système d'archivage.
Archivage de longue durée	Par archivage de longue durée, on entend la saisie et la conservation à long terme d'informations. C'est surtout l'archivage de longue durée d'informations existant sous forme digitale (digital preservation) qui pose de nouveaux problèmes.
Archival Storage	Processus dans le modèle OAIS: stockage des objets digitaux dans le système d'archivage.
Bases de données	Les bases de données sont des systèmes permettant de stocker et de gérer de grandes quantités de données. Elles se composent d'une base de données dans laquelle les données sont déposées selon une structure définie, et de programmes de gestion qui stockent les données, les cherchent ou exécutent d'autres opérations avec les données. Généralement on appelle aussi base de données un programme d'application pour l'organisation et la gestion structurées de données. Les contenus des bases de données sont saisis par des interfaces utilisateurs interactives (principalement propriétaires). Les utilisatrices/utilisateurs transmettent des demandes de recherches individuelles et les résultats individuels requis sont préparés.
BN	Bibliothèque nationale suisse
Cache	Le cache est une mémoire tampon rapide, c'est-à-dire une mémoire dont l'accès aux contenus est spécialement rapide et dans laquelle les données utilisées à plusieurs reprises sont déposées temporairement.
Crawler Traps	Un crawler trap (piège à robot) désigne des applications ou fonctions problématiques sur des sites web qui peuvent générer de nouvelles URL à l'infini, et que le robot va suivre (à l'infini). Ainsi, le harvesting d'un tel site ne peut jamais être terminé. Les crawler traps les plus connus sont : calendriers, plans de localité, linkfarms, wikis, forums et contenus générés par CMS.
Data Management	Processus dans le modèle OAIS: gestion et maintien des données d'archives par des métadonnées bibliographiques, techniques et administratives.
Digicoord	Digicoord est une plateforme d'information sur les projets de numérisation suisses. Ce service est mis à disposition par la Bibliothèque nationale suisse et RERO, le réseau des bibliothèques de Suisse occidentale.
Domain	Un domain est un domaine défini par des caractéristiques communes; on entend généralement par là un groupe d'ordinateurs ayant des éléments de nom communs (host name). Le plus petit élément commun du nom se nomme Top-Level-Domain, p.ex. l'abréviation du pays ".ch" ou l'abréviation ".com". Les domaines peuvent être divisés en sous-domaines qu'on appelle Second-Level-Domain, Third-Level-Domain etc.
e-Helvetic	Service de la Bibliothèque nationale suisse qui est responsable de la collection, du catalogage, de l'archivage de longue durée et de la mise à disposition des publications numériques originales.
e-Helvetic Access	Le terme désigne le système d'accès par lequel les publications numériques originales peuvent être recherchées en plein texte et affichées.
E-Mail (Electronic Mail)	Les e-mails sont des messages (textes, fichiers) qui sont transférés dans un réseau entre expéditeurs/trices et récepteurs/trices, par le service "Simple Mail Transfer Protocol" (SMTP).

FTP (File Transfer Protocol)	FTP (File Transfer Protocol) est un protocole de transfert de fichiers sur Internet. FTP désigne également le service Internet sur lequel ce protocole est basé. SFTP = Secure File Transfer Protocol.
Harvesting	Le harvesting, aussi appelé webharvesting, est un système de sélection automatique de sites web via un harvester.
Helvetica	On entend par Helvetica les publications qui paraissent en Suisse, se rapportent à la Suisse ou à des personnes de nationalité suisse ou qui ont été créées par ou avec le concours d'auteurs ayant un lien avec la Suisse.
Helveticat	Catalogue de la Bibliothèque nationale suisse.
Homepage	La homepage est la page d'accueil ou de présentation d'un fournisseur de pages Internet sur le World Wide Web. On désigne également sous ce nom le site web d'un particulier. Aujourd'hui, le mot ne désigne donc plus uniquement comme à l'origine l'écran de présentation d'un site mais l'ensemble du matériau offert sur le site.
HTTP (Hypertext Transfer Protocol)	HTTP est un protocole de transfert pour le transfert de données dans le WWW, basé sur le protocole de transport TCP/IP.
Hyperlien	L'hyperlien, ou simplement lien, est le lien entre un élément dans un hypertexte ou sur une page Internet avec un autre élément sur la même page ou avec une autre page. Le terme dont part le lien est souligné et mis en évidence en couleur ; le lien est parfois aussi représenté par un graphique (un symbole).
Ingest	Processus dans le modèle OAIS: prise en charge de données du producteur, préparation à l'archivage et intégration dans le système de stockage pour l'archivage de longue durée.
Internet	L'Internet est le plus grand réseau d'ordinateurs du monde, il se compose de nombreux réseaux reliés entre eux et met à disposition une infrastructure de communication variée. Les prestations de l'Internet – on parle aussi de services – sont entre autres: le mail, le courrier électronique; World Wide Web (WWW), un système de documents basé sur l'hypertexte; Usenet, un forum de discussion; FTP, un système de transfert de toutes sortes de fichiers; IRC, le chat; Gopher, Telnet, Wais, Archie, quelques anciens services dont l'importance décroît. Pour beaucoup d'utilisateurs privés, le WWW est la contribution la plus importante, c'est pourquoi cet élément égale souvent l'Internet.
Intranet	L'intranet est un réseau privé (propre à une entreprise) qui travaille avec la technologie de l'Internet. Contrairement à l'Internet, l'intranet n'est accessible qu'à un cercle restreint et local d'utilisateurs.
MARC	Machine Readable Cataloging
Méta-tag robots	Un méta-tag robots est l'une des méta-indications qui peuvent être stockées dans le code source d'une page web. Il donne au crawler de la machine de recherche des indications sur le comportement d'indexation. Par exemple que seule la première page doit être intégrée à l'index de la machine de recherche.
Métadonnées	Les métadonnées sont des données sur des données. Autrement dit, les métadonnées désignent des données qui décrivent d'autres données. Il s'agit donc d'informations sur des données qui facilitent au maximum l'accès aux données souhaitées ainsi que l'échange et le traitement de ces données. Ces informations de base renferment p. ex. des indications sur l'auteur du document, la date de publication, des renvois à des documents sur le sujet, etc.
Ninive	Nom du serveur d'archivage de la Bibliothèque nationale suisse.
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting.
OAIS	Open Archival Information System. Standard pour l'archivage de longue durée de données électroniques (ISO 14721:2012; Space data and information transfer systems). Voir sous https://www.iso.org/fr/standard/57284.html
Page web	Une page web est une page d'un fournisseur d'information sur la Toile (ou World Wide Web). La page web est à distinguer du site web. Un site web se compose habituellement d'un groupe de pages web ordonnées hiérarchiquement et dont la page principale est appelée homepage ou page d'accueil.

Persistent Identifier	Les Persistent Identifiers sont des identificateurs uniques, et indépendants de l'emplacement, d'objets numériques, qui garantissent en même temps l'accès constant à des ressources électroniques (p.ex. DOI, URN etc.).
Preservation Planning	Processus dans le modèle OAIS: planification et mise en œuvre des mesures de maintien de longue durée sur les objets digitaux dans le système d'archivage.
Publications électroniques	Par publications électroniques, on entend aussi bien les publications en ligne que les publications hors ligne. On dit aussi "publications numériques".
Publications en ligne	Les publications en ligne (ou online) forment un sous-ensemble des publications électroniques. Les publications en ligne sont publiées et diffusées sur Internet sans support physique. Les publications en ligne apparaissent sous toutes sortes de formats de données et de formes de présentation. Exemples: revue électronique, banque de données, newsletter par e-mail, etc. Autres expressions utilisées: publications en réseau, publications online.
Publications numériques (born-digital)	Les publications numériques sont des publications numériques originales. Fabriquées sur support numérique, elles n'existent que sous cette forme.
Publications numérisées (digitized)	Les publications numérisées sont des publications converties de l'analogique (papier p. ex.) au numérique.
RDA (Resource Description and Access)	Règles internationales de catalogage
Site web (site)	Un site web se compose généralement de plusieurs pages web classées hiérarchiquement.
Sites web dynamiques	Dans le contexte bibliothéconomique, les sites web dynamiques sont des publications qui ne sont pas achevées au moment de leur publication initiale. Elles peuvent être modifiées et augmentées dans des mesures et à des fréquences variables (ressources intégratrices).
Sites web statiques	Dans le contexte bibliothéconomique, les sites web statiques sont des publications achevées au moment de leur première parution et qui ne subiront donc plus aucune modification a posteriori (ressource monographique). Les "versions" de sites web statiques modifiées sur le fond sont considérées comme de nouvelles publications en ligne.
Stylesheet	Un stylesheet (feuille de style) est plus ou moins comparable à un type de format. Le langage standard du stylesheet dans le World Wide Web est CSS. L'idée de base est de séparer le plus possible les modèles de représentation des contenus : la représentation des contenus est définie dans des fichiers CSS séparés (par ex. la mise en page, les couleurs et la typographie).
Tableaux (Mailbox, Bulletin Board System)	Les tableaux sont des pages définies dans un mailbox (Bulletin Board System) ou un newsgroup dans lesquels les utilisateurs/trices peuvent échanger des informations – comme un avis accroché au tableau noir.
URL	L'URL (Uniform Resource Locator) est l'adresse du site web dans l'Internet.
URN (Uniform Resource Name)	Un URN (Uniform Resource Name) est un Persistent Identifier. Les Persistent Identifiers peuvent remplacer les URL (Uniform Resource Locator ; „lien“ dans l'Internet) dans le catalogue ou dans d'autres systèmes de recensement, ou être utilisés comme références stables dans les documents eux-mêmes, ce qui permet d'avoir des liens stables. La mise à jour des références prend moins de temps, puisque les URL sont mis à jour de façon automatisée en un seul endroit. Les liens peuvent être intégrés dans plusieurs services de recensement. Les publications digitales ont un identificateur unique au monde et peuvent donc être citées de manière fiable. L'URN garantit un accès durable à un objet. L'accès durable est garanti par l'archivage de longue durée ou l'archivage des objets, ainsi que par la haute disponibilité technique du service URN. Un URN renvoie au moins à un URL par lequel un objet est adressé. Un URN peut également gérer plusieurs copies du même objet, donc plusieurs URL, ainsi que différents formats de présentation des objets.

Usenet (News-groups)	Usenet est un réseau de forums ou groupes de discussion qui est en principe indépendant de l'Internet. Une autre désignation pour Usenet est Newsgroup, ce qui peut toutefois créer des malentendus. Un newsgroup n'est en réalité qu'une partie – peut-être très petite – de Usenet, qui traite d'un thème précis. Mais il existe aussi des news-groups en dehors de Usenet à proprement parler, par exemple dans les intranets propres à une entreprise ou dans des écoles ou universités.
Wayback Machine	La Wayback Machine est un environnement web fermé qui a été développé par Internet Archive. Elle permet de rechercher des versions archivées de sites web et d'y naviguer.
WebDAV (Web-based Distributed Authoring and Versioning)	WebDAV est un standard ouvert permettant de mettre à disposition des fichiers dans l'Internet.
Weblogs (Blogs)	Un weblog, en général simplement appelé blog, est un site web qui contient périodiquement de nouvelles entrées. Un weblog est une sorte de journal intime virtuel dans lequel l'auteur (blogger) note des pensées personnelles et des liens commentés à des sources Internet qu'il/elle a visitées.
World Wide Web (WWW)	Le World Wide Web est un système hypermédia pour l'Internet, développé en 1989 au Centre européen de recherche nucléaire CERN. Le WWW permet d'accéder à toutes sortes de documents stockés sous forme digitale sur un ordinateur qui se trouve quelque part dans le monde et qui est relié avec le WWW. Les documents sont des textes normaux, des hypertextes, mais aussi des fichiers de musique et d'images ainsi que des films. Le World Wide Web et l'Internet ne sont pas identiques, mais le World Wide Web se base sur l'Internet. Toutefois, la plupart des surfeurs et surfeuses sur Internet utilisent exclusivement le World Wide Web pour leurs activités. Le World Wide Web utilise le protocole de transmission http; il existe encore d'autres protocoles dans l'Internet, p.ex. FTP. La majorité des documents texte dans le WWW est aujourd'hui créée dans la langue de description de pages HTML. Sa caractéristique principale, c'est les hyperliens, qui permettent d'appeler à partir d'un document n'importe quel autre document qui est accessible dans le WWW, à l'aide d'une procédure d'adressage standardisée, et ce par le protocole HTTP. A côté de HTML, il existe encore une multitude d'autres formats de données pour les textes et autres types de documents.
XML	Extensible Markup Language
Z39.50	Z39.50 est un protocole réseau utilisé comme standard en bibliothéconomie pour la consultation de systèmes d'information bibliographique.