

L'Euro 08 pour l'éternité

L'Euro 08 est terminé, le champion est connu. Sur le web helvétique, Google recense plus de trois millions d'entrées sur la compétition. La Bibliothèque nationale suisse archive les plus importants de ces sites. Pour l'éternité.

Aujourd'hui, l'ambiance de l'Euro est encore dans toutes les têtes, mais demain elle sera oubliée.

Dans une génération, les chercheurs voudront savoir comment les choses se sont passées pendant le championnat d'Europe de football, comment la Suisse a appréhendé cet événement, comment elle a réagi. A côté des livres, des journaux et des périodiques, les sites web sont devenus une source d'information importante. Les collections de la Bibliothèque nationale ne proposeront donc pas seulement les publications imprimées d'aujourd'hui, mais aussi un choix de sites web intéressants.

La sélection

Dans le domaine de l'imprimé, la Bibliothèque nationale suisse ambitionne une collecte quasi complète des Helvetica. Le concept d'Helvetica désigne, en Suisse, les publications éditées en Suisse, sur la Suisse ou sur des personnes avec droit de cité suisse ou domiciliées en Suisse, ou produites par des auteurs suisses ou liés à la Suisse, sans jugement de valeur. Ce principe vaut également pour les collections d'Helvetica numériques.

En raison des quantités pharaoniques de données publiées sur Internet, une politique de sélection restrictive s'impose toutefois.

Une spécialiste en information documentaire de la Bibliothèque nationale a défini un principe de sélection pour constituer la collection des pages touchant l'Euro 08. Le critère de base était le suivant: l'Euro 08 devait être le thème central des sites retenus et non un point marginal. L'objectif était également de couvrir un large éventail de thèmes autour de l'Euro 08.

La collection comprend désormais près de 40 sites, documentant l'Euro 08 sous divers angles:

- Le site officiel de l'UEFA, par exemple, propose les résultats ainsi que les rapports sur les matchs, des informations sur les faits marquants et des photographies.
- Les pages web officielles des villes hôtes (Bâle, Genève, Zurich) présentent notamment les sites de leur villes les possibilités d'hébergement et signalent même des cours intensifs de langue pour les fans étrangers.
- Les pages des CFF présentent des attractions touristiques comme des informations sur les transports publics pendant l'Euro 08.
- Le Musée du sport suisse expose sur son site le projet culturel «1924 – nous allons reconquérir le titre».
- Le site Ecosport.ch propose aux organisateurs des conseils environnementaux pour la gestion des «fanzones».
- Divers sites ont été conçus spécialement pour cette occasion par les fans de football pour les fans de football. Ils contiennent, par exemple, des informations sur les fan-shops ou les possibilités de camping mais on y trouve aussi des pages web pour ceux que l'on appelle les «ronchons du foot». Pour tous ces sites, une sélection s'impose.
- Le site de La Poste témoigne des stratégies de marketing mises en œuvre lors des grands événements et présente les principales étapes que doit franchir une entreprise si elle veut prendre le train de l'Euro 08 en marche et en tirer profit.
- Les pages de l'Office fédéral de la police (fedpol) et de l'armée suisse informent sur les mesures de sécurité prises durant l'Euro 08.
- La «campagne Euro 08 contre la traite des femmes» se greffe sur le championnat de football et donne des informations sur ses objectifs.
- Enfin, tous les jours des pages web journalistiques donnent les informations les plus importantes autour de l'Euro 08 et les documentent.

Le traitement de l'information

Avant que ne soient téléchargée et collectée la sélection de sites web, tous les exploitants de sites ont été informés par écrit sur le projet de la Bibliothèque nationale suisse. Les réactions ont été positives.

La moisson en elle-même fut opérée par la fondation European Archive (<http://www.europarchive.org>).

European Archive s'est spécialisée dans l'archivage du web et mène des opérations de récolte sélective pour diverses institutions patrimoniales en Europe.

Durant l'Euro 08, un moissonneur – un programme qui collecte les documents sur Internet de manière systématique – télécharge les pages sur un serveur. Tous les liens sont répertoriés sur une page web, qui montre les offres d'information et sur quelles pages les obtenir.

La plupart des sites ont été téléchargés une seule fois durant la compétition. Les pages du site de l'UEFA ont été relevées deux fois par semaine.

Lors du téléchargement, les pages web sont en quelque sorte gelées et documentent avec précision l'état de l'information au moment auquel a eu lieu l'opération. Par conséquent, les informations saisies après la moisson ne sont pas incluses et sont donc perdues pour cette archive.

La volatilité des informations sur Internet est une réalité avec laquelle les institutions patrimoniales doivent aussi apprendre à vivre.

A partir des quelque 40 sites retenus, le moissonneur a récolté 52 millions de documents. Cela représente un volume de données d'environ 470 Gigabyte soit 100 couches de DVD ou 700 CD.

Le contrôle de qualité des sites collectés est actuellement en cours. Cette opération est lourde, car elle est faite manuellement et visuellement. Il n'existe pas à ce jour de moyens pour automatiser ce contrôle de qualité.

Les erreurs qui ont été détectées concernent, par exemple, des graphiques manquant (logos externes) ou des fonctions telles que Pulldowns qui ne réagissent plus correctement.

Dès que le contrôle de qualité est achevé, les pages sont disponibles. Elles peuvent être cherchées et consultées via Helvetikat, le catalogue de la Bibliothèque nationale (<http://www.nb.admin.ch/helveti...>).

La Bibliothèque nationale utilise l'URN (Uniform Resource Name) pour identifier de manière univoque les objets. L'URN garantit la permanence de l'accès aux pages web archivées à la Bibliothèque nationale.

Les sites sont tous sauvegardés dans l'archive à long terme de la Bibliothèque nationale et rendues disponibles de manière pérenne.

Archive Web Suisse

Les sites de l'Euro 08 forment une partie de l'Archive Web Suisse, opération développée par la Bibliothèque nationale avec les bibliothèques cantonales et d'autres institutions intéressées.

Outre les sites d'événements majeurs (comme les élections fédérales de l'automne 2007), l'Archive Web Suisse contient les productions des cantons et des chefs-lieux de district, ainsi que des pages touchant des thèmes suisses particulièrement importants. L'objectif de l'Archive Web Suisse est de construire une grande collection de sites importants sur la Suisse et de garantir leur conservation et leur disponibilité à long terme. Il s'agit là d'une utilisation optimale du potentiel de synergie entre les bibliothèques suisses patrimoniales.

e-Helvetica

Conserver et rendre accessible à long terme les sites Internet est un défi relevé au sein du projet e-Helvetica de la Bibliothèque nationale (<http://www.nb.admin.ch/e-helve...>).

Depuis 2001, une équipe de professionnels de l'informatique et de spécialistes en bibliothéconomie élabore les bases pour la collecte, l'indexation, la mise à disposition et la conservation à long terme des publications électroniques en provenance de Suisse ou ayant un rapport fort avec la Suisse

La collection comprend actuellement, outre les pages web, des livres et des revues en ligne d'éditeurs suisses, des travaux des universités et des hautes écoles suisses, ainsi que les publications en ligne de l'administration fédérale.

La plate-forme pour l'archivage des contenus est d'ores et déjà en service. Une partie des publications commerciales, des travaux universitaires et des pages web peut être déposée par voie électronique. Actuellement, un moteur de recherche pour l'interrogation de cette archive est en cours de développement. Cette plateforme permettra aussi bien aux usagers d'aujourd'hui que de demain d'accéder aux collections numériques de la Bibliothèque nationale suisse.



Barbara Signori

**Projektleiterin Webarchiv Schweiz an der Schweizerischen
Nationalbibliothek**