

Les archéologues du numérique

Autor(en): **Saraga, Daniel**

Objektyp: **Article**

Zeitschrift: **Horizons : le magazine suisse de la recherche scientifique**

Band (Jahr): **27 (2015)**

Heft 105

PDF erstellt am: **11.09.2024**

Persistenter Link: <https://doi.org/10.5169/seals-771909>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern. Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden. Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Les archéologues du numérique

Numériser les objets historiques ne les préserve pas du déclin. Un projet suisse veut sauver les données produites par la recherche.

Par Daniel Saraga

Internet n'oublie jamais», assèment les experts inquiets de voir notre vie privée étalée sur le web. Mais rien n'est moins sûr: le changement constant des supports de l'information numérique et des formats des fichiers laisse derrière lui des montagnes d'informations qui, inexorablement, deviennent illisibles. Livrées à elles-mêmes, elles sont condamnées à l'oubli.

Cette problématique touche de plein fouet les résultats scientifiques, exprimés de plus en plus souvent sous forme de bases de données. Car celles-ci vieillissent très rapidement: les langages de programmation se révèlent obsolètes et les systèmes d'exploitation ne sont plus compatibles avec le nouveau serveur.

«C'est un vrai problème, souligne Lukas Rosenthaler, responsable du Centre de données et services pour les sciences humaines (DDZ), un projet consacré à assurer la pérennité des résultats scientifiques en sciences humaines (voir «Sauvegarder le sauveur»). Il est très rare que les chercheurs continuent à maintenir les infrastructures numériques une fois un projet - et son financement - terminé. Une base de données inaccessible est inutilisable, et ne pas l'entretenir revient à détruire une

production scientifique. Paradoxalement, ce type de recherche numérique peut s'avérer bien plus fragile qu'un résultat publié sous forme d'article.»

Les copistes du III^e millénaire

Le directeur des Digital Humanities Lab à l'Université de Bâle a pu sauver l'une des plus importantes bases de données consacrée à la mythologie grecque, le *Lexicon Iconographicum Mythologiae Classicae*, qui s'est arrêté en 2009, après trente ans de développement: «Elle était complètement hors service, et l'entreprise qui l'avait programmée était en faillite, poursuit Lukas Rosenthaler. Nous avons même dû pirater le site, car les mots de passe avaient disparu. Notre travail ressemble parfois à une sorte d'archéologie du numérique.» Ces efforts redonnent une seconde vie aux résultats de la recherche. L'Université Harvard, par exemple, s'intéresse aujourd'hui à intégrer le *Lexicon* dans son commentaire d'Homère, en utilisant les «linked open data», une composante du web 3.0 qui permet de relier les informations en ligne de manière directe et dynamique.

Avec sa petite équipe, l'ancien physicien s'est appuyé sur la technologie sémantique pour créer une plateforme générique capable de structurer des données provenant de plateformes très variées. «J'estime que nous pouvons traduire 99% des bases de données utilisées en sciences humaines, et même certains projets en biologie. En trois ans, nous avons migré une trentaine de projets, de la mythologie grecque à une collection de photographies historiques de montagne.»

Le DDZ suit le concept de l'Open Archival Information System: copier les données régulièrement et les retranscrire dans un nouveau format plus actuel.

Un processus difficile et coûteux, qu'il faut répéter régulièrement - le pendant digital du travail des moines copistes du Moyen Âge. «La plupart des groupes de recherche n'ont pas les moyens de créer des outils stables, poursuit Lukas Rosenthaler. L'idéal est de travailler avec eux dès le début pour élaborer une base de données durable qui puisse ensuite être facilement mise à jour et migrée.»

Le mouvement de l'Open Research Data encourage les chercheurs à générer des résultats scientifiques accessibles à chacun et de manière interactive. Mais comme toute machine, l'outil exige un entretien régulier. Un travail de l'ombre essentiel pour éviter de voir cette connaissance disparaître à jamais.

Daniel Saraga est responsable de la communication scientifique du FNS.

Final report for the pilot project «Data and Service Center for the Humanities» SAGW, 2015

Le patrimoine culturel à l'ère numérique, Nike Bulletin, juin 2014

<http://openresearchdata.ch/>

Sauvegarder le sauveur

Voué à la préservation des archives numériques, le Centre de données et services pour les sciences humaines (DDZ) est lui-même en danger. «Nous nous battons depuis 2008 pour mettre en place une plateforme stable, déclare Markus Zürcher, secrétaire général de l'Académie suisse des sciences humaines, qui a lancé le projet. Toutes les personnes concernées soutiennent cette plateforme. La seule chose à régler, c'est le financement.»

Le DDZ reste un projet pilote qui touche à son terme. «Nous avons déposé en mars 2015 une demande portant sur 2 millions de francs pour 2017-2020 auprès du SEFRI (Secrétariat d'Etat à la formation, à la recherche et à l'innovation). D'ici là, nous sommes prêts à assurer le financement jusqu'en 2017, car une interruption du projet serait très néfaste.» Par comparaison, quelque 30 millions sont dépensés annuellement pour les bases de données dans les sciences humaines.

◀ Pages 19 et 20. La cartographie crée une forme de réalité augmentée en ajoutant des données à la géographie. Ces vues globales permettent d'identifier rapidement des informations cachées, comme ici l'explosion d'achats alimentaires avant Pâques et celle de la restauration sur la Costa Tropical au sud de l'Espagne.

Images: Valérie Chételat (p. 19);

MIT Senseable City Lab - Carlo Ratti, Director -

Pedro Cruz: Visualization (p. 20)