

Zeitschrift: Horizons : le magazine suisse de la recherche scientifique
Herausgeber: Fonds National Suisse de la Recherche Scientifique
Band: - (1998)
Heft: 36

Artikel: Archives sur Internet: attention, sites mouvants!
Autor: Giussani, Bruno
DOI: <https://doi.org/10.5169/seals-556027>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 05.05.2025

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Bruno Giussani
(bg@giussani.com) est journaliste,
spécialisé dans les nouveaux
médias, et chroniqueur Internet
du «New York Times» et de la
«Neue Zürcher Zeitung».



Archives sur Internet: attention, sites mouvants!

Une partie des sources de la recherche scientifique sont menacées par une phrase que les utilisateurs du réseau internet connaissent fort bien: «404 document not found.» Il s'agit du message d'erreur qui s'affiche sur l'écran de l'ordinateur lorsqu'une information n'est plus disponible sur le web, la portion multimédia de l'internet. Cela peut signifier que le document en question a été déplacé ailleurs dans le réseau, dans un autre ordinateur. Le plus souvent toutefois, «404» indique que l'information a été effacée. Qu'elle est probablement perdue.

L'internet est un médium en constant renouvellement. On estime que la durée moyenne d'un document sur le réseau est de 75 jours. Des milliers de pages web sont mises à jour toutes les heures, ou même toutes les quelques minutes.

Cette nature fluide donne à l'internet sa vitalité et sa fraîcheur. Elle n'est toutefois pas sans conséquences pour la recherche, notamment en sciences sociales (mais pas uniquement).

Depuis un an et demi, Brewster Kahle, un Américain de 37 ans, dirige à San Francisco une organisation appelée Internet Archive (www.archive.org) dont l'activité consiste, comme le dit son nom, à figer le renouvellement continu du réseau en archivant l'information publiquement accessible. «Ce qui se passe sur le réseau et dans le domaine des technologies numériques est tout simplement de la culture contemporaine», dit-il. Non seulement les textes, les images et les sons du web, mais également les documents disponibles dans les autres

portions publiques de l'internet (pas de panique: votre boîte aux lettres électronique ne sera pas touchée), ainsi que les logiciels.

L'initiative a le mérite de soulever la question de l'urgence de la sauvegarde d'information et de documents qui, de plus en plus souvent, sont exclusifs à l'internet, c'est-à-dire qui n'ont pas d'«édition papier». Comme on a conservé les enregistrements vidéo des messages publicitaires politiques et des débats transmis durant la campagne présidentielle américaine de 1996, par exemple, l'Internet Archive a également pris soin de faire une copie complète des sites web de tous les candidats. Pour l'historien, ou le politologue, il est évident qu'il s'agit de deux sources également précieuses.

L'archivage de l'information disponible sur l'internet pose tout d'abord des problèmes techniques. Le réseau est immense, multiforme, et continue de croître. Son contenu n'est pas structuré. Kahle reconnaît qu'il lui sera impossible d'enregistrer tous les changements, mais il espère être en mesure de prendre au moins une première «photo» de l'état actuel de l'internet, pour ensuite la mettre à jour à intervalles réguliers.

Ensuite, si nous pouvons lire aujourd'hui sans difficulté un livre imprimé par Gutenberg, cela n'est souvent pas le cas pour une disquette informatique vieille de dix ans. La conversion régulière des fichiers dans de nouveaux formats devrait résoudre le problème en partie. La sauvegarde des logiciels qui permettent de lire ou interpréter ces informations, ainsi que la conservation ou la simulation des machines sur lesquelles ils fonctionnent, paraît également primordiale.

Avant de pouvoir ouvrir ses archives aux chercheurs et, éventuellement, au public, Brewster Kahle doit toutefois résoudre aussi toute une série de questions juridiques sur la protection de la sphère privée et le droit d'auteur, comme copier une page web ne signifie-t-il pas contrevenir au droit d'auteur? (Oui).

En attendant de pouvoir formuler des réponses, Kahle continue d'archiver – pour éviter que ces «traces précieuses de l'aventure humaine» ne soient perdues à jamais.

B.G.