

Quelques cas de littérature grise ayant fait les manchettes

Disparition de pans entiers de données gouvernementales sous le régime Harper

- « Physicist Raymond Hoff, who published more than 50 reports on air pollution in transport and toxic chemicals in the Great Lakes—including pioneering work on acid rain—at Environment Canada between 1975 and 1999, doesn't seem to exist, either. “Nothing comes up when I type my name into the search engine on [Environment Canada's] website,” says Hoff, now a professor emeritus at the University of Maryland. Also gone are internal reports on the oil sands experiments of the 1970s. “That research was paid for by the taxpayer. Now, the people who need to protect Canada's environment can't get access.” »
- « Protecting Canadians' access to data is why Sam-Chin Li, a government information librarian at the University of Toronto, worked late into the night with colleagues in February 2013, frantically trying to archive the federal Aboriginal Canada portal before it disappeared on Feb. 12. The decision to kill the site, which had thousands of links to resources for Aboriginal people, had been announced quietly weeks before; the librarians had only days to train with web-harvesting software.

The need for such efforts has taken on new urgency since 2014, says Li, when some 1,500 websites were centralized into one, with more than 60 per cent of content shed. Now that reporting has switched from print to digital only, government information can be altered or deleted without notice, she says. (One example: In October 2012, the word “environment” disappeared entirely from the section of the Transport Canada website discussing the Navigable Waters Protection Act.) »

- Source : Kingston, A. (2015). [Vanishing Canada: Why we're all losers in Ottawa's war on data.](#) *Maclean's*

Cas très médiatisés autour de la volonté de rendre grise de la littérature blanche ou noire

- Du noir au gris :
 - Julian Assange et Bradley Manning par le biais d'informations sensibles dévoilées sur WikiLeaks
 - Edward Snowden, un ancien consultant de la NSA (service de renseignement américain), a révélé que les États-Unis espionnent les individus et les autres nations (le célèbre « Yes we can » d'Obama est devenu « Yes we scan »)
 - L'affaire Panama Papers au sujet de l'évasion fiscale
- Du blanc au gris :
 - Aaron Swartz, un jeune cybermilitant ayant téléchargé sur le Web de très nombreux articles à partir de bases de données à accès réservé par souci de diffuser l'information souvent financée à même des fonds publics, s'est suicidé étant donné les poursuites intentées contre lui
 - À voir : le documentaire *The Internet's Own Boy: The Story of Aaron Swartz* (2014) par Brian Knappenberger

Qu'est-ce que la certification HON?

Health On the Net Foundation (HON) :

- Voir les informations ajoutées en commentaires à la diapo 17 de la présentation
- « [Procédé d'adhésion](#) :

La demande de certification est gratuite et valable pour une période de un an mais liée au respect et à l'acceptation des [conditions générales de demande de certification](#). Le renouvellement et l'extension du certificat HONcode implique une cotisation en fonction du type et de la popularité du site. Vous êtes en accord avec le fait que votre application peut être rejetée et que la Fondation HON n'est pas obligée d'apporter de justifications quant aux motifs du rejet. De même, votre certification HONcode peut être suspendue ou retirée à tout moment par HON. Vous acceptez de respecter les décisions de HON et d'enlever le lien actif du HONcode si HON vous demande de le faire. »

Au sujet du Web

Est-ce que les sites .gov et .edu sont vraiment fiables?

- Le Internet Assigned Numbers Authority (IANA) est un département de la Internet Corporation for Assigned Names and Numbers (ICANN, ou Société pour l'attribution des noms de domaine et des numéros sur Internet, une société à but non lucratif fondée en 1998). L'ICANN « rassemble, au sein d'un partenariat à but non lucratif, des personnes du monde entier qui œuvrent au maintien de la sécurité, de la stabilité et de l'interopérabilité d'Internet. Elle encourage la concurrence et développe des politiques d'identifiants Internet uniques. » ([Quel est le rôle de l'ICANN ?](#))
- Le IANA est responsable de coordonner certains éléments clés permettant à Internet de fonctionner sans heurt, et ce, notamment en allouant et en maintenant les codes uniques et les systèmes de numérotation utilisés dans les standards techniques d'Internet (ex. : noms de domaine, adresses IP, protocoles)
 - Source : IANA, [Introducing IANA](#) et [Root Zone Database](#)
- Il existe une liste des domaines de premier niveau d'Internet, notamment ceux relatifs aux nations (*country code Top-Level Domain* ou *ccTLD*). Pour chaque pays, une entité coordonne l'attribution du domaine de premier niveau national
 - « L'inclusion d'un domaine de premier niveau dans un nom de domaine ne confirme pas la nationalité de l'organisation qui le possède ni sa présence sur le territoire mentionné. Pour utiliser un domaine de premier niveau national, il suffit de s'enregistrer auprès du bureau d'enregistrement correspondant. Certains bureaux d'enregistrement exigent une adresse postale dans le territoire correspondant, ce qui n'est pas très contraignant car il peut s'agir d'une simple boîte postale. D'autres bureaux d'enregistrement n'exigent même pas que le demandeur possède une telle adresse postale. » (Wikipedia (s.d.). [Domaine de premier niveau](#))
 - Au Québec, c'est « [l]e service de gestion des noms de domaine (DNS) du CSPQ [qui] prend en charge l'enregistrement et la gestion opérationnelle des sous-domaines Internet de **.gouv.qc.ca** et **.gov.qc.ca**, de même que des sous-domaines intranet de **.qc**. » ([Gestion des noms de domaines](#))
 - CSPQ = Centre de services partagés du Québec
- En juin 2011, l'ICANN a adopté une réforme pour libéraliser les extensions Internet
 - « En 2008, l'ICANN décide de lancer un **programme de libéralisation des extensions internet**. Le but étant de permettre des candidatures -presque- totalement libres pour créer de nouvelles extensions et par conséquent de nouveaux noms de domaine dès la fin 2013. Finalement c'est **1 930 candidatures** qui ont été déposées malgré des frais de dossier

s'élevant à 185 000 US Dollars. Il en reste actuellement 1 814 sachant que parfois plusieurs candidatures concourent pour la même extension : par exemple le « .app » qui compte 11 candidats. [...] En donnant son feu vert pour la création de nouvelles extensions, l'ICANN cherche surtout une solution à l'essoufflement dont sont victimes certaines extensions génériques (et historiques) telles que le .com ou le .net. En effet, ces extensions font face à un marché saturé où la disponibilité des termes s'amoinde. » (Muller, S. (2014). [Dossier : la libéralisation des extensions Internet \(1/3\)](#). *Économie numérique*)

Comment et pourquoi un site ou une page Web peut ne pas être indexable/repérable ?

- Le contenu numérique n'est pas moissonnable si certaines balises sont utilisées → ex. : <meta name="robots" content="noindex" />
 - index ou noindex : page indexable ou non
 - follow ou nofollow : page que l'on peut suivre ou non
 - archive ou noarchive : page archivable ou non
- Pourquoi? Pour prévenir l'indexation...
 - d'un contenu reproduit (ex. : version imprimable d'une page)
 - de pages incomplètes ou en développement
 - de pages ou de fichiers « confidentiels »
 - de ses données personnelles sur un site social
- Bref, tout le Web n'est pas visible!

Au sujet de la recherche

La méthode boule de neige

- « Snowballing is sometimes referred to as *reference harvesting* or *pearl growing search method*; here are some other terms that evoke similar kinds of search strategies:
 - forward citation searching, footnote chasing, reference scanning, reference harvesting, hand-searching & powerbrowsing
 - backward chaining, forward chaining, digital browsing, footnote chasing
 - pearl growing, reference harvesting, reference lists, reference searches, 'cited by searching' ...and so on » (HLWIKI International. [Snowballing](#), incluant une description des techniques)

Au sujet de Google et de Google Scholar

Peut-on exporter des références de Google Scholar vers un logiciel de gestion bibliographique comme Zotero ?

Oui, mais une référence à la fois.

Élément supplémentaire qui distingue l'interrogation de bases de données comme MEDLINE/PubMed et Google Scholar

- Une requête lancée dans une base comme MEDLINE est interprétée telle quelle, en fonction de la chaîne de caractères et des opérateurs saisis (il est possible d'élargir la recherche en activant des mises en correspondances par le moteur). En parcourant les résultats obtenus, on peut déceler des éléments qui causent du bruit et les rectifier (ex. : une troncature trop large ou un acronyme qui entraîne des résultats non pertinents).
- L'équipe derrière PubMed, un moteur gratuit grand public, a mis en place un algorithme qui repêche d'emblée des résultats en fonction de la chaîne de caractères et des opérateurs saisis et de

mises en correspondances automatiques. Il est important de vérifier comment une requête a été traduite à l'aide du *Search details* (à droite des résultats) et de la rectifier s'il y a lieu.

- Les résultats obtenus dans Google et dans Google Scholar dépendent bel et bien d'un algorithme secret sur lequel on a peu de prise. Il tient compte de plus de 200 critères (dont les pages consultées antérieurement, le lieu où on se situe, etc.), ce qui rend difficiles les ajustements. En outre, il existe plus d'un Google (Google.com, .ca, .fr...).

Pour ceux qui veulent en lire plus à ce sujet :

- HLWIKI Canada. [Bibliographie Google Scholar](#)