



HAL
open science

Archiver le Web

Nicolas Thély

► **To cite this version:**

Nicolas Thély. Archiver le Web. Read/Write Book n°2: introduction aux humanités numériques / dir. Pierre Mounier, Open Edition Press, pp.179-188, 2012. hal-01801166

HAL Id: hal-01801166

<https://univ-rennes2.hal.science/hal-01801166v1>

Submitted on 28 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Archiver le Web*

Nicolas Thély

Archiver le Web

179

En avril 1996, Brewster Kahle, brillant ingénieur en informatique à l'esprit franc-tireur formé au MIT, dévoile à un petit cercle de chercheurs et de journalistes Archive.org¹, un site ressource qui a pour objet de sauvegarder toutes les pages du Web : celles encore opérantes, mais surtout celles qui n'existent plus, abandonnées ou remplacées.

Ambitieux, ce projet humaniste sent bon la fibre libertaire qui agite encore le Net. Pourtant, pour y parvenir, Brewster Kahle s'appuie sur son savoir-faire et son sens des affaires. La bonne santé de son entreprise non lucrative repose sur un transfert de technologie : parallèlement à la création du site Archive.org, Brewster Kahle crée Alexa Internet, une société spécialisée dans l'analyse du trafic sur le Web. Les robots qu'il développe, appelés « Crawle » et « Spider », auxquels le film *Matrix* a donné un aspect inquiétant et menaçant, permettent de connaître instantanément le taux de fréquentation d'un site, le temps passé sur une page... Non content de vendre à AOL son premier moteur de recherche, Brewster Kahle cédera en 1999 pour 250 millions de dollars² la société Alexa à Amazon, un autre géant naissant. Comme le remarque Jean-Marie Chauvet, rédacteur du site spécialisé

* La version originale de ce texte : « Archiver le Web » in *Le tournant numérique de l'esthétique*, publie.net, coll. « Critique & essais », 2012, p. 104-126, <http://www.publie.net/fr/ebook/9782814505582/le-tournant-num%C3%A9rique-de-l-esth%C3%A9tique>

1. Archive.org, <http://archive.org>
2. Source : http://en.wikipedia.org/wiki/Alexa_Internet

*ITRmanager.com*³ : « chaque fois que l'on cherche un ouvrage sur Amazon, et maintenant, un extrait dans l'indexation plein texte de ces ouvrages, ou un produit sur le site commercial, on utilise le glaneur Alexa et son immense index⁴ ». Ce sont ces mêmes robots qu'Archive.org lance sur le Web ; ils ont pour tâche, tous les deux mois, de prendre des clichés (des *snapshots*) de sites publics. Les données récoltées servent uniquement à la constitution de la mémoire du Web ; sur Archive.org le temps commence toujours en 1996.

Pour accéder à ces fameuses pages disparues du Web, l'expérience est digne d'un film de science-fiction : une application est mise à la disposition de l'internaute, elle a pour nom « The Wayback Machine », autrement dit « la machine à remonter le temps de l'internet ». Son principe d'utilisation est simple comme celui d'un moteur de recherche : il suffit d'entrer l'adresse d'un site ; si celui-ci est référencé, s'offrent alors à l'internaute les différentes étapes de son évolution. Deux logiques de consultation sont à l'œuvre : la première, la plus fréquente, consiste à consulter des clichés de ces sites constitués essentiellement des pages dont les liens sont inopérants ; la seconde conduit à naviguer dans le site qui a été aspiré. C'est le cas par exemple de The File Room⁵, un site initié par Antoni Muntadas en 1994 à l'occasion d'une exposition au Centre culturel de Chicago⁶ : les liens étant opérants, l'internaute peut se déplacer dans le site même, si bien qu'au bout d'un moment il se trouve désorienté, ne sachant plus dans quelle époque du site il navigue, se demandant même s'il se trouve véritablement sur une version archivée plutôt que sur le site lui-même.

L'entreprise de Brewster Kahle, qui n'est du reste pas la seule dans le domaine de l'archéologie du Web⁷, a été mûrement réfléchie, programmée : elle associe des chercheurs, des ingénieurs et des spécialistes du monde de l'informatique ; elle repose sur un calcul prévisionnel de l'expansion du Web, sur la budgétisation du coût du matériel de stockage (mémoire, RAM, disque dur, etc.). Brewster Kahle affichait dès le départ une telle assurance qu'il a procédé par effet d'annonce en publiant ces informations et le programme de développement du site dans un article destiné à la revue *Scientific American*⁸. Mais la bonne volonté et la rigueur de l'inventeur de la machine à remonter le temps de l'internet ont été confrontées très rapidement à la libéralisation

3. *ITRmanager.com*, <http://www.itrmanager.com>

4. Jean-Marie Chauvet, « Amazon souhaite Joyeux Noël à Google, Yahoo et Microsoft », *ITRmanager*, 18 décembre 2005, http://www.itrmanager.com/article_tribune.php?oid=62, consulté le 20 juillet 2006. À la date du 13 septembre 2011, cette page n'est plus accessible.

5. The File Room, <http://www.thefileroom.org>

6. Exposition du 20 mai au 4 septembre 1994.

7. On peut trouver une liste complète des différentes entreprises de constitution d'archives à l'adresse suivante : <http://www.ifs.tuwien.ac.at/~aola/links/WebArchiving.html>

8. Brewster Kahle, « Archiving the Internet », *Scientific American*, 1997, http://web.archive.org/web/19980113025831/www.archive.org/sciam_article.html, consulté le 20 septembre 2011.

des mentalités qui colonisent le Web. Dans un article publié dans le quotidien *Le Monde*⁹, Hervé Morin fait référence aux vicissitudes de l'entreprise :

La collecte automatique fondée sur l'utilisation des automates est loin d'être parfaite : les robots ne peuvent passer outre les multiples formulaires qui agrémentent un nombre croissant de pages d'accueil. Les sites payants ou à abonnement, même gratuit, restent en dehors de leur portée. Les gestionnaires des sites eux-mêmes peuvent les interdire d'accès (par une commande nommée « robot.txt »). Cette procédure s'est multipliée à l'encontre des automates d'Internet Archive, après que le site a été ouvert au public : les éditeurs en ligne – notamment les journaux – ont soudain pris conscience que leurs contenus pouvaient être « aspirés » pour devenir accessibles gratuitement, alors qu'eux-mêmes peinent à commercialiser ce fond de commerce¹⁰.

Ainsi existe-t-il quelques pages référencées du *Monde* de 1996 à 2000, et après plus rien. Ainsi également, au lendemain du 11 Septembre, les sites de la Maison-Blanche et de la NASA ont-ils disparu. Premiers signes de troubles de la mémoire du Web, d'autant plus inquiétants quand un site critique envers la scientologie disparaît aussi des archives¹¹...

Dix ans après sa création, les services offerts par Archive.org se sont étoffés, profitant de donations de documents vidéo, audio, et proposent une bibliothèque de *softwares*, un cimetière technologique apprécié des artistes du *software art*. Cependant l'esprit d'internet a changé ; il offre un paysage qui ressemble davantage à une prolifération de centres commerciaux qu'à une autoroute de l'information. Le moteur de Google règne sur la maîtrise du réseau. L'impertinence humaniste du projet de Brewster Kahle prend l'eau face à Google Print qui se veut une nouvelle puissance d'organisation du savoir¹². Par ailleurs, fin 2005, Amazon a mis à disposition des internautes la technologie développée par Alexa afin que ceux qui sont à la recherche des références d'un ouvrage soient prioritairement conduits sur le site d'Amazon¹³.

Alors, quid d'Archive.org, si le site ne peut répondre à ses ambitions humanistes et si les grands acteurs économiques du Web opèrent une rétention des données ? La Wayback Machine est-elle aujourd'hui reléguée au rang

9. Hervé Morin, « Internet cherche à se préserver de l'amnésie », *Le Monde*, 6 avril 2002.

10. *Ibid.*

11. Lisa M. Bowman, « L'archive de l'internet fait taire les critiques de la scientologie », *Antisectes.net*, 24 septembre 2002, <http://www.antisectes.net/archive.htm>, consulté le 20 septembre 2011.

12. Le projet a été présenté le 14 décembre 2004 et a pour objectif de numériser une quinzaine de millions de livres imprimés.

13. Jean-Marie Chauvet, *art. cit.*

des curiosités technologiques, devenue une sorte de Hubble du Web, voire pire, une simple attraction de fête foraine qui invite l'internaute à regarder rétrospectivement le réseau dans un miroir déformant ?

Une mise en forme des données de l'époque

Les déconvenues de l'entreprise humaniste de Brewster Kahle en font aujourd'hui un objet philosophique qui éclaire sous un angle inattendu notre manière d'être, notre contemporanéité. Plus précisément, le site apparaît comme un objet esthétique qui permet d'exposer autrement la perception des « données de notre époque ». L'expression est de Pierre-Damien Huyghe, qui depuis quelques années interroge les modes de perception à travers les appareils techniques. Dans *Faire place*, un court essai publié en 2006, l'auteur, s'interrogeant sur la notion d'héritage et de « chez soi », revient sur le texte de Walter Benjamin, *Expérience et pauvreté*, dans lequel le philosophe allemand explique que le « déploiement de la technique plonge les hommes dans une pauvreté tout à fait nouvelle¹⁴ » : nous sommes devenus « pauvres en expérience communicable¹⁵ ». « Cette pauvreté ne porte pas seulement sur nos expériences privées, mais aussi sur les expériences de l'humanité tout entière¹⁶. » À la question de Walter Benjamin, « Que vaut en effet tout notre patrimoine culturel, si nous n'y tenons pas, justement, par les liens de l'expérience¹⁷ ? », Pierre-Damien Huyghe répond par un audacieux renversement de point de vue : « On peut aussi affirmer, pour être plus positif, que c'est autrement qu'il [l'homme moderne] s'enrichit en expérience : il pousse la sociabilité bien plus en avant que ses ancêtres, il partage immédiatement nombre d'informations sur le monde¹⁸. » En d'autres termes, Pierre-Damien Huyghe reformule à sa manière la notion de barbarie ambiante diagnostiquée par Walter Benjamin en essayant de comprendre comment cet enrichissement est possible et comment il se manifeste. De la sorte, il accorde un vif crédit aux entreprises artistiques qui viennent mettre en lumière les procédures invisibles de la technologie, la manière dont elles s'infiltrèrent et se dissimulent sans résistance ni contre-pouvoir dans notre sphère privée. L'enjeu est donc celui « d'une mise en forme différente des données de l'époque¹⁹ ».

14. Walter Benjamin, « Expérience et pauvreté », *Œuvres II*, trad. Pierre Rusch, Paris, Gallimard, coll. « Folio Essais », 2000, p. 365.

15. *Idem*.

16. *Ibid.*, p. 366.

17. *Idem*.

18. Pierre-Damien Huyghe, *Faire place*, Paris, Éditions MIX, 2006, p. 13.

19. *Ibid.*, p. 16.

Cette réflexion conduite par Pierre-Damien Huyghe permet de suspendre un temps la puissance prédictive des deux illustres rapports sur l'informatisation de la société publiés en 1978 par Simon Nora et Alain Minc d'une part, et en 1979 par Jean-François Lyotard d'autre part, qui structurent le champ théorique et l'imaginaire collectif lié aux technologies de l'information. Il est vrai cependant que Jean-François Lyotard avait insisté sur l'importance de relativiser la portée d'application de leurs analyses :

Le scénario de l'informatisation des sociétés les plus développées permet de mettre en lumière, au risque même de grossir excessivement, certains aspects de la transformation du savoir et de ses effets sur la puissance publique et sur les institutions civiles, effets qui resteraient peu perceptibles sous d'autres perspectives. Il ne faut donc pas lui accorder une valeur prévisionnelle par rapport à la réalité, mais stratégique à la question posée²⁰.

Pourtant, vingt-cinq ans plus tard, à l'heure de la « révolution numérique », sous la pression de Google, les esprits se concentrent exclusivement sur la question de l'archive et de la transmission du savoir, corroborant ainsi les hypothèses de Jean-François Lyotard :

Comme les États-nations se sont battus pour maîtriser des territoires, puis pour maîtriser la disposition et l'exploitation des matières premières, il est pensable qu'ils se battent à l'avenir pour maîtriser des informations. Ainsi se trouve ouvert un nouveau champ pour les stratégies industrielles et commerciales et pour les stratégies militaires et politiques²¹.

Comment ne pas penser à la première passe d'arme publique en France entre Google et le président de la Bibliothèque nationale de France, Jean-Noël Jeanneney :

Voici que s'affirme le risque d'une domination écrasante de l'Amérique dans la définition de l'idée que les prochaines générations se feront du monde. [...] Toute entreprise de ce genre implique [...] des choix drastiques, parmi l'immensité du possible. Les bibliothèques qui vont se lancer dans cette entreprise sont certes généreusement ouvertes à la civilisation et aux œuvres des autres pays. Il n'empêche : les critères de choix seront puissamment marqués [...] par le regard qui est celui des Anglo-Saxons, avec ses couleurs spécifiques par rapport à la diversité des civilisations²².

20. Jean-François Lyotard, *La Condition postmoderne*, Paris, Les Éditions de Minuit, 1979, p. 17-18.

21. *Ibid.*, p. 15.

22. Jean-Noël Jeanneney, « Quand Google défie l'Europe », *Le Monde*, 23 janvier 2005.

Deux mois après la tribune du président de la Bibliothèque nationale de France dans *Le Monde*, ce dernier a été reçu à l'Élysée en compagnie du ministre de la Culture et de la Communication afin de mettre en place un projet de bibliothèque virtuelle impliquant l'ensemble des pays européens. Mais retenons un point essentiel du scénario de l'informatisation de la société : « Avec la télématique, l'archivage change de dimension et de nature. [...] La constitution des banques de données va être à l'origine d'une restructuration rapide des connaissances suivant des contours difficiles à définir. »

Cette remarque vaut également pour Archive.org, mais en prenant la voie ouverte par Pierre-Damien Huyghe, on peut interroger différemment l'entreprise de Brewster Kahle. En effet, quel type d'expérience le site Archive.org et sa fabuleuse application la Wayback Machine proposent-ils ? Quelle expérience du monde nous invitent-ils à découvrir ? Archive.org apparaît comme une entreprise singulière de collectage d'informations qui ne manifeste aucune velléité hégémonique sur l'organisation et la transmission du savoir. Quel « goût de l'archive » et conscience des données de notre époque nous procure-t-elle ?

Le gage de notre engagement

Pour reprendre la terminologie du sociologue Laurent Thévenot²³, l'expérience du site Archive.org manifeste une manière d'être très ténue dans la sensibilité contemporaine, qui concerne notre engagement dans le monde contemporain, notre manière d'être attaché à des objets, à des données. Elles sont le produit de la manipulation d'objets incitatifs qui nous accompagnent dans notre vie quotidienne. Ces données sont le gage de notre rapport au monde, elles nous permettent de le saisir comme tel, d'avoir confiance. Archive.org permet de faire l'expérience de cette résistance psychologique difficilement explicable qui fait qu'on ne veut pas reconnaître que ces données soient faussement durables. En effet, depuis la fin des années soixante-dix, avec la révolution de la micro-informatique et le marché de masse de l'industrie des médias, de ce que Patrice Flichy nomme « l'industrie de l'imaginaire », on se trouve confronté à une vague de questions existentielles provoquées par la fréquentation des biens d'équipements audiovisuels et informatiques et le produit de leur manipulation. Comment expliquer notre attachement à ces appareils, à ces machines à communiquer (ordinateurs, magnétoscopes, caméra vidéo) ? Comment gérer chez soi un parc de machines le plus souvent en état de marche mais obsolètes ?

23. Je fais référence à une intervention de Laurent Thévenot dans le cadre de l'émission radiophonique *La Suite dans les idées* animée par Sylvain Bourmeau et diffusée le 21 février 2006 sur France Culture.

Comment profiter d'un stock de supports riches en données difficilement consultables, laissées à l'abandon, à l'image d'une friche stérile où rien ne pousse? Comment faire avec tous ces objets et leurs données, présentés comme durables par les arguments de vente, et que l'on voudrait comme tels? Que faire de ce délaissé sans avenir²⁴?

Ces questions terriblement banales et contemporaines témoignent de l'expérience du faussement durable qui se retrouve dans la navigation proposée par Archive.org. Le site se présente comme un site archéologique des sites web. Il affiche sans aucune rigueur scientifique les différentes évolutions d'un site web capturées de manière aléatoire par les robots œuvrant pour la Wayback Machine. C'est une invitation à la déambulation au sein de données, sans importance ou marquantes, sur lesquelles s'est fixée notre attention liant parfois notre conscience personnelle à la conscience collective du temps²⁵. Archive.org propose aux internautes une expérience similaire à celle du visiteur d'un site archéologique qui découvrirait non pas des ruines, mais une accumulation de vestiges de temps différents.

Sous son allure de machine de guerre hypersophistiquée qui aurait pu être l'invention de William Gibson ou de Bruce Sterling, la Wayback Machine transforme l'entreprise de collectage en ramassage de pages mortes, donnant à certains l'impression d'avoir affaire à une bien piètre décharge du Web, laissant de côté tous les mystères du Web invisible, soit 95 % du trafic réel de l'internet. Il ne faut donc pas compter sur Archive.org pour faire surgir à travers ces pages zombies les grandes heures du cyberpunk. Il existe certes sur le Web des travaux liés à cet imaginaire mais ils sont le plus souvent présentés de manière très innocente, ni critique ni radicalement utopique dans leur approche. Archive.org est né à la fin du cyberpunk et l'esprit artistique qui domine aujourd'hui le Web est celui de la génération des années quatre-vingt-dix, c'est-à-dire celle qui a grandi dans les années quatre-vingt²⁶. Cette génération s'applique à déjouer l'apparente transparence du Web à l'image d'Alex Galloway, jeune professeur en écologie des médias à l'université de New York et membre fondateur du Radical Software Group, un groupe d'artistes aux contours informels dont le nom fait directement référence à une revue historique de la « *digital culture* ». Avec Golan Levin et Cory Arcangel, Alex Galloway a attiré l'attention depuis 2002 par le développement du logiciel Carnivore Personal Edition, un logiciel inspiré

24. Je fais référence à la notion de « délaissé » développée dans Gilles Clément, *Manifeste du tiers paysage*, Paris, Éditions Sujet-Objet, 2004.

25. Voir Bernard Stiegler, « Les enjeux de la numérisation des objets temporels » in Franck Beau, Philippe Dubois, Gérard Leblanc (eds.), *Cinéma et dernières technologies*, Paris/Bruxelles, INA/De Boeck Université, 1998.

26. Voir à ce sujet Bruce Sterling, « Cyberpunk in the Nineties », <http://www.streettech/bcp/BCPtext/Manifestos/CPInThe90s.html>, consulté le 24 juillet 2006. La page n'est plus accessible.

du programme éponyme utilisé par le FBI au lendemain des attentats du 11 Septembre : un superlogiciel dit « renifleur » capable de scanner le réseau internet et notamment le trafic des mails afin de connaître la moindre étape des pérégrinations d'un internaute sur le Web...

Le logiciel de Galloway n'est pas un *crack* (une version piratée) du programme inventé par le FBI, ni une émulation. C'est une ré-interprétation artistique, explique David-Olivier Lartigaud²⁷. Galloway et les autres membres du Radical Software Group ne sont pas des *hackers*. Leur idéologie n'est pas héritée du « pas de portes fermées sur le réseau » mais vient plutôt de mouvements artistiques comme l'art conceptuel ou le Pop Art.

L'objectif du Radical Software Group est d'apporter à la lumière du grand public le principe de Carnivore. Aujourd'hui n'importe quel internaute peut télécharger le logiciel afin d'analyser et de diagnostiquer son trafic, la version Carnivore PE fonctionnant sur n'importe quel système (Windows, Mac et Linux).

De cette manière Alex Galloway prend place dans le débat public, mais pas à la manière d'un « hacktivisme » un peu cliché, souligne David-Olivier Lartigaud. Avec Carnivore PE, il tente de casser le principe manichéen héros/terroristes qui est à l'œuvre dans l'opinion publique américaine depuis les attentats. Au lieu de faire de l'espionnage avec ce genre de logiciel, il en montre le potentiel créatif.

Durable ?

À côté de Carnivore et de son double artistique, la Wayback Machine paraît appartenir à un autre monde, à un temps déjà très dépassé du Web, obsolète. Elle a toutefois un effet attachant car dans son fonctionnement, elle fait penser au bloc-notes magique analysé par Sigmund Freud, cet appareil mnésique qui a « une capacité indéfinie de recevoir des perceptions toujours nouvelles et pourtant il en fournit des traces mnésiques durables, même si elles ne sont pas inaltérables²⁸ ». Cette étude a nourri la théorie du cinéma dans les années soixante-dix, notamment la pensée de Thierry Kuntzel qui voyait en la vidéo un nouveau bloc-notes magique capable de produire des traces, de les conserver sans que celles-ci ne viennent perturber l'apparition

27. David-Olivier Lartigaud est responsable de la ligne de recherche « Art & programmation » conduite par la DAP et le Laboratoire des arts et des médias de l'université Paris 1. Les propos sont extraits de l'article « Radical Software Group: FBI vs RSG » publié dans *Les Inrockuptibles*, octobre 2004.

28. Sigmund Freud, « Note sur le "Bloc-notes magique" » (1924/1925), trad. J. Laplanche et J.-B. Pontalis, *Résultats, idées, problèmes II*, Paris, Presses universitaires de France, 1985.

de nouvelles traces²⁹. Thierry Kuntzel s'intéressait au processus d'animation et de réanimation des formes et à leur persistance, mais ce qui habite fondamentalement ce bloc-notes magique, c'est sa capacité à rendre durable quelque chose qui en apparence doit disparaître. Le durable apparaît comme une notion fondamentale qui structure la pensée de Sigmund Freud dans la description qu'il fait par analogie de l'appareil psychique. Par durable, on entend aussi une interprétation plus matérielle, liée aux objets qui nous entourent et à leur obsolescence programmée. Pour Pierre-Damien Huyghe, ce « durable »-là est bien loin d'être une obsession, une garantie qui nous permettrait de s'engager dans le monde; il est un leurre, un piège : « Le durable est devenu une terre d'ancrage pour les illusions. Se le proposer sérieusement comme perspective, c'est se condamner à une féroce guerre contre les forces économiques et la puissance de progrès lovée au cœur de ces forces³⁰ », car « l'expérience moderne est faite de cet effacement du bien durable, effacement qui peut se conjuguer avec l'obsolescence programmée des marchandises³¹. » En d'autres termes, notre rapport aux choses, aux données sous l'angle du durable est une impasse.

Archive.org est aujourd'hui un objet sympathique, un lapsus de la logique économique et industrielle qui anime les grands opérateurs informatiques comme Microsoft, IBM, Apple, Google et les producteurs d'appareils de prises de vues et d'enregistrement tels que Sony, JVC, Canon, Samsung et Ericsson. Alors que tous les esprits se concentrent sur l'archivage sous l'angle de la gestion des savoirs, Archive.org produit un effet dissonant car la Wayback Machine offre une expérience imparfaite. Si le fait de retourner en arrière, d'extraire de la mémoire du Web les différents états d'un site web, impressionne, Archive.org ne prolonge aucune sensation : rien n'aspire à griser l'utilisateur comme ce pourrait être le cas si un émulateur venait animer la navigation, la rendre plus attractive, la matérialisant en trois dimensions comme dans un film de science fiction – comme dans *Tron* –, à défaut d'avoir chez soi la technologie de Tom Cruise utilisée dans *Minority Report* lui permettant de fouiller dans les limbes d'une pythie futuriste. L'expérience se réduit à l'attente de l'actualisation de la page, à la découverte des restes de données – un texte, une image, des liens inactifs –, parfois au vide. L'internaute fait l'expérience décevante de la présence perdue de ce qui est soi-disant durable, là où il ne trouve qu'une empreinte, une trace. Cette impression de promenade dans un paysage « peu bavard mais qui bredouille » contraste avec l'ivresse de la pratique tout azimut du

29. Thierry Kuntzel, *Tittle TK*, Paris/Nantes, Éditions Anarchive/Musée des Beaux-Arts de Nantes, 2006.

30. Pierre-Damien Huyghe, *Faire place*, *op. cit.*, p. 21.

31. *Ibid.*, p. 12.

surf, de la profusion d'informations, du sens affûté de la lecture d'écran³², de l'effet d'immédiateté du Web. Pour autant, la déambulation n'est pas le terme avoué de l'expérience suggérée par l'interface du site qui induit une approche plus scrupuleuse, plus méthodique, plus attentive : un cheminement à travers les traces, les empreintes, correspondant davantage aux attentes d'un esprit scientifique.

Le salut de cette entreprise de résurrection anarchique du Web provient de manière insolite du côté de la recherche, et plus précisément de la recherche en sciences humaines et sociales. Archive.org devient un outil de travail pour la recherche, un outil rudimentaire, presque archaïque, tout aussi efficace qu'une brosse à dents sur un champ de fouilles archéologiques. L'initiative de Brewster Kahle produit un effet inattendu, qui s'impose aujourd'hui par défaut : dès le début il a une approche de la collecte des données du Web, une approche non sectaire des différents types de données à archiver, le seul critère pour faire partie des archives du site étant d'être un site public. À ce jour, il annonce près de 55 milliards de pages archivées³³. Contrairement aux autres entreprises d'archivage et d'indexation du Web, il n'y a pas de mépris pour l'expression qui n'est pas de la culture savante. C'est là un gisement de données précieuses pour les sciences humaines et sociales dont les techniques de travail sont autant discutées que les objets d'étude.

Le ratage involontaire d'Archive.org permet de revenir sur la production d'une culture *a priori* non savante, sur les pratiques ordinaires, les rites et les rituels contemporains suggérés par un environnement d'appareils informatiques et numériques. Du moins Archive.org nous permet d'en saisir les manifestations car la méthode de collecte de données des robots crée de l'oubli, mais un oubli non dommageable, du registre de l'absence, du vide, plutôt que de celui de l'amnésie. Le site Archive.org offre une expérience imparfaite, mais il montre que l'expérience d'un manque d'informations, pour reprendre Walter Benjamin, est plus stimulante que celle d'un simple transfert de données.

32. Voir Franck Ghitalla, Dominique Boullier, Pergia Gkouskou-Giannakou, Laurence Le Douarin et Aurélie Neau, *L'outre-lecture, manipuler, (s')appropriier, interpréter le Web*, Paris, Bibliothèque publique d'information/Centre Pompidou, 2003.

33. La dernière consultation avant publication de l'article date du 22 juillet 2006.