



**HAL**  
open science

# Le Web, du recours ponctuel à la source principale en histoire

Sophie Gebeil

► **To cite this version:**

Sophie Gebeil. Le Web, du recours ponctuel à la source principale en histoire. Atelier du Dépôt Légal du Web de l'INA n° 3 “ Archives et disciplines : à chacun son archive ”, Claude Mussou; Louise Merzeau; Institut National de l'Audiovisuel, Feb 2013, Paris, France. hal-04876494

**HAL Id: hal-04876494**

**<https://amu.hal.science/hal-04876494v1>**

Submitted on 9 Jan 2025

**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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License

# Le Web, du recours ponctuel à la source principale en histoire

---

L'explosion du Web<sup>1</sup> à l'échelle mondiale dans les années 1990 a démocratisé l'utilisation du réseau Internet<sup>2</sup>, jusqu'alors réservé aux universitaires américains. Devenu en quelques années un média incontournable, Internet constitue un bouleversement dans les modalités d'accès au savoir et à la connaissance. Les historiens ont intégré Internet dans leurs tâches quotidiennes, que ce soit dans le cadre de communications interpersonnelles (messagerie électronique), événementielles (sites de diffusion de colloques) ou de publication en ligne. Les premiers jalons d'une réflexion ont été posés par des « pionniers » tels que Rolando Minuti<sup>3</sup>, Serge Noiret<sup>4</sup>, Philippe Rygiel, Mayalen Zubillaga<sup>5</sup> et Marin Dacos<sup>6</sup>. Dans la diversité des approches, tous soulignent la même défiance des historiens à l'égard du Web, liée en grande partie à l'instabilité et à la volatilité du réseau. A la lumière de ces travaux et d'une pratique personnelle, l'enjeu est ici de comprendre comment le Web, utilisé au départ comme source ponctuelle ou espace de publication, peut constituer une source principale pour l'historien du temps présent.

## Utiliser le Web sans le dépôt légal

### La Toile comme source complémentaire

Nombreux sont les historiens et les étudiants en Histoire qui sollicitent le Web comme source complémentaire dans leurs recherches sans pour autant placer la Toile au cœur de

---

<sup>1</sup> Le Web est ici défini comme une application d'Internet dans lequel des milliards de pages cohérentes (protocole standard tel qu'http) sont reliées entre elles par le principe de l'hypertexte, accessibles via un navigateur (Mozilla Firefox ou Internet explorer) et une adresse URL (de type [www.exemple.fr](http://www.exemple.fr)). Le Web ou World Wide Web ou www ou 3w est un système mis au point par Tim-Berners Lee et Claude Caillot (CERN de Genève) au début des années 1990. Le Web a permis la démocratisation d'Internet.

<sup>2</sup> Internet désigne ici le réseau mondial dans lequel les ordinateurs (ou autre support numériques) sont interconnectés et qui permet une large circulation de flux d'information. Internet comprend au moins le Web, le courrier électronique, la messagerie instantanée et Usenet (forums ou listes de discussion).

<sup>3</sup> Minuti Rolando, *Internet et le métier d'historien : réflexions sur les incertitudes d'une mutation*, traduction de l'italien de Nadia Mansouri, PUF, Paris, 2002

<sup>4</sup> Noiret Serge et Rygiel Philippe (dirs.), *Les historiens, leurs revues et internet*, Publibook Université, Paris, 2005

<sup>5</sup> Zubillaga Mayalen, *Internet et les historiens*, DEA d'Histoire contemporaine, sous la direction de Maryline Crivello, Université de Provence, 2004.

<sup>6</sup> Marin Dacos (dir.), *Read/Write Book*, Marseille, Cléo (« Coll. Edition électronique »), 2010, intégralement disponible sur le site du CLEO, consulté le 12.01.12, <<http://cleo.revues.org/128>>

leurs travaux. Personnellement, il m'a été donné l'occasion d'intégrer des sites Web dans mon corpus de sources pour un Master 2 Recherche sur *Le patrimoine audiovisuel en Méditerranée dans le cadre du processus de Barcelone*<sup>7</sup>. L'objectif de ce travail était d'analyser les projets de coopération audiovisuelle mis au point dans le cadre du Partenariat euro-méditerranéen<sup>8</sup>. Le premier d'entre eux était « Capmed, le réseau méditerranéen des archives télévisuelles nationales ». Les quinze partenaires<sup>9</sup> de ce projet souhaitaient restaurer les archives des télévisions nationales, menacées de disparition du fait des mauvaises conditions de conservation. « Capmed » s'est déroulé de 2000 à 2005 et a abouti à la sauvegarde de plus de 3000 heures d'archives télévisuelles. Le second projet étudié a commencé en 2009 et demeure en cours d'achèvement. Il s'intitule « Med-Mem, partager les mémoires audiovisuelles de la Méditerranée » et réunit seize institutions<sup>10</sup>. L'objectif de « Med-Mem » est la restauration, la numérisation et la diffusion des archives télévisuelles via un site Internet sur le modèle des fresques interactives de l'INA (*Jalons pour l'histoire du temps présent*<sup>11</sup>, *Repères méditerranéens*<sup>12</sup>). Dans l'analyse de ces projets, deux types de sites Internet ont été utilisés comme sources complémentaires : ceux des institutions partenaires et ceux des projets eux-mêmes (en particulier celui de « Capmed » puisque celui de « Med-Mem » n'a pas encore vu le jour). Concernant les sites institutionnels, une attention particulière a été portée aux programmes européens de financement des projets (Euromed Audiovisuel<sup>13</sup> pour « Capmed » et Euromed Heritage<sup>14</sup> pour « Med-Mem »), avec en toile de fond, une analyse des pages du site de la Commission européenne consacrées au partenariat euro-méditerranéen. Pour l'U.E., le Web sert ici principalement

---

<sup>7</sup> Gebeil Sophie, *Le patrimoine audiovisuel en Méditerranée dans le cadre du PEM*, Master 2 Recherche d'Histoire contemporaine sous la direction de Maryline Crivello, Université de Provence, 2010

<sup>8</sup> Les pays membres du partenariat réunis à Barcelone en 1995 étaient les quinze pays membres de l'U.E. et dix partenaires méditerranéens (Algérie, Autorité palestinienne, Chypre, Egypte, Israël, Jordanie, Liban, Malte, Maroc, Syrie, Tunisie, Turquie).

<sup>9</sup> L'Italie (RAI), la Grèce (ERT), la Turquie (TRT), Chypre (CyBC), la Jordanie (JRTV), la Palestine (PBC), l'Égypte (ERTU), Malte (PBC), la Tunisie (ERTT), l'Algérie (ENTV), le Maroc (TVM - 2M) et la France (INA). MMSH et IOM.

<sup>10</sup> La France (INA), l'Italie (RAI et COPEAM), l'Algérie (ENTV), le Maroc (SNRT), la Jordanie (JRTV), la Grèce (ERT), l'Égypte (ERTU), et la Palestine (PBC), la Lybie (LBJC), la Croatie (HRT) et Israël (IBA), COPEAM, le CMCA et l'UER., MMSH

<sup>11</sup> Jalons pour l'histoire du temps présent, INA, consulté le 20.02.2012, <<http://www.ina.fr/fresques/jalons/accueil>>

<sup>12</sup> Repères méditerranéens, INA, consulté le 20.02.2012, <<http://www.ina.fr/fresques/reperes-mediterraneens/accueil>>

<sup>13</sup> Euromed Audiovisuel, consulté le 26.02.2012, <<http://euromedaudiovisuel.net/2012/>>

<sup>14</sup> Euromed Heritage, consulté le 26.02.2012, <<http://www.euromedheritage.net/>>

d'outil de communication autour des actions culturelles menées dans le cadre du partenariat. Le recours à ces sites a permis l'analyse du discours sur la Méditerranée émanant des institutions de l'U.E. investies dans le partenariat. Berceau des civilisations et des religions monothéistes, le foisonnement culturel y est valorisé dans sa diversité. Dans ce discours, le patrimoine est mis au service du dialogue entre les cultures et du rapprochement entre les peuples du bassin méditerranéen. Concernant les sites des projets, l'analyse s'est principalement fondée sur celui de « Capmed », [www.cap-med.net](http://www.cap-med.net), dont l'ambition était de présenter des extraits d'archives sauvegardées et des dossiers scientifiques rédigés par les chercheurs des institutions partenaires du projet (MMSH et IOE). Du point de vue formel, le site reprenait les codes du partenariat : un site trilingue (arabe, anglais et français) et une dimension humaniste valorisée (photographie d'un visage en gros plan présent en page d'accueil). Le site témoignait également des possibilités qu'offrait le Web pour valoriser et rendre accessible un patrimoine méconnu. Sur le fond, l'analyse de [www.cap-med.net](http://www.cap-med.net) a permis de mieux cerner comment, par l'absence de traitement des aspects politiques conflictuels, le projet véhiculait une représentation culturaliste et consensuelle de la Méditerranée. Dans le cadre de ce travail, l'étude des sites répertoriés s'est fondée sur les méthodes de l'analyse critique traditionnelle applicable à tout document questionné par la recherche. Les sites Web complétaient donc le corpus au même titre que les sources imprimées ou audiovisuelles, sans faire l'objet d'une réflexion particulière sur la spécificité des documents en ligne. Les contenus utilisés ont été sauvegardés par un processus d'archivage personnel sous forme de captures d'écran<sup>15</sup> et reproduits en annexe du mémoire. Dans le même temps, l'adresse URL des sites interrogés, accompagnée de la date de consultation, était indiquée au lecteur en notes de bas de page. L'archivage des pages n'avait alors pas été l'objet de questionnements particuliers. Tout au plus, le site [www.archive.org](http://www.archive.org) a été utilisé pour retrouver les pages du site [www.cap-med.net](http://www.cap-med.net) supprimées en 2008, mais sans véritablement étudier l'association *Internet Archive* à l'origine de ce site d'archives en ligne.

---

<sup>15</sup> Une capture d'écran est un procédé technique permettant la sauvegarde des contenus affichés sur l'écran sous la forme d'un fichier image (.jpg ou .btm par exemple) grâce à un logiciel.

### **La Toile comme espace de publication**

Ce travail a également été l'occasion de participer aux notices contextuelles des vidéos du projet « Je me souviens... de la Méditerranée » mis au point par Maryline Crivello (TELEMME, MMSH, Aix-Marseille Université) en collaboration avec Jacques Sapiéga (SATIS, Aix-Marseille Université). Il s'agit d'un projet de rencontre entre historiens et étudiants en sciences audiovisuelles autour du témoignage, de la mémoire et de l'histoire sur le principe des « Je me souviens.. » de Georges Perec. Intégrées dans « Med-Mem », les vidéos de cette première série portaient sur des Marseillais d'origine libanaise. Par la rédaction de notices historiques permettant aux internautes de mieux s'approprier les films mis en ligne, Internet est ici utilisé en tant qu'espace de diffusion du savoir historique.

La publication en ligne<sup>16</sup> est sans nul doute l'un des aspects les plus traités concernant les relations entre Internet et les historiens. Malgré les réticences de la profession, le nombre de publications scientifiques augmentent et une réflexion est menée sur l'impact de l'écriture en ligne sur le récit historique. Par exemple, les blogs de recherche<sup>17</sup> témoignent de l'évolution des pratiques mêmes d'écriture de l'histoire du fait des publications scientifiques en ligne.

Dans le cadre de ce travail, Internet a fait l'objet d'une utilisation relativement classique du réseau : tantôt espace de publication, tantôt source ponctuelle venant approfondir un corpus de sources imprimées. La loi DADVSI de 2006, instituant un dépôt légal du Web confié à la BNF et à l'INA constitue de notre point de vue, un tournant dans les possibilités d'usages du Web en histoire et plus largement en sciences humaines et sociales.

### **Le dépôt légal ou la possibilité d'utiliser le Web comme source principale.**

Après quelques ébauches d'études à partir d'Internet, le Doctorat est l'occasion d'approfondir ce travail à travers deux axes. Le premier est l'analyse des représentations dans le cadre de l'ANR ECRIN coordonné par Yves Gastaut (URMIS, Nice) sur *l'image des*

---

<sup>16</sup> Dacos Marin et Mounier Pierre, *L'édition électronique*, La Découverte Paris, 2010

<sup>17</sup> A titre d'exemple, le blog de l'ANR « Imasud. Suds imaginaires, imaginaires des Suds » hébergé par [www.hypotheses.org](http://www.hypotheses.org) (plateforme de carnets de recherche en Sciences humaines) du portail de ressources électroniques en SHS [www.openedition.org](http://www.openedition.org), consulté le 28.02.12, <<http://dakirat.hypotheses.org/>>

Sophie Gebeil, 17 février 2012, INA, Paris. Intervention lors de la séance des ateliers du dépôt légal du web, org. Claude Mussou et Louise Merzeau

« Arabes » dans les médias français de 1962 à nos jours<sup>18</sup>. Le second axe est la considération du Web comme espace médiatique singulier et ainsi l'analyse d'un corpus de sites tiré des archives du Web audiovisuel français. Cet aspect qui sera ici développé, est directement lié à l'existence du dépôt légal du Web à l'INA. L'usage du Web comme source principale est à l'origine de questionnements méthodologiques mais ouvre dans le même temps des perspectives d'étude pour l'historien du temps présent.

### **L'historien face aux déroutantes traces du Web**

Fruit de la révolution numérique, la nature même d'Internet est à l'origine de nombreux questionnements méthodologiques. Média de liens, de traces, dont la mémoire semble mouvante, quelle place peut être accordée à Internet en tant que source historique ? Marc Bloch avait, à partir de l'expression de François Simiand, défini l'histoire comme une « connaissance par traces ». Dans *Apologie pour l'histoire*, il s'interroge sur la notion de documents historiques :

« Qu'entendons-nous en effet par *documenta* sinon une « trace », c'est-à-dire la marque, perceptible aux sens, qu'a laissé un phénomène lui-même impossible à saisir<sup>19</sup> ? »

L'historien s'appuie donc sur l'ensemble de traces héritées du passé et les constitue en documents historiques à travers son questionnement scientifique. Comme le rappelle Alexandre Serres dans ses réflexions sur la trace en sciences humaines<sup>20</sup>, tout outil de communication tel qu'Internet produit et laisse des traces, écrites ou non écrites. Quelles sont donc les traces du passé produites par le Web ?

Des traces numériques avant tout, malgré la présence de publications en ligne ayant une existence sur support imprimé, ces traces sont bien souvent exclusivement digitales. Avec l'information numérique, le document perd sa matérialité et peut se réduire, du point de

---

<sup>18</sup>L'ANR ECRIN regroupe des partenaires issus des laboratoires de recherche URMIS, TELEMME, LCP et ETIC autour des questions de stéréotypes liés aux origines. Présentation de l'ANR sur le site de l'URMIS, Université de Nice Sophia Antipolis, consulté le 27.02.12, <<http://www.unice.fr/urmis/spip.php?article546>>. L'ANR ECRIN regroupe les laboratoires de recherche URMIS, TELEMME, LCP et ETIC.

<sup>19</sup> Bloch Marc, *Apologie pour l'histoire*, P.U.F, Paris, 1949, p. 21

<sup>20</sup> Serres Alexandre, « Quelle(s) problématique(s) de la trace ? », Texte d'une communication prononcée lors du séminaire du CERSIC, 13 décembre 2002, sur la question des traces et des corpus dans les recherches en Sciences de l'Information et de la Communication. Disponible en ligne sur la plateforme d'archives ouvertes en SIC « @sic », consulté le 28.10.11, <[http://archivesic.ccsd.cnrs.fr/sic\\_00001397/en/](http://archivesic.ccsd.cnrs.fr/sic_00001397/en/)>.

vue technique, à un ensemble d'opérations binaires composées de 0 et de 1 combinées à l'infini. Cependant la dématérialisation du document n'est pas une nouveauté puisque la télévision est également devenue intégralement numérique ce qui n'empêche pas les historiens de l'exploiter en tant que source<sup>21</sup>. Deux principes essentiels, à l'origine des spécificités du réseau, complexifient la notion de trace sur Internet : l'interactivité et l'interconnectivité (ou hypertextualité). L'interactivité se caractérise par le fait que tout internaute puisse réagir sur le réseau en y laissant sa propre empreinte. De ce fait, de nombreuses pages Web sont modifiées régulièrement par le fondateur du site et les réactions des internautes. Le Web 2.0. a accru les possibilités de publication de contenus avec des outils tels que l'encyclopédie en ligne *Wikipédia*, les plateformes de partage et d'échanges de fichier multimédia (*Youtube* ou *Dailymotion* pour les vidéos, *Flickr* pour les images) ou encore les réseaux sociaux (*Facebook*, *Twitter*). Avec ces nouvelles possibilités, les traces se sont démultipliées et complexifiées. La seconde spécificité d'Internet réside dans l'interconnectivité : toutes les données sont reliées entre elles grâce à un système d'hyperliens qui permet de « naviguer » à l'intérieur d'un site (hyperliens internes) ou entre différents sites (hyperliens externes). En plus des pages Web, l'activation des liens génère aussi des traces « invisibles » aux yeux de l'internaute. Ces traces techniques sont l'ensemble des opérations, essentiellement de copie, nécessaires au bon fonctionnement du Web et dont une partie est parfois mémorisée dans les ordinateurs des usagers (sous forme de *cookies* par exemple).

Face à ces traces du Web, mouvantes, interconnectées, l'historien est confronté à la volatilité des informations publiées sur le réseau, celle-ci étant accrue par l'évolution rapide des techniques de publications en ligne. Le créateur d'un site Web, désireux par exemple d'enrichir son site en proposant des services plus interactifs, devra choisir un hébergeur adapté à ses ambitions. Pour cela, il changera l'adresse<sup>22</sup> de son site. Avec cette

---

<sup>21</sup> Crivello Maryline, «Télévision et histoire», in Christian Delacroix, François Dosse, Patrick Garcia et Nicolas Offenstadt (dir.), *Historiographies, Concepts et débats*, Gallimard, Paris, 2010, p. 330

<sup>22</sup> Les adresses Web sont appelées également adresses URL (Uniform Resource Locator). Système unifié d'identification des pages Web, les URL sont des chaînes de caractères codés qui permettent d'indiquer à un logiciel comment accéder à une information sur le réseau. Leur forme est désormais assez connue et se présente bien souvent dans la forme suivante : [www.exemple.fr](http://www.exemple.fr). Niel Xavier et Roux Dominique, *Les 100 mots de l'Internet*, Que sais-je ?, P.U.F., Paris, 2010, p. 17

modification, tous les liens fondés sur l'ancienne adresse se transformeront parfois en « liens morts ». Le chercheur qui aurait noté l'ancienne adresse URL du site verrait alors apparaître sur l'écran un désagréable message d'erreur au lieu de la page Web sur laquelle il travaillait. La volatilité est telle qu'on estime à cent jours la durée moyenne de vie d'une page Web. Dans ces conditions, il demeure difficile pour un chercheur en sciences humaines, de travailler sur un corpus de sources qui risque de disparaître ou de se perdre dans l'immensité de la Toile. Si pour l'historien, toute empreinte du passé constitue une trace exploitable, elle doit pouvoir être sauvegardée. Elle fera ensuite l'objet d'une analyse et sera réutilisée lors de la mise en récit du discours historique, caractérisé par des notes de bas de page qui permettent au lecteur d'accéder lui-même à la source s'il le désire. Citer une page Web comme source, c'est prendre le risque de citer une trace qui s'est effacée ou n'existe plus, rendant ainsi la vérification impossible pour le lecteur. Cela revient à briser le « contrat de vérité<sup>23</sup> » qui lie l'historien et son lecteur. Ce rapport de confiance se fonde sur la communication et l'accessibilité des sources, preuves de la fiabilité de l'information citée. Ces obstacles expliquent, dans une large mesure, le petit nombre de travaux s'appuyant sur un corpus de sites Internet.

Face à ces difficultés méthodologiques, les historiens qui utilisent des sources en ligne – souvent d'ailleurs en tant que sources complémentaires et non principales<sup>24</sup> – ont proposé une sitographie ainsi que des captures-écrans des pages Web du site en annexe. Une capture-écran est un procédé technique qui permet de « photographier » un instantané d'une page de l'ordinateur à un moment donné. Il s'agit d'un fichier au format image (.jpg ou .btm par exemple) qui ne rend compte ni de l'interactivité ni de l'interconnectivité caractéristiques du Web. Ces procédés de stockage personnel, bien que nécessaires, ne sont pas assez fiables pour faire du Web une trace pérenne et donc une source privilégiée pour l'historien.

---

<sup>23</sup> Ricœur Paul, *Histoire et Vérité*, Paris, Le Seuil, 1955, p. 30.

<sup>24</sup> Atelier du dépôt légal du Web à l'INA, *Un web archive, pour quoi faire ?*, enquête en ligne réalisée auprès des chercheurs, saison 2010-2011, consulté le 12.01.12, <<http://atelier-dlweb.fr/wp-uploads/2011/06/enqu%C3%AAtes.pdf>>

## **La mise en archive, étape fondamentale dans la prise en compte du Web en tant que source historique.**

Les contraintes spécifiques d'Internet ont nécessité la mise en place d'un processus technique innovant permettant la création d'archives du Web. Pour l'INA, le processus technique mis au point en interne associé à une quantité de sites relativement restreinte, ont abouti à une archive particulièrement complète et fine.

La mise en archive, des traces désordonnées et volatiles laissées par le Web, semble constituer une alternative rassurante pour l'historien désireux d'utiliser des sources issues d'Internet. En effet grâce à l'archivage, les traces du Web, intrinsèquement désordonnées et volatiles, sont désormais pérennes. De surcroît, cette conservation est garantie, depuis 2006, par les institutions depositaires du dépôt légal du Web français. Dans ce contexte, le risque de briser le contrat de vérité entre l'historien et son lecteur est caduc, dès lors que le lecteur peut lui-même accéder aux sources via les bornes de consultation du dépôt légal. Sa recherche d'information sera d'ailleurs facilitée à l'INA par un moteur de recherche multicritère. Cette stabilité de l'archive réside à la fois dans la préservation du contenu des pages Web et dans la mise en œuvre des procédés de sauvegarde de l'archive. En effet le dépôt légal garantit également la préservation et rend ainsi possible dans le futur, une éventuelle consultation des sources Web citées dans les travaux, dans un souci de vérification ou d'analyse. Ainsi, le dépôt légal du Web assure la stabilité nécessaire pour travailler sur un média dont la volatilité intrinsèque pose problème. Il convient cependant de mener une réflexion sur la représentativité de l'archive compte tenu des particularités du processus d'archivage.

La mise en archive des traces du Web via un processus technique automatisé pose la question de l'authenticité de l'archive par rapport au Web, document original. Techniquement, l'archive n'est pas sauvegardée par enregistrement d'un flux continu comme cela est le cas pour le dépôt légal de la télévision numérique. Une captation brute des traces aboutirait à une série de photographies instantanées de pages Web figées et non connectées entre elles. Pour reproduire l'interconnectivité (hyperliens) qui caractérise le Web, il a été nécessaire à l'INA de reproduire la mise en réseau des pages. Résultant des étapes techniques de la chaîne de traitement, les archives sont des contenus reconstitués à

partir des traces originales du Web. Il s'agit donc d'une « re-création » d'archive et non d'une sauvegarde de l'original. Le paradoxe résidant dans le fait que la qualité de l'archive croît avec le nombre de modifications apportées. Pour autant, l'authenticité d'un document ne réside pas seulement dans l'absence de manipulation ou de transformation. Un document authentique est un document dont l'exactitude est garantie, ce qui est le cas dans le cadre des archives du dépôt légal. Les manipulations techniques apparaissent ainsi comme des moyens permettant la reconstitution des traces mises en réseau dont l'ambition est de garantir au mieux l'authenticité du contenu.

Travailler sur les archives du Web implique donc d'avoir conscience de la nature du fonds qui est consulté afin de mieux en cerner les limites. Pour cela, il est nécessaire pour l'historien de comprendre les techniques d'archivage. A titre d'exemple, les robots de capture peuvent rater des éléments entre deux passages. Pour un chercheur qui travaille sur les archives du Web, l'absence d'une information ne signifie donc pas qu'il n'existe aucune trace de celle-ci sur le Web mais simplement que la trace laissée n'a pas été suffisamment diffusée, ou l'a été durant un laps de temps trop court, pour que les robots d'indexation ne la sauvegarde. Il ne s'agit pas de prendre prétexte de ses limites pour rejeter tout emploi des archives mais davantage d'intégrer l'ensemble de ces contraintes lors de l'utilisation du dépôt légal du Web. A ce jour, l'enjeu n'est donc pas d'espérer de l'archive une reproduction à l'identique du Web audiovisuel francophone mais davantage de réfléchir aux apports scientifiques et méthodologiques d'un tel outil.

## **Perspectives offertes par le DL Web en Histoire**

Au vu de travaux d'historiens pionniers en la matière, l'étude du Web ouvre de nombreuses perspectives.

### **L'Histoire contemporaine sur le Web italien, les apports d'une expérience précurseur**

Dès les premières années de la décennie 2000, un groupe d'historiens contemporanéistes italiens s'est interrogé sur le type d'histoire présent sur le Web italien entre 2001 et 2002<sup>25</sup>.

---

<sup>25</sup> Criscione A., Noiret S., Spagnolo C., Vitali S. (dir.) : La Storia a(l) tempo di Internet : indagine sui siti italiani di storia contemporanea, (2001-2003), Bologna, Pàtron editore, 2004, cité par Noiret Serge, « Histoire et

Dans ce cadre, plus de 250 sites d'histoire contemporaine traitant de l'histoire militaire, du fascisme, de l'holocauste, des femmes ou encore des années 1970 marquées par les tensions et le terrorisme, ont été analysés. Le résultat rend compte d'une histoire sur le Web dominée par les non-professionnels avec une qualité très inégale de l'information. Les publications d'historiens concernent essentiellement des sujets très précis tandis que les pages rédigées par des amateurs rendent davantage compte d'une vision personnelle et plus traditionnelle de l'histoire avec ses batailles, ses anecdotes et ses héros. Ainsi, ce type de travaux ouvre la voie vers un enrichissement du champ des usages du passé par des groupes organisés, ou par des individus qui trouvent dans le Web un espace privilégié d'expression mais aussi de contact entre passionnés. L'étude de l'histoire sur Internet permet également d'appréhender différemment la place de la mémoire dans les constructions identitaires contemporaines. En effet les moyens techniques dont disposent les éditeurs de sites Web permettent de recréer une vision imaginaire d'une période comme le constate Serge Noiret pour le fascisme : « On obtient ainsi une vision historique du fascisme qui jamais n'exista, mais que le Web permet de recréer de manière virtuelle<sup>26</sup> ». Pour appréhender ces sources nouvelles, une réflexion méthodologique a abouti à l'élaboration de nouveaux critères permettant d'adapter la démarche historique aux spécificités du Web. Dans cet objectif, trois critères novateurs ont été ajoutés dans la démarche analytique : les moyens d'identification des sites (mots clés et moteurs de recherches sollicités), l'architecture du site Web (interconnectivité fondée sur les liens hypertextuels) et enfin les outils techniques permettant l'interactivité (forum, espaces de publication réservés aux membres par exemple). Ces critères permettent de prendre en compte les spécificités du réseau tout en s'ajoutant aux critères d'expertise scientifique traditionnels fondés sur le jugement critique du chercheur. Ces travaux novateurs, permettent d'entrevoir les formidables possibilités offertes par la toile en tant qu' « immense dépôt de sources nouvelles<sup>27</sup> », tout en notant la nécessaire adaptation des méthodes dans l'analyse du document en ligne. Dix ans plus tard, la pérennité de ces travaux précurseurs, fondés sur des sites qui auront inévitablement

---

mémoire dans la toile contemporaine italienne », in Noiret S. et Rygiel Philippe, *Les historiens, leurs revues et Internet*, Paris, Publibook, 2005, p. 25-79

<sup>26</sup> Noiret Serge, « Histoire et mémoire dans la toile contemporaine italienne », in Noiret S. et Rygiel Philippe, *Les historiens, leurs revues et Internet*, Paris, Publibook, 2005, p. 51

<sup>27</sup> *Ibid.*

évolué ou qui auront disparu aujourd'hui, se pose d'autant plus qu'il n'existe pas de dépôt légal du Web en Italie. Dans le même temps cette incertitude sur l'évolution des sites renforce le caractère indispensable de l'archive pour l'historien désireux de travailler sur le Web.

### **Les représentations de «l'Arabe » sur le Web. Questionnements, hypothèses et potentialités.**

Dans le cadre de l'ANR ECRIN citée précédemment, nous nous proposons d'interroger le stéréotype de l'« Arabe » et les représentations dont il fait l'objet, sur le Web audiovisuel français. L'image de l'immigré maghrébin et de ses descendants dans la production médiatique et culturelle a déjà fait l'objet d'analyses poussées. Tous font le constat d'une inégalité de traitement, souvent défavorable mais parfois positive sous forme de contre-stéréotype. Interroger le Web audiovisuel en tant qu'écran médiatique singulier est porteur de nombreuses questions. Les représentations du Web sont-elles identiques à celles de la télévision ou du cinéma ? Quelles sont les formes particulières que prend le discours diffusé sur le Web du fait des outils d'interactivité et de publication ? Comment les professionnels du secteur audiovisuel utilisent-ils le Web dans leurs stratégies de communication ? Autant d'hypothèses dont la vérification passe par l'étude du dépôt légal du Web à l'INA.

De plus, le travail de Cyril Domanico sur le film *Indigènes*<sup>28</sup>, montre que le Web peut également permettre de mieux cerner la réception des œuvres cinématographiques et audiovisuelles. A travers les réactions des internautes mais aussi les détournements de l'objet original sous formes de vidéo-montages plus ou moins critiques à l'égard de l'œuvre, l'étude du Web devrait permettre de mieux cerner la réactivité des internautes face aux représentations médiatiques et culturelles. A travers ces possibilités, l'hypothèse d'une plus grande appropriation du discours médiatique sur le Web sur une thématique aussi sensible que la représentation d'un stéréotype ethnicisant devra être examinée. Pour répondre à ces questions, le corpus de recherche sera constitué principalement de sites Web. Cela implique

---

<sup>28</sup> Domanico Cyril, *Quand Indigènes fabrique de la mémoire : Pratiques, enjeux et réappropriations mémorielles de notre temps*, Master 2 d'Histoire des Mondes Méditerranéens, Européens et Africains, sous la direction de Maryline Crivello, Université de Provence, 2011.

Sophie Gebeil, 17 février 2012, INA, Paris. Intervention lors de la séance des ateliers du dépôt légal du web, org. Claude Mussou et Louise Merzeau

un questionnement méthodologiques en particulier sur les relations entre les moteurs de recherche du Web actuel et la recherche sur le logiciel de consultation de l'archive mis au point à l'INA. Autant de portes ouvertes que l'historien peut désormais ouvrir grâce aux garanties d'authenticité et de stabilité apportées par le dépôt légal.