

RESUMES ET PRESENTATIONS PERSONNELLES

16^{es} Journées des Archives
Pérenniser l'éphémère. Archivage et médias sociaux

Louvain-la-Neuve, 9 et 10 mai 2016



Sources primaires en flux : collecter, analyser, archiver, pérenniser. Enseignements d'un projet de recherche sur les échos du Centenaire de la Première Guerre mondiale sur Twitter

Frédéric CLAVERT, Université de Lausanne, faculté des Lettres et sciences humaines, section d'histoire

Présentation

Docteur en histoire contemporaine, Frédéric CLAVERT a étudié les sciences politiques et l'histoire contemporaine à Strasbourg et à Leeds. Après une thèse et des travaux d'histoire monétaire et d'histoire des relations internationales, ses recherches s'orientent aujourd'hui sur les sources de l'historien.ne à l'ère numérique. Après cinq ans comme chercheur (Histoire de l'intégration européenne et Humanités numériques) au Centre virtuel de la connaissance sur l'Europe (Luxembourg), et deux ans comme ingénieur de recherche au LabEx Ecrire une Histoire Nouvelle de l'Europe (Paris Sorbonne), il est désormais maître-assistant en section d'histoire de l'Université de Lausanne, où il s'attèle à collecter et analyser les échos des commémorations du Centenaire de la Première Guerre mondiale sur Twitter.

Résumé

Depuis avril 2014, près de deux millions de tweets contenant des hashtags liés au Centenaire de la Première Guerre mondiale ont été émis sur le réseau socio-numérique en ligne *Twitter*. Stockés dans une base de données, ces tweets peuvent être analysés par l'historien pour, entre autres, comprendre les relations entre mémoire et histoire, pour prendre conscience du rôle des politiques de communication des institutions mémorielles, pour scruter les usages du passé par les citoyens, pour comparer des aires linguistiques et leurs mémoires respectives.

Outre les éléments juridiques qui ne seront pas ici abordés, exploiter une telle base de données, de telles sources primaires, posent des questions méthodologiques très sérieuses, que nous tenterons de passer en revue, et qui sont toutes reliées à une question centrale : comment traiter une source primaire dont la principale caractéristique principale est d'être un flux constant d'informations ?

Les notions relativement récentes de « lecture distante » ou de « microscope » permettent pour partie de répondre à ces enjeux méthodologiques en insistant sur l'usage des sciences informatiques dans les sciences humaines et sociales.

Les relations entre historien.ne.s, archivistes et bibliothécaires ont également toutes les chances d'en être renouvelées : l'historien.ne seul.e ne peut faire face aux sources primaires en flux, quel que soit son savoir-faire en sciences informatiques. En conséquence, la masse d'informations qu'engendrent ces sources en flux va-t-elle faire évoluer le rôle de l'historien.ne dans la chaîne de production de l'histoire ? Qu'attend l'historien.ne des centres d'archives (et bibliothèques) dans le cadre d'une mise en données de plus en plus étendue de la pratique de l'histoire ?

Enfin, étudier les réseaux socio-numériques en ligne nécessite une grande connaissance de ces derniers. Leur pratique par celui ou celle qui les analyse est souvent un préalable, ce qui force l'historien.ne à engager la discussion avec les historiens non-professionnels, souvent citoyens soucieux de mémoire de leurs ancêtres.

Du point de vue juridique, qui est propriétaire, qui est responsable des contenus véhiculés sur les réseaux sociaux ? Quelles limites pose le respect de la vie privée ?

Cécile de TERWANGNE, Université de Namur, CRIDS et
Séverine DUSOLLIER, Université de Namur et Sciences Po, Paris

Présentation personnelle

Cécile de TERWANGNE est professeur à la Faculté de Droit de l'Université de Namur, elle dispense notamment les cours de « Vie privée et TIC » et de « Droits de l'homme et société de l'information ». Elle dirige le Master de spécialisation en Droit des Technologies de l'Information et de la Communication (D.T.I.C.). Elle est également directrice de recherche au sein de l'Unité « Libertés et Société de l'information » du Centre de Recherches Information, Droit et Société (CRIDS) de l'Université de Namur. Elle a participé à de nombreuses recherches européennes et nationales et a publié de nombreux articles et ouvrages sur la protection de la vie privée, la protection des données, le droit à l'information, la transparence de l'administration, l'e.government, la réutilisation des données publiques.

Séverine DUSOLLIER est docteur en droit et Professeure à SciencesPo Paris et à l'Université de Namur. Elle enseigne également le droit européen de la propriété intellectuelle au CEIPI (Université de Strasbourg). Elle fut directrice du CRIDS (Centre de Recherches Information, Droit et Société) de 2010 à 2014. Elle a notamment effectué des recherches en droit d'auteur et en droits intellectuels pour l'OMPI, l'UNESCO, le Conseil de l'Europe et la Commission européenne. Elle fut chercheuse associée à la Faculté de Droit de l'Université de Berkeley (États-Unis) en 2001, Jean Monnet Fellow à l'Institut Universitaire Européen de Florence en 2005-2006 et Visiting Fellow à l'Université de Columbia. Ses recherches actuelles portent sur les biens communs, le domaine public, les exceptions au droit d'auteur et la protection du droit d'auteur sur Internet.

Elle est membre du Conseil belge de la Propriété Intellectuelle et directrice de la revue juridique belge, Revue du Droit des Technologies de l'Information. Elle est titulaire d'une bourse de recherche ERC sur la notion de droits inclusifs en propriété (2014-2019). Elle est membre du Comité exécutif de l'ATRIP (International Association of Teachers and Researchers in IP) et membre fondateur de l'European Copyright Society.

Les publications de Séverine Dusollier sont disponibles sur :

Bepress page : http://works.bepress.com/severine_dusollier/

SSRN page : http://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=632493

Résumé

Si les contenus postés sur Internet sont généralement en libre accès, cela ne signifie pas pour autant que leur utilisation est libre de droits. Droit d'auteur et protection de la vie privée constituent notamment deux obstacles légaux à l'archivage et à la réutilisation des contenus apparaissant sur les médias sociaux. D'une part, tout texte, image, film ou autre contenu posté sur le net est susceptible de faire l'objet d'une protection par le droit d'auteur. À quelles conditions et dans ce cas qui en est l'auteur ? Peut-on dans ce cas procéder à leur archivage sans l'accord de leur auteur et dans quelles limites ? Ou faut-il demander une autorisation aux auteurs de ceux-ci et comment faire en pratique ?

Et quelles limites à la reprise et à l'archivage pose le respect de la vie privée ? Qui peut archiver et communiquer ces contenus ? Quelles sont les implications du droit du public à l'information et du droit à l'oubli sur la gestion des réseaux sociaux ?

En définitive ces protections légales rendent-elles possible la conservation de la mémoire d'Internet ?

Comment travailler la mémoire sur Twitter ? Le cas de la Grande Noirceur et Révolution tranquille 2.0 au Québec

Alexandre TURGEON, Université d'Ottawa

Présentation

Alexandre TURGEON vient de compléter un doctorat en histoire à l'Université Laval. Financées par le Fonds de recherche du Québec – Société et culture, ses recherches doctorales ont porté sur le rôle du caricaturiste Robert La Palme dans la conception et la diffusion du mythe de la Grande Noirceur duplessiste au Québec. Il s'intéresse plus particulièrement au rôle du caricaturiste dans la Cité et aux mythes et mythistoires du Québec contemporain. Il a dirigé un numéro thématique sur la caricature au Québec pour *Recherches sociographiques* et il en prépare actuellement un autre sur la Grande Noirceur pour *Mens : revue d'histoire intellectuelle et culturelle*. Il est maintenant chercheur postdoctoral FRQSC au Département d'histoire de l'Université d'Ottawa où il travaille sur les liens entre l'histoire, la mémoire collective et Twitter.

Résumé

Tout au long du conflit étudiant québécois de 2012, mieux connu sous le nom de printemps érable, et de la campagne électorale qui suivit, ils ont été plusieurs milliers à évoquer dans les médias sociaux le souvenir de la Grande Noirceur et de la Révolution tranquille. Ces deux mythistoires occupent une place de choix dans l'imaginaire collectif des Québécois. Ils articulent une conception dichotomique du passé où l'année 1960 apparaît comme une fracture entre deux temps et deux mondes. Du 16 mai au 12 septembre 2012, j'ai recueilli plus de 6000 tweets où ces deux mythistoires sont convoqués pour traiter du conflit étudiant ou de la campagne électorale. Ce phénomène, que j'ai appelé la Grande Noirceur et Révolution tranquille 2.0, nous renseigne sur la manière dont les utilisateurs se souviennent du passé et comment ils s'en servent pour donner un sens au présent et au futur, sans qu'ils aient eu à remplir un questionnaire. À cet effet, Twitter est un média privilégié pour étudier ces questions sensibles touchant à la mémoire, aux usages et au rapport au passé. Limités à 140 caractères, les utilisateurs se doivent d'être synthétiques, d'aller à l'essentiel. D'où le recours, dans le cas présent, à ces images fortes de la Grande Noirceur et de la Révolution tranquille. Or, un problème de taille se pose. Comment travailler la mémoire sur Twitter ?

Revenant sur mon propre parcours, je compte répondre à cette question dans le cadre de ce colloque. Pour ce faire, je dégagerai les différentes étapes qui m'ont permis de circonscrire ce phénomène, de ramasser ce corpus, pour ensuite parvenir à l'étudier. Une telle réflexion est importante, car la Grande Noirceur et Révolution tranquille 2.0 ne s'est pas déployée sur Twitter par l'usage de hashtags. L'étude d'un tel phénomène requiert au contraire une connaissance suffisante de ces mythistoires autant que des débats sur Twitter, en cela qu'elles se renseignent l'une l'autre. C'est en procédant de la sorte qu'il est possible d'établir le lexique mémoriel de ce phénomène, tel qu'il est utilisé alors, dans ce contexte. C'est à partir de celui-ci qu'il est possible d'étudier un phénomène tel que la Grande Noirceur et Révolution tranquille 2.0 au Québec. Encore faut-il pouvoir archiver les tweets obtenus grâce à ce lexique. À la lumière des expériences tentées en 2012, les unes fructueuses, les autres un peu moins, j'en tire quelques enseignements qui pourront être utiles aux chercheurs intéressés par ces questions.

Le travail des Archiveilleurs : Une veille participative mise au service de la communauté archivistique

Alexandre GARCIA, Division des archives et de la gestion de l'information du Comité international de la Croix-Rouge

Présentation

Alexandre GARCIA est chef de projet au sein de la Division des archives et de la gestion de l'information du Comité international de la Croix-Rouge, où il conduit la mise en œuvre de plusieurs solutions informatiques (gestion d'archives audiovisuelles, records management), et mène des analyses sur les activités de gestion de l'information et d'archivage numérique.

Il est membre du groupe de travail « Records management et archivage électronique » de l'AAS et cofondateur du collectif « les Archiveilleurs ». Il intervient également dans la formation en information documentaire délivrée par la Haute école de gestion de Genève, sur les thématiques du records management.

Il a lancé en 2006 l'un des premiers blogs archivistiques francophones, et s'intéresse depuis aux possibilités offertes par les médias sociaux pour le développement professionnel et le partage d'expériences au sein de la communauté archivistique.

Résumé

Au printemps 2010, inspiré par une démarche similaire de bibliothécaires français, un groupe d'archivistes francophones lance les « Archiveilleurs », une plate-forme de veille collaborative à destination des archivistes. Depuis, une sélection de liens susceptibles d'intéresser la communauté professionnelle des archivistes et gestionnaires d'information est diffusée quotidiennement à plusieurs milliers d'abonnés.

Ce projet s'inscrit pleinement dans le développement du web social. Non seulement ses initiateurs ont fait connaissance sur les réseaux sociaux, et l'ont géré complètement en ligne, mais surtout les solutions sur lesquelles repose la plate-forme pour la gestion des contenus et leur diffusion appartiennent à l'écosystème des blogs et réseaux sociaux avec ses avantages (gratuité, rapidité de mise en œuvre) et ses inconvénients (risque de fermeture des services).

Cette présentation reviendra sur la genèse du projet, exposera le fonctionnement technique et les principes éditoriaux de la plate-forme, mettra en perspective les difficultés rencontrées et ouvrira la discussion sur l'apport d'un tel service au développement professionnel de ses usagers, dans le contexte d'une auto-formation reposant sur les ressources disponibles en ligne.

Capter l'éphémère pour conserver la mémoire d'un événement : La construction d'un fonds d'archives sur les attaques contre Charlie Hebdo en janvier 2015 par la bibliothèque d'Harvard

Lidia UZIEL, Université de Harvard et Lola MIRABAIL, Université Paris 8

Présentation

Lidia UZIEL est directrice de la Division des langues occidentales et conservatrice pour l'Europe de l'Ouest à la Bibliothèque de Harvard. Au sein de sa division, elle dirige la création et la mise en œuvre d'une vision et d'une stratégie de développement des collections en lettres et en sciences humaines provenant de pays de l'Europe de l'Ouest et de pays anglophones. Elle détient un doctorat en littérature comparée de l'Université de Montréal et également de l'Université Jean Moulin Lyon 3 ainsi qu'une Maîtrise en Sciences de l'Information de l'Université de Montréal.

Lola MIRABAIL est diplômée de Sciences Po Bordeaux et conservateur des bibliothèques. En février 2015, dans le cadre de sa formation de conservateur, elle réalise une mission à la bibliothèque d'Harvard et participe au lancement de ce projet d'archivage. Elle est chargée de relayer l'information auprès du grand public et de tisser des liens avec des organismes culturels ou de recherche français (universités, bibliothèques, associations, archives...). Elle est actuellement responsable du département des services au public de la bibliothèque de Paris 8.

Résumé

Affiches, articles, photos, vidéos, blogs, tweets, pages Facebook, tracts, dessins... Protéiformes, les réactions aux attaques des 7, 8 et 9 janvier 2015 fleurirent partout : dans la rue et sur internet, en France et à l'étranger. Et si tous ces témoignages pouvaient intéresser les chercheurs et le grand public dans un futur proche ou lointain ? L'Université d'Harvard en est persuadée : elle s'est rapidement mobilisée pour collecter toutes ces traces éphémères avant qu'elles ne disparaissent.

Ainsi, depuis février 2015, un groupe de travail international, composé de conservateurs et de chercheurs, travaille sur la constitution d'un tel fonds. Plusieurs canaux sont utilisés pour collecter tout type de matériel afin de documenter de la manière la plus exhaustive et fidèle possible les événements de janvier 2015. L'université d'Harvard a tout d'abord lancé un grand appel à contribution auprès du public. Un site internet a été créé pour expliquer la démarche et faciliter cette collecte. Il contient notamment un formulaire de contribution. Son adresse est : <http://library.harvard.edu/charlie>. Cet appel a été relayé par les médias français. Des liens ont également été tissés avec plusieurs organismes, privés et publics, ayant collecté des matériaux sur ces événements. L'Université Harvard a notamment noué des partenariats avec les archives municipales de Toulouse, de Rennes et de Saint-Étienne. Elle est aussi en contact avec la BNF. Parallèlement, Harvard archive des pages web pour conserver les multiples débats, fractures et nuances d'opinions qui ont émergé sur internet à la suite des attentats contre Charlie Hebdo.

La réussite de ce projet repose, d'autre part, sur un important travail de tri et de catalogage de tous les éléments collectés. Certaines métadonnées furent collectées en même temps que les documents qu'elles décrivent. La bibliothèque de Harvard doit toutefois les compléter.

Le groupe réfléchit, enfin, à la diffusion et à la valorisation de ce fonds. La majorité des documents collectés sera accessible à la communauté scientifique internationale, mais également au grand public. La volonté du groupe est, en effet, que ce fonds puisse être consultable non seulement au sein de la bibliothèque, mais aussi sur internet lorsque les dispositions juridiques le permettent. Pour chaque document collecté par don, une autorisation de diffusion est ainsi demandée au donateur.

Ce chantier ambitieux mené par l'Université Harvard pour conserver la mémoire des événements de janvier 2015 revêt ainsi plusieurs dimensions. Très stimulant pour ses participants, il décloisonne et donne un souffle nouveau au travail de conservation de la bibliothèque de Harvard.

La collecte des réseaux sociaux : Un enjeu pour la constitution des collections de dépôt légal du web à la BnF

Annick Le FOLLIC, Bibliothèque nationale de France

Présentation

Depuis 2008, Annick Le FOLLIC est chargée de collections numériques au sein du service du Dépôt légal numérique à la BnF. Elle pilote les activités de collecte du web en collaboration étroite avec des ingénieurs pour le paramétrage et la supervision des robots, afin de couvrir au mieux la diversité des domaines du périmètre français. Elle participe à l'évolution des applications informatiques existantes et mène également des activités de définition de corpus avec des bibliothécaires et des chercheurs.

Résumé

L'apparition des réseaux sociaux, sites web qui permettent à des communautés d'utilisateurs de se réunir par affinités et d'interagir en échangeant différents types de contenus, est contemporaine de la démocratisation du web, dans les années 2000. Au cours de la dernière décennie, l'audience et la fréquentation des réseaux sociaux, au premier rang desquels Facebook et Twitter, n'ont cessé de croître, jusqu'à supplanter celles de médias traditionnels.

Soucieuse de la préservation de ce patrimoine documentaire nouveau, la Bibliothèque nationale de France a mené dès 1999 plusieurs expérimentations en matière d'archivage du web. Elle a entrepris à partir de 2004, de réaliser de grandes campagnes de collecte annuelles, puis de compléter ses collections par des acquisitions auprès de la fondation américaine Internet Archive (pour la période 1996-2000 en particulier), et surtout par des opérations de collectes ciblées avec un réseau de sélectionneurs (la première lors des scrutins présidentiels et législatifs de 2002).

Depuis 2006, le code du Patrimoine mentionne que le champ d'application du dépôt légal est étendu aux « signes, signaux, écrits, images, sons ou messages de toute nature faisant l'objet d'une communication au public par voie électronique ». Cette mission de dépôt légal du web du domaine français a naturellement été confiée à la Bibliothèque nationale de France, dans la continuité de ses autres missions de dépôt légal.

Co-fondatrice de l'International Internet Preservation Consortium en 2003, qui réunit des institutions impliquées dans l'archivage du web, la Bibliothèque nationale de France a contribué au développement et à la diffusion d'outils OpenSource conçus à des fins de moissonnage automatisé, comme NetArchiveSuite ou Heritrix. Confrontée à l'évolution rapide des technologies et des modes de communication sur le web, elle veille à adapter en continu ses outils, mais également sa politique documentaire.

Les pages publiques des réseaux sociaux constituent un lieu d'expression unique, comme on a pu l'observer sur Twitter avec la réaction aux récents attentats parisiens ou lors de la dernière campagne électorale, mais aussi plus régulièrement, dans l'utilisation des plateformes de blogs ou de vidéos par des internautes de plus en plus nombreux. La volatilité, la diversité et la masse de ces documents impliquent une adaptation permanente des outils et des méthodes : un enjeu de taille pour offrir une photographie instantanée des médias d'aujourd'hui aux chercheurs de demain.

Pourquoi les archives de Twitter déposées à la Bibliothèque du Congrès restent inaccessibles ? Une analyse des obstacles techniques et archivistiques

Jean-Daniel ZELLER, consultant en archivage et gestion des documents d'activités

Présentation

De 1983 à 1988, Jean-Daniel ZELLER a dirigé le Service Courrier et Archives du Département de justice et police du canton de Genève. De 1989 à 2016, il occupe le poste d'archiviste principal des Hôpitaux Universitaires de Genève. Depuis février 2016, est consultant en archivage et gestion des documents d'activités.

Parallèlement, il a enseigné de 1995 à 2014 en tant que professeur vacataire à la Haute école de gestion de Genève (département information documentaire), les modules « Archivage des documents électroniques » et « Gestion de projet GED » et est blogueur sous le nom « Le regard de Janus » (<http://regardejanus.wordpress.com/>).

Membre de l'Association suisse des archivistes (<http://www.vsa-aas.org/>) ; du Groupe de travail sur l'Archivage des documents électroniques de l'AAS (1995-2002), puis de la Commission eArchives (dès 2002) ; de la Délégation à la formation (AAS /BBS/ASD) (1999-2008) ; de la Commission de formation de l' AAS (1999-2008) ; de la Commission régionale des cours d'introduction pour les apprentis assistant(e)s en information documentaire (1999-2008) et de la commission norme et standards de l'AAS (dès 2016).

Résumé

Tweeter et LOC

Le 1er avril 2010 (sic), la Bibliothèque du Congrès (LOC) et Twitter ont annoncé le projet commun d'archiver tous les tweets publics de la plateforme Twitter. En 2012, la LOC a annoncé avoir reçu les premiers versements historiques couvrant la période 2006-2008 (env. 133 Terabytes) puis des versements réguliers en constant accroissement.

Même si la LOC est une des institutions expérimentées dans le traitement des documents électroniques, les objets de nature particulière que sont les tweets posent des défis inédits qu'elle n'a pas encore entièrement résolus à ce jour, ce qui rend ces fonds encore inaccessibles aux chercheurs plus de 5 ans après l'annonce initiale.

Ces obstacles de nature techniques ou archivistiques sont les suivants.

La grande volumétrie

La LOC, en acceptant le versement des archives de Tweeter a doublé d'un seul coup le volume de ses archives numériques. Cela représente une charge de travail non négligeable, d'autant plus que cette masse est, pour l'instant, en constante augmentation.

L'abondance des métadonnées

Une des particularités des archives de Twitter est que le volume des métadonnées associées aux tweets est plus important que les tweets eux-mêmes (plus de 100 champs de métadonnées pour chaque tweet).

La question de la protection de la vie privée

Bien que Twitter ait fixé un délai de 6 mois pour la communication des tweets archivés, ils ont été accusés d'utiliser abusivement des données privées sans consentement. En effet, cette opération d'archivage ne fait pas partie des clauses contractuelles de l'abonnement à Twitter et la majorité des utilisateurs n'utilisent pas l'option « privé » pour leurs messages. Comme il n'y a pas eu divulgation à ce jour, il n'y a pas encore de procédure en cours, mais la question reste ouverte.

La question des suppressions rétrospectives de tweets

Dans la même veine, des auteurs de tweets ont supprimé certains tweets sujets à controverse, suppression qui n'est pas effective dans les tweets archivés et qui ne tient pas compte du « remords » de leurs auteurs.

Les chercheurs « bona fide »

Une autre clause de l'accord de versement stipule que les archives seront communiquées à des chercheurs de « bonne foi » ce qui exclut un accès libre à tout public, en contradiction avec l'esprit des archives publiques. Il n'y a pas actuellement d'information de la LOC sur comment ces chercheurs seront désignés et habilités.

A cela s'ajoute l'annonce récente de Twitter visant à augmenter considérablement le nombre des caractères des tweets à terme (de 140 à 1000). Les problèmes de volumétrie cités plus haut ne feront donc que s'accroître.

La communication développera ces problématiques et tentera d'apporter des hypothèses pour en résoudre certaines. Elle posera la question de la création d'un nouveau champ de recherche engendré par la problématique de l'archivage des réseaux sociaux : *l'algorithmique archivistique*.

Du tract au clic : Réflexions sur la collecte, la conservation et l'analyse du matériel électoral

Odile GAULTIER-VOITURIEZ, Centre de recherches politiques de Sciences-Po, Paris
(et Thierry VEDEL, CNRS)

Présentation

Responsable du centre de documentation du CEVIPOF, de la documentation, des archives, des données et de l'édition scientifique, Odile GAULTIER-VOITURIEZ est notamment chargée de la conception, de la gestion et de l'animation du fonds documentaire du CEVIPOF. Elle participe à plusieurs groupes de travail transversaux au sein de Sciences Po et collabore avec les différentes directions. Elle est aussi chargée d'enseignement en master à Sciences Po et dans d'autres établissements universitaires.

Thierry VEDEL est chercheur CNRS au CEVIPOF. Après avoir travaillé sur les politiques publiques des nouveaux médias, il s'intéresse désormais à trois thèmes de recherche : les mutations de la communication politique dans les pays industrialisés ; les relations entre internet et le politique, notamment lors des campagnes électorales ; la régulation des médias dans un contexte de mondialisation. Il a coordonné le projet *Mediapolis « Information politique et citoyenneté à l'ère numérique »* soutenu par l'Agence nationale de la recherche (2009-2013).

Thierry Vedel enseigne la communication politique et la sociologie des médias à l'IEP de Paris et à l'Institut français de presse (Université de Paris 2). Il a participé à la création du réseau Démocratie électronique (DEL) qui rassemble environ deux cents chercheurs travaillant sur les relations entre démocratie et internet.

Il est membre des comités éditoriaux de plusieurs revues dont : *Information Communication and Society ; Réseaux ; Communications & Strategies ; Policy & Internet*. Thierry Vedel a participé à de nombreux projets internationaux, dont récemment le projet *Mapping Digital Media in the World* soutenu par l'Open Society Foundation, et est chercheur associé à l'Oxford Internet Institute où il a séjourné en 2005.

Résumé

Depuis une vingtaine d'années, le matériel de communication (ou propagande) électorale a considérablement évolué. Si les tracts, brochures, professions de foi et affiches, militantes ou commerciales, continuent d'être utilisés, internet est devenu un puissant vecteur d'information au travers d'une multitude de sites web, forums, blogs et autres plateformes de partage ou réseaux sociaux. Cette évolution soulève des problèmes inédits pour les chercheurs travaillant sur le matériel de campagne électorale. La constitution de corpus, au sens traditionnel du terme, devient difficile dans la mesure où les données disponibles sur internet sont hétérogènes, constamment mouvantes, et parfois personnalisées en fonction de chaque récepteur. De plus, une large part de ces données circule sur des plateformes gérées par des opérateurs commerciaux qui tendent à les éditorialiser, les filtrer, voire à limiter leur libre accès par les chercheurs. Plus généralement, les campagnes électorales semblent passer d'une logique de persuasion, point à masse et s'appuyant sur des contenus stables, à une logique d'activation de réseaux d'individus, relayant en chaîne des contenus dynamiques. Dans ces conditions, l'analyse des campagnes électorales doit peut-être abandonner l'analyse des contenus pour se concentrer sur la réception et l'interprétation de l'information électorale par les citoyens.

Réflexions sur le rôle des réseaux sociaux pour un service d'archives audiovisuel : Acquisition, documentation et valorisation

Florian DELABIE, Secteur Gestion de l'information et Archives de la RTBF

Présentation

Détenteur d'un Master en Histoire à finalité histoire et archives de l'UCL et d'un master complémentaire en gestion de la Solvay Brussels School of Economics & Management, Florian DELABIE depuis 2011 au sein du service des archives audiovisuelles de la RTBF.

Son département se définit de plus en plus comme un service de record management avec une perspective historique et patrimoniale. Dans ce cadre, il travaille à la formalisation et l'amélioration des processus d'archivage et intervient lors des réflexions sur les workflow de production et la gestion des métadonnées.

En tant que service intégré à la direction technologique (DGTE), il officie également de plus en plus en qualité de consultant interne pour la gestion de l'information et a pu également prendre part à des projets de recherche et développement visant l'utilisation des technologies sémantiques et du *linked data* comme GEMS (projet interne) ou Mediamap+. Le département est aussi partie prenante du projet européen de valorisation d'archives audiovisuelles EUscreen, sous-ensemble d'Europeana.

Résumé

« Notre concurrent n'est plus RTL mais Facebook ». C'est en ces termes que Jean-Pierre Jacqmin, directeur de l'Information et des Sports de la RTBF, définissait en ce début d'année 2016 la nouvelle ligne éditoriale des chaînes publiques belges francophones. Cette citation montre la place et l'importance qu'ont prises les réseaux sociaux, figures de proue du web collaboratif imaginé par Tim Berners-Lee, dans le monde de la communication et des médias en particulier.

Les réseaux sociaux influencent de façon grandissante l'évolution des méthodes de création et de production des contenus audiovisuels tant dans leurs aspects narratifs et éditoriaux que techniques. Ils sont également devenus des canaux de communication à part entière concurrençant voire se substituant aux canaux traditionnels que sont la radio et la télévision.

Dans ce contexte passionnant, un service d'archive audiovisuel se doit d'adapter ses processus et tenter d'anticiper les évolutions futures qui toucheront inmanquablement notre discipline. À travers l'analyse de diverses expériences, l'exposé tentera d'apporter un regard critique quant à l'utilisation des réseaux sociaux dans le domaine de l'archivage audiovisuel. Dès lors, notre propos s'articulera autour des fonctions d'acquisition, de documentation et de valorisation de nos documents d'archives et l'apport que nous pouvons tirer de ces outils de communication.

Les métadonnées des images fixes et les médias sociaux

Patrick PECCATTE, Laboratoire d'Histoire visuelle contemporaine, EHES

Présentation

Ancien professeur de mathématiques, Patrick PECCATTE est informaticien, spécialiste des bases de données d'images numériques. Il a créé avec une équipe de passionnés le projet collaboratif PhotosNormandie. Il est actuellement chercheur associé au Laboratoire d'histoire visuelle contemporaine (Lhivic/EHES). Ses recherches portent sur la redocumentarisation, les objets informationnels, les théories de la fiction, les images dans les pulp magazines et les comics, les dispositifs de représentation de la guerre dans la culture ordinaire. Il tient le carnet de recherche Déjà Vu sur la plate-forme Hypothèses

Résumé

Après un rappel succinct sur les différents standards de métadonnées embarquées dans les images fixes (IPTC/IIM, Exif, XMP), l'exposé examinera la façon dont certaines plates-formes sociales (Dropbox, Facebook, Flickr, Instagram, Pinterest, Tumblr, Twitter, etc.) se comportent avec ces métadonnées. La plupart d'entre elles en effet ignorent les métadonnées embarquées et suppriment carrément celles-ci, ce qui est inacceptable tant du point de vue documentaire qu'en ce qui concerne les droits associés aux images. L'intervention conclura sur les (rares) possibilités d'exploitation des métadonnées sur les réseaux sociaux, en particulier à travers l'expérience du projet PhotosNormandie qui consiste à améliorer les légendes d'un fonds de photos historiques en s'appuyant sur les métadonnées embarquées des images.

Techniques de collecte et d'archivage des tweets : Partage de pratiques et d'outils

Hubert NAETS, CENTAL, Université catholique de Louvain

Présentation

Après avoir travaillé pendant quatre ans au Commissariat à l'Énergie Atomique (Fontenay-aux-Roses, Paris), Hubert Naets rejoint en 2007 le Centre de traitement automatique du langage (CENTAL) de l'UCL en tant que linguiste informaticien. Son travail et ses recherches portent sur des domaines très variés liés au traitement automatique des langues, dont notamment le traitement de corpus en langues anciennes, l'acquisition de langue seconde ou encore l'extraction d'informations sémantiques.

Résumé

Depuis plusieurs années, le Centre de traitement automatique du langage (CENTAL) de l'Université catholique de Louvain est périodiquement sollicité pour collecter et analyser des tweets portant sur des sujets divers. Il s'agit de récolter les messages postés par des musées belges ou luxembourgeois, ceux écrits par des parlementaires européens ou des candidats aux élections américaines, ou encore des tweets en français ou en italien à des fins d'études linguistiques. Ces demandes traduisent la difficulté d'obtenir aisément un ensemble limité de messages ciblés, accompagnés de leurs métadonnées.

Il existe aujourd'hui essentiellement deux moyens de constituer automatiquement des collections de tweets : il faut soit passer par des prestataires, agréés par Twitter ou lui appartenant, spécialisés dans la revente en masse de messages, soit utiliser des interfaces de programmation (API) gratuites et limitées fournies par Twitter, souvent employées pour intégrer des flux de tweets au sein de pages web. Les interfaces de programmation étant avant tout destinées aux développeurs informatiques, des outils d'analyse qualitative ou statistique de données ont intégré des connecteurs permettant d'utiliser de façon transparente ces interfaces et d'importer ainsi plus facilement des tweets.

Dans cet atelier, nous reviendrons dans un premier temps sur le fonctionnement de base de Twitter et nous passerons en revue les différentes interfaces de programmation proposées par le réseau social, en mettant en évidence leurs avantages et leurs limitations. Nous présenterons ensuite plusieurs outils qui s'appuient sur ces interfaces de programmation et qui permettent de collecter et d'archiver des tweets sous différentes modalités.

Archives participatives, quelles stratégies ? L'expérience du portail Wikipasdecalais

Ivan PACHEKA, Archives départementales du Pas-de-calais
(et Sophie LEGER, Comité d'histoire du Haut-Pays)

Présentation

Sophie LEGER anime le *Comité d'Histoire du Haut-Pays*, une association de 400 membres qui possède son Centre de documentation et de recherche, et développe des actions de valorisation de l'histoire et du patrimoine local sur 107 communes dans le Pas-de-Calais. En 2011, elle devient présidente de *Wikipasdecalais*, une association de bénévoles portant un projet éponyme d'encyclopédie en ligne collaborative destinée au grand public, projet qu'elle sert par son expérience dans l'édition et sa pratique des archives, et autour duquel elle fédère un réseau de professionnels et d'amateurs.

Ivan PACHEKA est le créateur du site Internet *Mémoires de pierre* qui depuis 2001 recense les lieux de mémoire des victimes civiles et militaires des conflits mondiaux dans le Pas-de-Calais, une référence pour les travaux universitaires du Centenaire de la Grande Guerre. En 2011, il devient administrateur de *Wikipasdecalais*, et poursuit ainsi son œuvre de collecte d'archives privées et publiques avec de nouvelles possibilités. Chef du service par intérim des archives contemporaines aux Archives départementales du Pas-de-Calais, il apporte à *Wikipasdecalais* son expertise en matière de méthodologie et d'indexation, tout en fournissant de nombreuses analyses de documents inédits.

Tous deux sont depuis 2015 membres du bureau de la Commission départementale d'Histoire et d'Archéologie du Pas-de-Calais.

Résumé

Depuis 2014, les grandes collectes organisées autour du Centenaire de la Grande Guerre ont rendu plus que jamais d'actualité la question de la restitution en ligne des archives rassemblées, et appelé un travail participatif pour tenter de les renseigner. Parmi les nombreuses formes que prennent ces éditions web, un outil essentiel reste encore peu usité : le wiki. Pourtant conçu comme un procédé d'archivage de données (textes, images, médias) collaboratif en ligne, le wiki est encore mal connu des créateurs de web. Son emploi par les géants Wikipedia ou Wikicommons livre une fausse image de l'outil wiki qu'il convient de démythifier.

Logiciel en open source, le wiki est d'abord un outil de mutualisation de connaissances dont les potentialités peuvent favoriser des projets de territoire participatifs. Sa mise en place et son administration, de même que le téléversement, ne sont pas réservés à des informaticiens chevronnés.

L'exemple de Wikipasdecalais est à même de le démontrer. Cette initiative de bénévoles, après 5 années de développement, propose aujourd'hui plus de 41000 articles et 16200 médias. Les partis pris par ses créateurs, son évolution, ses succès et ses difficultés constituent autant d'indices riches d'enseignements pour les porteurs de projets collaboratifs de réseau social numérique de contenu. En partageant l'expérience d'un wiki de territoire, les créateurs de Wikipasdecalais, militant pour la libéralisation des savoirs, visent à convaincre les professionnels et amateurs que le wiki est, non seulement abordable à tous moyennant un peu de pratique, mais aussi un fer de lance pour l'enrichissement perpétuel de nos connaissances.