



Édito

Michèle Baussant, DAS InSHS, et François-Joseph Ruggiu, DAS InSHS

Depuis 2010, la direction de l'Institut des Sciences Humaines et Sociales du CNRS a résolument affiché le soutien

aux aires culturelles comme un des axes prioritaires de sa politique scientifique au même titre que le genre, les méthodes quantitatives ou encore le numérique [...]

VIE DES RESEAUX

FRANTIQU : 30 ans d'expertise documentaire en archéologie

Qu'un réseau documentaire fête cette année ses trente ans n'est pas rien. Sachant que la vie d'une telle organisation ne tient souvent qu'à la solidité des relations humaines entre les partenaires, pouvoir fêter cet anniversaire est déjà un succès [...]

A PROPOS

Le développement durable et les sciences de la conservation du patrimoine

Le Centre de recherche sur la conservation fêtait récemment son 50e anniversaire. C'est en 1963 que l'unité de recherche fut créée sous l'impulsion de la Bibliothèque nationale, au MNHN, pour trouver un remède aux altérations biologiques qui menaçaient les collections patrimoniales [...]

LA TRIBUNE D'HUMA-NUM

Un grand corpus de SMS en français : 88milSMS ; Enquête DARIAH sur les pratiques numériques ; deux nouveaux articles de la TGIR Huma-Num [...]

A SIGNALER

Appels à idées et à contributions pour le colloque scientifique international *Our common future under climate change* (7 au 10 juillet 2015 à l'UNESCO à Paris) [...]

NOUVELLES DE L'INSTITUT

L'InSHS accueille trois nouveaux membres [...]

FOCUS

Aires culturelles. Vers une géométrie variable de la mondialisation

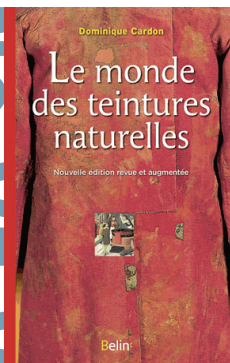
L'Institut des Amériques, Groupement d'Intérêt Scientifique créé en 2007, réfléchit depuis plusieurs années sur les aires culturelles et leurs reconfigurations dans un environnement global. Fort de ses réflexions, l'IdA a décidé de consacrer son colloque annuel 2014 à cette thématique [...]

ZOOM SUR...

Crowdsourcing, participation, démocratie : le citoyen comme ressource et comme acteur de la connaissance

L'arrivée des *Big Data* — la production massive d'informations par des expériences scientifiques, des capteurs ou des populations munies d'équipements de communication — n'est pas seulement une révolution technologique, c'est une nouvelle méthode de découverte : elle ne remplace pas les autres méthodes (hypothético-déductive, expérimentale, par modélisation), mais y superpose une approche radicalement empirique, qui trouve la connaissance dans les données sans passer par l'échantillonnage, la règle, le modèle ou le dispositif causal hypothétique [...]

LIVRE



Le monde des teintures naturelles, Dominique Cardon, Belin, 2014

De la pourpre impériale extraite de coquillages marins aux étoffes teintées aux écorces et à la boue, les teintures naturelles étaient, jusqu'à la fin du XIX^e siècle,

les seules sources de couleurs de tous les textiles utilisés par l'Homme. Elles servaient aussi à teindre les peaux, les poils [...]

[voir toutes les publications](#)

REVUE



Fondée en 1991, *Genèses* est une revue trimestrielle de sciences sociales et d'histoire. Comprendre nos sociétés contemporaines à la lumière de l'histoire, restituer les processus qui les ont façonnées, contribuer à une

histoire de nos disciplines sont les ambitions de *Genèses*. Cette revue est un lieu de rencontre et d'échanges [...]

[voir toutes les revues](#)

PHOTO



Essai de navigation avec le bateau Gypts, au départ de la rade de Marseille et avant un périple le long des côtes desanciens comptoirs massaliotes.



Édito

Michèle Baussant, DAS InSHS
François-Joseph Ruggiu, DAS InSHS

Depuis 2010, la direction de l'Institut des Sciences Humaines et Sociales du CNRS a résolument affiché le soutien aux aires culturelles comme un des axes prioritaires de sa politique scientifique au même titre que le genre, les méthodes quantitatives ou encore le numérique. Elle s'est ainsi placée dans la continuité de l'action du CNRS qui, de longue date, a fait une large place dans ses recrutements aux recherches menées sur les aires africaines, américaines, asiatiques et océaniques et qui a suscité de nombreuses manifestations scientifiques, par exemple en 2003, et plusieurs rapports consacrés à cette notion, comme en 2010. La constance avec laquelle cette priorité donnée aux aires culturelles a été maintenue depuis quatre ans est cependant nouvelle et les actions récentes de l'InSHS dans ce domaine sont nombreuses. Des coloriages de concours chercheurs entre 2011 et 2014 ont été faits au service des « aires culturelles ». Trois Groupements d'Intérêts Scientifiques (« Etudes africaines en France » ; « Etudes asiatiques » ; « Moyen-Orient et mondes musulmans ») et deux Groupements de Recherche (« Mondes britanniques » ; « Europe médiane ») ont été créés pour faciliter les synergies entre les centres de recherches consacrés à ces terrains, développer entre eux une politique scientifique concertée et servir d'instruments de conseil et de prospective aux organismes et aux établissements liés à l'Enseignement Supérieur et à la Recherche. Le réseau des 29 UMIFRE, en partenariat avec le MAE, et USR, piloté par l'InSHS, a été affermi de manière à fournir à nos collègues de solides centres d'échanges scientifiques et académiques. Enfin, la direction de l'InSHS a toujours cherché à préserver, dans un contexte budgétaire contraint, les ressources et les personnels des unités dont la recherche est entièrement ou partiellement tournée vers les aires culturelles.

La notion d'aires culturelles date de l'époque même de la naissance des sciences sociales. Apparue d'abord au sein de l'anthropologie allemande, elle s'est épanouie dans les areas studies développées par les sciences sociales américaines dans la première moitié du ^{xx}e siècle. Elle a connu en France un apogée académique lorsqu'elle a entraîné, en particulier au moment de la création de la VI^e section de l'Ecole Pratique des Hautes Etudes par Fernand Braudel, la formation de grands centres de recherche pluridisciplinaires consacrés à une région du monde. L'idée sous-jacente était que la présence dans un même lieu de spécialistes de disciplines différentes mais s'intéressant à un même objet, fréquentant un même terrain et parlant les mêmes langues, ouvrait à des collaborations scientifiques mutuellement enrichissantes et encourageait à la réflexivité de chaque chercheur sur son travail. La forte impulsion donnée à la recherche par les laboratoires d'aires culturelles a permis dans le même temps de développer une réflexion critique, conduisant à une refondation épistémologique ou méthodologique de la notion loin de tout europocentrisme.

L'organisation de la recherche en aires culturelles pose, en effet, une série d'écueils qui ont peu à peu diminué sa puissance heuristique. Le premier a été une certaine tendance au cloisonnement des aires, voire des sous-régions. Les échanges, par exemple, entre les spécialistes des Amériques et ceux de l'Asie ou entre les spécialistes de l'Afrique et de l'Asie, à l'heure où, pourtant, la mondialisation bouleverse les ordres anciens, peuvent paraître moins fluides que ceux développés

au sein d'une même aire. Elle tend aussi à minorer la place de certaines régions, comme la Caraïbe ou l'Océanie, qui sont de formidables observatoires des changements, en particulier sociaux et environnementaux, à venir. La seconde difficulté vient d'un progressif appauvrissement des croisements disciplinaires. A l'heure actuelle, ils impliquent majoritairement les historiens, les anthropologues, les sociologues et les spécialistes des sciences politiques, et beaucoup moins les géographes, les juristes, les économistes, voire les philosophes, les linguistes et les spécialistes de littérature, qui préfèrent souvent ancrer leurs recherches dans des unités disciplinaires. Le troisième problème est le relatif partage qui s'est installé entre, d'un côté, le CNRS et les grandes écoles qui accueillent la plupart des unités d'aires culturelles et, d'un autre côté, les Universités qui tendent à privilégier les unités disciplinaires souvent en lien avec les formations qu'elles proposent. Ce constat n'est naturellement pas vrai pour toutes les aires et pour tous les domaines des SHS. Il pose néanmoins la question d'un large accès, partout sur le territoire national, des étudiants de licence à des enseignements diversifiés sur l'Afrique, sur les Amériques, sur l'Asie, sur l'Europe médiane et orientale, sur le Moyen-Orient ou sur le Pacifique. Et le potentiel de médiation que représentent les spécialistes des aires culturelles entre la recherche française et européenne et la recherche menée par les communautés scientifiques des pays concernés, quelle que soit la discipline, n'est pas pleinement exploité.

C'est avec ces questions présentes à l'esprit que l'Institut des Sciences Humaines et Sociales du CNRS travaille depuis quatre ans à soutenir les recherches sur les aires culturelles. Sa préoccupation première a été de maintenir et de développer le potentiel de recherche dans ce domaine au sein des UMR qu'il anime avec les grands établissements d'enseignement et de recherche. Les spécialistes des aires culturelles ont, en effet, un parcours différent de leurs collègues spécialistes de la France et de l'Europe occidentale. La formation en langues joue pour eux un rôle fondamental comme les années passées sur des terrains qui ne sont pas toujours facilement accessibles, ni aisés à fréquenter. Inscrits dans des réseaux spécifiques, ils passent souvent par des circuits différents de ceux de leurs contemporains, en particulier les agrégations dans les disciplines où elles existent. Les aider à s'insérer dans le dispositif de la recherche en France est donc un élément clef de notre politique scientifique. La vision que nous avons des pays qui composent ces aires culturelles a été complètement bouleversée depuis une vingtaine d'années car ils ont, pour la plupart, vu l'émergence de nouveaux modèles politiques, économiques ou culturels, étroitement liés à l'intensification de la globalisation. Et les spécialistes des aires culturelles sont particulièrement bien outillés pour analyser ces phénomènes et ainsi accroître notre appréhension du monde. A travers les recherches sur une aire, ce sont toujours les enjeux les plus généraux des sciences humaines qu'ils cherchent à dégager et, en ce sens, ils participent de manière centrale à l'avancement de la recherche fondamentale tout en s'inscrivant au cœur des débats traversant nos sociétés.

**Michèle Baussant, DAS InSHS
& François-Joseph Ruggiu, DAS InSHS**

L'InSHS accueille trois nouveaux membres



Marie Gaille

Marie Gaille est nommée chargée de mission pour la section 35 (Sciences philosophiques et philologiques, sciences de l'art) au sein de l'InSHS. Elle est philosophe, directrice de recherches au CNRS (SPHERE, UMR 7219). Ses travaux portent sur la manière dont la philosophie politique et morale est investie depuis les présocratiques par la question de ses rapports avec la médecine. Elle s'intéresse aux frontières communes entre discours philosophique et discours médical, à l'histoire de leurs interactions et à ses enjeux conceptuels. De la Renaissance à nos jours, en contexte occidental, elle a exploré trois modalités de cette relation, qui évolue au fil du temps : les emprunts de la pensée politique à la théorie médicale, en particulier à la Renaissance ; la dimension anthropologique, morale et politique de certaines parties du corpus médical, à la Renaissance et dans l'œuvre de Cabanis ; les décisions médicales contemporaines relatives à la vie, la santé, la fin de vie et plus particulièrement la procréation comme enjeu collectif et individuel. Elle examine comment la philosophie peut tirer profit d'un dialogue avec la sociologie, l'anthropologie, la psychanalyse et la psychologie clinique, pour comprendre ces décisions et identifier les interrogations éthiques et juridiques qu'elles suscitent. A l'heure actuelle, elle développe par ailleurs un travail de recherche sur la santé humaine face au risque environnemental, ce qui la conduit à explorer une nouvelle interaction entre les dimensions épistémologique et politique. Marie Gaille poursuit enfin en filigrane de ces recherches une interrogation en philosophie politique, centrée sur les thèmes du droit de cité et du conflit civil.

marie.gaille@cnrs-dir.fr



Sylvie Ollitrault

Sylvie Ollitrault est nommée chargée de mission pour la section 40 (Politique, pouvoir, organisations). Directrice de recherche au CNRS (CRAPE, UMR 6051), ses domaines de compétences relèvent de la sociologie de l'engagement, des mobilisations et son terrain est — depuis plus de vingt ans — celui des mouvements écologistes, des mobilisations environnementales. Elle a participé à de nombreux programmes de recherche financés par le Ministère de l'environnement, de l'écologie ou encore l'ANSES, l'ADEME. Elle a intégré des réseaux de recherche tant nationaux comme le réseau Développement durable et Territoires (Lille) qu'internationaux (membre du Standing group de l'ECPR : « *Green politics* »). Elle travaille sur les questions de protestations, les usages militants de l'expertise et s'intéresse aux effets concrets de la transnationalisation sur les représentations et les pratiques des individus. Elle s'oriente de plus en plus vers les questions de l'appropriation profane des sciences. Elle a obtenu son habilitation à diriger les recherches sur les questions d'espaces des mobilisations en novembre 2012. Ses publications principales ont trait principalement aux questions d'environnement soit dans les revues à comité de lecture, soit par des ouvrages. Elle dirige l'équipe « Mobilisation, citoyennetés

et vie politique » du CRAPE à Rennes, est membre du conseil du laboratoire. Elle participe au pôle santé sur le volet santé environnementale de la Maison des Sciences Humaines de Bretagne. Elle préside le GIS Institut des Amériques Rennes qui regroupe 13 UMR et EA en SHS, Arts et Lettres et enseigne à Sciences Po Rennes, à l'Université Rennes 1, à l'École Supérieure de Chimie et à l'université de Rennes 2.

s.ollitrault@yahoo.fr



Didier Trony

Didier Trony est nommé directeur adjoint scientifique de l'InSHS, chargé des sections 36 (Sociologie et sciences du droit) et 40 (Politique, pouvoir, organisations) et de l'Information Scientifique et Technique. Il succède à Florent Champy.

Directeur de recherche à l'INRA (UR RiTME), ses travaux principaux portent sur la production normative publique et privée d'inspiration sanitaire, c'est-à-dire des normes cherchant à imposer la santé des personnes ou des animaux comme un objectif légitime de régulation des activités.

S'appuyant sur une approche pragmatique comparée (nutrition, sécurité sanitaire des aliments, santé animale, sécurité sanitaire des produits de santé), il analyse les opérations argumentatives et critiques à l'œuvre dans la définition et l'écriture de ces normes, depuis les simples recommandations jusqu'aux accords internationaux. Il mène par ailleurs des travaux sur l'évaluation de la recherche et sur l'économie politique de la publication scientifique, notamment la fabrique et la critique d'instruments d'évaluation (classifications de revues, métrologie de citations ou d'audience) et les transformations de l'évaluation par les pairs.

Didier.TORNY@cnrs-dir.fr

FRANTIQ : 30 ans d'expertise documentaire en archéologie



Qu'un réseau documentaire fête cette année ses trente ans n'est pas rien. Sachant que la vie d'une telle organisation ne tient souvent qu'à la solidité des relations humaines entre les partenaires, pouvoir fêter cet anniversaire est déjà un succès. Il n'en reste pas moins que le consensus sur les objectifs du réseau et une remise en question continue de ses outils ont joué une part non négligeable. Retour sur une histoire originale où archéologues, documentalistes et informaticiens se partagent les rôles.

C'est en 1984 que se réunissent quatre centres de recherche sur l'Antiquité (Centre Pierre-Paris de Bordeaux, Centre de recherches archéologiques de Valbonne, Maison de l'Orient méditerranéen de

Lyon et Centre d'histoire ancienne de Besançon) dans le souci de mutualiser les catalogues informatisés de leurs bibliothèques. En créant ce premier catalogue commun, alimenté par versements périodiques sur la base d'un format minimal de compatibilité des données, ils comblent une lacune. En effet, aucun *Bulletin signalétique* ni bases de données bibliographiques Francis du CNRS-CDSH¹ ne couvrent ces périodes. Cet embryon de réseau est appelé FRANTIQ, Fichiers de Recherche sur l'ANTIQUité.

La fin des années 1980 marque la période de construction du réseau. Avec l'arrivée d'une documentaliste, Brigitte Lequeux, le réseau se structure et ses outils se diversifient. Les produits FRANTIQ s'organisent autour de trois objets documentaires : le CCI, Catalogue collectif indexé ; l'ETS, Ensemble thématique Signé qui rassemble des bibliographies de chercheurs ou de thématiques de laboratoires ; IMAGES, base de données d'images numérisées et de scans de tables de matières. Les contenus font l'objet d'une répartition entre les centres, notamment en ce qui concerne le dépouillement des revues et des colloques. Une cellule de pilotage est mise en place avec Bruno Helly comme administrateur (auquel Pierre Rouillard succédera en 2003), Brigitte Lequeux comme coordinatrice et les directeurs successifs du Centre de recherche archéologique à Valbonne — devenu CEPAM (Cultures et Environnements. Préhistoire, Antiquité, Moyen Âge, UMR7264, CNRS/Université Nice Sophia Antipolis) — et de la Maison de l'Orient à Lyon. Sa mission est de faire respecter normes et formats du catalogue et de développer un thesaurus. Le nombre croissant de bibliothèques partenaires nécessite l'organisation de formations pour la prise en main des outils documentaires successifs². Le catalogue évolue vers un format Unimarc et ses thématiques sont réaffirmées. Les mots-clés utilisés dans le catalogue sont alors organisés en un vocabulaire contrôlé pour les Anthroponymes et la Chronologie, prémisses des six micro-thesaurus constituant aujourd'hui le thesaurus PACTOLS (Peuples et cultures, Anthroponymes, Chronologie relative, Toponymes, Œuvres, Lieux, Sujets).

Structure et contenu sont validés par consensus entre des commissions de spécialistes et la cellule de pilotage, alors que le logiciel Thesaplus est choisi pour ses capacités de gestion du multilinguisme. Un effort particulier est fait pour rendre l'ensemble accessible à tous : l'hébergement des bases par le SUNIST, serveur du ministère de l'enseignement supérieur, permet dès 1987 un accès par minitel (3615 FRANTIQ pour une recherche publique ; 3614 FRANTIQ pour un accès professionnel) puis sur Internet par protocole WAIS. En même temps, un audit commandité par le Département des sciences de l'homme et de la société du CNRS, conduit par le groupe « Réseau-Bibliothèques » sous la direction de Lucie Fossier³, conclut au maintien du réseau, à sa prise en charge financière par la direction scientifique du Département SHS et précise sa thématique sur l'archéologie des temps anciens.

Ainsi, la philosophie qui animera le réseau jusqu'à aujourd'hui est posée dès ces premières années : le signalement de ressources documentaires spécialisées par consensus sur les contenus entre archéologues et spécialistes de l'information, des choix informatiques respectueux des normes et de l'accès direct à l'information, avec une volonté continue de suivre au plus près les évolutions des techniques de l'information.

La consolidation du réseau se fait cependant dans l'adversité. Il faudra toute l'énergie et l'obstination des responsables de FRANTIQ pour faire face aux problèmes techniques qui vont se succéder pendant une décennie, alors que le réseau, avec la quantité des données à traiter, ne cesse de croître. Le SUNIST d'abord, puis le centre de calcul de Strasbourg-Cronenbourg ne peuvent plus héberger les bases FRANTIQ, qui sont finalement accueillies en 1993 sur les serveurs de l'université Lyon 2. L'interrogation sur Internet y est confortée en plus du minitel. Peuvent aussi s'y développer quelques avancées informatiques : au moteur de recherche SILEX (Système d'Indexation LEXical) réalisé pour FRANTIQ qui permet notamment d'unifier les formats du catalogue en 1996, vient s'ajouter deux ans plus tard un moteur Z39-50. En parallèle, le logiciel documentaire Texto doit être progressivement abandonné faute de maintenance, au profit du seul WinISIS. Le nombre des bibliothèques « productrices » grimpe : avec une moyenne de 3 nouvelles adhésions par an, de 11 en 1994, il passe à 27 en 2000. Mais l'accroissement exponentiel du catalogue qui s'ensuit (il dépasse les 100.000 notices en 1997, 150.000 en 2001 et 180.000 en 2003), impose un changement radical de méthode et

1. Le Centre de documentation en sciences humaines, fusionné avec le CDST, Centre de documentation scientifique et technique, deviendra en 1988 l'INIST, Institut de l'information scientifique et technique.

2. Texto puis TextoW et CDS/ISIS puis WinISIS.

3. Constitué de Jean Bourdon, Gilles Dagand, Alain Degenne et Antoinette Hesnard, sous la responsabilité de Lucie Fossier.

d'outil : à cette échelle, l'alimentation par versements et surtout l'important travail manuel de dédoublement et de normalisation des autorités, ne sont plus envisageables. Un véritable catalogage partagé doit se mettre en place. Confronté à la nécessaire évolution de ses outils et à leur financement, FRANTIQ se tourne vers le logiciel libre Koha pour la migration de son catalogue. A cette époque, la fusion des trois bases initiales dans le seul CCI incite FRANTIQ à changer de nom : les Fichiers de Recherche deviennent Fédération et Ressources sur l'ANTIQUité. L'évolution marque aussi l'importance progressive des bibliothèques productrices dans le fonctionnement du réseau qui, peu à peu, abandonnent leur système local et les versements périodiques au profit d'une alimentation directe client-serveur du catalogue.

Le recrutement d'un ingénieur informaticien participera à stabiliser la situation. Affecté à la plateforme des réseaux documentaires de la Maison de l'Orient, il adapte le moteur SILEX en XML, puis le remplace par un moteur MySQL. Son implication est désormais essentielle au développement des outils du réseau : il participe à la rédaction du cahier des charges pour le changement de logiciel de catalogage avec la cellule documentaire⁴ installée depuis 1998 à la Maison Archéologie et Ethnologie à Nanterre, puis assure la migration du catalogue de Texto sur Koha. Il développe pour FRANTIQ le logiciel OpenTheso pour gérer le thesaurus PACTOLS. Son expertise permet alors à FRANTIQ d'être plus ambitieux dans ses actions, notamment au niveau de l'ouverture de ses données et de sa maîtrise en indexation spécialisée.

Le début des années 2000 marque de ce point de vue un tournant : le réseau propose ou répond à des projets de recherche et de développement en informatique documentaire. Toujours dans le but de travailler en complément d'autres bases de données bibliographiques existantes, une ACI du réseau des Maisons des sciences de l'homme conduite avec l'INIST et l'équipe du *Bulletin analytique d'histoire romaine* permet de monter le portail DAPHNE pour l'interrogation simultanée du CCI, du BAHR et des bases bibliographiques de Francis concernées par la thématique archéologique. A l'occasion, un important travail d'équivalence entre les trois vocabulaires PACTOLS, index des bases FRANCIS et mots-clés du BAHR a été entrepris. Ajouté à une organisation informatique originale, il a permis à DAPHNE d'offrir des temps de réponse remarquables. Son excellent référencement lui offre un succès visible notamment à l'étranger, d'où les consultations sont majoritaires. Les moyens pour sa maintenance et sa mise à

Thesaurus

34.000 mots-clés au thesaurus PACTOLS : Peuples et cultures/Anthroponymes/Chronologie relative/Toponymes et sites archéologiques/Ceuvres/Lieux/Sujets

Multi-hiérarchique, multilingue (français, allemand, anglais, italien, espagnol, néerlandais, arabe), format SKOS, mis à jour en continu

► <http://pactols.frantiq.fr>

OpenTheso

Logiciel libre de gestion de thesaurus, compatible norme ISO 25964-1: 2011

► <http://opentheso.frantiq.fr/>

niveau informatique ont malheureusement manqué et le portail a dû interrompre son service début 2014.

D'autres projets de valorisation du catalogue sont entrepris dès 2007 dans la perspective d'ouvrir et de partager ses données. D'abord, il s'agissait d'étudier l'intégration du CCI dans le catalogue de l'Enseignement supérieur et de la recherche. Ainsi, un rapprochement, entrepris avec les documentalistes du réseau des MSH, vers l'Agence bibliographique de l'enseignement supérieur pour l'intégration des bibliothèques des MSH et du CCI dans le Sudoc n'a pu encore aboutir. Bien que des avancées aient pu être notées (possibilité d'ajouter des notices Sudoc à un autre réseau documentaire non commercial, acceptation des notices de dépouillement et des tirés-à-part...), le coût des licences est resté un obstacle majeur, ainsi que le refus d'intégrer le thesaurus PACTOLS aux outils d'indexation disponibles. Ensuite, il fallait tenir compte de la montée en puissance de l'Internet qui offrait un accès direct, via des portails de revues ou des archives ouvertes, aux textes de nombreux documents scientifiques. Un accord avec Persée, le portail de revues scientifiques en sciences humaines et sociales, est conclu en 2009. Il concerne un enrichissement automatique et réciproque des notices d'articles présents dans les deux bases, avec le lien vers le texte intégral pour les notices du CCI et avec les mots-clés PACTOLS pour les articles numérisés dans Persée. Les essais conduits sur des revues tests ont donné alors trop peu de réponses pertinentes. Le projet doit être pourtant maintenu en élargissant les partenaires à d'autres réservoirs numériques. Aujourd'hui, le catalogue se prépare à la révolution Frbr/RDA par un contrôle accru et par des recommandations particulières aux catalogueurs sur les relations entre notices.

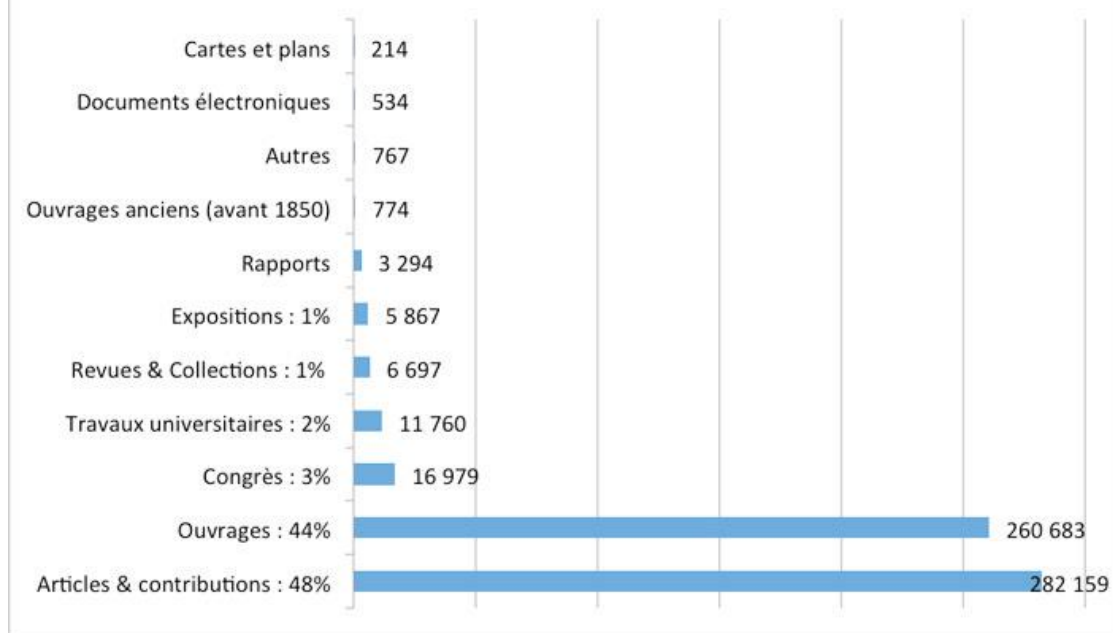
Le thesaurus PACTOLS, seule terminologie française en accès libre pour l'archéologie et les sciences de l'Antiquité⁵, constitue par ailleurs un ambassadeur de choix : sa mise à jour continue, son multilinguisme et sa structure multi-hiérarchique et normalisée en font un produit attractif, pour des services documentaires nationaux et dans le cadre de programmes internationaux. Celui-ci lui donne des opportunités d'évolution qui le renforce comme outil incontournable dans son domaine. Si le projet soumis dans le cadre du 6e PCRD européen "Accès multilingue au Patrimoine" (AMP), sur une ébauche de concordance de vocabulaires spécifiques et de thesaurus, n'aboutit pas, celui sur les "Innovations et développement du réseau de ressources documentaires FRANTIQ" (IdR2D) proposé en 2007 au TGE Adonis permet de transformer en ontologie une première partie du thesaurus sur les objets mobiliers. L'adaptation des PACTOLS au format SKOS⁶ lui permet en outre d'être intégré aux référentiels du moteur de recherche en sciences humaines et sociales ISIDORE. Dans le

4. Un ingénieur documentaliste seconde Brigitte Lequeux depuis 2001 dans sa mission de coordination du réseau.

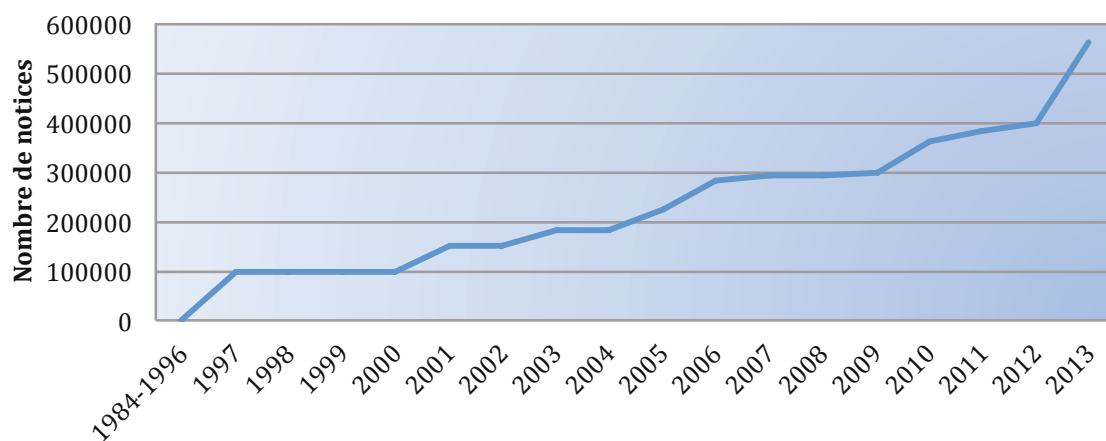
5. Le vocabulaire de la base "Patrimoine" est exclusif aux services du ministère de la Culture et de la Communication.

6. *Simple Knowledge Organization System* (Système simple d'organisation des connaissances), recommandation du W3C pour l'adaptation des vocabulaires d'indexation et des thesaurus au web sémantique.

Typologie du catalogue



Progression du catalogue



Catalogue

600.000 documents signalés au catalogue collectif
 60 titres de revues et collections dépouillées
 50 bibliothèques de localisation

► <http://catalogue.frantiq.fr>

même esprit, la signature toute récente de la participation de FRANTIQ au projet européen d'infrastructure de recherche pour l'archéologie Ariadne, l'intègre parmi les meilleurs outils terminologiques de la discipline en Europe. La traduction des PACTOLS et les nécessaires révisions sont prises en charge à la fois par le TGE Adonis (anglais, allemand, italien, espagnol) et par le programme MULTITA piloté par les Musées royaux de Belgique (néerlandais). Puis, le consortium MASA de la TGIR HumaNum participe à la traduction en arabe, menant à sept le nombre de langues disponibles. Le thesaurus constitue ainsi l'outil idéal pour la traduction de termes spécialisés, pour l'indexation de publications, d'archives scientifiques et de bases de données de recherche. De plus, HumaNum a choisi le gestionnaire OpenTheso comme outil de transformation en format SKOS des vocabulaires non normalisés utilisés par ISIDORE. Une aide au développement du logiciel permettra d'offrir fin 2014 une version *OpenSource*, *full web* et respectueuse de la norme ISO 25964-1: 2011 sur l'interopérabilité des thesaurus.

L'effort de FRANTIQ pour adapter ses outils à Internet, salué dès 1994 par Anne-Marie Guimier-Sorbets dans son rapport au département des SHS, s'est maintenu jusqu'à suivre les exigences actuelles du web de données. Il ne doit cependant pas minimiser l'aventure humaine qui reste le ciment du réseau. De ce point de vue, 2010 est l'année du changement. Après plusieurs demandes de GDR et de Centre de Ressources Thématiques, la création par l'InSHS du Groupement de services FRANTIQ marque enfin sa reconnaissance institutionnelle. Au moment où l'équipe de direction aborde le virage de la retraite, c'est aussi le moyen de pérenniser la structure et de lui assurer un financement récurrent. En même temps, elle oblige FRANTIQ à revoir son organisation. Les "producteurs" deviennent "adhérents" et, si leur nombre a varié au cours des décennies en fonction des regroupements ou des disparitions d'équipes CNRS, leur statut s'élargit aujourd'hui à tous les acteurs publics de l'archéologie en France par le biais de conventions de partenariat. Des groupes de travail permettent à ceux qui le souhaitent de s'approprier les outils et de s'impliquer dans l'amélioration du catalogue et de son contenu, dans l'enrichissement du thesaurus et le développement d'OpenTheso, dans la communication et la valorisation, dans la formation... Le fonctionnement des groupes est organisé par une charte validée en assemblée générale. Le Conseil scientifique, constitué d'archéologues et de spécialistes de l'information scientifique, s'est réuni pour la première fois en 2014, indépendamment de l'Assemblée annuelle. Une nouvelle dynamique d'échanges et de partage s'installe désormais qui laisse une large place à l'implication volontaire de chacun. La structure administrative de FRANTIQ n'autorisant l'affectation d'aucun personnel, la participation aux actions du GDS reste dépendante du bon vouloir et de la compréhension des directeurs d'unités adhérentes. Alors qu'il est aisé pour un bibliothécaire de faire valoir sa participation dans l'alimentation du catalogue partagé, bien commun s'il en est, il est parfois difficile de justifier de son implication dans des projets annexes tout aussi essentiels. Il faut cependant saluer l'appui historique de la Maison de l'Orient et de la Méditerranée à Lyon pour son infrastructure informatique et celui de la Maison Archéologie et Ethnologie à Nanterre pour l'environnement administratif, favorables à la continuité des missions de FRANTIQ.

Alors qu'en archéologie tout particulièrement la documentation reste cumulative, les publications y sont encore largement imprimées et le rétrospectif difficilement numérisable au regard notamment des droits liés à l'image. Les archéologues apprécient désormais l'accès direct à l'information scientifique et Internet est devenu incontournable dans leur pratique documentaire. Pourtant, ils continuent de fréquenter les bibliothèques qui assument toujours un rôle privilégié d'espaces de travail et de centres de ressources. Alors celles-ci doivent faire évoluer leurs catalogues : d'abord vers une diversification des ressources proposées, ensuite vers des formats de données enrichis et ouverts qui

les rendront plus visibles/lisibles et leur permettront de participer à la construction du web de données. Ce sont les objectifs que FRANTIQ s'est donné pour les années à venir.

Le réseau en 2014

35 adhérents représentatifs de l'archéologie publique française : 17 unités du CNRS ; 6 musées et services nationaux, 8 services régionaux de l'archéologie du Ministère de la Culture et de la communication ; 4 musées et services des collectivités territoriales et autres établissements publics

Au moins 6000 chercheurs, enseignants-chercheurs et étudiants avancés

90 professionnels de l'information + 1 ingénieur informaticien

20 personnels formés aux outils Frantiq chaque année

► <http://www.frantiq.fr>

contact&info

► Blandine Nouvel, Frantiq
Blandine.NOUEVEL@frantiq.fr

► Pour en savoir plus
<http://www.frantiq.fr/>

Aires culturelles. Vers une géométrie variable de la mondialisation

L'Institut des Amériques, Groupement d'Intérêt Scientifique créé en 2007, réfléchit depuis plusieurs années sur les aires culturelles et leurs reconfigurations dans un environnement global. Cette problématique, travaillée par certaines disciplines, s'interroge donc plus spécifiquement sur les articulations entre mondialisation et aires culturelles.

Fort de ses réflexions, l'IdA a décidé de consacrer son colloque annuel 2014 à cette thématique.

Organisé du 5 au 7 novembre prochain, au siège du CNRS, ce colloque sera l'occasion d'une réflexion scientifique internationale à trois niveaux : la genèse des aires culturelles ; les institutions scientifiques du champ des aires culturelles ; les enjeux et perspectives pour les aires culturelles.

Il réunira des chercheurs venant d'autres réseaux de recherche en aires culturelles, y compris issus d'autres pays.

L'InSHS, qui a fait des aires culturelles l'une de ses priorités scientifiques, a décidé ici de publier un entretien mené et proposé par Arezki Cherfaoui, coordinateur du colloque, avec Jean-Michel Blanquer, président de l'Institut des Amériques, qui nous livre sa position sur les aires culturelles et la mondialisation.

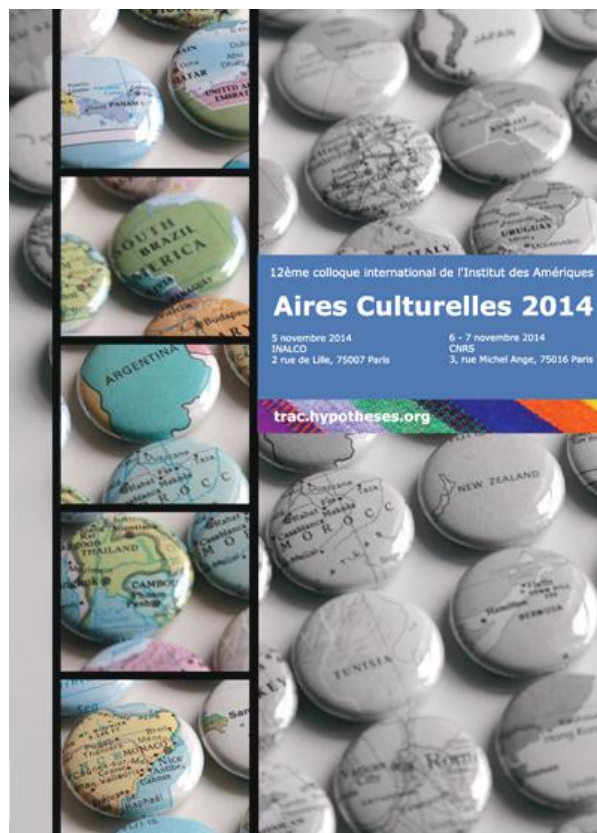
Jean-Michel Blanquer, vous êtes président de l'Institut des Amériques, un Groupement d'Intérêt Scientifique (GIS) créé en 2007, qui a vocation à penser de manière novatrice les liens entre l'Amérique du nord, l'Amérique latine et les Caraïbes. D'une façon plus générale, vous proposez à partir de l'expérience de l'Institut des Amériques, de réfléchir à la notion d'aires culturelles en mettant l'accent sur les liens, les connexions entre les différentes régions du monde. En quoi le cadre actuel de la globalisation nécessite de revisiter la notion d'aires culturelles ?

La notion d'aires culturelles est questionnée par la mondialisation dans le meilleur sens de ce terme, c'est-à-dire que les différentes caractéristiques de la mondialisation ont conduit, au cours des dernières années, à remettre en question le côté un peu figé qu'il y a dans la notion d'aires culturelles et, en même temps, à se rendre compte qu'on a besoin d'un ancrage territorial et concret pour analyser les différents aspects de la mondialisation. Du coup, cela a conduit forcément à remettre en cause cette notion, tout en montrant son caractère indispensable, quitte à réinventer le mot et ce qu'il recouvre comme réalité(s).

Concernant ce qu'il recouvre, quels sont justement les enjeux du point de vue épistémologique et méthodologique ?

Les enjeux sont considérables : sur le plan épistémologique, ils sont purement scientifiques et, en même temps, ils articulent des enjeux scientifiques à des enjeux autres, notamment des enjeux politiques.

Le colloque permettra de retracer une grande histoire, en évoquant notamment la façon dont on a découpé le monde dans les sciences humaines et sociales telles qu'elles ont pu croître au cours du vingtième siècle, elles-mêmes héritières de différentes traditions qui précédaient les sciences sociales. L'histoire de ce dé-



coupage est donc, en soi, une histoire des perceptions du monde qui peuvent être différentes d'un pays à l'autre, d'une région à l'autre mais aussi d'une discipline à l'autre. Ce qui fait que finalement, il n'y a jamais eu une histoire figée des aires culturelles ; c'est quelque chose qui a évolué sans arrêt.

Si on regarde quelques éléments forts — je pense particulièrement aux universités aux Etats-Unis au lendemain de la seconde guerre mondiale — on voit très clairement la nécessité qu'il y a à découper le monde pour l'appréhender. On le fait non seulement en étant héritier de certaines traditions européennes, mais aussi en étant dépendant de certains découpages imposés par le pouvoir politique aux Etats-Unis à cette époque, dans l'ambiance de la guerre froide.

Il y a des découpages du monde qui s'imposent du fait de cette logique et l'on voit bien comment cela a donné de la force aux aires culturelles, non seulement d'un point de vue conceptuel — parce qu'il y a un véritable lien entre la dimension scientifique et la dimension politique —, mais aussi sur le plan économique parce que, dans ces cas là, des moyens financiers sont donnés par les Etats aux chercheurs.

Tout ceci interroge les modes de fonctionnement académiques, les approches scientifiques. Il n'en demeure pas moins que c'est à travers ces aires culturelles que de véritables spécialistes ont pu se déployer et être très complémentaires sur des approches plus globales comme, par exemple, les approches venues des relations internationales *stricto sensu* ou de l'anthropologie et des diffé-

rentes sciences sociales. Un certain nombre de logiques s'enchevêtrent donc et créent un caractère vivant et évolutif des aires culturelles qui ont eu des hauts et des bas.

Mais ce qu'on a observé très vite, c'est qu'il a été commode de critiquer les aires culturelles, dès les années cinquante ou soixante, non pas parce qu'on pouvait y voir des impuretés — celles que j'ai signalées, politiques et économiques — mais pour leur éventuel fixisme, l'essentialisme culturel qu'elle pouvaient refléter. Il y avait donc des raisons philosophiques, politiques, scientifiques de critiquer les aires culturelles. Mais ces critiques ne sont jamais venues à bout des aires culturelles car on en a besoin, ne serait-ce que pour structurer les choses et puis parce que dès qu'on s'en abstrait, on sent vite le besoin d'y revenir.

Le cas typique, c'est l'intervalle de temps qu'on a vécu entre 1989 et 2001 : finalement, en 1989, avec la chute du mur de Berlin et les années qui ont suivi, il y a eu un engouement pour le concept de mondialisation et des idées comme la fin des territoires, la globalisation, la fin de l'histoire sont devenues directrices durant cette période. Cela a contribué à ce qu'on pourrait appeler une « ringardisation » supplémentaire des aires culturelles. Puis, arrive le 11 septembre 2001 et, sans être trop schématique, le rappel au réel. Par ce que signifie le 11 septembre 2001 sur le plan politique, on se rend compte de nouveau, un peu comme après 1945, qu'on a besoin de spécialistes, de compétences académiques capables d'éclairer la décision politique, économique, sociale et le champ culturel.

De ce point de vue là, la notion, si ce n'est le mot « aires culturelles », a connu un regain d'intérêt ces dernières années. Mais nous sommes désormais en présence d'un phénomène que j'appelle *la géométrie variable de la mondialisation* : finalement, quand on cherche à découper le monde, à l'articuler en catégories territoriales, on constate qu'une même partie du monde appartient à plusieurs sous-ensembles du monde. C'est cet enchevêtrement, ces intersections, ces problèmes de flux qui représentent une géométrie variable de la mondialisation, ce qui est très stimulant sur le plan intellectuel et scientifique et qui correspond au réel. De ce fait, il faut voir comment, aujourd'hui, le milieu académique est capable de prendre en compte cela pour avoir non seulement les outils institutionnels mais aussi les outils méthodologiques pour appréhender une telle réalité.

Du point de vue méthodologique, des pistes nouvelles s'ouvrent, notamment sur le comparatisme, le renouvellement de la cartographie ou la traduction. Or, les spécialistes opposent souvent disciplines et aires culturelles. Comment dépasser cette opposition aujourd'hui ?

Je ne pense pas que disciplines et aires culturelles s'opposent, mais plutôt qu'elles se questionnent mutuellement dans un sens favorable au deux. L'étude des aires culturelles ou ce qu'il faut appeler ainsi — le but du colloque, c'est aussi de déboucher sur un autre mot, de le faire évoluer — vient croiser très utilement la question disciplinaire parce qu'elle représente un des principaux vecteurs d'une pluridisciplinarité qui est de plus en plus indispensable dans un monde complexe où tout pousse le monde académique vers l'hyperspécialisation.

Je voudrais rappeler que le mot « aires culturelles » a, en français, une connotation qu'il n'a pas dans d'autres langues. Quand on parle d'*Area Studies*, on n'a pas plus le mot « aires culturelles » et tout ce qu'il charrie. Pour ma part, si je réfléchis à ma propre

pratique, j'ai été amené à croiser des disciplines parce que j'étais un spécialiste de l'Amérique latine. Je viens du Droit public. Il était évident que je ne pouvais pas rester étranger à des problématiques de science politique, que je devais tenir compte de la géographie et de la sociologie dans tout ce que j'ai pu faire. Si je suis arrivé, par exemple, en étudiant la Colombie, au concept de *société de droit*, c'est évidemment parce qu'au-delà du droit public qui était mon point de départ, j'ai eu des approches de science politique et de sociologie indispensables pour traiter ce que j'étudiais et comprendre ce que je voyais de la part du juge constitutionnel colombien ou des différents acteurs du droit dans la société colombienne. Donc, cette pluridisciplinarité indispensable vient aussi du regard extérieur qu'on est capable de porter quand on étudie les aires culturelles.

Je crois qu'il faut assumer ce regard extérieur parce que s'il y a un reproche courant fait aux aires culturelles, c'est finalement celui de l'impérialisme du regard. Mais regard il y aura toujours de part et d'autre. C'est-à-dire que celui qui veut substituer un regard à un autre va toujours introduire une nouvelle subjectivité, ce qui est inévitable et sans doute positif dès lors qu'on sait mettre tout cela en perspective. Donc, de ce point de vue là, ce qui me paraît intéressant dans la question des aires culturelles, c'est qu'elle invite au moins à un regard sur autrui, le moins impérialiste et le plus subtil possible. Il est évidemment souhaitable qu'il soit exempt de subjectivité, mais c'est un idéal mais qui ne sera jamais atteint. C'est plutôt à l'intersubjectivité qu'il faut inviter, au service d'idéaux à la fois intellectuels et éthiques.

D'autres groupements d'intérêt scientifiques sur les aires culturelles ont été créés l'année dernière. Quel impact peut-on en attendre pour la structuration de la recherche, voire de l'enseignement supérieur ?

Dans le cas de l'Institut des Amériques, l'idée a commencé à poindre à la fin des années 90, alors que nous étions convaincus qu'un certain nombre de phénomènes de la mondialisation conduisaient non pas à supprimer la notion d'Amérique latine, mais à la mettre en perspective et à être capables d'avoir une approche plus complexe de l'aire que nous étudions. La frontière Mexique – Etats-Unis devenait, par exemple, un objet d'étude extrêmement important depuis un certain nombre d'années. Des phénomènes comme l'immigration latino-américaine à l'intérieur des Etats-Unis, l'intégration économique à l'échelle du continent, le nouveau rôle en matière de relations internationales des grands pays émergents, les différents flux de personnes de biens, de marchandises devenaient des objets d'études qui invitaient à avoir une approche transaméricaine. Sans parler de l'histoire et de la géographie qui invitent à une telle approche. Ce faisant, nous avions plusieurs objectifs dont celui de renouveler nos problématiques scientifiques, de créer des effets de masse critique entre différents milieux scientifiques, de créer de l'interdisciplinarité, du comparatisme, bref de dynamiser les approches scientifiques et de recréer des logiques collectives en matière scientifique.

C'est évidemment un travail de longue haleine qui n'est pas achevé aujourd'hui et qui a conduit à de grands progrès autant institutionnels que scientifiques dans notre champ. Cela nous a permis de fédérer plus d'une cinquantaine d'établissements, de nous organiser en régions, d'identifier des priorités scientifiques, d'encourager des jeunes chercheurs dans de multiples manifestations, au travers de bourses, de pôles sur le continent américain... Autant d'éléments qui ont été des stimuli, sans pour autant abandonner l'Amérique latine, les Etats-Unis et le Canada,

ni même les autres sous ensembles, par exemple les Caraïbes qui sont elles-mêmes totalement transaméricaines puisqu'elles correspondent à différentes aires culturelles au sens classique du mot. Si vous prenez l'exemple des Caraïbes, il est tout à fait pertinent d'avoir une approche en terme de géométrie variable de la mondialisation.

Avec l'Institut des Amériques, on a donc une série de dynamiques qu'il nous a semblé pertinent d'ordonner au travers d'un GIS et nous l'avons toujours fait en dialogue étroit avec le CNRS. C'est d'ailleurs pourquoi le CNRS a prêté attention à ce qu'on a fait et a eu une aspiration comparable avec d'autres aires culturelles. Nous sommes ainsi très heureux d'avoir des collègues regroupés à travers un GIS Asie, un GIS Afrique et d'autres GIS qui sont eux-mêmes dans ce type de logique où tout le monde a à gagner. Parce qu'une autre vertu de ce mouvement, c'est de pouvoir comparer les aires culturelles entre elles. Si vous prenez, par exemple, le déploiement considérable des relations Asie-Amérique latine, d'ailleurs multiforme, qui ne concerne pas seulement les questions de commerce et de politique, c'est assez pertinent de pouvoir faire communiquer des milieux entre eux.

Nous venons d'évoquer le cas de la France qui évolue dans un milieu académique en restructuration. Comment les autres systèmes nationaux de recherche évoluent ? Se pose-t-on les mêmes questions ?

Un des buts du colloque est évidemment de faire le point là-dessus, sachant qu'il est difficile d'être exhaustif puisqu'il y a de multiples systèmes nationaux de recherche. Nul ne peut se considérer omniscient sur une telle question. Pour ma part, ce que j'observe, c'est que d'autres pays se posent des questions semblables. En Angleterre, s'est créé un Institut des Amériques à partir de questions comparables aux nôtres. A une échelle européenne, nos collègues latino-américanistes ont regardé avec beaucoup d'intérêt ce que nous avons fait. De façon plus générale, je pense que le monde entier se pose des catégories de questions à peu près comparables : quelles conséquences de la mondialisation sur notre manière de percevoir le monde ? Quelle alternative pour des approches nouvelles ? Comment s'approprier ces visions ? Comment s'inspirer des autres dans notre manière de découper le monde ? Comment prendre en compte non seulement des réalités géographiques qui s'imposent à nous, par exemple l'étude d'un continent, mais aussi des réalités culturelles plus subtiles, plus mouvantes, des réalités économiques, géopolitiques ? Comment avoir ainsi différents cercles qui se dessinent et sont en intersection.

Pour conclure, la globalisation invite davantage à un renforcement de la notion d'aire culturelle plutôt qu'à sa dissolution. A quelles conditions ?

Je pense que les conditions sont d'ordre scientifique et éthique. D'abord, il s'agit d'étudier scientifiquement la globalisation dans toutes les disciplines des sciences sociales, mais pas seulement. Il est évident que la globalisation n'est qu'un mot si elle ne s'appuie pas sur une approche par les territoires et sur l'étude des allers-retours entre les réalités locales, régionales et les réalités globales, parce que ce sont des phénomènes d'interaction qui se manifestent.

Quand vous avez tel type de musique qui devient une musique globalisée, elle se transforme en se globalisant mais elle a un point de départ territorial qui ne vient pas que des Etats-Unis.

Quand elle est globalisée et qu'elle revient sur le territoire, elle transforme aussi ce territoire. Donc, cette dialectique du local et du global est absolument fondamentale sur le plan scientifique.

Puis, il y a une dimension éthique et politique qui fait que, derrière cela, se pose la question du monde qui se dessine. Quelle dialectique y-a-t-il entre l'approche académique et les autres approches, politiques et économiques en particulier ? Dans quelle mesure l'approche académique contribue-t-elle à un monde qui ne soit pas complètement standardisé, à un monde qui édicte de la norme, juridique ou morale, venue d'un seul point de domination ? Comment, au contraire, permet-elle la compréhension de la diversité, la préservation et le dialogue dans cette diversité ? Ce sont des mots qui sont davantage utilisés dans la sphère politique que dans la sphère scientifique mais, en réalité, il ne faut pas se voiler la face sur le fait que dans le passé, le monde scientifique a pu avoir un impact sur ce genre de questions — si on prend l'histoire de l'anthropologie, par exemple, c'est une évidence — plus souvent dans le sens de la standardisation que dans le sens de la diversité. Demain, compte tenu des grands enjeux mondiaux qui sont devant nous, de tous ordres mais dont beaucoup se rassemblent autour de la question de la diversité humaine, de la coexistence pacifique entre ces diversités et aussi, éventuellement, de leur capacité d'interaction pacifique, il est évident que l'enjeu académique fait partie de cela. Faire semblant de ne pas voir cela est en fait ascientifique.

Les propos tenus ici sont bien entendu de la responsabilité de leur auteur.

Propos recueillis par Arezki Cherfaoui, coordinateur du colloque Aires Culturelles

contact&info

► Arezki Cherfaoui,
Université de Versailles
Saint-Quentin-en-Yvelines
arezki.cherfaoui@yahoo.fr

► Pour en savoir plus
<http://trac.hypotheses.org/298>

ZOOM SUR...

Crowdsourcing, participation, démocratie : le citoyen comme ressource et comme acteur de la connaissance



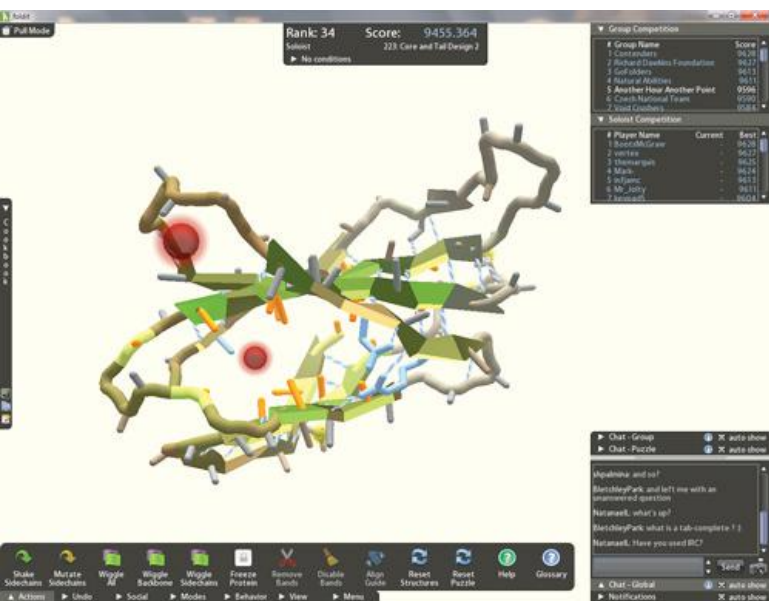
L'arrivée des *Big Data* — la production massive d'informations par des expériences scientifiques, des capteurs ou des populations munies d'équipements de communication — n'est pas seulement une révolution technologique, c'est une nouvelle méthode de découverte : elle ne remplace pas les autres méthodes (hypothético-déductive, expérimentale, par modélisation), mais y superpose une approche radicalement empirique, qui trouve la connaissance dans les données sans passer par la règle ou les modèles imposés, les généralisations ou dispositifs causaux hypothétiques... L'extraction de savoirs, la sélection et l'agrégation de données, la visualisation et la navigation dans de grands espaces de données sont autant de modalités nouvelles de la connaissance, de la perception des phénomènes, et de l'action. De nombreuses recherches scientifiques (en santé, sciences du vivant, linguistique, sciences de l'univers, chimie quantique, études de populations, imagerie...) produisent désormais des volumes considérables de données hétérogènes nécessitant non seulement le développement de nouvelles méthodes pour leur visualisation, leur sélection et leur expression, mais aussi un examen épistémologique attentif des transformations qu'elles engagent et de ce nouveau paradigme, dont le *crowdsourcing* (recueil de données en ligne à partir d'une sollicitation ouverte et en masse) constitue une spectaculaire exemplification.

Les *Big Data* impliquent un investissement important dans des outils de recueil, d'archivage, d'analyse et de visualisation et donc une intégration accrue des sciences de l'information dans le travail des chercheurs, devenant « *data scientists* ». Mais ils mettent en œuvre aussi bien les données fournies par nous tous,

involontairement (par les traces qu'il/elle laisse sur le réseau, l'identité numérique constituée progressivement par chacun-e et qui fournit le thème du film *HER* de Spike Jonze) et parfois volontairement, dans des programmes de recueil de données par les citoyens ordinaires. Il s'agit de nouvelles ressources, et de formes d'ouverture et de démocratisation de la production des savoirs qui suscitent des débats et des inquiétudes : comment garantir la qualité de données et savoirs ainsi produits ? Comment éviter l'exploitation, y compris financière, des données et de ceux qui les offrent à titre gratuit ? Que penser de l'intervention de nouveaux acteurs non professionnels dans le processus de constitution de la connaissance ?

Ces nouvelles modalités de l'innovation et de la découverte sont en tout cas déjà à l'œuvre dans de nombreux programmes et à l'international, notamment aux Etats-Unis, et permettent d'emblée la compréhension de nouveaux phénomènes. Les analyses des *Big Data* donnent un rôle essentiel autant à l'examen des tendances lourdes qu'à des faits mineurs émergeant dans les masses de données ; elles changent la donne de la décision sociale et scientifique, par la prise en compte de données négligées ou d'indicateurs reconfigurés par de nouvelles exigences de la société ; et, dans une période où l'on évoque souvent les « changements des comportements » à imposer, elles permettent d'envisager autrement les phénomènes et transformations en cours dans la société, les changements et processus réellement à l'œuvre mais pas toujours perçus, et la complexité des facteurs et indicateurs à prendre en compte dans les décisions et politiques publiques.

Les *Big Data* et l'analyse empirique de grandes masses de données ouvrent ainsi sur une mise en capacité (*empowerment*) de chacun d'agir en fonction de l'information disponible et de faire prendre en compte les enjeux qui comptent pour lui. Les analyses purement rationalistes des processus qui conduisent de la science à la décision et l'action (raisonnement, délibération, cognition, procédures de choix publics, théorie des jeux, choix rationnel...) font désormais face aux limites intrinsèques de la connaissance (*known unknowns/unknown unknowns*), aux constructions de l'ignorance, aux difficultés de la prise de décision en contexte d'incertitude et de controverse. La décision publique doit désormais reconnaître la compétence des citoyens : la démocratie se définissant comme gouvernement de la participation égale de tous, sans distinction relative à leur niveau présent de connaissances, dans les domaines qui font l'objet des décisions collectivement prises. Elle doit prendre en compte le développement de dispositifs de consultation multiples et concurrents et inclure dans l'établissement d'indicateurs pertinents des responsables publics, des experts scientifiques et des citoyens ordinaires. Les méthodes en développement à partir



Interface du jeu vidéo Foldit
© Animation Research Labs, University of Washington

des *Big Data* (réseaux sociaux, *crowdsourcing*, design des processus de décision et de la participation, mesures et indicateurs, interprétation des signaux faibles et alertes, etc.) sont désormais, pour les sciences humaines et sociales, une source de nouveaux outils démocratiques autant que rationnels.

Ces procédures d'analyse empirique et d'ouverture des données conduisent à modifier le rôle des institutions qui doivent fournir et diffuser ces données, comme la place des chercheurs qui, devenant analystes et diffuseurs des données, entrent en dialogue avec des communautés interdisciplinaires constituées en des réseaux internationaux élargis. Wikipedia est le meilleur exemple de ce type de fonctionnement et de la valeur ajoutée de la participation du plus grand nombre. L'apport des *Big Data* va de pair avec une approche pragmatique, interdisciplinaire, décloisonnée, des faits scientifiques, sociaux et culturels. Ce tournant épistémologique conduit à la production de logiques qui sont moins définitionnelles et essentialistes que descriptives, transactionnelles, expérimentales.

C'est dans ce nouveau paradigme de la production des connaissances que s'inscrit le projet des *Citizen Sciences* — la production et le recueil de données de masse par des citoyens, des non-scientifiques-professionnels. *Foldit* et *Seti@home* en sont de parfaits exemples.

La collecte en ligne de données massives à l'échelle internationale par *crowdsourcing* constitue une forme de révolution scientifique et politique, mettant en place une action participative qui propose à des individus d'origines et de caractéristiques diverses de participer à la science. La collaboration, pilier du web 2.0, est à la base du *crowdsourcing*, qui ne se fait pas aisément. Pour garantir la qualité des données, le réseau doit non seulement se construire avec des échanges entre les participants mais aussi avec les organisateurs et les scientifiques qui animent le projet. La question de l'évaluation, de la présentation et de la pertinence des données est ainsi l'enjeu central du *crowdsourcing*.

D'autres projets, par leurs objets, invitent les citoyens à avoir un engagement plus fort : par exemple, la collecte citoyenne des données de mesure de la radioactivité au Japon après Fukushima (par les stations de mesure, l'équipement et l'instrumentation

individuelle), l'apparition de dispositifs autonomes permettant, à partir du partage de matériel de mesure, de faire circuler des données indépendantes et de générer des cartographies de la contamination de l'air, des sols, de l'alimentation et de l'eau sur les réseaux sociaux, les mesures citoyennes de la radiation sont un exemple de *crowdsourcing* qui conjugue l'*empowerment* citoyen et la production d'une contre-connaissance contestant les affirmations officielles des experts.

Il faut sans doute tenter de distinguer la méthode *crowdsourcing* (externalisation du recueil des données, acceptable par tous) de la collecte plus militante et contestataire, et des sciences citoyennes (*Vigie-nature*) qui élaborent la connaissance à partir des données recueillies par des citoyens ordinaires. Il n'en reste pas moins que cette méthode, en valorisant les capacités innovantes et scientifiques de la société elle-même, ébranle certains privilèges attachés aux scientifiques professionnels et permet une forme de démocratisation de la science qui est compatible avec son objectivité, voire la renforce.

Le CNRS a déjà soutenu un certain nombre de projets de *crowdsourcing* — financés par le programme Mastodons de la Mission pour l'interdisciplinarité (biodiversité, *Crowdhealth*), dans le cadre du PEPS HuMaIn, ou de programmes de critique littéraire ou d'analyse de manuscrits en ligne (*folksonomies*). Mais rares sont encore les projets de *crowdsourcing* proprement SHS.

Il revient aux SHS de s'engager dans un examen attentif des enjeux sociaux, éthiques et politiques du *crowdsourcing*, afin de lancer et d'enrichir la réflexion au CNRS autour des questions touchant au caractère public et ouvert de la science, encore relativement tabou ; de présenter et de discuter les compétences, protocoles et approches en sciences humaines et sociales dans le domaine de l'étude des méthodologies scientifiques et de la recherche ancrée dans les communautés et les groupes sociaux ; d'analyser et de consolider les dispositifs collaboratifs, la prise en compte des parties prenantes dans les processus de décision ; de déceler et d'analyser les différentes formes d'exploitation du travail gratuit ou nano-rémunéré que crée le *crowdsourcing* ; les techniques de soumission ou de contrôle qui sont aussi parfois à l'oeuvre dans les dispositifs participatifs ; de tracer, donc, les contours, normes et méthodes d'une science citoyenne à partir des capacités avérées du public à organiser et réguler le recueil et le partage des données et d'information — du droit de tous à intervenir dans le traitement de questions qui les concernent, le domaine que le philosophe John Dewey appelle « Le Public » et auquel renvoie aujourd'hui le nom de *démocratie*.

Ainsi, à partir du développement du *crowdsourcing* comme méthode scientifique de découverte, on pourra ouvrir le débat sur le rôle déjà bien réel des citoyens dans la recherche scientifique et sur les possibilités, multiples aujourd'hui dans notre économie de la connaissance et de l'information, de faire progresser conjointement l'ouverture des données, la participation citoyenne ET l'autonomie de la recherche.

Sandra Laugier, InSHS



© CNRS Photothèque/LMGE - ROUDEL Mathieu

Le crowdsourcing en biodiversité : l'expérience Vigie-Nature

Professeur au Muséum national d'histoire naturelle, Denis Couvet est membre du Centre des Sciences de la Conservation (CESCO). Il nous livre ici un exemple de ce que peuvent être les sciences participatives en biodiversité.

Dans le domaine de l'environnement, les recherches scientifiques bénéficient largement de l'apport des citoyens : plus de 500 000 expériences participatives sont répertoriées dans le monde (Conrad et Hilchey, 2011), notamment dans le domaine de la biodiversité. En France, le Muséum national d'Histoire naturelle a mobilisé les citoyens volontaires depuis sa création, notamment pour constituer ses collections et ses bases de données sur le patrimoine naturel : des naturalistes amateurs organisés en sociétés savantes remontent spontanément données et spécimens. Plus récemment, le programme Vigie-Nature propose depuis plus de 20 ans à des volontaires bénévoles de participer à des programmes scientifiques de suivi des populations animales et végétales, afin d'évaluer l'état et la dynamique de la biodiversité qui nous entoure : des chercheurs académiques proposent de suivre des protocoles pour répondre à une question explicite. Proposé au départ uniquement à des naturalistes, le programme Vigie-Nature s'adresse depuis presque 10 ans à des observateurs citoyens non avertis, en proposant des protocoles de suivi adaptés à un public néophyte.

Toujours caractérisés comme une relation entre des scientifiques et des non-scientifiques, les dispositifs de sciences participatives varient ainsi entre deux extrêmes :

- ▶ Un type '*transformatif*', qui s'appuie sur les parties prenantes non scientifiques, qui construisent ou verbalisent des questions avant de contacter des équipes de recherche susceptibles de pouvoir les aider à y répondre. C'est le modèle actuel d'associations comme la Fondation Sciences Citoyennes.
- ▶ Un type '*consultatif*', dans lequel les scientifiques définissent les problématiques et les protocoles de suivi, puis recrutent les

observateurs (individus ou groupes sociaux) susceptibles d'être intéressés. C'est le modèle actuel de Vigie-Nature.

Néanmoins, dans le cadre de ce type consultatif, il semble plus ambitieux de considérer les participants comme des interlocuteurs exigeant des laboratoires, demandant des analyses en retour, compréhensibles des non spécialistes, susceptibles en conséquence d'avoir un large impact politique. Dans cette vision idéale, voire irénique, les observateurs sont des éléments moteurs de la diffusion des résultats, de leurs implications, dans les différents groupes sociaux, les données et les résultats leur appartenant (Couvet et al. 2008). Cette appropriation des résultats, leur diffusion par les participants, doivent enclencher un cercle vertueux conduisant au recrutement de nouveaux observateurs.

Organisation consultative : une efficacité démontrée quand les observateurs sont des naturalistes

Les observatoires citoyens de biodiversité sont particulièrement adaptés pour aider aux expertises sur la biodiversité. De la même façon que pour le climat ou l'économie, ces observatoires permettent aux scientifiques de disposer d'une représentation synthétique de la biodiversité et de sa dynamique sur l'ensemble des écosystèmes et à une échelle spatiale pertinente (nationale ou régionale). En effet, ces observatoires permettent de mobiliser un grand nombre de bénévoles, sur un maillage fin du territoire national.

Dans ces observatoires pilotés par les chercheurs, les volontaires collectent des données (présence ou abondance d'espèces, phé-

nologie) selon des protocoles standardisés construits par des scientifiques (Couvet et al. 2011). De cette standardisation des données dépend la qualité des informations recueillies et des indicateurs et scénarios produits.

Le programme Vigie-Nature est un bon exemple de la qualité de la recherche (e.g. Julliard et al. 2006, Devictor et al. 2008) et de l'expertise produites à partir de ces observatoires, en particulier le programme STOC (Suivi temporel des oiseaux communs) qui s'adresse à la communauté des ornithologues amateurs (revue dans Jiguet et al. 2012).

En effet, cette organisation montre toute son efficacité dans un cadre scientifique normalisé qui regroupe une formalisation théorique et une méthode d'analyse issues des sciences écologiques. Elle est relativement facile à mettre en œuvre avec des observateurs naturalistes, qui ont les mêmes objectifs de connaissances écologiques que les chercheurs. Les indicateurs et scénarios produits répondent aux exigences normatives de l'écologie scientifique et sont proposés aux décideurs. Mais dans quelles conditions répondent-ils réellement à la demande sociale ?

Ouverture aux observateurs non naturalistes

Une première ouverture des échanges entre scientifiques et non scientifiques est entamée par la construction d'observatoires de biodiversité pour un public non naturaliste. C'est le cas de plusieurs observatoires de Vigie-Nature, dont le SPIPOLL (Suivi photographique des insectes pollinisateurs ; Deguines et al. 2012). Dans ces observatoires, les observateurs sont intéressés par les questions d'environnement, mais ils ne sont pas spécialistes des groupes d'espèces qu'il est proposé de suivre. Beaucoup ne font qu'une visite brève dans le processus, mais quelques-uns sont fidèles : ceux-là témoignent volontiers d'un changement de regard progressif de leur part, non seulement vis-à-vis des espèces qu'ils suivent, mais aussi plus largement de la biodiversité de leur quotidien (Cosquer et al. 2012).

Au-delà de la collecte de données de suivi de biodiversité (qui sont utilisées comme celles récoltées par les naturalistes), ces observatoires permettent un changement d'attitude des citoyens volontaires vis-à-vis de la nature qui les entoure (par l'acquisition de connaissances et un nouvel intérêt), et peut-être, ensuite, vis-à-vis des enjeux de conservation de la biodiversité. Cependant, la position du discours scientifique reste prépondérante, les volontaires devant toujours se conformer à une norme établie par la science. La communauté scientifique dans son ensemble tarde d'ailleurs à reconnaître la valeur des données récoltées par des non-naturalistes, ce qui retarde l'ouverture d'un dialogue avec ces derniers.

Ouverture vers les citoyens : co-construire les questions et protocoles de suivi

Une façon d'ouvrir ce dialogue est d'imaginer des procédés qui permettent aux non-scientifiques de participer à plusieurs étapes du processus de recherche, notamment à la construction de nouvelles questions de recherche.

Il ne s'agit pas ici de recueillir simplement des idées de questions intéressant les citoyens, pour ensuite les formaliser dans le cadre théorique de l'écologie scientifique, mais plutôt de proposer un processus qui permette d'utiliser la biodiversité comme levier pour combler le fossé qui existe entre scientifiques/experts et citoyens.

Par le biais d'animations participatives développées dans d'autres cercles (éducation populaire par exemple) telles que le forum ouvert, le programme Vigie-Nature a pour ambition d'intégrer les différentes communautés participantes dans la construction de nouveaux observatoires.

En bref, il y a sans doute un aller-retour à faire, une synergie à obtenir, entre les démarches transformatives et consultatives.

En encourageant le dialogue entre les scientifiques et les non-scientifiques, ces démarches auraient bien sûr un rôle d'éducation. Cependant, elles demanderont également aux scientifiques d'adopter une démarche réflexive pour bien comprendre pourquoi et comment ils s'engagent dans ces activités, comme scientifiques mais aussi comme citoyens.

**Anne-Caroline Prévot, Lisa Garnier,
Romain Julliard et Denis Couvet**

► Références

- Conrad, C.C., and Hilchey, K.G. 2011. *A review of citizen science and community-based environmental monitoring: issues and opportunities*. Environ. Monit. Assess. 176, 273-291
- Cosquer, A., R. Raymond, and A.-C. Prévot-Julliard. 2012. *Observations of everyday biodiversity: a new perspective for conservation?* Ecology and Society 17(4): 2.
- Couvet, D., Jiguet, F., Julliard, R., Levrel, H. and A. Teyssède. 2008. *Enhancing citizen contributions to biodiversity science and public policy*. Interdisciplinary Science Reviews 33, 95-103
- Couvet, D., Devictor, V., Jiguet, F., and Julliard, R. 2011. *Scientific contributions of extensive biodiversity monitoring*. C. R. Biologies 334, 370-377.
- Deguines N, Julliard R, De Flores M, Fontaine C (2012) *The whereabouts of flower visitors: contrasting land-use preferences revealed by a country-wide survey based on citizen science*. PLoS ONE. 7: e45822.
- Devictor V., Julliard R., Couvet D., and Jiguet F. 2008b. *Birds are tracking climate warming, but not fast enough*. Proc Roy. Soc. B 275, 2743-2748
- Jiguet F, Devictor V, Julliard R, Couvet D (2012) *French citizens monitoring ordinary birds provide tools for conservation and ecological sciences*. Acta Oecologica. 44, 58-66.
- Julliard, R., Clavel, J., Devictor, V., Jiguet, F. and Couvet, D. 2006. *Spatial segregation of specialists and generalists in bird communities*. Ecology Letters 9, 1237-1244

contact&info

► Denis Couvet,
CESCO

couvet@mnhn.fr

► Pour en savoir plus
<http://vigienature.mnhn.fr/>

Le crowdsourcing et la recherche en France

Sihem Amer Yahia est directrice de recherche au sein du Laboratoire d'Informatique de Grenoble (LIG), rattaché à l'Institut des sciences de l'information et de leurs interactions (INS2I) du CNRS. Dans les recherches qu'elle mène sur le crowdsourcing, elle insiste sur la nécessité de faire travailler les disciplines entre elles. Oui, informatique et sciences sociales peuvent interagir !

La recherche dans le domaine du *crowdsourcing* est en pleine effervescence aujourd'hui. Wikipedia, un exemple particulièrement brillant de ces méthodes collaboratives, définit le *crowdsourcing* comme la production participative qui consiste en l'utilisation de la créativité et du savoir-faire d'un grand nombre de personnes, en sous-traitance en quelque sorte, pour réaliser certaines tâches traditionnellement effectuées par un employé ou un entrepreneur. Les personnes impliquées dans une tâche sont parfois rémunérées. Il existe deux types de tâches dans le *crowdsourcing* : des tâches binaires qui sont effectuées par chaque individu indépendamment et en parallèle et des tâches collaboratives qui nécessitent l'implication de plusieurs personnes à la fois pour être accomplies de façon complète.

Les tâches peuvent être aussi simples que la reconnaissance d'un monument sur une photo, ou plus sophistiquées telles que l'enrichissement de méta-données (annotations textuelles d'images et de vidéos dans Flickr ou YouTube, la description d'un aliment dans *OpenFoodFacts*), une opinion (critique de restaurant ou de film), un travail collaboratif (alignement d'anciennes cartes géographiques dans <http://maps.nypl.org/warper/>, identification d'étoiles dans GalaxyZoo, reconnaissance de caractères dans Recaptcha qui a pour caractéristique d'utiliser le travail des usagers de Google pour déchiffrer des manuscrits) et, enfin, l'édition collaborative (Wikipédia encore).

Le *crowdsourcing* a la particularité d'intéresser des chercheurs de domaines très variés — des STIC aux SHS en passant par la santé, les sciences de l'univers, etc. — et a le potentiel de susciter des recherches interdisciplinaires, ne pouvant réellement être analysées sans prise en compte de multiples facteurs, informatiques bien sûr, mais aussi sociaux en raison de la diversité des acteurs impliqués.

Les problématiques abordées en informatique et en sciences sociales sont tout particulièrement complémentaires, voire inséparables. En informatique, les compétences nécessaires vont de la modélisation des utilisateurs en ligne au développement d'algorithmes performants pour la gestion des utilisateurs et des tâches en passant par les statistiques et l'apprentissage automatique. En sciences sociales, l'étude de la dynamique des communautés en ligne et des relations sociales qu'elles créent et recomposent, ainsi que la compréhension des mécanismes de compensation, d'engagement et d'encouragement des individus pour effectuer des tâches sont autant de sujets qui nécessitent une attention particulière.

En informatique, les défis abordés concernent le développement d'algorithmes à grande échelle pour apprendre les compétences des personnes impliquées dans la réalisation de tâches afin de mieux cibler les demandes de participation et également de mieux rémunérer les personnes impliquées. De nouveaux algorithmes sont aussi développés pour le problème d'assignation de tâches qui permettent d'optimiser la qualité des résultats obtenus, en particulier pour les tâches nécessitant des experts dans des domaines différents. Un bon exemple d'une telle créativité est le cas des tâches de *funsubbing* (sous-titrage décalé inventé par les fans) qui consiste à sous-titrer une vidéo déjà existante ou connue. En effet, cette tâche nécessite de comprendre les langues utilisées dans la vidéo pour créer les sous-titres appropriés. Elle requiert également d'appliquer des techniques de segmentation de la vidéo en plusieurs images successives et des techniques de synchronisation du son et de l'image. Enfin, lorsque les sous-titres souhaités sont dans une autre langue, il est impératif d'impliquer des personnes qui maîtrisent la langue cible. Ainsi, un choix ciblé des personnes impliquées et une coordination entre ces personnes sont indispensables pour la réussite de cette tâche complexe. Les systèmes actuels sont très limités et n'offrent pas la possibilité d'automatiser ce processus, qui demande l'intervention de sujets compétents et impliqués.

Le CNRS a organisé une [journée sur le crowdsourcing](#) le 15 septembre 2014 afin de réunir les chercheurs de toutes disciplines qui, en France, s'intéressent à cette thématique. Cette journée a regroupé en particulier des chercheurs en sciences sociales et en informatique. Elle était organisée dans le cadre de MASTODONS, une action CNRS pilotée par la Mission pour l'interdisciplinarité, pour la recherche interdisciplinaire sur de grandes masses de données. Les conférences invitées ont couvert des aspects de calcul impliquant des humains, une description de *Crowd4U*, une plateforme académique pour le *crowdsourcing* qui se lance en France, et un exposé sur les jeux en ligne.

Dans mon récent [SIGMOD blog post](#), j'ai évoqué le besoin de développer une plateforme ouverte et générique pour le *crowdsourcing* afin de permettre aux chercheurs de différentes disciplines de tester leurs modèles, hypothèses et algorithmes et de confronter les questions que soulève le *crowdsourcing* dans chaque domaine. En effet, malgré la popularité du *crowdsourcing*, le monopole en est quasiment détenu par Amazon, à travers sa plateforme *Amazon Mechanical Turk* : elle permet de faire exécuter des tâches de tout type et d'inviter les clients d'Amazon à participer à leur réalisation moyennant une compensation sous forme de bons d'achat Amazon. Cette plateforme est toutefois privée et fermée. Par exemple, elle ne permet pas de tester de nouvelles approches pour l'assignation de tâches aux travailleurs ou d'étudier la dynamique entre les travailleurs dans le cadre de tâches collaboratives. De plus, la compensation monétaire n'est pas toujours une bonne approche pour la participation, car elle se fonde sur les inégalités de situation, elle ne garantit ni la qualité du résultat, ni les droits des travailleurs et surtout, elle produit un profit important pour l'entreprise pour un salaire minimum. On touche ici aux limites du *crowdsourcing*, qui de pratique désintéressée pour les uns peut rapidement tourner à l'exploitation rentable pour les autres. Mais une réponse est, en ce domaine comme en d'autres, l'ouverture des données. Il est donc essentiel aujourd'hui de mettre à disposition des chercheurs une plateforme ouverte gratuite.

contact & info

► Sihem Amer-Yahia,
LIG

Sihem.Amer-Yahia@imag.fr

Des sciences participatives aux sciences citoyennes

Alexandre Gefen est chercheur au Centre d'étude de la langue et de la littérature françaises des XVII^e et XVIII^e siècles (CELLF). Spécialiste des humanités numériques, il rappelle à quel point le développement de nouveaux médias a permis au grand public de contribuer à la construction du savoir.

Internet 2.0, centré sur l'interaction et la contribution des internautes, a ouvert la voie à des formes originales de participation : forums de discussion, démocratie participative, écriture collaborative, financement collectif (*crowdfunding*), consommation coopérative, journalisme amateur, etc. : des centaines de milliers de contributeurs de Wikipedia aux dizaines de milliers de programmeurs au noyau du logiciel libre Linux (qui alimente la majorité des serveurs de l'internet) coécrit par des milliers de contributeurs, le monde et la culture numérique sont à la fois producteurs et produits de formes de travail participatives. Le site GalaxyZoo.org propose au grand public de contribuer au classement d'un million de photographies de galaxies prises par des instruments professionnels, certains scientifiques utilisent la plateforme de microfinancement Kickstarter pour le lancement de projet de recherche dans un contexte d'économies budgétaires¹ ; en France, le Muséum national d'Histoire naturelle a lancé des **projets pionniers**, une équipe de l'université de Rouen, le Centre Flaubert, fait appel à la communauté régionale des amateurs de l'écrivain normand pour la transcription de 4600 folios des manuscrits de *Madame Bovary*, le CNRS envisage le lancement de plusieurs projets faisant appel à des bénévoles pour l'étude de la pollution et du climat : les modèles et modes de fonctionnements sont riches et variés². Mais ces formes de connaissance et d'intelligence collectives, que l'on peut regrouper dans le domaine scientifique sous le terme de *crowdsourced sciences*, « sciences participatives », ont pour point commun d'être ouvertes et gérés par des protocoles de recherche eux-mêmes convenus ensemble et donc de se rattacher à une conception désintéressée de la recherche : tantôt désignées comme une révolution scientifique majeure³, tantôt critiquées⁴, elles s'appuient sur des modèles directs, « désintermédiés », de coproductions des savoirs en réseau, dont il importe de mesurer les ambitions autant que les difficultés.

La science des amateurs

La première remarque que nous invitent à faire la sociologie et l'histoire des sciences, c'est que, dans le temps long, de telles pratiques sont anciennes : c'est par milliers que les sociétés savantes du XIX^e siècle ont contribué à la construction du savoir moderne ; la distinction entre amateurs et professionnels de la science étant récente et parfois encore contredite — comment ne pas rappeler que Charles Darwin a dû financer une partie de ses propres recherches ? Cette contribution reste aujourd'hui essentielle pour la promotion, la diffusion et la communication scientifique, les clubs, regroupements d'amateurs et autres collectivités participant à la vulgarisation, terme à prendre en bonne part, du savoir, rôle accru par l'hyperspécialisation disciplinaire contemporaine. Est-il utile de rappeler que le Web est une invention de chercheurs du CERN cherchant à partager des données expérimentales ? Fa-

cilités par les outils sociaux comme par le travail des « curateurs », les savoirs scientifiques essaient en dehors de leurs niches et peuvent se féconder de manière transversale et transdisciplinaire sans même passer par ces intermédiaires traditionnels que sont les revues — non seulement au bénéfice de la société civile, mais aussi parfois des scientifiques eux-mêmes.

Par ailleurs, les amateurs ont retrouvé avec les modes numériques d'organisation du travail scientifique un rôle important de collecte ou de repérage⁵. « Encore aujourd'hui, les amateurs apportent une contribution notable, parfois essentielle, dans les disciplines ou sous-disciplines où l'équipement nécessaire n'est pas trop complexe ou coûteux, et où les connaissances préalables ne demandent pas plusieurs années d'apprentissage. C'est pourquoi les amateurs se rencontrent surtout dans des domaines impliquant des activités naturalistes, d'observation ou de découverte sur le terrain », rappelle Patrick Matagne⁶. Par delà ces travaux de repérage et d'inventaire, on peut considérer plus largement que la science traditionnelle peut s'appuyer sur les possibilités de larges communautés à traiter par des calculs ou une intelligence distribuée les grandes quantités de données, les « big data » des sciences contemporaines. Du côté des SHS, le « surplus cognitif » des individus peut servir à contribuer à la transcription ou à la traduction⁷ tandis que la puissance non utilisée des ordinateurs peut être mise à la disposition des scientifiques pour permettre de gigantesques calculs distribués, par exemple pour l'étude du génome ou le déchiffrement de signaux, l'un des exemples les plus célèbres étant le projet SETI, qui utilise le processeur des ordinateurs individuels des participants pour analyser les données de radiotélescopes de la NASA à la recherche de possibles signaux extra-terrestres. Ces entreprises passent parfois par la « gamification » c'est-à-dire la transformation de problèmes scientifiques en jeu : ainsi *Quantum Moves*, jeu proposé par l'université d'Aarhus, vise à constituer une communauté de joueurs pour aider à la résolution de problèmes propres à l'informatique quantique. Chacune à sa manière, ces initiatives illustrent « des passions cognitives »⁸ largement partagées et partageables.

Des sciences citoyennes ?

La conversion numérique des publications et des débats scientifiques a conduit non seulement à une accélération des processus d'évaluation et de diffusion, mais aussi à la mise à disposition d'un plus large public des données scientifiques : dans le domaine médical, ce mouvement a par exemple contribué à faire des patients des acteurs de leurs propres soins, avec parfois d'immenses bénéfices, mais aussi des conflits (pensons à l'activisme parfois virulent des associations de patients), en ménageant des espaces d'informations semi-scientifiques si l'on pense aux fo-

1. Voir <http://www.nytimes.com/2011/07/12/science/12crowd.html>

2. Un collectif pour aider au montage de tels projets axés sur la biodiversité a été récemment lancé : <http://www.naturefrance.fr/sciences-participatives/le-collectif-national-sciences-participatives-biodiversite>. Aux USA, voir le métaprojet <http://www.citizensciencealliance.org>

3. Voir l'essai de Michael Nielsen, *Reinventing Discovery: The New Era of Networked Science*, Princeton University Press, 2014

4. Voir la synthèse de <http://scholarlykitchen.sspnet.org/2013/05/10/the-limits-of-crowdsourcing-in-the-scientific-disciplines/>

5. Sur la question globale de la place des amateurs dans le monde numérique on peut consulter l'essai de Patrice Flichy, *Le sacre de l'amateur*, Le Seuil, 2010 et les débats sur Internetactu : <http://www.internetactu.net/2011/03/30/le-role-des-amateurs-12-quest-ce-quun-amateur/>

6. Patrick Matagne, *Pour la science*, n° 350, décembre 2006, p. 142.

7. Voir le projet « Transcrire Bentham »

8. Voir Florian Charvolin, « Comment penser les sciences naturalistes « à amateurs » à partir des passions cognitives », *Natures Sciences Sociétés* n° 17, 145-154, 2009.

rums consacrés à la santé. Bien sûr, il y a loin des banques de données scientifiques maîtrisées ou des recherches organisées en direction d'une évaluation ouverte (ce qu'on appelle l'*open peer review*⁹) et la manière dont des non-spécialistes s'emparent de la « science ouverte » pour la transformer en informations, souvent pratiques, entre savoirs, connaissances et données, mais il existe une convergence entre des initiatives dont le point commun est la production en réseau et collaborative de savoir et de compétence, de manière asynchrone délocalisée, au profit d'une plus grande efficacité. C'est dire à quel point, dans les sciences de la culture et les SHS, où les recherches s'articulent souvent à des enjeux mémoriaux ou symboliques collectifs et peuvent directement interférer avec l'ordre de l'action politique, ces sciences ouvertes ne peuvent rester réservées qu'aux seuls chercheurs dûment reconnus en tant que tels. En témoignent par exemple des projets de documentation collaboratifs comme l'entreprise de collectes et récits de la Première Guerre mondiale menée par [Europeana](#).

Le point de rencontre entre les initiatives d'optimisation de la recherche et d'*empowerment* des communautés a été le concept de « sciences citoyennes » (*citizen sciences*) qui recouvrent non seulement les travaux accomplis avec l'aide des amateurs, mais aussi des sciences d'initiative citoyenne, si ce n'est la reconnaissance d'une intelligence collective démocratiquement construite. C'est précisément dans les domaines sociétaux que ces sciences citoyennes ont émergé : la question des écosystèmes, de la pollution, du climat, champ où les collaborateurs scientifiques écovolontaires sont plus que des assistants, des « [parataxonomistes](#) » pour des unités de recherches officielles. D'autres projets impliquant fortement leurs acteurs « amateurs » sont également apparus, comme le projet « [Genographic](#) » de cartographie génétique des migrations humaines : il interroge l'assise sociétale des programmes de recherche autant que la structure des communautés scientifiques et de leurs décisions.

Riche de succès aussi beaux que l'expérience [Foldit](#) où les joueurs ont contribué dans un « jeu sérieux » à la découverte de la structure d'une protéase rétrovirale et en ont été crédités dans des journaux spécialisés, la science contributive est au moins aussi riche de perspectives en SHS qu'en sciences dures — que l'extrême qualité des articles de l'encyclopédie scientifique Wikipédia sans laquelle cet article n'aurait pas été possible en témoigne. Elle pose néanmoins de considérables questions éthiques (que penser par exemple du financement participatif du projet anglais *Cancer Research*, qui demande aux contributeurs appelés en appui au financement de la science de choisir le type de cancer sur lesquels il apparaît nécessaire de travailler ?), politiques (dans quelle mesure le recours au *crowdsourcing* pourrait masquer le sous-financement de la recherche ?), et bien entendu, épistémologiques (la science peut-elle être un espace égalitaire ? à qui appartient l'expertise ?), que l'on vise le rêve d'une cognition sociale distribuée ou simplement celui d'une démocratisation du savoir.

9. Voir <http://blogs.lse.ac.uk/impactofsocialsciences/2013/08/20/libre-project-open-peer-review-perakakis/> et le mouvement « [Altimétriques](#) »

contact&info

► Alexandre Gefen,
CELLF
alexandre.gefen@gmail.com

Mobilisations, controverses et recueil des données à Fukushima

Trois ans et demi après l'accident nucléaire de Fukushima, on entend avancer l'argument selon lequel il n'aurait en tant que tel causé « aucun mort » — les victimes étant à attribuer au tremblement de terre et au tsunami — et qu'il y a peu de chance que ce nombre augmente à l'avenir. La réticence des populations à retourner vivre dans les environs de la centrale serait due à « la peur » irrationnelle des radiations. Cette psychologisation du risque s'appuie sur des argumentations qui tendent, contre l'évidence, à rendre « controversée » la question du danger avéré auquel sont exposés les habitants d'une zone contaminée et à minimiser les conséquences sanitaires et humaines de l'explosion de la centrale.

On peut prendre pour exemple la multiplication du nombre des cancers de la thyroïde¹ qui, selon les experts de la commission départementale de Fukushima, ne serait pas la conséquence du désastre nucléaire et de la contamination de la région, mais serait liée au perfectionnement des outils radiologiques actuels.

Le Ministère de l'environnement soutient par exemple, dans un rapport rendu public le 17 août 2014 via la chaîne du gouvernement TV-internet², 5 journaux nationaux et deux journaux locaux, qu'en deçà de 100 msv/an, aucune conséquence ne serait visible sur la santé des populations. De nombreux épidémiologistes, à l'intérieur et à l'extérieur du Japon, se sont naturellement mobilisés pour contester cette position.

La population, ballotée au gré des expertises et contre-expertises, s'est alors mobilisée pour produire des bases de données considérables, la plupart du temps restées inutilisées ou négligées. Le nombre d'associations de citoyens, sur l'ensemble du territoire, qui offrent des services liés à la mesure par *crowdsourcing* de la radioactivité dans l'environnement ou dans l'alimentation se multiplie depuis trois ans. Les données produites sont pour la plupart cartographiées de façon extrêmement précise. Le CRMS (*Citizen's Radioactivity Measuring Station*) a été l'une des premières ressources de ce genre, proposant également la mesure de la contamination des aliments (classifiés en fonction de leur région de provenance) ou encore des services comme le *Whole Body Counter*, importé de Tchernobyl, chaise mesurant la radioactivité dans le corps humain. D'autres associations, comme IRMC (*Iwaki Radioactivity Medical Center*) à Iwaki, en plus des mesures environnementales et alimentaires, ont ouvert des consultations gratuites faites par des médecins conventionnés pour les personnes désirant passer une échographie de la thyroïde. L'association Memoli (*Kodomo Mirai Sokuteijō*) présente des données précises tant environnementales qu'alimentaires sur l'ensemble du territoire. On peut enfin citer l'ambitieux projet international *Safecast* qui vise à donner à l'ensemble de la population la capacité à fournir des données sur la radioactivité et à y accéder.

Le *crowdsourcing* peut ainsi être vu comme une prise en main de la situation par la population, une forme d'*empowerment* citoyen, et comme source de production de données aptes à mettre à disposition du public, local et international, une vision alternative et des informations transparentes. Mais il est aussi parfois analysé comme une forme de déresponsabilisation de l'Etat qui amène à faire peser la responsabilité de la radioactivité sur les individus. Ces derniers deviennent les fournisseurs d'un travail gratuit à un Etat qui refuse de les prendre en charge, et contraints de surcroît à organiser la co-gestion de leur vie en environnement gravement contaminé.

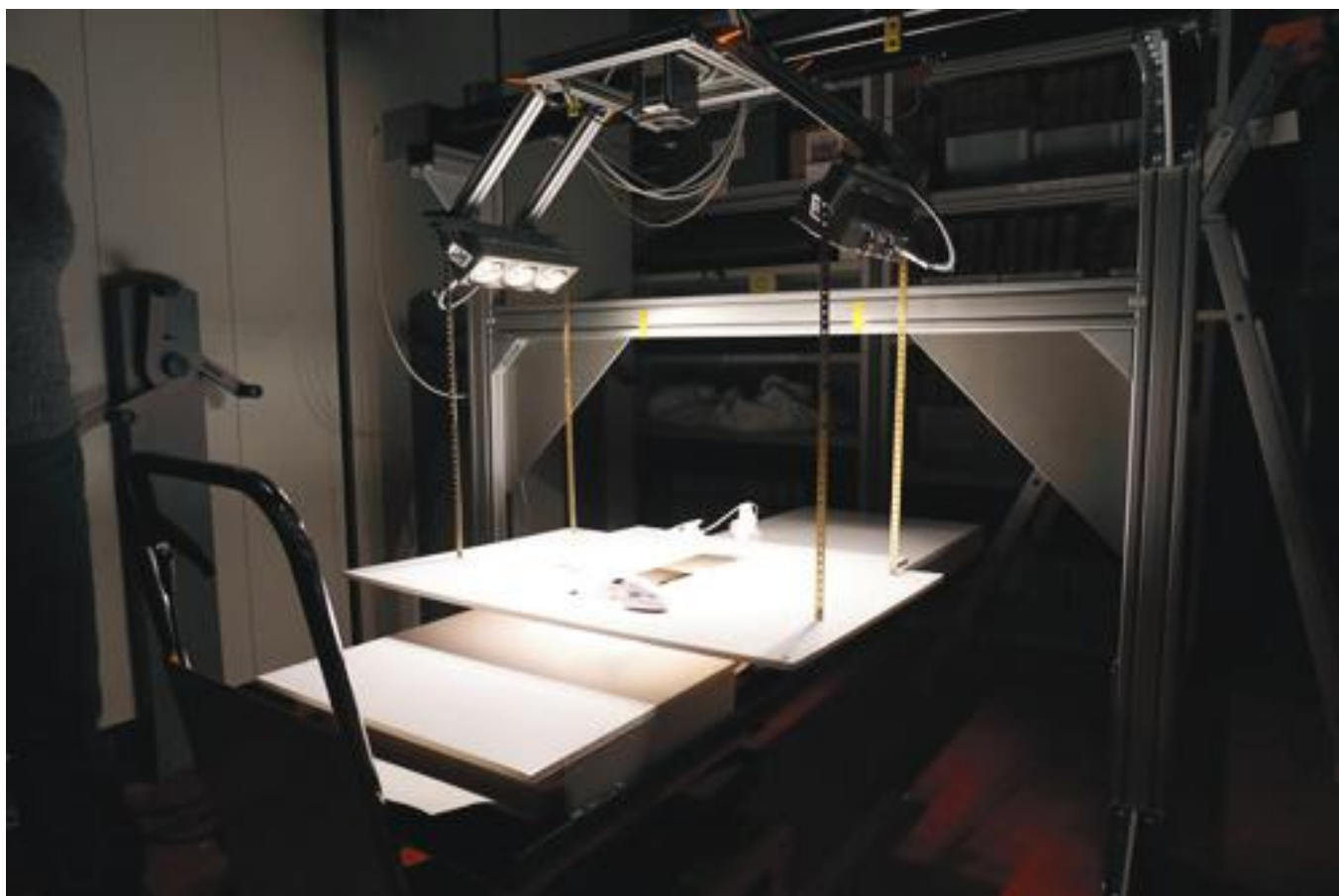
Cécile Asanuma-Brice, Maison franco-japonaise, Tokyo

1. Asahi, 24 août 2014 : *Cancer de la thyroïde, 104 personnes, enquête sur 300 000 enfants de Fukushima*.

2. Dr. Keiichi Nakagawa (Associate Professor, Tokyo University Hospital)

A PROPOS

Le développement durable et les sciences de la conservation du patrimoine



Analyse d'un manuscrit ancien par imageur hyperspectral
© USR3224-CRC

Le Centre de recherche sur la conservation (CRC, USR 3224, CNRS / Muséum national d'histoire naturelle / Ministère de la culture et de la communication) fêtait récemment son 50e anniversaire. C'est en 1963 que l'unité de recherche fut créée sous l'impulsion de la Bibliothèque nationale, au Muséum national d'histoire naturelle, pour trouver un remède aux altérations biologiques qui menaçaient les collections patrimoniales. Au cours de ces cinq décennies, les priorités en matière de recherche sur la conservation du patrimoine culturel ont évolué au fil des changements qui s'opéraient dans la société. Cet événement fut l'occasion de réfléchir aux nouvelles orientations scientifiques et aux priorités de recherche, dans le cadre d'une *conférence internationale*.

Les diverses évolutions et crises — technologiques, économiques, énergétiques — ont influé sur les thématiques de la recherche dédiée à la conservation du patrimoine culturel ; il est possible d'en discerner plusieurs depuis les années 1950. La première s'est concentrée sur les actions curatives, les travaux visaient alors à élaborer des traitements éprouvés. La seconde marque l'essor, dans les années 1980, de la conservation préventive : les recherches tendent alors à déterminer le rôle de l'environnement dans les dégradations et les moyens de les prévenir. Enfin, depuis la fin du xx^e siècle, on intègre l'aspect du « développement durable ». L'objectif n'est pas à proprement parler l'utilisation « durable » des biens culturels : la conservation préventive s'y consacre déjà en rendant le patrimoine accessible à un grand nombre, le plus

longtemps possible. Ce qui est nouveau, c'est que pour y parvenir on prenne en compte des enjeux qui dépassent le cadre strict du patrimoine culturel. Il s'agit, maintenant, dans les solutions qu'apporte la recherche pour préserver le patrimoine culturel, d'appréhender les impacts et aspects écologiques, sociaux ou économiques, sinon d'en faire une priorité¹. L'expression *green conservation* a fait son apparition au début des années 2000 et suscite un intérêt croissant. La prise en compte de ces nouveaux impératifs amène à réexaminer un certain nombre de pratiques et d'usages en matière de conservation et de restauration, elle conduit à explorer et exploiter de nouveaux outils, pour mesurer et chiffrer l'impact des mesures de conservation ou réévaluer des savoir-faire traditionnels.

1. Catherine Antomarchi, The Iccrom Forum on conservation sciences. <https://www.youtube.com/watch?v=gpBjBH9LqSk>



En Juillet 2007, les Archives bouddhistes de la photographie ont été installées dans un bâtiment monastique dans les locaux de Vat Souvanna Khili à Luang Prabang, Laos. Cette salle d'archive principale, ici fraîchement rénové, sera bientôt un atelier et un lieu de stockage. Il détient aujourd'hui environ 20 000 photographies de la vie monastique et de la société bouddhiste au Laos, couvrant une période d'environ 100 ans. © Martin Jürgens, 2007

Mesurer et chiffrer pour mieux anticiper

Dès l'instant où la société reconnaît qu'un bien précieux, participant au bien-être collectif, n'est pas inépuisable, il lui incombe d'une part d'évaluer au mieux son état et, d'autre part, d'estimer sa durabilité. Pour le patrimoine naturel, le recensement des ressources et de leur évolution est une activité scientifique indispensable qui comprend la mise en place de multiples **indices et indicateurs** (indicateur de la biodiversité, empreinte écologique, bilan carbone, indice de durabilité environnementale, indice de qualité de l'air, etc.) et d'outils économiques (comme la bourse du carbone dans l'industrie), ainsi que le développement de modèles prédictifs. Il en va de même dans le domaine des biens culturels où l'activité scientifique vise à répondre à deux questions légitimes et primordiales : quel est l'état de notre patrimoine et est-il pérenne ?

Le recensement des biens culturels et la création de bases de données sont encouragés depuis de nombreuses années en France et résultent d'une politique clairement affichée par le ministère de la Culture et de la Communication avec l'existence d'un service de l'inventaire. Dans l'objectif du développement durable, cette action comptable des œuvres dans l'espace doit être complétée par une meilleure appréhension de leur état et de leur évolution. Pendant plus de cent ans, la photographie a permis d'enregistrer un état de conservation apparent des œuvres ; elle fut, à ce titre, un

outil indispensable à leur documentation. Des analyses chimiques ont conduit à déterminer leur composition et à comprendre certaines altérations. La recherche vise aujourd'hui à développer des techniques et méthodologies analytiques plus fines et non destructives (ou micro-destructives) par le biais de grands équipements, comme le synchrotron Soleil, et d'instruments d'analyse portables performants. Les laboratoires d'analyse mobiles (développés dans le cadre du projet européen Charisma (« Molab ») ou des investissements d'avenir « Equipex Patrimex ») mettent ces outils au service des œuvres et non l'inverse. Parmi ceux-ci, l'imagerie hyperspectrale ouvre des perspectives nouvelles qui remplacent et complètent avantageusement la photographie « argentique », en ajoutant à l'imagerie des possibilités d'analyses colorimétriques et spectroscopiques. On peut ainsi espérer suivre et mesurer les transformations et les altérations des biens culturels de façon « non invasive » par de la simple imagerie dans une large gamme spectrale qui s'étend de l'ultraviolet à l'infrarouge. L'analyse est à la fois locale et globale, elle permet de corréliser des données spectrales des matériaux à leur distribution spatiale pour l'étude de matériaux présents en quantités variables et distribués au sein de mélanges hétérogènes, de documenter la composition chimique et l'apparence d'une œuvre sur chaque pixel numérisé et sur l'ensemble de l'objet photographié. C'est sans doute avec ces techniques que sera recueilli, dans les décennies à venir, le corpus d'informations le plus riche et le plus per-

minent sur les œuvres et leur devenir ; il permettra une approche chimométrique et ainsi de mieux évaluer les cinétiques d'altération et d'anticiper certains dommages. La masse de données ainsi recueillie est cependant considérable, leur gestion et leur pérennisation², l'interopérabilité soulèvent encore des questions et suscitent de multiples recherches.

En période de contraintes budgétaires, il convient de développer des outils à même d'éclairer sur la portée des choix et de fonder les priorités sur une analyse aussi objective que possible. Une méthode d'analyse des risques pour les biens culturels a été introduite et développée, en particulier grâce aux efforts de l'Institut canadien de conservation et de l'Iccrom (Centre international d'études pour la conservation et la restauration des biens culturels) à Rome³. Elle s'accompagne de modèles prédictifs introduits à la fin du xx^e siècle pour les archives et les bibliothèques. Grâce aux diagrammes d'isopermanence et aux indices de permanence, il est possible de quantifier l'impact d'une modification d'un environnement climatique sur la durabilité des matériaux. Donald Sebera (*Library of Congress*, États-Unis), comme James Reilly⁴ (*Image Permanence Institute*, États-Unis) appliquent des modèles simples fondés sur la théorie de l'état de transition (équation d'Eyring-Polanyi) ou la loi d'Arrhenius. Le dispositif de James Reilly, qui se présentait sous forme d'abaque destiné aux collections de films ou de photographies, a été informatisé et appliqué à l'ensemble des collections de musées, d'archives et de bibliothèques. Un indice de préservation, le TWPI — *time weighted preservation index* : indice de préservation pondéré — est calculé à partir de relevés hygrométriques continus. Ce dispositif s'avère un outil d'aide à la décision⁵ lorsqu'il s'agit de déterminer ou de comparer, dans un bâtiment, l'impact du climat de diverses salles sur la pérennité des œuvres. Il l'est également pour établir le meilleur compromis entre les conditions de conservation et la durabilité des matériaux (rapport coût/efficacité). L'effet des variations thermo-hygrométriques sur les processus d'altération des matériaux organiques fait également l'objet de travaux, avec comme objectif celui d'assouplir les recommandations en vigueur, sans pour autant compromettre la pérennité des objets.

Ces diverses approches permettent de mieux appréhender le rôle du climat. Les œuvres sont cependant soumises à d'autres dangers, comme ceux liés à l'exposition à la lumière ou à la pollution. Les laboratoires de recherche ont mis au point un moyen

d'évaluer, in situ, la fragilité à la lumière des couleurs⁶. Certaines parties représentatives sont exposées quelques minutes sous un micro-faisceau lumineux (la zone exposée a un diamètre inférieur à 1 mm²), un spectrophotocolorimètre suit et enregistre l'évolution des couleurs. Le rapport dose/dommage est ainsi calculé et utilisé pour établir une politique d'exposition s'appuyant sur ces données objectives et quantifiables⁷. Les recherches actuelles se focalisent sur l'étude de la non-réciprocité de l'action lumineuse et de la sensibilité spectrale des matériaux. Avec un objectif similaire, on travaille à établir les seuils critiques ou DMENO (dose minimale avec effets nuisibles observables) des polluants sur les biens culturels⁸.

Ces divers développements sont des aides indispensables pour prévoir le comportement de certains objets dans les musées, les archives et les bibliothèques. Ils permettent d'évaluer au mieux les risques en matière d'exposition des œuvres ou de conservation et de chiffrer l'impact de toute action sur leur pérennité, et soutiennent, de ce point de vue, une politique de développement durable.

La prise en compte de l'impact environnemental de la conservation

Le patrimoine culturel est à l'origine d'une activité humaine qui a un impact non négligeable sur l'environnement, quantifiable en termes de bilan carbone ou d'empreinte écologique. L'émergence récente du concept de *green museum*⁹ traduit une responsabilisation des institutions muséales qui souhaitent concilier et associer la préservation du patrimoine culturel avec la protection et la mise en valeur du patrimoine naturel, d'autant que la frontière entre ces deux types de patrimoine s'avère ténue : jardins et paysages étant parfois intégrés au patrimoine culturel. Les musées, archives et bibliothèques deviennent le lieu de promotion du développement durable¹⁰. Ce souci du respect de l'environnement se traduit également, au niveau de la conservation matérielle des œuvres, par des conceptions novatrices de bâtiments de réserve et d'exposition moins énergivores et par la réévaluation de certaines recommandations normalisées, en se fondant sur les modèles décrits précédemment¹¹. L'étude de l'éclairage, et en particulier des possibilités offertes par l'usage des diodes électroluminescentes (LED), ouvre à de nouvelles applications qui dépassent le simple fait d'éclairer, mais peuvent simuler les condi-

2. Rudolf Gschwind, Monolith: longterm digital archiving without migration. <https://www.youtube.com/watch?v=xgFNFUcQxI0>

3. Stefan Michalski, Cost, effectiveness, and cost-effectiveness: Three linked but often conflicting measures in the evaluation of options for collection risk reduction. <https://www.youtube.com/watch?v=kbosRJITPeM>

4. James Reilly, Sustainable environmental management for cultural heritage. <https://www.youtube.com/watch?v=0QxPDSOKCXM>

5. Cassar M., Evaluating the benefits of cultural heritage preservation: An overview of international initiatives, www.ucl.ac.uk/sustainableheritage/evaluating_benefits.pdf

6. Whitmore, P.M., Pan X., Baillie C. (1999). Predicting the fading of objects: Identification of fugitive colourants through direct non-destructive lightfastness measurements, *J. Am. Inst. Cons.*, 38, 395-409.

7. Bruce Ford, Measurement is better than rules: microfading and light exposure risk assessment without the guesswork. <https://www.youtube.com/watch?v=c59HGihE9SQ>

8. Anne-Laurence Dupont, The impact of volatile compounds released by paper on cellulose degradation in ambient hygrothermal conditions. <https://www.youtube.com/watch?v=a8nwYlzCArI>

Tetreault J. (2003). Polluants dans les musées et les archives : évaluation des risques, stratégies de contrôle et gestion de la préservation. Ottawa : ICC.

9. <http://sustainablemuseums.blogspot.fr/>

10. Stefan Simon, The museum environment - Monitoring, safeguarding and developing standards at the National Museums Berlin. <https://www.youtube.com/watch?v=neWi0Salxaw>

Naoko Sonoda, Sustainable and environmental friendly museum environment: a case study from the National Museum of Ethnology after the Great East Japan Earthquake. <https://www.youtube.com/watch?v=2vS57VkrPqk>

Véronique Rouchon, La recherche en conservation: du patrimoine culturel au patrimoine naturel. <https://www.youtube.com/watch?v=XYkjcHc7ZaQ>

11. Tim Padfield, Containing the discussion about museum environmental standards. <https://www.youtube.com/watch?v=whugvdkwgEM>

Jochen Käferhaus, Sustainability and energy saving in museums, depots, churches and historic buildings. <https://www.youtube.com/watch?v=vFO4rVfoFgk>

tions originales d'œuvres décolorées ou véhiculer des données numériques¹².

Du point de vue des traitements destinés à combattre la contamination biologique (mites, insectes xylophages, etc.), il a fallu délaisser l'usage de certains insecticides toxiques (bromure de méthyle, lindane, DDT, etc.) au profit de solutions identifiées comme « vertes ». La recherche en conservation a permis de valider diverses méthodes comme l'anoxie (privation d'oxygène), la congélation, le traitement thermique, l'usage de substances alternatives (pyréthres, huiles essentielles) plus respectueuses de l'environnement et de la santé, dans une démarche guidée par la prévention. En Corée du Sud, des chercheurs ont analysé certaines reliques insérées dans des sculptures centenaires en bois qui avaient été singulièrement préservées des attaques biologiques ; ils ont identifié la présence d'anis étoilé et des clous de girofle. Ces huiles essentielles ont été proposées, afin de lutter contre les insectes xylophages et les moisissures dans les institutions patrimoniales et également chez les particuliers¹³. Certes, le fait d'être naturels ne signifie pas que ces produits soient dépourvus d'effets indésirables pour les hommes comme pour les œuvres, mais ces exemples témoignent de nouvelles recherches et pratiques intégrant la dimension du développement durable.

L'étude et l'intégration de savoirs traditionnels

La recherche se consacre également à l'étude et au développement de solutions alternatives qui reposent sur l'analyse de l'habitat ancien, des matériaux traditionnels ou hybrides et la mise en valeur de sites naturels plus propices à une bonne conservation¹⁴. L'équipe de recherche du *Getty Conservation Institute* (*Getty Center*, Los Angeles, États-Unis) dirigée par Shin Maekawa¹⁵, s'est longtemps consacrée à ces études de climat dans des bâtiments anciens et sur divers sites. Les recherches dans ces domaines se multiplient¹⁶. En Corée, l'analyse approfondie d'un édifice qui conserve des matrices en bois pour l'impression de textes sacrés (*Tripitaka Koreana* de la dynastie Goryeo au temple d'Haeinsa) a révélé l'intérêt des microclimats et de la ventilation naturelle qui permettent d'assurer une conservation optimale depuis plus de mille ans. Ces techniques et principes traditionnels sont alors appliqués ou adaptés à des constructions modernes dans des pays occidentaux. Ce sont non seulement les savoir-faire ancestraux dans l'habitat, mais aussi les pratiques et matériaux tradition-

nels qui sont étudiés et valorisés, comme le sont les médecines traditionnelles et les substances naturelles dans la recherche en phytothérapie¹⁷. D'autres aspects sont abordés comme les pratiques de manipulation, le stockage des objets et des collections, le nettoyage des objets et des espaces de stockage (pratiques rituelles et entretien du matériel)¹⁸. Cette ré-évaluation et ré-appropriation du savoir et des savoir-faire locaux participent tout autant du développement durable que d'une approche intégrée de la conservation, telle qu'elle fut encouragée par les travaux de la Conférence de Nara sur l'authenticité. Ces savoir-faire sont non seulement sources de recherches académiques, mais ils impliquent aussi de nouveaux acteurs et ouvrent des perspectives nouvelles de développements économiques locaux.

Dans le domaine de la conservation des biens culturels, la prise de conscience de leur aspect non renouvelable, au même titre que celui du patrimoine naturel, a induit les mêmes conséquences : ces ressources s'épuisant, il nous faut apprendre à les gérer au mieux pour le profit de tous, en prenant garde que la préservation de l'un ne se fasse pas au détriment de l'autre. Les quelques exemples présentés ici et abordés lors de la conférence anniversaire du CRC montrent que, de fait, la recherche s'attache déjà à apporter des solutions en matière de conservation qui ne sont plus examinées du seul point de vue de l'œuvre : elles prennent en compte les impacts économiques, écologiques et sociaux.

contact&info

► Bertrand Lavédrine,
CRC

lavedrin@mnhn.fr

► Pour en savoir plus

<http://crrc50.sciencesconf.org/>

12. Françoise Viénot, Les diodes électroluminescentes : qualité de la lumière et confort visuel. <https://www.youtube.com/watch?v=1zdAn-0pagA>

Paul Whitmore, Better preservation through solid-state lighting ? <https://www.youtube.com/watch?v=n4Fxm400-to>

13. Yong-Jae Chung, Conservation and control of biodeterioration for cultural heritage using essential oil. <https://www.youtube.com/watch?v=A013YHC4ehc>

14. Larsen Poul Klens, Energy efficient museum stores and archives. <https://www.youtube.com/watch?v=VPbhggOLqY>

Birgit Dulski, Innovation versus preservation - different approaches how to combine sustainable development with architectural conservation and restoration. <https://www.youtube.com/watch?v=RPGHhKwHSY>

Philippe Goergen, Climat tropical et musées : comportement des bâtiments de musées, des vitrines et des collections en milieu muséal non climatisé. <https://www.youtube.com/watch?v=vXfO8Z3tVIA>

Gisèle Laure Lecomte-Nana, Development of a green clay-fiber composite for fire protection. <https://www.youtube.com/watch?v=HgYe41v4baQ>

15. Maekawa, S., and Toledo F. (2003). "Sustainable climate control for historic buildings in hot and humid regions," Management of environmental quality: An international journal, Vol. 14, No. 3 [Also presented at the 18th International Conference on Passive Low Energy Architectures (PLEA) Conference, November 2001, Florianópolis, Brazil.]

16. Andrea Luciani, Good practices from the past? The application of heritage science to museum environments in the 1930s. <https://www.youtube.com/watch?v=ey3gLoM4mh8>

17. Seki Masazumi, Applicability of Japanese bast fibers (koko, mitsumata, gampi) as strengthening fibers for "fleece" method. <https://www.youtube.com/watch?v=O1DmBWZEKsg>

Claire Chahine, Le cuir : enjeux pour la conservation et l'environnement. <https://www.youtube.com/watch?v=Bzd6fk6FIO>

18. Martin Jürgens, The buddhist approach: conservation of photographs in Luang Prabang, Laos. <https://www.youtube.com/watch?v=Kv5NfjZaZ8g>

Aurélien Vissac, Ann Bourgès, Protection et conservation du patrimoine architectural en terre par des stabilisants naturels, d'origine animale ou végétale. Interactions argiles/biopolymères. Projet PaTerre+. <https://www.youtube.com/watch?v=krUbjAAEuNo>

Isabelle Brianso, Une muséologie participative issue de la coopération comme processus de développement durable à Angkor (Cambodge) : enjeux et limites. <https://www.youtube.com/watch?v=u7CK5A5QNbS>

Partage d'expériences

Un grand corpus de SMS en français : 88milSMS



Contexte historique

Pour mieux comprendre comment la langue évolue, le linguiste recueille des données, les observe et les analyse. L'informaticien peut lui venir en aide en minimisant les traitements manuels ; la rencontre entre linguistique et informatique — ou le traitement (partiellement ou entièrement) automatique de la langue (TAL) — a lieu. Depuis plusieurs décennies, l'accès facilité aux données numériques permet la constitution aisée de corpus informatisés. Les données nativement numériques sont de plus en plus nombreuses et les recherches en humanités numériques foisonnent : le discours écrit des courriels, forums, *chats*, blogs, réseaux sociaux, etc. est passé au peigne fin par les chercheurs, car il révèle des usages émergents, novateurs, différents, spécifiques. Pour mieux comprendre les pratiques et les usages, les chercheurs s'intéressent aussi à la communication et aux discours qui circulent entre les individus se servant d'ordinateurs — la *communication médiée par ordinateur* ainsi que le *discours électronique médié* (Panckhurst, 1997, 2006¹). Avec l'arrivée du téléphone portable, puis des SMS (services de messages succincts) au début des années 1990, un autre outil de médiation ouvrait aux chercheurs un grand terrain d'enquête, mais une difficulté *a priori* insurmontable persistait : comment recueillir ces données authentiques — écrites spontanément en situation réelle — sans qu'elles subissent aucune re-saisie au moment du recueil ?

Constitution d'une grande base de données mondiale de SMS

Jusqu'au début des années 2000, les analyses de SMS étaient restreintes, faute de collectes de quantité significative. En 2004, un groupe d'universitaires belges a lancé un projet international, intitulé *sms4science*, afin de recueillir, organiser (en une base de données mondiale) et analyser des SMS authentiques (Fairon et al. 2006, Cougnon, 2014²). S'en sont suivies d'autres collectes

de SMS et l'initiative la plus récente pour le français est le projet *sud4science LR*. En trois mois, à l'automne 2011, plus de 93 000 SMS authentiques ont été recueillis auprès du grand public par un groupe de chercheurs dans la région Languedoc-Roussillon (Panckhurst et al. (2013), Panckhurst & Moïse (2014)³). Plus de 88 000 SMS seront finalement conservés et mis à disposition après divers prétraitements décrits ci-dessous.

Le téléphone comme outil de recueil



Le projet montpelliérain s'est distingué des collectes précédentes par la méthode de récolte. Après inscription et consentement légal en ligne, les participants donateurs de SMS, au moment de l'envoi de leur texto à autrui, pouvaient l'envoyer en copie aux chercheurs. Il était également possible de réexpédier des SMS (précédemment envoyés et contenus dans la mémoire du téléphone du scripteur) aux chercheurs.

Le moyen utilisé ? Un smartphone⁴, grâce auquel l'ensemble des textos a été recueilli pendant 13 semaines. Ce dispositif a été un véritable pari technique, car personne ne savait à l'avance si l'iPhone allait permettre un recueil de SMS très important, sans défaillir. En définitive, aucun problème n'est survenu. Chaque semaine, les SMS ont été copiés sur un disque dur externe, déconnecté d'Internet (pour des raisons juridiques). La grande base de

1. Panckhurst R. (1997), « La communication médiatisée par ordinateur ou la communication médiée par ordinateur ? », *Terminologies nouvelles*, 17, 56-58.

Panckhurst R. (2006), « Le discours électronique médié : bilan et perspectives », in A. Piolat (Éd.), *Lire, écrire, communiquer et apprendre avec Internet*. Marseille : Éditions Solal, p. 345-366.

2. Fairon C., Klein J.-R., Paumier S., (2006), *SMS pour la science. Corpus de 30.000 SMS et logiciel de consultation*, Presses universitaires de Louvain, Louvain-la-Neuve, Manuel+CD-Rom.

Cougnon L.-A. (à paraître, 2014) *Langage et sms. Une étude internationale des pratiques actuelles*. Presses universitaires de Louvain.

3. Panckhurst R., Détrie C., Lopez C., Moïse C., Roche M., et Verine B. (2013). « Sud4science, de l'acquisition d'un grand corpus de SMS en français à l'analyse de l'écriture SMS ». *Épistémé — revue internationale de sciences sociales appliquées*, 9 : Des usages numériques aux pratiques scripturales électroniques, 107-138.

Panckhurst R. et Moïse C., (2014), « French text messages. From SMS data collection to preliminary analysis », in *SMS Communication. A Linguistic Approach*, éd L.-A. Cougnon, C. Fairon, John Benjamins : Amsterdam/Philadelphia, p. 141-168.

4. L'entreprise *iTribu*, très enthousiaste à l'idée de participer à un projet de recherche universitaire, a prêté un iPhone aux chercheurs pour la durée de la collecte.



Interface de Seek&Hide

données (BD) en constitution devait rester dans son intégralité sur le téléphone également jusqu'à la fin de la récolte, afin d'assurer que la BD soit entière et homogène, avant transfert final. Depuis le début des collectes de SMS en 2004, il était très important que la méthode de recueil ne passe pas par une (re)saisie des données : seul ce type de transfert technologique était possible afin d'assurer que les données demeurent réellement authentiques.

Un logiciel pour l'anonymisation

Après l'étape de la collecte, en raison des aspects juridiques liés à la protection de la vie privée, tous les SMS du corpus « 88milSMS » incluant des prénoms, noms, surnoms, adresses, lieux, numéros de téléphone, codes, URL, marques, courriels, etc. ont été anonymisés de manière semi-automatique, en plusieurs étapes. Le logiciel *Seek&Hide* (Accorsi et al., 2014⁵), ayant pour double tâche d'anonymiser le corpus et de fournir aux annotateurs humains une interface en ligne agréable à utiliser, a été élaboré par des étudiants.

Ce logiciel s'appuie sur des méthodes de traitement automatique des langues. Il propose une page web sécurisée accessible pour les annotateurs. Le but du logiciel est de faciliter l'expertise et de traiter une quantité importante de données. L'approche développée se décline en trois phases :

► Phase automatique : traitement automatique des données (mots) qui ne présentent *a priori* aucune ambiguïté quant à leur interprétation (à anonymiser ou non). Par exemple, un prénom comme *Cédric* serait automatiquement anonymisé ; un nom commun comme *crayon* serait automatiquement écarté de l'anonymisation. Notons que des approches d'apprentissage automatique

ont également été proposées (Patel et al. 2013⁶).

► Phase semi-automatique : traitement manuel de l'information nécessaire pour les SMS qui présentent des mots ambigus (*Pierre*=prénom, *pierre*=nom commun) ou inconnus (*Namrata* = prénom inconnu du dictionnaire utilisé). Ceci s'effectue à travers un système qui met en relief les éléments nécessitant une expertise. Cette mise en valeur facilite significativement le travail de l'annotateur.

► Phase de validation : relecture et validation des SMS anonymisés automatiquement ou modification d'une anonymisation appliquée par l'outil lors de la phase automatique (cf. cas 1 à 3 ci-dessous).

Seek&Hide a automatiquement anonymisé 72% du corpus « 88milSMS » ; les 28 % restants ont été soumis à une phase semi-automatique. Le logiciel propose une interface web sécurisée permettant aux annotateurs-experts linguistes de mener à bien la phase suivante, qui permet de désambigüiser les SMS et de décider si l'anonymisation doit ou non être effectuée. Par exemple, un prénom comme *Pierre* qui est également un nom commun, en minuscules, serait traité pendant cette deuxième phase, pour qu'un annotateur humain puisse décider, en fonction du cotexte/contexte si l'occurrence doit être ou non anonymisée.

La troisième phase de validation consiste en la lecture de tous les SMS (72% du corpus) anonymisés de manière automatique par *Seek&Hide*, afin de vérifier si tous les textos l'ont bel et bien été correctement. Trois cas de modification éventuelle ont été repérés par les annotateurs.

5. Accorsi P., Patel N., Lopez C., Panckhurst R., Roche M. (2014), « *Seek&Hide* : Anonymising a French SMS corpus using natural language processing techniques », in *SMS Communication. A Linguistic Approach*, éd L.-A. Couston, C. Fairon, John Benjamins : Amsterdam/Philadelphia, p. 11-28.

6. Patel N., Accorsi P., Inkpen D., Lopez C., Roche M. (2013) "Approaches of anonymisation of an SMS corpus", in *Computational Linguistics and Intelligent Text Processing*, pp. 77-88, Springer Verlag, Berlin, Heidelberg.

Cas 1 : anonymisation automatique à enlever

grace a lui on comprend trop bien franchement ke kiffe la physique cette *anne* meme si cest bien dur

Dans cet exemple, *grace* et *anne* ont été anonymisés, mais ce n'est pas une erreur du logiciel. Si le scripteur avait ajouté l'accent circonflexe, *Seek&Hide* n'aurait pas procédé à l'anonymisation en prénom pour *grâce* ; l'autre occurrence est *anne* au lieu d'*année*, qui n'est donc pas un prénom dans ce contexte.

Cas 2 : anonymisation manquante à insérer

Excuse pour c texto si tard c'était pour t dire q *mat* a u l permis bisous bisous

Mat est absent du dictionnaire de prénoms, puisqu'il s'agit d'un diminutif (e.g. *Mathieu*), et le nom commun (le *mat*) existe. Le logiciel avait donc ignoré cet élément.

Cas 3 : balises d'anonymisation à remplacer

Une *clio* noir phase 2 vendue par une amie d'une collègue de boulot.

Clio est ici un nom de voiture de la marque Renault, et non un prénom. Il faut donc changer la balise de l'anonymisation.

Les annotateurs humains peuvent donc retirer, ajouter, modifier les étiquettes précédemment insérées de manière automatique par le logiciel⁷. À ce stade, ils peuvent également décider de noter certains SMS comme devant être supprimés du corpus si ceux-ci contiennent des propos inacceptables au regard de la loi.

L'opération totale d'anonymisation a nécessité 21 mois et a été réalisée grâce au travail de nombreux étudiants stagiaires. Si le travail avait été mené de manière entièrement manuelle, la périodicité aurait été augmentée de manière significative et le travail accompli aurait été très certainement moins fiable, tant la lecture humaine des SMS inflige une réelle surcharge cognitive.

Transcodage, alignement, annotation

Une fois l'anonymisation terminée, les SMS sont prêts à être transcodés en français standardisé afin de permettre d'éventuels traitements ultérieurs en linguistique-informatique (incluant des analyseurs morpho-syntaxiques). L'idée est de restituer l'orthographe et la grammaire afin de faciliter la compréhension, mais non d'« injecter » des éléments supplémentaires (cf. l'exemple 1 ci-dessous). Tous les SMS bruts anonymisés, un échantillon de 1000 SMS transcodés et un échantillon de 100 SMS annotés sont téléchargeables. Le transcodage est utile pour le grand public, ou pour ceux qui veulent lire et comparer rapidement les SMS bruts

anonymisés et transcodés, à des fins de recherche. Cependant, d'un point de vue linguistique, il est extrêmement difficile de procéder à un transcodage qui convienne à tous, car les interprétations sont nombreuses et variées.

Exemple 1 : passage du SMS brut anonymisé au SMS transcodé

SMS brut anonymisé (n° 22446 du corpus 88milSMS) :

En fait c rien de spécial, jprends juste un peu de recul et jcomprends pas ce que jfous là, fac, psycho, montpellier, pourquoi simplement je vis, enfin bref rien de grave. Qu'est ce qui cloche chez toi?

SMS anonymisé et transcodé :

En fait c'est rien de spécial, je prends juste un peu de recul et je comprends pas ce que je fous là, fac, psychologie, Montpellier, pourquoi simplement je vis, enfin bref rien de grave. Qu'est-ce qui cloche chez toi ?

Dans l'exemple 1, on n'ajoutera pas la particule de négation, *ne/n'*. On n'« injectera » pas non plus des éléments prépositionnels ou des déterminants (« à la fac », « en psychologie », « à Montpellier »), car le traitement automatisé demeure possible sans ces informations. En revanche, pour des formes abrégées, agglutinées, etc., on transcode en français standardisé pour qu'un analyseur morpho-syntaxique soit à même de traiter automatiquement la phrase. Dans cet exemple, l'apocope⁸ « fac » demeure telle quelle dans la version transcodée, car les chercheurs ont décidé de valider le transcodage en lien avec les informations apparaissant au sein du *Petit Robert* en ligne 2014 (PR14) : si une entrée dictionnaire existe, elle n'est pas transcodée dans sa forme entière (« fac » demeure intact, mais « psycho » sera transcodé en « psychologie », car si l'élément « psycho- » existe effectivement dans le PR14, l'apocope qui renvoie à « psychologie » n'y figure pas). Par ailleurs, lorsque la ponctuation est présente, les normes typographiques sont rétablies pour le français, ici l'espace absente avant le point d'interrogation final.

Ces choix ne conviennent pas nécessairement à tous, d'où l'importance de maintenir le lien visible entre la consultation du SMS brut anonymisé et le SMS anonymisé.

Des étudiants ont également travaillé sur le transcodage et l'alignement en explorant la faisabilité d'une méthode d'alignement des SMS pour faciliter le passage du SMS brut anonymisé au SMS transcodé en français standardisé. Ils ont ainsi proposé un modèle pour une interface en ligne afin d'aider le travail de l'annotateur humain. Le modèle d'alignement incluant une interface s'intitule AlignSMS (cf. Lopez et al. 2014⁹).

Les chercheurs ont renoncé à effectuer le transcodage sur l'ensemble du corpus pour deux raisons :

- 1) le temps très important exigé par cette tâche ;
- 2) la façon d'effectuer le transcodage susciterait vraisemblablement des désaccords théoriques.

7. Sur un échantillon de 20 000 SMS, seules 358 modifications ont dû être effectuées : 66 % (cas 1), 29 % (cas 2), 5 % (cas 3).

8. Chute d'un ou de plusieurs phonèmes à la fin du mot par suite d'une évolution phonétique ou d'un abrègement.

9. Lopez C., Bestandji R., Roche M., Panckhurst R. (2014) « Towards Electronic SMS Dictionary Construction: An Alignment-based Approach », Actes du colloque LREC, Reykjavik, Islande, 26-31 mai, p. 2833-2838.

Par la suite, un extrait de 100 SMS du corpus « 88milSMS » a été annoté, à l'aide de 8 balises : TYPographie, MODification, GRammaire, BINettes, ABSence, LANGue, ORThographe, DIVers. Il ressort que les phénomènes de *typographie* sont les plus saillants, suivis par les *modifications* (substitutions, réductions, ajouts, etc.). La balise qui concerne la *grammaire* arrive en troisième position, suivie, dans l'ordre, par les *binettes*, l'*absence*, la *langue*, l'*orthographe* et la balise *divers*. Il est également intéressant de constater que 70 % de l'extrait des 100 SMS n'a subi aucune modification.

Comme pour le transcodage, il est extrêmement difficile de proposer une annotation standardisée. Lors du projet *sud4science LR*, les chercheurs ont invité les acteurs des collectes précédentes, dans le cadre de *SMS4science*, à présenter les balises utilisées pour l'annotation de leurs corpus de SMS. Une harmonisation générale a ensuite permis aux chercheurs *sud4science* de réduire le nombre de balises précédemment utilisées, afin d'envisager le balisage éventuel du corpus 88milSMS. Tout bien pesé, ils ont décidé de fournir un petit échantillon d'annotation de 100 SMS. Mais ils ont renoncé à l'annotation de l'ensemble du corpus 88milSMS car, d'une part, la tâche aurait été gigantesque et, d'autre part, les chercheurs n'auraient pas été nécessairement en accord avec le choix des balises. Si cet échantillon permet de fournir des pistes de recherche, il nous est apparu au final que le plus important est d'abord de mettre à disposition le corpus anonymisé de telle sorte que chacun puisse le catégoriser et l'annoter en fonction de sa propre problématique de recherche.

Diffusion : mission de service public

L'objectif du dépôt sur la grille de services d'Huma-Num est de mettre sans tarder à la disposition de la communauté scientifique et, plus largement, de tous ceux qui sont intéressés par les mutations sociales, comme les responsables des politiques publiques en matière d'éducation et d'intégration sociale, une base de données directement [téléchargeable](#). Les chercheurs du projet proposent donc au public, via un téléchargement direct, le corpus intitulé « 88milSMS » dans son intégralité, deux échantillons (100 SMS annotés, 1 000 SMS transcodés en français standardisé), ainsi qu'un questionnaire sociolinguistique soumis aux donateurs, et leurs réponses.

Les chercheurs ont débuté l'*observation*, la *fouille*, la *description*, le *traitement* et l'*analyse* du grand corpus « 88milSMS », mais beaucoup de recherches doivent encore être menées. Le corpus de SMS pourra être exploité afin d'élaborer des applications informatiques variées (par exemple, élaboration de lexiques transcodés français standardisé -> SMS ou SMS -> français standardisé

consultables en ligne, mise en place de systèmes de vocalisation des SMS à l'usage de personnes déficientes visuelles ou de personnes momentanément empêchées de consulter leur écran de téléphone – en situation de conduite, etc.). Par ailleurs, il serait envisageable de rendre l'outil d'anonymisation disponible afin qu'il soit réutilisé dans d'autres projets de recherche, voire dans des activités professionnelles soumises à la confidentialité.

L'intérêt de la mise à disposition du corpus 88milSMS sur la grille de services d'Huma-Num (et un éventuel archivage au Centre Informatique National de l'Enseignement Supérieur) relève d'une véritable mission de service public : permettre à un grand nombre de chercheurs et d'étudiants, toutes disciplines confondues, ainsi qu'à des personnes du grand public de tous horizons, de fouiller, d'analyser, d'approfondir nos connaissances à propos des pratiques contemporaines de la textualité numérique pendant de nombreuses années.

► Référence officielle du corpus
"88milSMS. A corpus of authentic text messages in French"
Panckhurst R., Détrie C., Lopez C., Moïse C., Roche M., Verine B. (2014), produit par l'Université Paul-Valéry Montpellier 3 et le CNRS, en collaboration avec l'Université catholique de Louvain, financé grâce au soutien de la MSH-M et du Ministère de la Culture (Délégation générale à la langue française et aux langues de France) et avec la participation de Praxiling, Lirmm, Lidilem, Tetis, Viseo.

Rachel Panckhurst, Catherine Détrie, Bertrand Verine
Praxiling (UMR 5267 CNRS / Université Paul-Valéry Montpellier 3)

► rachel.panckhurst@univ-montp3.fr

► catherine.detrie@univ-montp3.fr

► bertrand.verine@univ-montp3.fr

Cédric Lopez

Objet Direct - VISEO

► cedric.lopez@viseo.com

Claudine Moïse

Lidilem, Université Stendhal Grenoble 3

► claudine.moise@u-grenoble3.fr

Mathieu Roche

Tetis. MTD (Maison de la Télédéttection) UMR TETIS

► mathieu.roche@cirad.fr

Remerciements

Nous remercions la MSH-M (Maison des Sciences de l'Homme de Montpellier), la DGLFLF (Délégation générale à la langue française et aux langues de France) et le CNRS (PEPS ECOMESS, HuMain) qui ont soutenu ce travail. Nous remercions chaleureusement le correspondant informatique et libertés, CIL, Nicolas Hvoinsky (SAJI, Université Paul-Valéry Montpellier 3) de nous avoir accompagnés et conseillés sur le plan juridique, tout au long de notre projet. Nous remercions vivement nos étudiants stagiaires : Anthony Stifani, étudiant en Master Information et Communication à l'Université Paul-Valéry Montpellier 3, qui a manuellement analysé une partie des SMS, permettant ainsi d'évaluer le système d'anonymisation ; Pierre Accorsi et Namrata Patel (étudiants en Master d'Informatique à l'Université de Montpellier 2), qui ont développé le système informatisé *Seek&Hide*, permettant d'anonymiser le corpus ; Michel Otell, Camille Lagarde-Belleville, Frédéric André et Yosra Ghliss (étudiants en Master de Sciences du Langage à l'Université Paul-Valéry Montpellier 3) qui ont procédé à l'anonymisation manuelle en ligne à l'aide de *Seek&Hide* et à la vérification de l'anonymisation automatique du corpus : Aghiles Lounes, Tarik Zaknoun, Zakaria Mokrani, Reda Bestandji, Takfarinas Sider, Ahmed Loudah, (Master I Informatique, Spécialité : « Informatique pour les sciences », Université Montpellier 2) qui ont travaillé sur un système de transcodage automatique.

Huma-Num#Actu

Enquête DARIAH sur les pratiques numériques



Au terme de plusieurs années préparatoires, l'ERIC DARIAH : l'infrastructure européenne de recherche visant à soutenir le développement des humanités numériques, a été créé le 15 août de cette année. Au printemps dernier, l'un des groupes de travail de l'axe « Recherche et enseignement » (VCC2) de DARIAH a élaboré une enquête sur les pratiques numériques des chercheurs européens en SHS. Coordonnée par le Digital Curation Unit (DCU), Athena Research Centre d'Athènes, cette enquête est avant tout le fruit d'une collaboration internationale. Costis Dallas nous en explique ici les objectifs et les enjeux.

Pourquoi cette enquête ?

L'objectif de cette enquête est de comprendre comment les chercheurs européens utilisent les technologies de l'information et du numérique dans leur travail. Au sein de DARIAH, nous misons beaucoup sur la capacité des technologies numériques à ouvrir de nouvelles perspectives pour la recherche en sciences humaines et sociales. Mais nous souhaitons également nous assurer que les gros investissements en infrastructure, dans chaque pays comme au niveau européen, sont justifiés et répondent aux besoins actuels des chercheurs. Dans le cadre du VCC2, nous cherchons en effet à comprendre les différents usages du numérique : comment s'organise la gestion personnelle de l'information dans la recherche scientifique, comment les chercheurs agencent, enrichissent, annotent, diffusent des ressources, ou encore utilisent des outils numériques. Nous souhaitons approfondir notre compréhension des pratiques de travail dans un large spectre de disciplines. Au-delà des pionniers des humanités numériques, nous nous intéressons également à ceux et celles qui s'initient à ces pratiques pour leur recherche.

Enfin, cela nous intéresse de voir en quoi ces pratiques et ces besoins diffèrent, ou sont similaires, dans les pays européens considérés, selon les générations ou bien encore selon les communautés de recherche. C'est la raison pour laquelle l'enquête a été conçue à une échelle européenne ; elle a été diffusée en plusieurs langues et nous espérons pouvoir la mener à intervalles réguliers dans les années à venir. Nous disposerons ainsi d'une vue longitudinale, permettant de comprendre comment cela évolue dans le temps et de comparer différents contextes.

Qui a participé à l'élaboration de cette enquête ?

C'est le Digital Curation Unit (DCU), Athena Research Centre, en Grèce, qui a coordonné et soutenu l'enquête. Depuis cinq ans, le DCU a développé des méthodologies et mené différents travaux

sur les rapports entre pratiques de recherche et environnement numérique. Sur le plan théorique, notre recherche s'inscrit dans le modèle SRAM (*Scholarly Research Activity Model*), que nous avons développé. Ce modèle repose à la fois sur une théorie historico-culturelle des pratiques, sur l'important travail consacré à l'étude des pratiques informationnelles des chercheurs et sur l'extension de la notion de *scholarly primitives*, qui a été proposée par John Unsworth pour saisir les micro-activités récurrentes et modulaires du travail en sciences humaines et sociales. Notre équipe s'est également inspirée d'enquêtes antérieures mises au point pour des projets tels que [EHRI](#) ou [Europeana Cloud](#).

Mais ce travail, à toutes les étapes de son élaboration (conception, mise au point, traduction des questions dans différentes langues, diffusion), résulte bien d'une collaboration internationale au sein de DARIAH (voir encadré).

En quoi cette enquête est-elle différente des enquêtes antérieures, du point de vue de sa portée, de ses cibles, de ses hypothèses et des résultats attendus ?

L'étude des pratiques informationnelles existe maintenant depuis plusieurs dizaines d'années et a fait l'objet de nombreux travaux. Mais les infrastructures numériques constituent un phénomène relativement récent et peu étudié. De plus, l'intérêt pour des questions telles que la gestion personnelle de l'information dans la recherche scientifique, la curation, l'utilisation de l'information scientifique n'est que très récent. DARIAH, par exemple, va bien au-delà du domaine de la recherche d'information (*information retrieval*), en intégrant des activités de curation, d'analyse, d'affichage, de diffusion et des activités collaboratives. La portée de notre travail est donc significativement différente. De plus, notre ambition est plus générale et comparative : la plupart des études antérieures sur les pratiques informationnelles ont été élaborées

Une collaboration internationale

La mise au point de l'enquête résulte d'une véritable collaboration internationale puisque les membres de DARIAH, en particulier de l'axe de travail « Recherche et enseignement », sont intervenus à chacune des étapes de son élaboration. À l'origine du projet, il y a Costis Dallas et Agiatis Benardou (Grèce). Les a rejoint Nephelie Chatzidiakou en 2013. Petra Links (Pays-Bas), Jurij Hadalin (Slovénie) et Ingrida Vosyliute (Lituanie) ont travaillé sur les processus de planification et de révision. Jane Olmeyer, Alex O'Connor (Irlande), Aurélien Berra (France), Stef Scagliola (Pays-Bas), Manfred Thaller (Allemagne), l'ont commentée à plusieurs reprises. Elle a été très fortement soutenue par les coordinateurs du VCC2 : Susan Schreibman (Irlande), Erik Champion et Marianne Huang Ping (Danemark). L'enquête a par ailleurs été traduite en plusieurs langues (notamment en allemand par Zoe Schubert, en serbe par Toma Tasovac, en français par Anne Joly), et sa diffusion a été relayée par les partenaires nationaux dans les différents pays de DARIAH (notamment Bente Larsen au Danemark, Walter Scholger en Autriche, Lorna Hughes au Royaume-Uni).

Si vous souhaitez participer, l'enquête est toujours ouverte et se trouve ici en français : <http://surveys.dcu.gr/index.php/196761/lang-fr>.

pour tester des hypothèses s'appliquant à un contexte national ou disciplinaire particulier. Or, l'environnement intégré qui a émergé en Europe, de même que l'interdisciplinarité croissante de la recherche en sciences humaines et sociales, nous placent presque devant le défi opposé.

Selon vous, quelles conséquences les résultats de l'enquête auront-ils sur les projets de DARIAH, en particulier sur les activités du VCC2, l'axe « Recherche et enseignement » ?

Nous espérons que l'enquête aidera à renforcer l'idée que les infrastructures numériques doivent être pilotées par la recherche. Cette ambition s'inscrit pleinement dans l'esprit des travaux de l'axe « Recherche et enseignement » de DARIAH. Nous souhaitons également que l'enquête fournisse des informations utiles aux chercheurs en SHS qui découvrent les méthodes et les outils du numérique, tout en encourageant la réflexivité méthodologique des chercheurs plus expérimentés dans le domaine — il s'agit d'être attentif à nos pratiques et de porter un regard critique sur notre façon de travailler.

En outre, cette enquête devrait fournir des recommandations précieuses et permettre d'étoffer les arguments en faveur des infrastructures numériques pour les SHS. Alors que plusieurs initiatives nationales sont en train de démarrer dans le cadre des projets ESFRI, elle devrait aider à fixer des priorités de recherche tant au niveau européen qu'au niveau national.

Les résultats de l'enquête pourront ainsi aider à répondre aux questions suivantes : quelles sont les priorités en matière de numérisation ? L'accent doit-il être mis sur les contenus ou sur les procédés ? À quels services ou outils des ressources doivent-elles être allouées ? Les systèmes doivent-ils être généralisés ou spécifiques à des disciplines ou des champs particuliers ?

Enfin, cette enquête s'inscrit dans une entreprise plus vaste. Les résultats de l'analyse quantitative seront en effet mieux compris, mieux interprétés et mieux utilisés s'ils sont mis en relation d'une part avec des « histoires » de chercheurs, précisant quel est leur usage des outils et des ressources numériques, d'autre part avec les ressources associées à l'ontologie que nous sommes en train d'élaborer et qui porte sur les méthodes numériques.

Dans ce contexte, cette enquête participe à la construction d'un observatoire européen des humanités numériques et d'une base de connaissances, qui pourront nourrir directement des initiatives visant à promouvoir les humanités numériques, à mettre en réseau les communautés et à développer des activités d'enseignement et de formation, tel que cela a été prévu par les autres groupes de travail du VCC2.

Entretien¹ mené et traduit par Aurélien Berra (Université Paris-Ouest Nanterre La Défense) et Anne Joly (TGIR Huma-Num).

1. La version intégrale de cet entretien en anglais est disponible sur le site DARIAH-FR

contact&info

► Anne Joly,
Huma-Num
Anne.Joly@huma-num.fr
Aurélien Berra,
ArScAn
aurelien.berra@u-paris10.fr

contact&info

► Nadine Dardenne
Chargée de la communication
et de la structuration des réseaux
nadine.dardenne@huma-num.fr
► Pour en savoir plus
www.huma-num.fr

la lettre de l'InSHS

- ▶ **Directeur de la publication** Patrice Bourdelais
- ▶ **Directrice de la rédaction** Michèle Baussant
- ▶ **Responsable éditoriale** Armelle Leclerc armelle.leclerc@cnrs-dir.fr
- ▶ **Conception graphique** Sandrine Clérisse & Bruno Roulet, Secteur de l'imprimé PMA
- ▶ **Graphisme Bandeau** Valérie Pierre, direction de la Communication CNRS
- ▶ **Crédits images Bandeau**
© Photothèque du CNRS / Hervé Théry, Émilie Maj, Caroline Rose, Kaksonen
- ▶ **Pour consulter la lettre en ligne**
www.cnrs.fr/inshs/Lettres-information-INSHS/lettres-informationINSHS.htm
- ▶ **Pour s'abonner / se désabonner**
com-shs@cnrs-dir.fr
- ▶ **Pour accéder aux autres actualités de l'INSHS**
www.cnrs.fr/inshs

Institut des sciences humaines et sociales CNRS

• 3 rue Michel-Ange 75794 Paris cedex 16 •