



# Journées EVEille (Exploration et valorisation Electroniques des corpus en Sciences Humaines et Sociales)

## *Constitution des données de la recherche*

Animées par Anne Réach-Ngô, Marine Parra et Régine Battiston

*janvier-juin 2021*

Organisé en cinq journées mensuelles de janvier à juin 2021, le colloque EVEille organisé par l'Institut de recherche en Langues et Littératures Européennes de l'Université de Haute-Alsace (UR 4363) vise à lancer, au sein de l'ILLE et en collaboration avec d'autres acteurs du monde des SHS et des humanités numériques, une réflexion collective sur les usages du numérique, les atouts et innovations mais aussi les obstacles à sa mise en œuvre, ainsi que les évolutions à venir des pratiques de recherche. Le projet EVEille entend prendre ses distances avec le discours du « tout-numérique » et revenir à une réflexion sur un humanisme numérique fondé sur la mise en commun des expérimentations intellectuelles, des explorations informatiques et leur appropriation par les usagers, qu'ils soient chercheurs, bibliothécaires et documentalistes ou encore responsables de centres culturels.

L'objectif de cet espace de réflexion n'est pas de défendre un nécessaire virage des SHS en direction des humanités numériques, malgré les fortes incitations des organismes de financement et d'évaluation de la recherche qui par la pression imposée en dénaturent les méthodes et les questionnements. Il s'agit en revanche de donner aux participants qui s'interrogent sur l'intérêt des humanités numériques les moyens d'examiner si une telle orientation peut – ou non – être pertinente

dans le cadre de leurs recherches ou dans le traitement des biens et objets patrimoniaux et culturels dont ils ont la charge. Il s'agit également de permettre aux porteurs de projets en humanités numériques de présenter leurs travaux et d'échanger sur la mise en réseau de ces différents projets, qui restent encore parfois mal intégrés aux dynamiques d'équipes ou à l'environnement culturel local.

Comment certains projets de recherche ou de valorisation patrimoniale et culturelle peuvent-ils gagner à s'orienter en direction du numérique ? Dans quelle mesure les méthodologies employées s'ancrent-elles dans des pratiques plus anciennes qui ont fait leurs preuves ? Tous les projets se prêtent-ils à une telle approche ? Comment s'y prendre pour engager, poursuivre, relancer ou enrichir un projet de recherche à l'aide des humanités numériques ?

### Contacts :

[Projet-eveille@uha.fr](mailto:Projet-eveille@uha.fr)

Inscription sur [Sciencesconf](https://www.sciencesconf.org/)

Carnet du projet [EVEille](https://www.eveille.uha.fr/)

Chaîne vidéo sur [Uha pod](https://www.uha-pod.com/)



[@Eveille HN](https://www.facebook.com/EveilleHN)



[@eveille\\_hn](https://www.instagram.com/eveille_hn)



[@Eveille\\_HN](https://www.twitter.com/Eveille_HN)



28 MAI 2021 JOURNÉE 5. VISUALISATION DES RÉSULTATS DE LA  
RECHERCHE, GRAPHES ET SCHÉMAS DE DONNÉES

9h RÉGINE BATTISTON, ANNE RÉACH-NGÔ et MARINE PARRA (Université de Haute-Alsace, ILLE), Ouverture de la journée

9h15-10h15 Conférence inaugurale

PIERRE CUBAUD (CNAM, CEDRIC), « Visualisation pour les bibliothèques numérisées »

Portée par les progrès de l'informatique graphique, la visualisation est devenue un outil incontournable pour l'analyse et le suivi de grands volumes de données. L'application de ces procédés aux grands corpus documentaires est ancienne mais son usage systématique est assez récent, et surtout restreint aux outils d'analyse de contenu des collections. Nous pensons — avec d'autres — que la visualisation peut avoir un rôle plus important encore dans les interfaces, pour un intermédiaire encore à inventer entre les outils de recherche et les outils de lecture. La visualisation peut ainsi apporter une réponse aux besoins de déambulation et d'appropriation rapide des collections numérisées. En appui de notre thèse, nous présenterons quelques travaux réalisés au CEDRIC depuis une trentaine d'année pour divers projets collaboratifs issus des sites ABU et Cnum : des interfaces 3D de déambulation, un mur de pages pour le contrôle qualité dans une usine de numérisation et un espace de conception pour les corpus de correspondances.

10h30-11h45 Retours d'expérience animés par ANNE RÉACH-NGÔ

LAURENT CAPRON (Centre Jean Pépin, CNRS), « Cartographie des philosophes antiques : voir moins, c'est voir mieux ! »

L'objectif premier du projet « [Cartographie des philosophes antiques](#) » était d'offrir une visualisation spatiale des philosophes de l'Antiquité qui montrerait la vacuité de la représentation inconsciente et commune des philosophes concentrées dans les villes d'Athènes puis de Rome, et mettrait en valeur la riche diversité des provenances et des centres culturels antiques. Mais il a rapidement fallu faire face à la maigreur des données disponibles et fiables en comparaison du nombre de noms de philosophes connus. La structuration des données selon des critères scientifiques bien définis, intégrant la part d'incertitude quant aux informations disponibles, nous a obligés à choisir certains modes de visualisation plutôt que d'autres. Plutôt que de pouvoir tout visualiser, nous avons préféré mieux montrer la réalité de nos connaissances.

THOMAS MOHNIKE (Université de Strasbourg, Mondes germaniques et nord-européens) et LUDOVIC STRAPPAZON (Université de Strasbourg, DNU), « Cartographier le savoir narratif – approches numériques aux mythes du Nord en circulation culturelle »

Le projet [Mythèmes du Nord](#) cherche à identifier les unités du savoir narratifs (mythèmes) dans des textes narratifs traitant ou venant du Nord de l'Europe et au-delà avec des approches computationnelles et à cartographier leur grammaire discursive. Nous utilisons des logiciels d'analyse développés ou déployés dans notre « laboratoire des mythèmes » et des logiciels de visualisation comme Gephi pour créer des cartes des univers narratifs, les horizons d'attente (Jauss) et les changements de la grammaire du savoir social dans le temps.

ALEXANDRE ZEITLER (Université de Strasbourg, Mondes germaniques et nord-européens), « Cartographier les réseaux du savoir dans le Finnmark norvégien : méthode(s) et enjeux »

À partir des savoirs écologiques locaux recueillis par les autorités norvégiennes pour déterminer les droits de propriété et d'usage sur les îles de Stjernøya et de Seiland en Norvège, j'esquisserai dans un premier temps un cadre méthodologique pour cartographier les réseaux de production du savoir. Pour ce faire, mon étude prendra appui sur les outils des humanités numériques : de la cartographie GIS à la visualisation Gephi en passant par les outils de *text mining*. Les premiers résultats obtenus avec cette méthode d'analyse pour le corpus primaire laissent entrevoir trois clusters de savoirs (le renne, l'activité minière et la pêche) produits par les communautés des deux îles pour construire leurs identités. Afin de donner une vue d'ensemble du potentiel de cette méthode, j'exposerai dans un second temps les résultats obtenus à partir d'un corpus secondaire : les déclarations du Conseil sámi.

LAURENT MILLET-LACOMBE (MetaindeX), « MetaindeX : statistiques d'un corpus avec Kibana »

L'outil open-source [MetaindeX](#) propose un environnement de travail en ligne avec les services de base que pourraient attendre une institution ou un chercheur pour le catalogage, la consultation et l'étude de ses données. Il se base sur ElasticSearch pour le stockage et l'interrogation des données, et intègre l'outil Kibana pour la création de graphiques statistiques complexes que l'utilisateur configure lui-même. La présentation portera sur la manière d'utiliser Kibana au travers de quelques exemples simples et didactiques, qui s'appuient sur un corpus de démonstration importé via MetaindeX.

11h30 Discussions

11h45-12h30 *Speed searching*

Cette session vise à donner la parole à de jeunes chercheurs ou à des chercheurs confirmés qui débutent un projet et s'interrogent sur les choix à mener en matière d'outils et de méthodologie numériques. Le défi d'une présentation de 5 minutes « montre en main » les incitera à caractériser la singularité de leur projet de recherche et à formuler leurs attentes numériques pour aborder le questionnement scientifique qui leur est propre.

MARIE-LOU SOLBACH (Université de Haute-Alsace, ILLE), « De la visualisation à la rédaction de thèse : cartographie d'enquête(s) au sein d'un corpus de polars polaires »

CLAIRE DAVRIL (Université de Strasbourg, ACCRA), « Une réflexion par l'image. Propos parallèle sans nombre de signes ? »

GEOFFREY KOENIG (Université de Strasbourg, Service historique de La Défense), « L'idéologie nazie et sa langue. La lexicométrie comme outil d'exploration de la *Lingua Tertii Imperii* »

Pause déjeuner

13h45-15h Atelier d'initiation aux outils numériques

Séance d'initiation à l'analyse de réseaux, animée par PABLO RUIZ FABO (Université de Strasbourg, LiLPa)

Les réseaux constituent une structure pratique pour représenter des données relationnelles de façon à donner un aperçu de groupes et interactions. Dans l'atelier nous regarderons quelques caractéristiques des réseaux ainsi que des outils qui permettent de les tracer et analyser, comme Cytoscape ou Gephi. Une application des réseaux à l'analyse textuelle avec l'outil CorText Manager sera également montrée.

15h15-16h30 Projets collaboratifs en humanités numériques

Table ronde « Genèse et visualisation », animée par ANNE RÉACH-NGÔ

En mettant au cœur de l'étude le processus plus que le produit fini et le geste créateur plus encore que l'effet sur le lecteur, l'analyse des phénomènes de genèse littéraire pose pleinement la question des modes de restitution des résultats de la recherche. Les méthodes et outils qui relèvent de la visualisation constituent à cet égard une entrée éclairante pour interroger la publication de ces enquêtes et la manière dont elles viennent renouveler le regard sur une œuvre, un corpus, un auteur, un fait artistique.

MATHILDE LABBÉ (Universités de Nantes, L'AMo), *Éditer L'Étoile du Sud*

Le projet « [Éditer L'Étoile du Sud](#) » se concentre sur un objet dont la genèse comporte des zones d'ombre : si l'on dispose de l'avant-texte rédigé par Paschal Grousset pour ce roman édité sous le nom de Jules Verne, le travail de ce dernier sur le texte du premier n'est accessible par aucun document connu. Il s'agit ici d'étudier un cas de pluriauctorialité contrainte et, par la suite, de rendre visible la complexité de ce palimpseste.

FRANZ JOHANSSON (Sorbonne Université, ITEM), *Robinson de Paul Valéry : édition génétique*

[Robinson de Paul Valéry : édition génétique](#) se propose de reconstituer le dossier de genèse d'un écrit demeuré, comme bien d'autres projets chez son auteur, inachevé : le conte Robinson. Elle aspire également à déterminer la cohérence des différents éléments et le jeu complexe de leurs interactions.

HAKIM USOOF (University of Peradeniya, Sri Lanka, Dept. of Statistics & Computer Science), CHRISTOPHE LEBLAY (Université de Turku, Finlande, Département des langues et Traduction) ET MINJING ZHONG (Université de Xiangtan, Chine, Département de français), *GenoGraphiX*

Le programme [GenoGraphiX](#) a été développé à partir de la génétique textuelle, de la théorie mathématique des graphes et des sciences de l'informatique. Ce programme met l'accent sur l'utilisation de la visualisation des données scripturales afin de rendre compréhensible l'abondance et la complexité de ces données pour l'apprentissage et l'enseignement de l'écriture ainsi que pour la recherche.

16h15 Discussions et clôture de la journée

