Maison des sciences humaines et sociales en Bretagne

Séminaire HN 2022 #1 - Utiliser des outils et des services numériques dans sa pratique d'analyse de discours (oraux et écrits)



Organisé par la Maison des sciences de l'homme en Bretagne depuis 2014, le séminaire Humanités numériques prend cette année la forme d'un cycle de trois séances qui propose une réflexion sur les pratiques de recherche dans un contexte de transformation numérique et de développement de la science ouverte. Il vise à interroger de manière critique et réflexive les conditions de production et de réutilisation des données dans la recherche en sciences humaines et sociales, mais aussi l'expérimentation de méthodes permettant de traiter et analyser ces données. Son objectif est d'ouvrir un espace de discussion sur ces différentes problématiques, en s'appuyant sur la présentation de projets de recherche en cours et sur la mise en partage de questionnements.

Cette séance se tiendra le mercredi 6 avril, 14h-17h, en visioconférence.

Programme

Cette séance sera animée par Jean-Baptiste Pressac (CRBC, CNRS/Université de Bretagne Occidentale) et Aurélie Hess (TEMOS, CNRS/Université Bretagne Sud).

Présentation de Lexico 3, logiciel de lexicométrie, Gudrun Ledegen (PREFics, Université Rennes 2)

Les logiciels de lexicométrie, ici présentés à travers Lexico 3, sont des outils d'exploration de grands corpus textuels qui permettent un va-et-vient entre la comparaison entre différentes variables qui viennent baliser les documents et un retour au texte. Si Lexico 3 présente des fonctionnalités proches d'un logiciel comme Iramuteq, il va être spécifiquement au service d'une analyse de discours (que d'une analyse de contenu avant tout). Le logiciel permet ainsi d'établir des liaisons transversales entre les mots fonctionnant en réseau : « le numérique complète-augmente la linéarité par des réticularités » (Mayaffre, 2015). Il s'agit d'une méthodologie d'étude du discours, dite exhaustive, systématique et automatisée, tout en laissant à l'analyste l'entière tâche d'interprétation des résultats obtenus. Les fonctionnalités complémentaires du logiciel Iramuteq – qui sera présenté par d'autres collègues lors de ce séminaire – seront ici particulièrement mises en lumière.

Interpréter avec Alceste (ou IramuteQ), Catherine Adam et Cécile Plaud (ENSTA Bretagne)

Une analyse de discours assistée par ordinateur (ADAO) réalisée à l'aide d'un logiciel d'analyse statistique de données textuelles, tel qu'ALCESTE ou IramuteQ semble particulièrement appropriée dans le cadre d'analyses transversales de corpus volumineux. L'ADAO permet non seulement de limiter la perte d'informations, inhérente à une analyse manuelle des données d'un vaste corpus, mais favorise aussi la distanciation du chercheur par rapport à ses données. Dans le cadre de cette communication, nous montrerons, au travers d'exemples d'usages de ce type de logiciel dans nos recherches en SHS, la valeur heuristique de la démarche. Elle offre la possibilité à la fois de découvrir de manière inductive d'autres aspects non analysés du corpus et d'obtenir une forme de confirmation ou d'infirmation des résultats obtenus manuellement. L'usage d'outils numériques dans des travaux de recherche, lorsqu'ils peuvent se révéler un soutien à la démarche, est à considérer. Bien entendu, un logiciel ne fournit pas à lui seul les résultats d'une recherche. Il est paramétré en fonction des hypothèses du chercheur. Ce dernier doit ensuite analyser et interpréter les données classifiées. Cependant, ce recours à l'outil numérique se révèle être une aide précieuse et un atout dans une méthodologie de la complexité.

Un entrepôt spécialisé pour les corpus de parole: la plateforme COCOON, Michel Jacobson (IR* Huma-Num, CNRS)

La plateforme Cocoon a été mise en place en 2006 et assure depuis plus de 15 ans des services de gestion pour les corpus oraux des chercheurs, des laboratoires et des projets en SHS. Elle donne accès actuellement à environ 6000 heures d'enregistrements de parole sur plus de 250 langues. Cette plateforme s'est montée autour d'une architecture d'entrepôt spécialisé sur les données orales enregistrées (audio, vidéo ou autres mesures physiologiques) accompagnées ou non d'annotations.

Les services assurés depuis le début sont ceux d'un entrepôt classique (stockage, accès, description, signalement, conservation). Au cours du temps, de nouvelles fonctionnalités sont venues enrichir les services parmi lesquelles nous pouvons citer à titre d'exemples : le versement dans un service d'archives intermédiaire (celui de l'opérateur CINES), l'alignement des métadonnées sur des référentiels largement partagés (VIAF, RAMEAU, GeoNames, Lexvo), une modélisation en EDM (Europeana Data Model), une exposition des métadonnées en mode web de données complétant celle en OAI-PMH, l'attribution d'identifiant de type DOI, etc.

Informations pratiques

Les inscriptions sont closes. Les codes de connexion à la visioconférence seront transmis par mail aux participant·es la veille ou le jour du séminaire.

Le séminaire Humanités numériques

Le séminaire Humanités numériques est organisé par la Maison des sciences de l'homme en Bretagne, avec la participation du groupe de travail Humanités numériques. Formé en 2016, ce GT interdisciplinaire rassemble à l'échelle régionale des chercheur·ses, ingénieur·es et professionnel·les de la documentation et de l'information scientifique et technique.

Le séminaire est soutenu par le Réseau national des Maisons des Sciences de l'Homme et l'association Humanistica.

Comité scientifique du séminaire

- **Aurélie Hess**, ingénieure en traitement et analyse de base de données, UMR TEMOS, CNRS/Université Bretagne Sud
- **Morgane Mignon**, coordinatrice de la plateforme Humanités numériques, UAR MSHB
- **Gwenaëlle Patat**, chargée d'édition de corpus numériques, UAR MSHB
- **Jean-Baptiste Pressac**, responsable de la plateforme CRBC Open Science, UAR CRBC, CNRS/Université de Bretagne Occidentale

Vidéo séminaire HN 2022 #1 | Utiliser des outils et des services numériques dans sa pratique d’analyse de discours (oraux et écrits)