

Séminaire de recherche « Les données à mille temps »

15 novembre 2019, Enssib, Villeurbanne

10h30 – 16h30

Titre de la séance « Rendre perceptible l'imperceptible : les données de mesures du temps »

Cette séance sera structurée essentiellement autour de la problématique des données dite temporelles ou « *time-oriented* ». Ces dernières sont liées essentiellement au temps physique ou chronométrique, et de ce fait, construites par des instruments de mesure du temps (montres à quartz, horloges atomiques, etc.). Leurs valeurs représentent les unités de base du temps (telle que la seconde dans le SI) et leurs sous-multiples et multiples (... , microseconde, milliseconde, minute, heure, jour, année, ...), et nous servent alors à dater et indexer par le temps, d'autres données représentant des phénomènes, événements, états, objets, etc.

10h30 – Introduction, Agnieszka Tona, Enssib

10h50 – « Une approche temporaliste du Web. Temporalités des données et temporalités d'un événement professionnel », Jean-Claude Domenget, Maître de conférences en sciences de l'information et de la communication, Eliiadd, Université Bourgogne Franche-Comté.

L'intervention se fera en deux temps. En premier lieu, certains principes d'une analyse temporaliste d'un objet de recherche en SHS seront rappelés, en s'intéressant particulièrement au cas du Web. Il s'agit en effet de passer d'un objet abstrait, le temps, à des objets concrets les temporalités, de s'appuyer sur la richesse notionnelle développée par les études sur les temporalités, d'analyser le Web comme un dispositif sociotechnique complexe ou encore d'en faire un objet d'étude foncièrement interdisciplinaire. En second lieu, une étude de cas portant les usages de Twitter par des professionnels du référencement Web permettra d'analyser les relations entre temporalités des données partagées et temporalités d'un événement professionnel. Les données partagées ont été recueillies à la suite d'une périodisation de l'événement. L'approche temporaliste d'une analyse structurale de réseaux permet alors d'analyser les modalités de reproduction des formes d'autorité entre professionnels, notamment lors de temps forts de l'événement.

11h40 – « Les différentes temporalités des données produites par les historien.ne.s », Pierre Vernus, Maître de conférences en histoire contemporaine, LARHRA, Université Lumière Lyon 2.

Les données numériques créées ou utilisées par les historien.ne.s ont à voir avec le temps, à plusieurs titres. En premier lieu, comme procédé de connaissance l'histoire repose sur des traces. Les données collectées pour construire les faits historiques le sont à partir d'informations chronologiquement situées. Mais la date ou la durée d'un événement n'est souvent connue que de manière incertaine ou approximative. Il convient donc de pouvoir enregistrer cette incertitude ou cette approximation. Par ailleurs, la phase d'analyse des données implique des opérations de mise en ordre, de rapprochement ou de comparaison chronologiques facilitant la construction des faits, la détermination des imputations causales et, in fine, la compréhension des processus étudiés. Enfin, l'essor des technologies du web et le mouvement des linked data permettent d'envisager l'ouverture des silos individuels dans lesquels les historien.ne.s stockent leur données au cours de leurs recherches et le passage à une collecte collaborative et cumulative. Cette perspective impose de réfléchir à la question du cycle de vie des données historiques et aux conditions autorisant une telle évolution des pratiques concrètes de la recherche. Cette communication repose sur la réflexion menée au sein du pôle histoire numérique du LARHRA.

12h30 – Déjeuner libre

14h00 – « **Dater et indexer le temps, données et métadonnées en bibliothèque** », **Marion Jonget**, Numelyo et Service catalogues et métadonnées, Bibliothèque Municipale de Lyon.

La notion de datation et l'indexation du temps constituent une pierre angulaire dans la description d'un objet en bibliothèque. Le traitement de ces données temporelles, aujourd'hui mises en ligne, répond à des normalisations qui varient en fonction du type d'objet décrit, des thésaurus utilisés, des pratiques bibliothéconomiques... Autant de possibilités faisant ainsi du temps un véritable sujet pour le bibliothécaire chargé de la gestion des données pour valoriser les collections sur le web....

14h50 – « **Conserver les données de la recherche dans un centre de calcul : tout est question de temps !** », **Jean-Yves Nief**, Ingénieur de recherche, Centre de calcul de l'IN2P3 Lyon.

Un centre de calcul est amené à stocker les données de milliers de chercheurs travaillant dans de nombreux projets scientifiques de l'échelle locale à internationale. Ces projets peuvent avoir une longue durée de vie pouvant s'étendre sur plusieurs décennies comme en physique des particules ou en astrophysique. Leurs données doivent être conservées sur des durées pouvant être encore plus longues afin de les exploiter soigneusement. Si on y ajoute l'injonction des pouvoirs publiques à rendre les données de la recherche en accès libre, cela implique une conservation de celles-ci sur le plus long terme. Mais est-il vraiment possible voire souhaitable de le faire en toute circonstance ? Un centre de traitement des données est confronté à des défis technologiques liés à l'évolution rapide dans le temps des supports matériels et des couches logicielles. L'accumulation des données dans le temps posent aussi des problèmes du point de vue économique. Enfin, des obstacles du point de vue humain peuvent aussi apparaître. Le temps est donc aussi un facteur important pour un centre informatique même si ce facteur est souvent minimisé voire ignoré par les acteurs de la recherche.

15h40 – 16h00 – Clôture de la séance, **Agnieszka Tona**, Enssib