

« GESTION NUMÉRIQUE DES SOURCES DE LA RECHERCHE EN SHS « Produire et utiliser des corpus de sources numériques »

Action de formation nationale CNRS 2010
Centre Paul Langevin du CAES-CNRS,
Aussois (Alpes), du 11 au 15 octobre 2010.

Présentation

Aujourd'hui les projets d'édition électronique des sources de la recherche (manuscrits, corpus iconographiques, audiovisuels, etc.) se multiplient. De plus en plus de chercheurs souhaitent diffuser et rendre accessible, sur internet, les matériaux originaux leurs ayant permis de produire leurs recherches. Des bibliothèques de recherche et des centres de documentation en SHS souhaitent valoriser leur patrimoine scientifique. Cette valorisation par l'édition électronique accroît les possibilités de diffusion des données de la recherche et permet de nouvelles investigations. La production et l'utilisation de sources numériques contribuent au développement du web de données, cette extension du web qui vise à améliorer la diffusion des savoirs.

Comment se lancer dans un projet d'édition électronique et de diffusion de sources ? Cela passe bien évidemment par le dialogue entre chercheurs et éditeurs ou informaticiens, par la prise en compte de problématiques spécifiques comme la pérennisation de l'information, l'interopérabilité des données ou de nouvelles possibilités de recherches et d'exploitation. Il y a de nouveaux usages qui s'instaurent.

Une première action de formation, l'école thématique « Préservation et diffusion numériques des sources de la recherche en sciences humaines et sociales » (octobre 2008) a déjà permis de construire une synthèse de ces nouvelles pratiques. Une université d'été, proposée par le centre pour l'édition électronique ouverte (CLEO, septembre 2009), a permis de mieux comprendre les enjeux de l'édition électronique de la production des chercheurs.

Dans ce contexte, nous proposons de donner aux porteurs de projet la connaissance des moyens possibles et des erreurs à éviter pour mettre en ligne leurs sources qu'ils utilisent. Pour cela, des acteurs du domaine des humanités numériques se sont associés au Service éditorial et publication électronique de l'Institut de recherche et d'histoire des textes (SEPE, IRHT-CNRS) afin de proposer une action nationale à gestion déconcentrée (ANGD), soutenue par le TGE Adonis, sur la production, l'édition, la gestion et la diffusion des sources numériques de la recherche en sciences humaines et sociales.

À partir d'exemples concrets et de mises en situation, cette formation retracera tout le parcours de gestion de projet, de l'analyse des besoins et du cahier des charges, jusqu'aux méthodes de codage et d'édition, la diffusion sur internet, etc. Ces ateliers seront organisés avec un nombre minimum d'intervenants pour privilégier le suivi des projets et la prise en main progressive des outils et méthodes. Ces ateliers pratiques avec analyse de documents, simulation de situation, travaux de codage, mise en forme et diffusion de documents apportés par les stagiaires (ex : un extrait du corpus qu'ils devront éditer ou un exemple d'archive ou de source qu'ils gèrent). Pour éviter une hétérogénéité importante de ces documents, des consignes préalables seront données aux participants. L'apport de ces documents sera un critère de participation.

Les ateliers suivront l'ordre chronologique du déroulement d'un projet numérique. Les stagiaires verront ainsi toute la chaîne du traitement des documents, les étapes et les procédures nécessaires, les outils de gestions (analyse de besoins, planning, etc.). On abordera des thématiques comme l'importance du cahier des charges, la définition et la production de métadonnées normalisées, le choix des méthodologies ou des outils de construction, la notion de « livrable », les procédures de travail et leur documentation, les outils collaboratifs (wiki, blogs), l'étape cruciale de l'indexation, les solutions d'interopérabilités des données, etc. Tout au long de la formation, l'accent sera mis sur les conditions et techniques de pérennisation des données numériques et leurs enjeux méthodologiques.

À travers l'organisation d'atelier progressif et qui se veulent surtout pratiques, nous voulons privilégier le dialogue et l'expression des participants. Nous pourrions ainsi confronter des témoignages et retours d'expérience qui affineront outils et méthodes utilisés. Cela permettra de faire émerger plus facilement les besoins réels d'intervention sur le matériau et les solutions à y apporter. Pouvoir analyser et synthétiser la confrontation à un corpus numérique ou à numériser est l'étape indispensable pour une bonne réalisation du projet.

« GESTION NUMÉRIQUE DES SOURCES DE LA RECHERCHE EN SHS « Produire et utiliser des corpus de sources numériques »

Action de formation nationale CNRS 2010
Centre Paul Langevin du CAES-CNRS,
Aussois (Alpes), du 11 au 15 octobre 2010.

Organisation :

Action organisée par le Service Éditorial et Publications électronique (SEPE) de l'IRHT avec les centres de ressources numériques Telma (Traitement électronique des manuscrits et archives) et CN2SV (Centre national pour la numérisation de sources visuelles), avec le soutien et la collaboration du TGE Adonis, du réseau Mutec et du Centre d'études supérieures de la Renaissance (CESR). L'ANGD est gérée par la délégation Paris A et soutenue par la Formation Professionnelle du CNRS.

Comité d'organisation

Responsables : Richard Walter, Stéphane Pouyllau.
Christine Berthaud, Thierry Buquet, Marjorie Burghart, Lou Burnard, Nadine Dufournaud, Véronique Ginouvès, Maud Ingarao, Michel Jacobson, Shadia Kilouchi, Jean Philippe Magué, Emmanuelle Morlock-Gerstenkorn, Cyril Masset, Olivier Marlet, Fabrice Melka, Pierre Mounier, Denise Ogilvie, Joséphine Simonnot, Delphine Usal, Anais Wion, Marie-Hélène Wronecki
Délégation Paris A : Chrystèle Guilloteau

Les publics

Prioritairement

- Chercheurs dans le domaine des Sciences humaines et sociales qui se destinent à la diffusion de leurs sources ou qui ont un projet d'édition de corpus
- Ingénieurs d'études ou de recherche ou équivalents, chargés de la mise en œuvre de projets de diffusion électronique de sources sur le web.

Autres publics potentiels

- Professionnels de la conservation du patrimoine, afin qu'ils prennent davantage conscience de certains enjeux et connaissent mieux certains outils (normes, concepts, logiciels) lorsqu'eux-mêmes décident de mettre en ligne des sources.

Pré-requis

Il sera recommandé d'avoir des notions de bases dans l'utilisation d'un système d'exploitation (Windows, Linux, Mac) et en matière de structuration de l'information en SHS. Des notions en matière de langage XML seraient un plus. Seront également appréciées des notions en matière de gestion de projets et d'informatisation des sources pour les projets de recherche. Il sera également attendu que le stagiaire soit impliqué, à court ou moyen terme, dans un projet d'édition de sources numériques.

Financements et prise en charge

- L'hébergement au centre Paul Langevin du CAES-CNRS est pris en charge par la délégation du CNRS de Paris 1 pour tous les participants.
- Les transports sont pris en charge par le service de formation de la délégation de rattachement pour les agents CNRS et par l'organisme de tutelle (Université ou autre) ou le laboratoire pour les agents non-CNRS.

Pour tout renseignement sur la formation : irht.sepe@cns-orleans.fr

**« GESTION NUMÉRIQUE DES SOURCES DE LA RECHERCHE EN SHS
« Produire et utiliser des corpus de sources numériques »**

Action de formation nationale CNRS 2010
Centre Paul Langevin du CAES-CNRS,
Aussois (Alpes), du 11 au 15 octobre 2010.

Comment s'inscrire ?

Les inscriptions seront closes le 15 juin 2010. Réponses avant le 1^{er} juillet 2010.
Envoyer un courrier électronique à <irht.sepe@cnrs-orleans.fr> avec dans le corps du message :

NOM
Nationalité
Diplômes
Statut
Affiliation
Laboratoire

Joindre en fichiers attachés deux documents distincts d'une page chacun impérativement :
Un CV

Une lettre de motivation avec

Une présentation du projet
Une présentation du corpus
L'indication du type de corpus :
– corpus textuels
– corpus images
– corpus oraux
L'état du corpus