

Exemple d'utilisation des outils du web dans un processus collaboratif de production de métadonnées

inra.fr

L'université de North Texas (UNT) et la Société d'Histoire de l'Oklahoma (OHS) collaborent dans un projet de numérisation et de publication de plus d'un million de photographies d'archives historiques (disponibles sur le portail The Gateway to Oklahoma History). Pour ce projet, ils sont mis au point un processus dans lequel ils utilisent les outils du web et pour vérifier et nettoyer les métadonnées produites au moment de la numérisation. Ce processus permet aux 2 institutions de se partager les tâches et de travailler simultanément sur les métadonnées, chacun s'attachant à un type de données et à un aspect différent du nettoyage (lieux, images, dates...etc.). Le processus comporte 5 étapes : préparation de l'image pour la numérisation, saisie de l'image et génération des métadonnées (groupées par lots de 40 000 images/métadonnées au format csv), nettoyage des métadonnées à l'aide d'Open Refine et Google Fusion Tables, mapping des métadonnées et conversion au format Dublin Core qualifié en XML (outil m2m développé localement), dernières modifications et contrôle de la qualité des métadonnées obtenues (via le portail).

Référence :

Mark Phillips, Hannah Tarver, and Stacy Frakes (2013). **Implementing a Collaborative Workflow for Metadata Analysis, Quality Improvement, and Mapping**. *code4lib Journal*, 23, article 9199. <http://journal.code4lib.org/articles/9199>