

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

Résumé

Projet sur l'histoire des femmes dans le domaine des médias imprimés de 1750 à 1830 –
Résumé du plan de gestion des données

Ce projet a permis de recueillir des données à partir de bibliographies imprimées et de bases de données numériques afin d'élaborer la première base de données bibliographiques complète sur la contribution des femmes à la culture de l'imprimé de 1750 à 1830. L'objectif de cette base de données accessible au public est de permettre de nouvelles interprétations quantitatives de la participation des femmes à la culture de l'imprimé. Ce plan de gestion des données est accompagné d'un schéma de métadonnées élaboré par les chercheurs pour décrire le contenu unique de la base de données (lien fourni à l'intérieur).

Chercheuse principale : Michelle Levy (Simon Fraser University)

Détails administratifs

Nom du projet : Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

Titre de la subvention : 31 – R639951

Chercheuse principale / chercheuse : Michelle Levy

Responsable des données du projet : cgraebne@sfu.ca

Description : Le projet sur l'histoire des femmes dans le domaine des médias imprimés de 1750 à 1830 constituera la première base de données bibliographiques complète de la contribution des femmes aux médias imprimés pendant l'une des périodes les plus mouvementées de l'histoire des femmes et des médias imprimés. Les données tirées de bibliographies imprimées et de bases de données numériques sont regroupées et normalisées afin de permettre de nouvelles interprétations quantitatives de la participation des femmes à la culture de l'imprimé. La base de données, qui sera mise à la disposition d'autres chercheurs,

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

permettra la réalisation d'une analyse quantitative rigoureuse des tendances au chapitre de l'histoire des femmes dans le domaine des médias imprimés, en faisant fond sur le vaste corpus d'ouvrages qualitatifs sur l'histoire des femmes dans le domaine de la littérature qui ont été produits au cours du dernier demi-siècle. Voici quelques-unes des questions à poser au sujet des données : Comment la production des femmes dans le domaine des médias imprimés a-t-elle changé au fil du temps? Dans quelle mesure les femmes étaient-elles susceptibles de publier plus d'un livre au cours de leur vie, et à quelle fréquence les livres ont-ils fait l'objet d'une deuxième édition ou plus? Certains éditeurs imprimaient-ils plus d'ouvrages produits par des femmes que d'autres? À quelle fréquence les femmes ont-elles changé d'éditeurs au cours de leur carrière? Quels réseaux sociaux ont été créés grâce à la collaboration des femmes avec les éditeurs, les imprimeurs et d'autres femmes? Comment pourrions-nous mieux comprendre la contribution des femmes à l'histoire des médias imprimés dans divers rôles (en tant qu'auteures, mais aussi en tant que coauteures, rédactrices en chef, traductrices, mécènes, graveuses ou libraires)?

Établissement : Simon Fraser University

Collecte des données

Quels types de données recueillerez-vous, créerez-vous, couplerez-vous, acquérez-vous ou consignerez-vous?

Lorsque vous recueillez des données de sources écrites, il est utile de fournir au lecteur des renseignements sur les documents à partir desquels vous recueillerez les données, ainsi que quelques exemples des ressources desquelles vous puiserez pour trouver ces documents.

Nous recueillerons des métadonnées bibliographiques sur les titres d'ouvrages écrits, les personnes et les entreprises comme le titre de l'ouvrage, le nom de l'auteur, de l'éditeur, la date de publication, etc. En ce moment, nous recueillons des données à partir d'environ 50 différentes sources, certaines étant imprimées, la plupart étant numériques. Citons, par exemple, le catalogue en ligne de la British Library et des bibliographies imprimées spécialisées comme *The English Novel 1770-1830*.

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

Il est utile d'indiquer d'où proviennent les données que vous recueillez en utilisant des exemples de bibliographies imprimées et de bases de données à partir desquelles vous réunirez cette information.

Dans quels formats de fichiers les données seront-elles recueillies? Ces formats permettront-ils la réutilisation et le partage des données, de même que l'accès à long terme à celles-ci?

Les données proviennent de certaines sources qui nous ont été fournies (des MARC¹ records, par exemple, d'ESTC²), et de quelques autres bibliothèques et projets. Toutefois, c'est nous qui trouvons et saisissons la plupart des données, en consultant des bibliographies et des catalogues imprimés et numériques, ainsi que des ressources en texte intégral comme ECCO³ et Google Books. Parfois, nous avons des données bibliographiques, mais nous créons souvent les données en examinant d'autres sources. Nous avons aussi vérifié manuellement ou examiné à vue d'œil un plus petit nombre de livres.

Toutes les données sont enregistrées dans une base de données MySQL⁴ à partir de laquelle les données peuvent être extraites sous forme de fichiers .csv.

Quelles conventions et procédures utiliserez-vous pour structurer et nommer vos fichiers, et pour contrôler leur version, afin de vous aider, ainsi que les autres, à mieux comprendre comment vos données sont organisées?

¹ Catalogage lisible par machine

² English Short Title Catalogue

³ Eighteenth Century Collections Online

⁴ MYSQL est un système de gestion de base de données ouvertes

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

L'équipe de projet a élaboré un ensemble de tableaux dans notre base de données MySQL. Chaque enregistrement a un identifiant unique et les tableaux sont liés par ces clés. Par exemple, le titre n° 3 est « Life of Carsten Niebuhr ». Il a été imprimé à l'emplacement n° 2643743 à Londres, en Grande-Bretagne. Sarah Austin (personne n° 9) est l'auteure.

Documents et métadonnées

Quels seront les documents nécessaires pour que les données soient lues et interprétées correctement à l'avenir?

Nous élaborerons une structure de carte de données ainsi que des détails sur les données pour chaque champ. Une description détaillée de la méthodologie sera consignée et mise à jour suivant les besoins.

Comment vous assurerez-vous que les documents sont créés ou saisis de façon uniforme tout au long de votre projet?

Nous tenons des réunions d'équipe régulièrement pour veiller à ce que des normes uniformes soient appliquées. Dans la notice du titre, nous avons à la fois une tentative de vérification finale (lorsque nous sommes incapables de confirmer deux sources pour nos entrées au titre, mais que nous avons tenté de le faire) et un mécanisme de vérification finale (lorsque nous avons été en mesure de confirmer deux sources indépendantes) pour veiller à ce que nos normes en matière de données soient respectées. Lorsque nous avons vérifié un titre manuellement (c'est-à-dire que nous l'avons examiné physiquement), nous le prenons en note et nous effectuons la vérification finale de l'entrée; dans ces cas, nous n'avons pas besoin de deux sources. Nous effectuons également une vérification finale des données de nos tableaux de personnes, pour veiller à ce que les notices aient été vérifiées par rapport à des sources fiables.

Si vous utilisez une norme ou des outils de métadonnées pour documenter et décrire vos données, veuillez les énumérer ici.

Si vous et votre équipe de recherche avez créé une norme des métadonnées qui s'applique à vos données et à vos documents, il est avantageux pour le lecteur de l'indiquer et de justifier cette décision.

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

Nous avons élaboré notre propre norme des métadonnées adaptée à la nature du projet. La description de nos normes des métadonnées, qui donne également des renseignements contextuels sur le rapport entre les tableaux, est fournie à titre d'Annexe dans le présent document.

Stockage et sauvegarde

Quelles sont les exigences prévues en matière de stockage pour votre projet, en ce qui concerne l'espace de stockage (en mégaoctets, gigaoctets, téraoctets, etc.) et la durée de stockage?

Notre programmeuse de projet maintient en ce moment l'espace serveur que nous avons avec la doyenne des Arts de la SFU⁵. Nous migrerons le serveur à la bibliothèque de la SFU, après avoir pris des dispositions pour obtenir une capacité de serveur adéquate plus tôt. La base de codes actuelle du projet sur l'histoire des femmes dans le domaine des médias imprimés est d'environ 1,285 Go d'espace disque, ce qui comprend des copies des fichiers de données sources fournis par ESTC et Osborne.

La base de données requiert un peu moins de 1 Go d'espace disque.

Les ressources système du serveur actuel sont les suivantes :

UC : 2XIntel(R) Xeon(R) UC E5-4650 0 @ 2,70 GHz

MEV : 4 Go

Disque : >250 Go

Comment et où vos données seront-elles stockées et sauvegardées pendant votre projet de recherche?

⁵ Web-service stack, named for Linux, Apache, MySQL, PHP/Perl/Python

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

À l'heure actuelle, les données sont stockées dans l'espace serveur maintenu par notre programmeuse de projet, puis dans le serveur de la bibliothèque de la SFU, après la migration.

Comment l'équipe de recherche et les autres collaborateurs pourront-ils accéder aux données, les modifier et y contribuer tout au long du projet?

Notre programmeuse a établi une interface Web de saisie de données sur une plateforme LAMP⁶.

Compte tenu du fait que l'accès au serveur est contrôlé par l'intermédiaire d'une liste d'accès administrée par la SFU, il est possible d'ajouter ou de supprimer des personnes sur demande. Il n'est pas nécessaire de partager les mots de passe pour l'administration du serveur puisque n'importe quelle personne qui figure sur la liste y a pleinement accès. La connaissance de l'administration Linux constitue le principal obstacle. Site Web

Les mots de passe d'accès (wphp/WPHP...) perdus ou oubliés peuvent être générés de nouveau sur la ligne de commande. Les mots de passe peuvent également être stockés dans des courriels précédents qui vous sont adressés ou qui sont adressés à d'autres personnes au cours de la discussion sur le site Web.

Pour le stockage général des renseignements administratifs, l'instance canadienne OSF⁷ assurera une redondance complète et permettra le contrôle des versions à mesure que les fichiers sont actualisés.

⁶ Ensemble de services Web, nommé pour Linux, Apache, MySQL, PHP/Perl/Python

⁷ Plateforme collaborative de partage des données ouvertes créée par le Center for Open Science

Plan de gestion des données modèle #1: Sciences Humaines Numériques

**Projet sur l’histoire des femmes dans le domaine des médias
imprimés (1750-1830)**

Preservation

**Où déposerez-vous vos données aux fins de conservation à long
terme et d’accès à la fin de votre projet de recherche?**

À la fin du projet, nous déposerons nos données dans le dépôt de données de recherche de la SFU, Radar, comme fichiers .csv accompagnés d’une carte de données pour indiquer notre structure de tableaux. Nous prévoyons également mettre toutes nos données à la disposition du public au moyen d’une interface Web mise au point par le Digital Humanities Innovation Lab de la SFU.

Le dépôt Radar est pris en charge par Islandora, un dépôt développé en code source ouvert. Sont attribués aux données déposées dans Radar des identifiants uniques (DOI), des rubriques et des mots-clés interrogeables, des renseignements pertinents sur les licences et d’autres métadonnées applicables.

Plan de gestion des données modèle #1: Sciences Humaines Numériques

**Projet sur l'histoire des femmes dans le domaine des médias
imprimés (1750-1830)**

Indiquez comment vous vous assurerez que vos données sont prêtes à être conservées. Envisagez des formats de fichiers favorables à la conservation, en assurant l'intégrité des fichiers, l'anonymisation et la dépersonnalisation, et l'inclusion des documents justificatifs.

Il est utile d'indiquer clairement les normes auxquelles vous renvoyez [normes institutionnelles, principes gouvernementaux, etc.] et de mentionner où seront les documents justificatifs et dans quel format. Si vous n'êtes pas certain des normes applicables au moment de la création de votre PGD, indiquez comment vous comptez harmoniser la gestion de vos données avec les pratiques exemplaires à l'avenir.

Nous utiliserons les meilleures normes, c'est-à-dire les fichiers .csv et la carte des données, car il appert que Radar n'accepte pas les bases de données MySQL. En consultant les bibliothécaires ou des professionnels de l'information, nous pourrions assurer notre conformité aux lignes directrices et aux pratiques exemplaires applicables en matière de conservation.

Partage et réutilisation

Quelles données partagerez-vous et sous quelle forme? (P. ex., brute, transformée, analysée, définitive.)

Nous partagerons les données définitives, selon vos descriptions, c'est-à-dire, les données régularisées que les utilisateurs peuvent convertir en format .csv.

Avez-vous envisagé le type de licence d'utilisateur final à inclure dans vos données?

Nous avons reconnu les fournisseurs de nos métadonnées bibliographiques, que nous mettrons à la disposition de tous. Les sources sont incluses et reconnues dans notre modèle de données.

Quelles mesures seront prises pour aider à faire connaître l'existence de vos données dans le milieu de la recherche?

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

Le fait de décrire la façon dont vous comptez communiquer vos données peut aider le lecteur à envisager plus concrètement votre plan de diffusion. Cela contribue aussi à renforcer votre engagement à partager vos données.

Il y aura un site Web public. Par ailleurs, nous présenterons le projet lors de conférences, nous citerons nos données dans des analyses publiées, et nous enverrons des avis par courriel aux serveurs de listes au moment de la diffusion.

Responsabilités et ressources

Déterminez qui sera responsable de la gestion des données de ce projet pendant et après le projet, et les principales tâches de gestion des données dont cette personne sera responsable.

Jenn Ross, notre programmeuse, gère nos données. À la fin de l'étape d'élaboration du projet, il est prévu que la SFU gèrera les données conformément aux modalités d'une entente d'hébergement en cours de formulation.

Comment les responsabilités relatives à la gestion des activités de données seront-elles gérées si des changements importants surviennent dans le personnel qui supervise les données du projet, y compris un changement de chercheur principal?

Mon gestionnaire de projet et mon étudiant au doctorat s'occuperont des données du projet lorsque je ne serai plus en mesure de le faire en tant que chercheuse principale.

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

De quelles ressources aurez-vous besoin pour mettre en œuvre votre plan de gestion des données? À combien estimez-vous le coût global de la gestion des données?

Actuellement, la mise à jour et la tenue de nos données coûtent environ 100 \$ par mois (deux heures de développement). Je m'attends à ce que ces coûts se poursuivent et à ce qu'ils soient assumés par la bibliothèque de la SFU advenant la conclusion d'une entente d'hébergement. J'examine également d'autres sources de financement permanent.

Éthique et conformité juridique

Si votre projet de recherche comprend des données de nature délicate, comment veillerez-vous à ce qu'elles soient gérées de façon sécuritaire et accessibles uniquement aux membres approuvés du projet?

S. O.

Le cas échéant, quelles stratégies entreprendrez-vous pour contrôler les utilisations secondaires des données sensibles?

S. O.

Comment allez-vous gérer les questions juridiques, éthiques et de propriété intellectuelle?

S. O.



Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)
Portage Network | portage@carl-abrc.ca | portagenetwork.ca

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

Annexe

Tableau des titres

Champ	Description
Identifiant du titre	Numéro d'identification unique pour cette entrée.
Titre	Titre complet tel qu'il apparaît sur la page de titre, y compris le sous-titre, l'auteur signataire et la mention d'édition, le cas échéant.
Attribution	Lien avec le tableau des personnes . Les rôles d'attribution comprennent le nom des auteurs (si connu) et, s'il y a lieu, des rédacteurs en chef, des traducteurs, des auteurs de documents d'introduction, des graveurs et des femmes qui travaillent dans les domaines de l'impression, de l'édition, de la vente de livres, etc. Les rôles attribués à une femme dont le nom est inconnu seront désignés « Inconnu ».
Auteur signataire	Attribution de l'auteur telle qu'elle apparaît sur la page de titre ou à la fin de la préface. (P. ex. « Par une dame », « Par l'auteure de . . . ».)
Pseudonyme	Si l'auteur a signé l'ouvrage en utilisant un nom qui s'est révélé faux, ce sera inscrit ici.
Entreprise	Lié au tableau de l'entreprise . Comprend l'ensemble des imprimeurs, éditeurs et libraires énumérés dans la marque d'éditeur et le colophon.
Volumes	Le nombre de volumes compris dans l'édition publiée, indiqué en chiffres arabes.
Pagination	La pagination de chaque volume. Chaque volume indiqué par une lettre majuscule en chiffres romains, la matière préliminaire indiquée par des chiffres romains minuscules, la pagination régulière indiquée en chiffres arabiques.
Date de publication	La date (année) telle qu'elle apparaît dans la marque de l'éditeur. S'il n'y a pas d'information figurant dans la marque de l'éditeur, ce renseignement correspondra à la meilleure estimation que nous avons.

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

Édition	L'édition telle qu'elle apparaît dans la mention d'édition sur la page de titre; les premières éditions ne sont pas indiquées sur les pages de titre. Dans le cas d'une impression la plus ancienne connue, on suppose qu'il s'agit de la première édition.
Date de la première publication	La date (année) de la première publication de l'ouvrage.
Marque de l'éditeur	Les renseignements sur les imprimeurs, les éditeurs, les libraires représentés sur la page de titre.
Lieu d'impression	Le lieu géomarké tel qu'il est indiqué par la marque de l'éditeur.
Format	La façon dont les feuilles ont été pliées et rassemblées en pages. Par exemple, in-folio, in-quarto, in-octavo, in-douze, in-seize, in-dix-huit, in-vingt-quatre, etc.
Taille	La longueur et la largeur mesurées en centimètres.
Prix	Le prix indiqué sur la page de titre ou dans les catalogues et les périodiques exprimés en livres, en shillings et en pence.
Genre	La catégorie dans laquelle le travail s'inscrit. [Lien vers la liste complète des genres?]
Source/N. d'identification de la source	Les sources consultées pour remplir les champs de saisie. Déterminez quelles données peuvent devoir être partagées afin de répondre aux besoins de l'établissement ou en matière de financement, et quelles données peuvent devoir être interdites aux fins de confidentialité.
Cote de placement	Les numéros d'appel pour l'emplacement dans diverses bibliothèques.
Remarques	Toute information importante qui ne s'inscrit pas dans les champs ci-dessus sera indiquée ici, ainsi que des liens vers des sources qui n'ont pas de numéro d'identification unique.
Vérification manuelle	Indique qu'une copie physique du texte a été consultée.

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

Vérification finale	Indique que deux sources ont été consultées ou que le texte a été vérifié manuellement.
----------------------------	---

Attribution/Rôles des personnes

Rôle	Description
Auteur	La personne qui a composé l'ouvrage. Voir Foucault.
Éditeur	Indiquez si l'entreprise est dirigée par une femme. Indique la personne qui dirige l'entreprise pour laquelle l'ouvrage a été imprimé.
Libraire	Indiquez si l'entreprise est dirigée par une femme. Indique la personne qui dirige l'entreprise qui a vendu l'ouvrage.
Imprimeur	Indiquez si l'entreprise est dirigée par une femme. Indique la personne qui dirige l'entreprise qui a imprimé le travail.
Rédacteur en chef	La personne qui a choisi, révisé et organisé l'ouvrage pour l'impression.
Traducteur	La personne qui traduit le texte d'une langue à une autre.
Graveur	La personne responsable de la transformation d'une illustration en gravure.
Introduction	La personne qui a composé l'introduction ou tout document préliminaire pour l'ouvrage. Il doit s'agir d'une personne autre que l'auteur.
Illustrateur	La personne qui a créé l'image à partir de laquelle des gravures ont été créées. Habituellement, il s'agit d'une personne autre que le graveur.
Compilateur	La personne qui a sélectionné les ouvrages contenus dans un texte. S'applique habituellement aux anthologies et aux collections.
Compositeur	La personne responsable de mettre un ouvrage en musique.
Artiste	

Rôles de l'entreprise

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l'histoire des femmes dans le domaine des médias imprimés (1750-1830)

Rôle	Description
Imprimeur	L'entreprise responsable de la composition et de l'impression de l'ouvrage. Habituellement indiqué par l'expression « imprimé par » dans la marque de l'éditeur et/ou le colophon.
Éditeur	L'entreprise qui assume la responsabilité financière de la révision, de l'impression et de la distribution de l'ouvrage. Habituellement indiqué par l'expression « imprimé pour » dans la marque de l'éditeur. Si un ouvrage a été autoédité, ce sera indiqué dans le champ « Autoédité ».
Libraire	L'entreprise auprès de laquelle l'ouvrage peut être acheté. Indiqué par l'expression « vendu par » dans la marque de l'éditeur.
Inconnu	Aucun rôle de l'entreprise indiqué dans la marque de l'éditeur.

Tableau des personnes

Champ	Description
Numéro d'identification de la personne	Numéro d'identification unique pour cette entrée.
Nom de famille	Nom de famille de la personne (la préférence sera accordée au nom par lequel elle est le plus communément connue).
Prénom	Prénom de la personne.
Titre	Le cas échéant, le titre aristocratique complet.
Sexe	Soit M ou F.
Date de naissance	Si connue, JJ-MM-AAAA.
Date du décès	Si connue, JJ-MM-AAAA.
Ville de naissance	Si connue; géomarkée.
Ville du décès	Si connue; géomarkée.

Tableau de l'entreprise

Plan de gestion des données modèle #1: Sciences Humaines Numériques

Projet sur l’histoire des femmes dans le domaine des médias imprimés (1750-1830)

Champ	Description
Nom	Le nom le plus complet connu de l’entreprise.
Adresse municipale	L’adresse municipale de l’entreprise, si elle est connue. Chaque fois qu’une entreprise déménage dans un nouvel emplacement, une nouvelle entrée est faite sous la rubrique de l’entreprise. Les adresses sont habituellement tirées des marques de l’éditeur. Géomarcuée.
Ville	Ville/ville/village où se trouve l’adresse municipale de l’entreprise; géomarcuée.
Date de début	Si connue, JJ-MM-AAAA.
Date de fin	Si connue, JJ-MM-AAAA.



Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

Portage Network | portage@carl-abrc.ca | portagenetwork.ca