



# LE PROJET « ENDINGS »

Un gage de pérennité  
en humanités numériques

Claire Carlin, University of Victoria, BC, Canada

# ENQUÊTES À L'ORIGINE DU PROJET

- Sondage du Projet « Endings » effectué en 2018, ouvert à la communauté HN international (127 réponses) : 60% des collègues dont le projet était toujours en cours n'avaient pas de plan pour le terminer.
- Analyse du catalogue « Digital Scholarly Editions », Cologne Center for eHumanities, U zu Köln : sur 466 éditions recensées en 2020, 376 avaient disparu dès 2019. Une édition numérique dure en moyenne 8.5 ans.

# LA NOTION D'ÉDITION

## MoEML v.7.0

You are looking at a static release of MoEML, v.7.0 released on May 4, 2022. See the [Release Notes](#) for an overview of new content and features in v.7.0.

Version 6.9, svn rev. 2646 2023-03-10 14:30:43 -0800.

# LES PRINCIPAUX DÉFIS À LA PRÉSERVATION NUMÉRIQUE

- la gestion des données et les résultats
- l'hébergement, la sauvegarde, le stockage
- la documentation et les métadonnées
- les institutions et les infrastructures
- les droits légaux et l'accès
- l'équipe/la collaboration
- les questions académiques (e.g. l'évaluation par les collègues)
- le financement

# LES PRINCIPES DE BASE DU PROJET ENDINGS

- Les données
- La documentation
- Le traitement
- Les produits
- La gestion des versions

# LES DONNÉES

- Les données sont stockées uniquement dans des formats conformes aux normes des données ouvertes et pouvant être traitées (TEI XML, GML, ODF, TXT).
- Les données sont soumises à un contrôle de version (Subversion, Git).
- Les données peuvent toujours être l'objet d'une validation et d'une analyse diagnostique.

# LA DOCUMENTATION

- Les modèles de données, y compris les noms de champs, les descriptions et les valeurs contrôlées, doivent être clairement documentés dans un document statique qui est conservé avec les données et fait partie des produits.
- Toutes les questions relatives aux droits et à la propriété intellectuelle doivent être clairement documentées. Dans la mesure du possible, les données et les produits doivent être diffusés sous licence libre (Creative Commons, GNU, BSD, MPL).

# LE TRAITEMENT

- Validation incessante : tout traitement comprend la validation/*linting* de toutes les entrées et sorties. Toutes les erreurs de validation doivent arrêter le processus et empêcher la poursuite de l'exécution jusqu'à ce que les erreurs soient résolues.
- Intégration continue : toute modification des données sources nécessite une reconstruction complète du site (déclenchée automatiquement dans la mesure du possible).
- Le code est contingent : bien que le code ne soit pas censé avoir une longévité importante, il doit, dans la mesure du possible, respecter les principes « Endings » pour les données et les produits.

# LES PRODUITS

- Pas de dépendance à l'égard des logiciels côté serveur : créer un site web statique sans bases de données, ni PHP, ni Python.
- Pas de technologies sur mesure ou à la mode : n'utiliser que des standards avec soutien à travers toutes les plateformes et dont la viabilité à long terme est assurée. Nos choix sont HTML5, JavaScript et CSS.
- Pas de dépendance de bibliothèques ou de services externes : pas de JQuery, pas d'AngularJS, pas de Bootstrap, pas de Google Search.

# LES PRODUITS

- Pas de chaînes de requête : chaque entité du site a une page unique avec une URL simple qui fonctionnera sur n'importe quel domaine ou adresse IP.
- Inclusion de données : chaque site devrait inclure une copie documentée des données sources, afin que les utilisateurs du site puissent facilement réutiliser le travail.
- Redondance massive : chaque page contient tous les composants dont elle a besoin, de sorte qu'elle puisse fonctionner sans le reste du site si nécessaire, même si cela implique de reproduire l'information sur l'ensemble du site.
- Défaillance en douceur : chaque page doit continuer à fonctionner efficacement même en l'absence de support JavaScript ou CSS.

# LES PRODUITS

- Lorsqu'un site statique fonctionne parfaitement, il peut être amélioré par l'utilisation d'autres services tels qu'un outil d'indexation côté serveur (Solr, eXist) pour faciliter la recherche et d'autres fonctionnalités similaires.
- L'utilisation d'une bibliothèque externe peut s'avérer nécessaire pour prendre en charge une fonction spécifique trop complexe pour être codée localement (comme la cartographie ou la cryptographie). Ces bibliothèques doivent être en accès libre et largement utilisées, et ne doivent pas avoir elles-mêmes de dépendances.

# GESTION DES VERSIONS

- Les versions doivent être périodiques et soigneusement planifiées. Le modèle *rolling release* ou publication continue doit être évité.
- Une version ne doit être diffusée que lorsque l'ensemble du produit est cohérent, homogène et complet (c'est-à-dire qu'il a passé avec succès tous les tests de validation et de diagnostic).

# GESTION DES VERSIONS

- À l'instar des éditions d'ouvrages imprimés, chaque version d'une ressource web doit être clairement identifiée sur chaque page par sa date de création et un numéro de version.
- Les ressources web doivent inclure des instructions détaillées pour la citation, afin que les utilisateurs finaux puissent citer sans ambiguïté une page spécifique d'une édition spécifique.
- Les URL des ressources individuelles d'une publication numérique doivent persister d'une version à l'autre. Toute ressource déplacée, retirée ou supprimée qui n'est plus disponible à une URL précédemment accessible doit être redirigée de manière appropriée.

# LES STRATÉGIES DE PRÉSERVATION

## Logiciels en libre accès sur GitHub : **staticSearch**

- Ce codebase fournit un moteur de recherche purement JSON pour votre site web HTML5, y compris des recherches booléennes et phrastiques, différents types de filtres de recherche, des résultats par mot-clé dans le contexte, et des résultats de recherche pouvant être mis en signet. Il n'utilise aucun code d'arrière-plan et respecte donc les principes Endings. Voir par exemple l'engin de recherche sur le site [Mapping Keats's Progress](#) pour un exemple simple et fonctionnel, et celui du site [The Colonial Despatches](#) un exemple plus complexe.
- Vous pouvez consulter la documentation complète de la dernière version [sur notre site](#).

# LES STRATÉGIES DE PRÉSERVATION

## Logiciels en libre accès sur GitHub : **Diagnostic TEI**

- Ce [codebase](#) un ensemble d'outils permettant de vérifier la cohérence, la consistance et la complétude de votre projet TEI XML à partir de l'éditeur XML Oxygen (bien que vous puissiez également l'utiliser à la ligne de commande). Il effectue un certain nombre de vérifications génériques à travers la collection, mais il est également facilement extensible pour vérifier d'autres caractéristiques qui peuvent être particulières à votre projet.
- Une [boîte à outils](#) similaire est disponible pour vérifier votre HTML.

# DES ÉTUDES DE CAS

- The Diary of Robert Graves
- The Map of Early Modern London
- Le mariage sous l'Ancien Régime

# CONCLUSIONS : UN RETOUR AU FUTUR ?

La simplicité avant tout : moins il y a de dépendances, moins il y a de technologie, plus il y aura de sites fonctionnels. En effet, c'est un retour aux années 1990 qui nous permettra d'assurer l'avenir de nos projets numériques.

Voir <https://endings.uvic.ca/> pour accéder à toutes nos données, tous nos logiciels, tous nos résultats.

Questions : contacter notre programmeur-développeur Martin Holmes ([mholmes@uvic.ca](mailto:mholmes@uvic.ca)).

# REMERCIEMENTS

- Martin Holmes
  - Programmeur-développeur consultant, University of Victoria
- J. Matthew Huculak
  - Directeur des services de recherche avancée et Bibliothécaire chargé de la recherche numérique, University of Victoria
- Le Conseil de recherches en sciences humaines du Canada (CRSH)