N° 1 | 2020
Faire dialoguer les disciplines via l'indexation des connaissances : la recherche interdisciplinaire en débat

Écrire les SHS en environnement numérique. L'éditeur de texte Stylo

Marcello VITALI-ROSATI *Professeur*
Département des littératures de langue française
Titulaire de la Chaire de recherche du Canada sur les écritures numériques
Université de Montréal
Nicolas SAURET *postdoctorant*
Huma-Num
HN Lab
French National Centre for Scientific Research, Paris
Antoine FAUCHIÉ *Doctorant*
Pavillon Lionel Groulx
Université de Montréal
Margot MELLET *Doctorant*
Pavillon Lionel Groulx
Université de Montréal

Édition électronique :
URL :
https://intelligibilite-numerique.numerev.com/index.php/articles/revue-1/2168-ecrire-les-shs-en-environnement-numerique-lediteur-de-texte-stylo
DOI : numerev_1697
Date de publication : 13/05/2020

Cette publication est sous licence **CC BY-NC-ND** (Attribution - No commercial - No derivatives).

Pour **citer cette publication** : VITALI-ROSATI, M., SAURET, N., FAUCHIÉ, A., MELLET, M. (2020) Écrire les SHS en environnement numérique. L'éditeur de texte Stylo. *Intelligibilité du numérique*, (1). https://doi.org/10.34745/numerev_1697

L’écriture et l’édition scientifiques en contexte numérique sont relativement peu interrogées ou remises en cause. Quelques outils, qui posent un certain nombre de questions et de problèmes, sont utilisés pour structurer, mettre en forme et produire des documents académiques. Face à ces solutions bien souvent propriétaires, Stylo est un outil de rédaction et d’édition qui répond aux exigences scientifiques, et actuellement utilisé par des revues savantes. Expérimenter avec Stylo ouvre la perspective de nouvelles pratiques d’écriture numérique tout en alimentant une réflexion théorique sur la structuration de notre pensée.

Scientific writing and editing in a digital context are relatively little interrogated or questioned. A few tools, which pose a number of questions and problems, are used to structure, format and produce academic documents. Faced with these often proprietary solutions, Stylo is a writing and editing tool that meets scientific requirements, and is currently used by academic journals. Experimenting with Stylo opens up the prospect of new digital writing practices while fueling theoretical reflection on the structuring of our thinking.

Mots-clés :
Édition numérique, Edition scientifique, Écriture numérique, Revue savante, Stylo, Academic journal, Digital edition, Scientific edition, Digital writing, Stylo

---

1. Introduction

Que signifie écrire en contexte numérique ? Quelle est l’influence des nouvelles technologies sur la forme de nos contenus scientifiques ? Quelles pratiques d’écriture les chercheurseuses en Sciences humaines (SH) doivent-elles développer afin de produire des documents qui garantissent à la fois la rigueur scientifique et la pérennité de leurs écrits ?

L’écriture en environnement numérique – destinée à l’impression ou à la diffusion

numérique – s'est aujourd'hui totalement banalisée : personne n'écrit un article à la main pour le faire ensuite dactylographier. Pourtant, la réflexion sur les enjeux liés à cette écriture demeure peu développée, et nos usages des technologies numériques s'avèrent limités au regard de leurs potentialités.
Par ailleurs, les outils disponibles pour écrire en contexte académique ne sont pas nombreux, ils demeurent souvent propriétaires et sont rarement conçus et adaptés aux besoins des chercheurs.euses en SH. Concrètement, la quasi-totalité de la production scientifique en SH passe aujourd'hui par un seul logiciel et un seul format : Word et le format docx, propriétés de Microsoft. Conçu pour tous les usages, ce duo se trouve mal adapté aux besoins spécifiques des auteures universitaires.

Stylo est un projet que la Chaire de recherche du Canada sur les écritures numériques a entamé en partenariat avec le diffuseur Érudit afin de palier ce manque de réflexion et de diversité. Il s'agissait – et il s'agit encore, étant donné que le projet est encore en phase de développement – d'une part d'analyser et de documenter l'impact et l'influence des technologies d'écriture sur la production de documents scientifiques en SH, et d'autre part de développer un outil d'écriture et d'édition spécifiquement destiné à la rédaction scientifique en SH en se fondant sur cette réflexion.
Stylo est à la fois un outil de rédaction de texte scientifique, et un outil d'édition de document scientifique, spécialement conçu pour répondre aux enjeux de l'édition numérique savante. Utilisé ou expérimenté par plusieurs revues savantes et en passe d'intégrer le panel de service d'Isidore.science, Stylo demeure encore un prototype, dont le développement reste fortement lié à une réflexion théorique sur l'écriture en environnement numérique. L'ambition du projet n'est pas tant de créer un ennième outil, mais de penser les enjeux épistémologiques des pratiques d'écriture en environnements numériques. Comment ces nouvelles écritures structurent-elles la pensée ?

2. Considérations préliminaires

Nos outils d'écriture exercent un impact majeur sur la production des connaissances. En effet :

1- Ils influencent la structuration des contenus.
2- Ils définissent les critères d'accessibilité des contenus.
3- Ils déterminent la pérennité ou l'obsolescence des contenus.

La quasi-totalité des contenus scientifiques reste produite avec un logiciel de traitement de texte propriétaire qui n'est pas pensé et conçu pour des usages académiques, à savoir Microsoft Word. Or cet outil grand public présente des lacunes problématiques pour l'édition scientifique en SH :

1- Le principe du WYSIWYG (_What You See Is What You Get_) entretient une confusion entre structuration et mise en forme : pour faciliter la lecture et le

traitement d'un article ou d'un ouvrage, il faut dissocier la valeur sémantique des fragments de leur rendu graphique, par exemple : un titre de niveau 2 et son rendu graphique dans une taille de police importante,

2- Le format *docx* ne permet pas une structuration scientifique des contenus. Les contenus produits en *docx* sont pauvrement balisés et difficilement accessibles, car peu indexés. À une époque où la production de textes n'a jamais été aussi foisonnante, il devient crucial que les documents soient correctement structurés pour améliorer leur indexation et permettre leur requête par des outils de recherche avancée.

3- Les éditeurs sont dans l'obligation de faire un long et fastidieux travail de reprise des textes afin de réintroduire une couche sémantique et de corriger les nombreuses erreurs introduites par ces traitements de texte.

4- La pérennité des documents de format *docx* dépend de la bonne volonté de Microsoft à maintenir ce format propriétaire. Or la pérennité ne constitue pas la mission et la priorité d'une entreprise commerciale. Rien ne garantit que les contenus au format *docx* seront accessibles dans le futur.

En étudiant plus précisemment les pratiques éditoriales des revues scientifiques en SH, il apparaît que, de l'auteur au diffuseur en passant par l'éditeur, l'article scientifique passe par des états et des formats fortement disparates. Ce processus engendre tout à la fois une perte considérable de temps, de données et de signification.

Il est donc urgent de penser et d'expérimenter de nouveaux modèles d'écriture, plus adaptés aux besoins de la production des connaissances en SH.

3. État de l'art

La dimension matérielle de l'écriture a été mise entre parenthèses aux XIXe et XXe siècles, notamment au sein de tendances théoriques rassemblées de façon parfois arbitraire sous le nom de "post-structuralisme". Ces tendances théoriques ont pu contribuer à véhiculer une conception idéalisée du texte, transcendant sa dimension proprement inscrite. Que l'on pense en particulier aux travaux sur l'intertextualité de Kristeva (1969), ou à l'idée de "texte" défendue par Derrida ou Barthes. Ces penseurs ont déterminé une omniprésence de la notion de texte qui a contribué à nuire à la compréhension et la considération de sa matérialité. La généralisation massive de l'écriture numérique dans les dernières décennies a suscité un nouvel intérêt pour cette thématique, et une prise de conscience aiguë de la matérialité du texte ainsi que de l'influence des technologies d'écriture sur la pensée.

Le rapport entre outil d'écriture et structuration de la pensée a fait l'objet de plusieurs recherches. Jack Goody a montré comment l'émergence de l'écriture a permis le développement des formes de raisonnement abstrait (Goody, 1979). Dans un même

ordre d'idée, Christian Vandendorpe a observé les changements de structures de pensée lors du passage du rouleau (volumen) au codex, puis à celui de l'hypertexte (Vandendorpe, 1999).

La naissance des *Software Studies* et des *Critical Code Studies* (Marino, 2020) témoigne de l'intérêt pour ces questions. Dans la tradition des *Media Studies* initiée par McLuhan (2013), des auteurs comme Friedrich Kittler, Lev Manovich, Alexander Galloway, Geert Lovink, Espen Aarseth, Matthew Fuller ont travaillé sur la signification et les enjeux culturels des outils numériques (Aarseth, 1997 ; Galloway, 2012). Les logiciels – et plus généralement le code – ne sont pas neutres : ils portent et promeuvent des valeurs, et conditionnent ainsi le développement de nos sociétés.

Malgré ces nombreuses recherches consacrées à l'effet culturel des technologies numériques, nous constatons une quasi-absence de travail ciblant spécifiquement les logiciels d'écriture. À l'exception notable de Matthew Kirschenbaum qui a dédié un ouvrage entier à l'usage des traitements de texte comme Word dans le domaine de l'écriture littéraire (Kirschenbaum, 2016), aucune étude ne traite de l'influence de ces logiciels sur la production scientifique. On peut tout de même citer l'initiative de Maxwell et al. (2019) recensant les plateformes d'édition et de diffusion en milieu académique, ou celle de Dehut (2018) proposant une analyse critique des traitements de texte.

4. Les principes de Stylo

L'hypothèse théorique à la base du développement de Stylo est la suivante : les technologies numériques, loin de constituer des outils neutres, conditionnent profondément les pratiques, mais aussi les valeurs et les cultures (Doueihi, 2008). Le traitement de texte (et plus spécifiquement le logiciel Microsoft Word) est l'outil par défaut pour la rédaction et l'édition scientifique. L'approche WYSYWYG et les formats associés promeuvent le rendu graphique à la structuration logique des contenus. Ainsi, au lieu de considérer une chaîne de caractères comme un titre ou comme le nom d'un auteur et de l'exprimer comme tel, le format n'associe à cette chaîne de caractères qu'un style visuel défini par un type de caractère, une taille, etc. Le sens de ces fragments stylisés n'est jamais exprimé sémantiquement, ni encodé en tant que tel dans le document. L'éditeur est ainsi contraint de concevoir son texte à partir de sa représentation graphique imprimée, et à le structurer en conséquence. Cela détermine l'illusion d'une désintermédiation et d'une continuité entre la représentation à l'écran et le document final imprimé. En réalité cette illusion cache la situation inverse. Plutôt que rapprocher l'écrivain du document final en diminuant le nombre de médiations, l'approche de Word implique une multiplication de médiations invisibles à l'utilisateur mais structurantes sur le contenu. Souvent converti et importé dans le logiciel Indesign en vue d'une restructuration et d'une mise en forme, le sens du texte se voit ainsi *délegué* à des instances plus ou moins opaques - et, dans le cas de Word - propriétaires

et fermées.

L'illusion de désintermédiation s'accompagne par ailleurs d'une illusion de littératie : Word donne l'impression que n'importe qui, sans compétences particulières, peut choisir un gabarit de page, définir ses marges, établir une police de caractères, etc., soit se faire éditeur au détriment de cinq siècles du savoir-faire éditorial et typographique de l'imprimé.

A contrario, des formats conçus selon une approche WYSIWYM (*What You See Is What You Mean*) encouragent la structuration logique avant le rendu graphique : l'éditeur de texte sémantique Stylo s'inscrit dans ce type d'iniative, en lien avec le mouvement des *Software Studies* (Fuller, 2008).
En tenant compte de ces considérations, la philosophie de Stylo repose sur trois principes :

> 1- Redonner aux chercheurseuses la possibilité d'exprimer la sémantique du texte. C'est au chercheur d'assurer la signification et la structuration de son texte et de ses métadonnées scientifiques. La chaîne éditoriale doit par ailleurs en garantir la continuité réelle.
> 2- Produire des données plus significatives. Dans un environnement où les écritures et les publications se multiplient et se concurrencent, les documents riches en données sont susceptibles d'une meilleure visibilité et accessibilité.
> 3- Utiliser des standards et des technologies simples et pérennes. Cette approche *lowtech* facilite la maintenance de l'éditeur mais aussi son interopérabilité avec d'autres environnements, garantissant l'indépendance des utilisateurs et des documents vis-à-vis d'un logiciel ou d'un format.

Publiés sur le web, les textes numériques sont destinés à être lus autant par des humains que par des machines. Ainsi, éditer un livre numérique ne consiste plus seulement à travailler sa mise en forme, mais requiert en premier lieu de le structurer selon un balisage sémantique. S'il existe aujourd'hui un certain consensus autour des formats et des protocoles numériques les mieux adaptés aux contenus savants (balisage XML, protocole OAI-PMH, etc.) la mise en oeuvre de bonnes pratiques reste l'un des plus grands défis de l'édition scientifique dans les humanités : écrire directement en XML se révèle souvent problématique pour les chercheurseuses en SH. La solution que propose Stylo est d'offrir un environnement simplifié pour produire un document au format HTML enrichi sémantiquement et convertible vers n'importe quel format standard de sortie.

Concrètement un document Stylo est basé sur trois fichiers sources, le corps de texte, les métadonnées du document et les références bibliographiques, respectivement aux formats Markdown/YAML/BibTeX. Le corps du texte est écrit en Markdown, une syntaxe simplifiée du HTML, permettant d'intégrer pendant la rédaction un balisage minimaliste pour spécifier le sens des différents fragments du texte. Ce format est de plus en plus adopté pour les écritures en ligne, et son usage se démocratise dans l'édition savante. Plus simple que le HTML, il a le mérite de rendre accessible une structuration rigoureuse

du texte. Les éléments principaux d'un texte peuvent ainsi être spécifiés : niveaux de titres, citations, notes, figures, etc. Un titre de niveau 1 est spécifié par un croisillon # en début de ligne, le niveau 2 par ##, et ainsi de suite. Une citation sera indiquée par un chevron > en début de ligne. Il est aussi possible d'enrichir sémantiquement un terme ou une phrase. Le balisage suivant spécifie que la phrase est une definition et le terme signe est un concept :

```
[Un [signe]{.concept id="1"} est un phénomène perceptible ou observable qui indique la probabilité de l'existence ou de la vérité d'une chose.]{.definition}
```

Cela donnera en HTML :
`Un signe est un phénomène perceptible ou observable qui indique la probabilité de l'existence ou de la vérité d'une chose.`

Le choix du HTML permet d'adapter l'écriture à une multiplicité de modèles et de structures de données. Stylo n'impose pas un modèle de données et permet au contraire d'augmenter un modèle selon les besoins d'un éditeur ou d'une plateforme de diffusion. Conformément aux principaux environnements de diffusion scientifiques, le document HTML peut être transformé via une feuille de style XSL dans un format XML, qui sera alors validé contre le schéma XML du diffuseur.
Le format YAML permet de définir les métadonnées du document. Stylo intègre nativement la plupart des métadonnées employées par les diffuseurs scientifiques et par les moteurs de recherche grand public ou spécialisés. Un travail est en cours pour aligner certaines de ces métadonnées, comme les mot-clés ou les auteur·e·s du document, avec des autorités, en exploitant notamment l'API mise à disposition par le moteur de recherche Isidore.science.

Le format BibTeX enfin permet de structurer les références bibliograhiques. Stylo offre la possibilité de synchroniser son document à une collection Zotero.

Ces trois fichiers sont ensuite utilisés dans une chaîne de conversion Pandoc et XSLT, pour produire plusieurs fomats de sortie.

Maintenir la transparence sur ces trois fichiers sources est au coeur de la philosophie de l'outil. Cette transparence et cette accessibilité à la source sont en effet davantage garantes de pérennité, de modularité et de littératie, invitant les usagers à exploiter eux-mêmes ces sources, hors de la chaîne Stylo. Ainsi, si Stylo facilite le processus d'édition, de *versionning*, d'indexation et de conversion vers les formats d'usages, le choix a bien été de conserver un accès direct à la matérialité du texte numérique. Outre la question de la pérennité des données, ce principe va dans le sens d'une technique ouverte, c'est-à-dire visible. Cette transparence du media est un premier pas pour une transmission d'une littératie du numérique, à contre-sens des formats propriétaires fermés.

Chaque version d'un document est accessible en ligne dans un format HTML de prévisualisation, exposé via une URL pérenne, permettant à un auteur ou un éditeur de partager son texte. La prévisualisation est outillée du module d'annotation de la plateforme Hypothesis, ouvrant la voie à divers protocoles éditoriaux de relecture et/ou d'évaluation.

5. Conclusion

Dans la chaîne numérique d'écriture et d'édition des revues savantes en SH, il est de plus en plus clair que l'éditeur ne possède plus la maîtrise sur l'existence culturelle, sociale ou technique de ses contenus dans un environnement dépendant de plus en plus des données liées au texte. En effet, faire exister un texte dans l'environnement numérique suppose d'exposer des données suffisament riches et pertinentes pour que les contenus soient indexés par les moteurs de recherche. Stylo défend la vision selon laquelle chacun des acteurs, de l'auteur au diffuseur, est susceptible d'intervenir dans un processus de production garantissant la continuité des données. Cela est possible grâce au partage d'un format pivot unique, simple d'accès et d'utilisation, et suffisament riche pour l'édition de textes scientifiques en sciences humaines.

Dans cette vision, l'écriture ne se résume plus seulement à écrire. *Écrire dans l'environnement numérique* s'inscrit déjà dans une dynamique d'édition, autrement dit l'acte d'écriture s'accompagne nécessairement d'un acte d'édition. Cette idée que *écrire = écrire + structurer* nous rapproche d'une culture de l'éditorialisation en tant que processus de production de l'espace numérique. Cette culture n'est rien d'autre que la littératie numérique qui ajoute au *savoir-lire-et-écrire* la maîtrise du milieu d'écriture et une capacité à *inscrire* l'espace autant qu'à le structurer. Il s'agit à la fois de savoir évoluer dans cet environnement (l'habiter) et de le faire évoluer (le designer).

Mené comme une recherche-action au plus près des praticiens de l'écriture et de l'édition savante, le développement de Stylo tend certes à une évolution et un enrichissement de ces pratiques, mais aussi à la réflexivité des praticiens sur ces dernières.

6. Bibliographie

Aarseth, Espen J. (1997). *Cybertext: Perspectives on Ergodic Literature*. UK ed. edition. Baltimore, Md: Johns Hopkins University Press.
Dehut, J. (2018). « En finir avec Word ! Pour une analyse des enjeux relatifs aux traitements de texte et à leur utilisation ». Carnet de recherche.
Doueihi, M. (2008). *La Grande conversion numérique*. Traduit par Paul Chemla. Paris, France: Éd. du Seuil.

Fuller, M., éd. (2008). *Software Studies: A Lexicon*. Cambridge, États-Unis d'Amérique: MIT Press.
Galloway, A. R. (2012). *The Interface Effect*. Cambridge, UK ; Malden, MA: Polity Press.
Goody, J. (1979). *La Raison Graphique. La Domestication de La Pensée Sauvage.* Le Sens Commun. Les Editions de Minuit.
Kirschenbaum, M. G. (2016). *Track Changes: A Literary History of Word Processing*. Cambridge, États-Unis d'Amérique: The Belknap Press of Harvard University Press.
Kristeva, J. (1969). *Σημειωτικὴ: recherches pour une sémanalyse*. Paris, France: Éditions du Seuil.
Marino, M. C. (2020). *Critical Code Studies: Initial(Methods)*. Cambridge, Massachusetts: The MIT Press.
Maxwell, J. W., Hanson, E., Desai, Tiampo, C. O'Donnell, K., Ketheeswaran, A., Sun, M., Walter, E., & Michelle, E. (2019). *Mind the Gap: A Landscape Analysis of Open Source Publishing Tools and Platforms*.
McLuhan, M. (2013). *Pour comprendre les médias: les prolongements technologiques de l'homme*. Traduit par Jean Paré. Paris, France: Points.
Sanders, E. B.-N., & Jan Stappers, P. (2008). « Co-Creation and the New Landscapes of Design ». *CoDesign* 4 (1):5-18. https://doi.org/10.1080/15710880701875068.
Vandendorpe, C. (1999). *Du papyrus à l'hypertexte: essai sur les mutations du texte et de la lecture*. Paris, France: La Découverte.