



Nouveaux cahiers de Marge

ISSN : 2607-4427

8 | 2024

Fiction & données

Directeur de publication Gilles Bonnet et Anaïs Guilet

🔗 <https://publications-prairial.fr/marge/index.php?id=814>

Référence électronique

« Fiction & données », *Nouveaux cahiers de Marge* [En ligne], mis en ligne le 24 mai 2024, consulté le 05 septembre 2024. URL : <https://publications-prairial.fr/marge/index.php?id=814>

Droits d'auteur

CC BY-NC-SA

DOI : 10.35562/marge.814



SOMMAIRE

Gilles Bonnet et Anaïs Guilet
Introduction : #datafictions

Données/documents

Servanne Monjour
De la data visualisation à la captafiction

Stéphane Lambion
Cléo de 7 à 9 : de la fiction au document à la fiction

Données/récits

Bertrand Gervais
Bases de données et lignes brisées

Françoise Chambefort
Rythmer le chaos : la programmation du récit dans *Smart Rat City*

Données/réel

Isabelle Gribomont et Anaïs Berck
Bases de données et littérature numérique (belge) : une perspective décoloniale

Serge Bouchardon
Enterre-moi, mon amour : le jeu sur la frontière entre réalité et fiction dans les fictions pour smartphones

Données/auctorialité

Erika Fülöp
Écrire-avec l'intelligence artificielle, ou l'esthétique de la sympoïèse

Françoise Cahen
Ta grand-mère dans la brume

Banques de données

Nathalie Lacelle et Eleonora Acerra
Lab-yrinthe : plateforme québécoise de diffusion d'œuvres numériques jeunesse

René Audet
Constituer un patrimoine littéraire numérique du Québec : méthodologie et défis (projet COLiN, Catalogue des œuvres littéraires numériques)

Christian Cote

Lifranum : Coconstruire un corpus et une archive de littérature nativement numérique

Plateforme Weblitt

Plateforme Weblitt

Plateforme COLiN

Plateforme COLiN – catalogue des œuvres littéraires numériques – construire le corpus du Québec

Plateforme Répertoire des écritures numériques

Plateforme Répertoire des écritures numériques

Plateforme Encodage

Plateforme Encodage

Cultures populaires/cultures savantes

Anne Besson

Quelle(s) légitimation(s) pour les littératures de genre et de grande diffusion ?

Introduction : #datafictions

Introduction: #datafictions

Gilles Bonnet et Anaïs Guilet

DOI : 10.35562/marge.891

Droits d'auteur

CC BY-NC-SA

TEXTE

- 1 Selon les dernières estimations, le volume de données numériques créées ou répliquées à l'échelle mondiale a été multiplié par plus de trente au cours de la dernière décennie, passant de 2 zettaoctets en 2010 à 64,2 zettaoctets en 2020, soit un peu plus de 64 mille milliards de milliards d'octets¹. La production est exponentielle et s'est vue largement aggravée par la crise sanitaire. Selon les prévisions de l'IDC (International Data Corporation), la production de données devrait ainsi augmenter en moyenne de 23 % par an jusqu'à 2025. « La quantité de données numériques créées au cours des cinq prochaines années sera supérieure à deux fois la quantité de celles créées depuis l'avènement du stockage numérique² », a alors précisé le cabinet.



Gilles Bonnet @GillesBonnet5 · 1 mars

...

Pré-colloque, question2: les données sont en si grand nombre, big [#data](#), qu'il faut les abriter dans des "banques", des "entrepôts" ou des "lacs". La datafiction se comprendrait-elle d'abord comme un rapport entre le quantitatif et le qualitatif ?



Source : X/Twitter

- 2 L'enjeu serait de ne pas laisser au marketing, au *data analytics* et autres *data scientists*, le *data storytelling*³, mais de le détourner à des fins critiques et littéraires. Nous aimerions ainsi interroger dans cet ouvrage, proposé à la suite d'un colloque organisé en partenariat avec LQM (Littérature québécoise mobile), Lifranum (Projet de cartographie du Web francophone) et le laboratoire LLSETI (Langages, littératures, sociétés, études transfrontalières et internationales) en mars 2023, les liens entre données et fictions dans le champ de la littérature numérique. De la protection des données au *big data* en passant par l'avènement d'une société de surveillance, les données sont un enjeu majeur pour nos sociétés contemporaines, dont les échos se répercutent autant dans les procédés d'écriture des œuvres littéraires numériques que dans les thématiques qu'elles abordent.



Gilles Bonnet @GillesBonnet5 · 1 mars

...

Pré-colloque, question 1: "datafication" = transformation en données. Par ex. un poème numérisé est transformé en données. Nous nous intéressons au phénomène inverse : partir de données et en proposer une création littéraire, ce que l'on pourrait nommer une... 📖 datafiction. [#data](#)

Source : X/Twitter

- 3 La donnée s'impose comme un mode de visualisation du réel. Selon Serge Abiteboul :

« [Elle] est une description élémentaire d'une réalité. C'est par exemple une observation ou une mesure. La donnée est dépourvue de tout raisonnement, suppositions, constatations, probabilités. Étant indiscutable ou indiscutée, elle sert de base à une recherche ou à un examen quelconque⁴ ».

- 4 Ce qui nous intéressera dans la donnée c'est tout à la fois le rapport étroit qu'elle entretient avec le réel, son statut factuel et la prémisse qu'elle constitue, son caractère inaugural, invitant à l'analyse, mais aussi à l'explication, la mise en récit.



Gilles Bonnet @GillesBonnet5 · 1 mars

...

Pré-colloque, question3: qu'est-ce que fictionnaliser des données ? La fiction incarne/traduit/performe/hacke/traduit-elle les [#data](#) ? quelle pertinence pour l'un des paradigmes qui traverse l'histoire des arts et de la littérature, celui de la représentation?

Source : X/Twitter

- 5 En effet, l'utilisation des données en contexte littéraire numérique engage bien des questions narratives et fictionnelles : quelle singularité, quel statut pour ces îlots référentiels d'un nouveau genre ?



Gilles Bonnet @GillesBonnet5 · 1 mars

...

Pré-colloque, question10: comment la [#datafiction](#) dialogue-t-elle avec les autres esthétiques numériques comme le data art ? La visualisation de la [#data](#), tentation constante pour la rendre lisible car visible, crée-t-elle un nveau lien entre texte et [#image](#) dans la datafiction?

Source : X/Twitter

- 6 Dans la lignée du data art, dont les artistes comme Nathalie Miebach, Aaron Koblin ou Peter Crnokrak s'approprient des données pour produire des représentations numériques ou plastiques⁵, les auteurs et autrices numériques ont recours à différentes données dans des contextes narratifs à visée non pas documentaire ou didactique, mais fictionnelle. C'est le cas, par exemple, du travail de Françoise Chambeffort⁶ dans *Lucette, gare de Clichy*⁷ (2017) qui se construit en temps réel sur le flux de données du réseau ferroviaire de la région parisienne mis en parallèle du récit de Lucette qui vit en face de la gare ; ou de J. R. Carpenter qui, dans *In Absentia*⁸ (2008), a recours aux cartes API-Google pour créer une narration interactive non linéaire, constituée d'histoires *cartes postales*, écrites selon le point de vue d'anciens locataires du Mile-End à Montréal, forcés de quitter leur logement à cause de la gentrification. Dans ces récits, données et fictions se mêlent de manière parfois inextricable, créant comme c'est le cas dans le transmédia *Netwars*⁹ (2018), un espace pour le moins paranogène qui sied à l'environnement informationnel contemporain hanté par les complotistes, comme par les *fake news*¹⁰.
- 7 L'utilisation des données en contexte littéraire numérique explore aussi les modes d'immersion fictionnelle en inscrivant le corps ou l'espace du sujet lecteur dans l'univers diégétique grâce à l'utilisation de la caméra, du gyroscope ou de la géolocalisation par exemple. S'il existe toute une tradition de cette insertion dans les jeux vidéo, notamment depuis le CD-ROM *In Memoriam*¹¹ en 2003, précurseur des jeux en réalité augmentée et autres *locative narratives*, ou encore *ambient* littérature, ce champ reste à explorer dans le cadre plus spécifique de la littérature numérique. À ce titre, l'œuvre de James Attlee, *The Cartographer's Confession*¹² (2017), se présente comme une combinaison de fiction et non-fiction, d'espaces imaginés et réels dont le récit repose en partie sur les données de localisations

de sa lectrice. Nous pouvons aussi citer *Breathe*¹³ (2018), l'application de Kate Pullinger, qui inclut le lecteur au fil de l'histoire grâce à la géolocalisation, au recours à la caméra comme aux données météorologiques et temporelles.



Gilles Bonnet @GillesBonnet5 · 1 mars

...

Pré-colloque, question8: Emmanuelle Pireyre ("Devenirs du romans" 2) définit la littérature comme "machine à transformer les données". Comment articuler dès lors calcul et #écriture? La #datafiction invente-t-elle une voie entre les deux?

Source : X/Twitter

- 8 Certes, la littérature n'a pas attendu le numérique pour avoir recours à des données. Les auteurs·rices peuvent faire des recherches documentaires pour leurs romans, ils·elles peuvent créer dans leurs récits des îlots référentiels ou jouer d'une intertextualité factuelle. Cette soif de réalité a très bien été analysée par David Shields dans *Reality Hunger – A Manifesto*¹⁴. La data-fiction apparaît alors comme une nouvelle forme de ce désir de réalité qui impose en retour d'interroger les spécificités de ces données numériques. Elles semblent, selon nous, reposer sur :

- leur accessibilité en même temps que leur caractère parfois caché (cookies, tracking, surveillance) ;
- leur immédiateté ;
- leur quantité ;
- leur grande reproductibilité, appropriation, copier-coller.



Gilles Bonnet @GillesBonnet5 · 1 mars

...

Pré-colloque, question4: la fiction est entendue ici comme une poïétique s'emparant de la #data par appropriation. Qui est un pilier de la culture numérique. Cette appropriation comporte-t-elle une dimension critique ? Se confond-elle avec le paradigme post-moderne du #remix ?

Source : X/Twitter

- 9 Posée ainsi, la donnée constitue donc un matériau tout à fait fertile pour les auteurs numériques.

 **Gilles Bonnet** @GillesBonnet5 · 1 mars ...
Pré-colloque, question6: assistons-nous à une réplique du "tournant documentaire" de la [#litterature](#) contemporaine, qui serait, pour la littérature [#numerique](#) un tournant de la [#data](#)? Quelles porosités entre les deux?

Source : X/Twitter

 **Gilles Bonnet** @GillesBonnet5 · 1 mars ...
Pré-colloque, question11: la [#datafiction](#), par son lien à la donnée qui est elle-même liée au monde sur le mode de l'enregistrement, est-elle par nature une "factographie", une de ces « littératures factuelles » définies par Marie-Jeanne Zenetti?

Source : X/Twitter

 **Gilles Bonnet** @GillesBonnet5 · 1 mars ...
Pré-colloque, question9: Ars datae=ars memoriae ? Ces [#datafictions](#), parce qu'ancrées dans la [#data](#), sont-elles des œuvres qui font de la littérature un nouvel art de la mémoire ? Cf. [@nt2bert](#) et ses inspirantes « architectures de la mémoire »

Source : X/Twitter

- 10 Si l'on pense également à toutes les œuvres numériques s'insérant sur les réseaux sociaux numériques, qu'il s'agisse de la twittérature ou des romans sur Instagram pour ne citer que ces exemples, c'est la frontière entre réel et fiction qui se voit sans cesse réinterrogée.

 **Gilles Bonnet** @GillesBonnet5 · 1 mars ...
Pré-colloque, question7: dialectique ouverture/clôture. Imaginaire de la donnée = clôture+pureté : il faut "préserver" la [#data](#) qui encapsule le réel. Mais la fiction revendique pluralité et impureté, pour « faire advenir des possibles par expérience de pensée » ([@NMurzilli](#))

Source : X/Twitter

- 11 Ces récits provoquent une intrusion de la fiction dans des espaces qui ne lui sont pas originellement destinés, quand bien même ils frayent souvent avec l'autofiction et la distance qu'implique toute

représentation de soi. Les réseaux sociaux numériques fonctionnent en effet comme autant de leurres documentaires que ces récits littéraires fictionnels viennent révéler. Les œuvres de littérature numérique sur ces réseaux induisent de nouveaux pactes fictionnels qu'il s'agira dès lors d'analyser.



Gilles Bonnet @GillesBonnet5 · 1 mars

...

Pré-colloque, question12/12: White Cube vs Black Box. Le contemporain (cf.Pascal Mougin) rompt avec l'insularité revendiquée par l'art moderne. Quid de la [#datafiction](#) qui semble revenir au monde mis en [#data](#), mais pour habiter le [#Web](#)...? Quels régimes de "recontextualisation" ?

Source :X/Twitter

- 12 Les œuvres numériques semblent donc sinon contribuer, jouer, notamment grâce aux recours aux données, au brouillage des frontières entre réel et fiction. Un brouillage qui correspond à une idée largement répandue que Françoise Lavocat dans *Faits et fiction* se sera largement attachée à déconstruire, cherchant à « montrer l'existence et la nécessité cognitive, conceptuelle et politique des frontières de la fiction¹⁵ » dans tous les arts de la fiction. Partant de ce constat ambivalent : entre la nécessité de fonder les frontières entre fiction et réel et le travail de brouillage perpétuel pratiqué par les auteurs de littérature numérique, cet ouvrage se propose d'explorer comment données et fictions interagissent, ainsi que d'analyser les conséquences narratives, esthétiques et théoriques de ces interactions.

BIBLIOGRAPHIE

ABITEBOUL Serge, « Sciences des données. De la logique du premier ordre à la Toile », *Leçon inaugurale à la chaire Informatique et sciences numériques du Collège de France*, 8 mars 2012.

CHAMBEFORT Françoise, *Mimèsis du flux, exploration des potentialités narratives des flux de données*, thèse de doctorat, sous la direction de I. Roxin et S. Bouchardon, université Bourgogne-Franche-Comté, 2020.

DAVENPORT Thomas H., *Big Data @Work*, Boston, Harvard Business school Press, 2014.

GROUPIERRE Karleen et GUILLET Anaïs, « In the dispositif / out of CTRL : *netwars* ou le *hacking* transmedia », *Itinéraires. Littérature, textes, cultures* [en ligne], 2016-2, 2017, DOI : [10.4000/itineraires.3495](https://doi.org/10.4000/itineraires.3495).

IDC, « Data Creation and Replication Will Grow at a Faster Rate than Installed Storage Capacity, According to the IDC Global DataSphere and StorageSphere Forecasts », *Internet Archive* [en ligne], 24 mars 2021, URL : <https://web.archive.org/web/20210324180326/https://www.idc.com/getdoc.jsp?containerId=prUS47560321> [consulté le 4 juin 2024].

GERVAIS Bertrand et DALLE Jean-Marie, « Architectures et palais de mémoire », *Architectures de mémoire* [en ligne], Nanterre, Presses universitaires de Paris-Nanterre, 2019, DOI : [10.4000/books.pupo.25797](https://doi.org/10.4000/books.pupo.25797).

LAVOCAT Françoise, *Faits et fiction : pour une frontière*, Paris, Seuil, 2016.

MOUGIN Pascal, *Moderne / contemporain – Art et littérature des années 1960 à nos jours*, Paris, Presses du réel, 2019.

MURZILLI Nancy, « La fiction ou l'expérimentation des possibles », in : GEFEN Alexandre (dir.), *L'effet de fiction* (2001) [en ligne], DOI : [10.58282/colloques.7710](https://doi.org/10.58282/colloques.7710).

NUSSBAUMER KNAFLIC Cole, *Datavisualisation. Utilisez le storytelling pour faire parler vos données*, Paris, Pearson France, 2019.

PIREYRE Emmanuelle, « Comment ne plus être data victim ? » in : Collectif, *Devenirs du roman 2 : écriture et matériaux* [en ligne], éditions Inculte, 2014, URL : <http://www.emmanuellepireyre.com/wp-content/uploads/2015/11/Art.-Data-victim.pdf> [consulté le 8 avril 2024].

PIREYRE Emmanuelle, « Témoins de l'horreur », *Le Monde des livres* [en ligne], 19 mai 2016, URL : https://www.lemonde.fr/livres/article/2016/05/19/emmanuelle-pireyre-temoins-de-l-horreur_4922023_3260.html [consulté le 8 avril 2024].

SHIELDS David, *Reality Hunger – A Manifesto*, London, Hamish Hamilton, 2010.

The Open Window, « Data art, la donnée comme matériau d'origine », *Usbek et Rica* [en ligne], 8 mars 2017, URL : <https://usbeketrica.com/fr/article/data-art-la-donnee-comme-matériau-d-origine> [consulté le 8 avril 2024].

ZENETTI Marie-Jeanne, *Factographies. L'enregistrement littéraire à l'époque contemporaine*, Paris, Classiques Garnier, 2014.

NOTES

1 IDC, « Data Creation and Replication Will Grow at a Faster Rate than Installed Storage Capacity, According to the IDC Global DataSphere and StorageSphere Forecasts », *Internet Archive* [en ligne], 24 mars 2021, URL : [h](https://web.archive.org/web/20210324180326/https://www.idc.com/getdoc.jsp?containerId=prUS47560321)

[c.jsp?containerId=prUS47560321](#) [consulté le 4 juin 2024]. L'article n'étant plus en ligne, nous avons ajouté le lien vers Internet Archive, où il est consultable.

2 Loc. cit.

3 Voir <http://www.datacomics.net> [consulté le 8 avril 2024] ; Cole Nussbaumer Knaflig, *Datavisualisation. Utilisez le storytelling pour faire parler vos données*, Paris, Pearson France, 2019 ; Thomas H. Davenport, *Big Data @ Work: Dispelling the Myths, Uncovering the Opportunities*, Boston, Harvard Business School Press, 2014.

4 Serge Abiteboul, « Sciences des données. De la logique du premier ordre à la Toile », *Leçon inaugurale à la chaire Informatique et sciences numériques du Collège de France*, 8 mars 2012.

5 Cf. The Open Window, « Data art, la donnée comme matériau d'origine », *Usbek et Rica* [en ligne], 8 mars 2017, URL : <https://usbeketrica.com/fr/article/data-art-la-donnee-comme-materiau-d-origine> [consulté le 8 avril 2024].

6 Se référer aussi à son travail de thèse : Françoise Chambefort, *Mimèsis du flux, exploration des potentialités narratives des flux de données*, thèse de doctorat, sous la direction de I. Roxin et S. Bouchardon, université Bourgogne-Franche-Comté, 2020.

7 Françoise Chambefort, *Lucette, gare de Clichy* [en ligne], 2007, URL : <http://fchambef.fr/lucette/> [consulté le 8 avril 2024].

8 J. R. Carpenter, *In Absentia* [en ligne], URL : <http://luckysoap.com/inabsentia/> [consulté le 8 avril 2024].

9 Thomas Tielsch, Saskia Kress, *Netwars*, filmtankGmbH [en ligne], 2018, URL : <https://www.netwars-project.com> [consulté le 8 avril 2024].

10 Karleen Groupierre et Anaïs Guilet, « In the dispositif / out of CTRL : netwars ou le hacking transmedia », *Itinéraires. Littérature, textes, cultures* [en ligne], 2016-2, 2017, DOI : [10.4000/itineraires.3495](https://doi.org/10.4000/itineraires.3495).

11 Eric Viennot, *In Memoriam* [CD-ROM], Lexis Numérique, 2003.

12 James Attlee, *The Cartographer's Confession* [en ligne], 2017, URL : <http://research.ambientlit.com/cartographersconfession> [consulté le 8 avril 2024].

13 Kate Pullinger, *Breathe* [en ligne], 2018, URL : <https://www.breathe-story.com> [consulté le 8 avril 2024].

- 14 David Shields, *Reality Hunger – A Manifesto*, London, Hamish Hamilton, 2010.
- 15 Françoise Lavocat, *Faits et fiction : pour une frontière*, Paris, Seuil, 2016.

AUTEURS

Gilles Bonnet

IDREF : <https://www.idref.fr/073426725>

ORCID : <http://orcid.org/0000-0002-1028-3898>

ISNI : <http://www.isni.org/0000000367655055>

BNF : <https://data.bnf.fr/fr/14493276>

Anaïs Guilet

IDREF : <https://www.idref.fr/178389382>

ORCID : <http://orcid.org/0000-0001-7247-7156>

ISNI : <http://www.isni.org/000000043641853X>

Données/documents

De la data visualisation à la captafiction

Autour d'Anna K de Catherine Lenoble

From data visualization to captafiction: around Catherine Lenoble's Anna K

Servanne Monjour

DOI : 10.35562/marge.894

Droits d'auteur

CC BY-NC-SA

PLAN

« Captafiction », un néologisme de la littérature numérique, encore ?
Pour une approche graphique de la littérature des données
Portrait de l'artiste en *dataset*
L'écriture littéraire des données : une poétique de la curation numérique
Anna K : folle à lier ?
Conclusion

TEXTE

- 1 « Fictions et données » : la tension entre ces deux termes – le premier étroitement associé à l'imaginaire, le second rattaché au réel (non sans problème, on y reviendra) – n'est que l'une des nombreuses manifestations possibles du brouillage entre réel et fiction qui caractérise le moment contemporain de la littérature. Décrit par une partie de la critique, dont Françoise Lavocat qui, dans *Faits et fictions*, pointe avec justesse la « nécessité cognitive, conceptuelle et politique des frontières de la fiction¹ », ce brouillage est pourtant l'une des caractéristiques majeures de la littérature numérique, et se trouve abondamment traité dans les travaux qui lui sont consacrés.
- 2 Il faut dire que les œuvres littéraires numériques, de par leur présence dans un espace en ligne, dont le statut ontologique n'est pas si simple à déterminer – tant il empiète sur notre espace dit « réel » ou, du moins, sur notre façon d'habiter le monde² – occupent une place particulière dans ce débat théorique. De fait, il est essentiel de souligner combien l'étude de la littérature numérique investit nécessairement une tension entre deux conceptions de la création

qui s'opposent bien au-delà du champ théorique littéraire depuis déjà plusieurs dizaines d'années : d'une part, une conception représentative où l'œuvre se pose en fenêtre sur le monde, d'autre part une conception performative dans laquelle le geste artistique participe à façonner et à structurer le réel. Cette dernière est centrale dans la théorie de l'éditorialisation³ régulièrement convoquée dans le champ des études littéraires numériques. Ainsi, dans des travaux antérieurs, nous avons même défendu un parti pris radical en affirmant l'idée d'une *fusion* entre le réel et l'imaginaire :

- 3 Le fait de penser le geste de production des contenus numériques comme un geste de production de l'espace nous permet justement d'aller au-delà de l'opposition imaginaire/réel. Les pratiques, les discours et les technologies impliquées dans un geste d'éditorialisation ont comme résultat l'agencement d'un espace tout à fait réel, à savoir l'espace dans lequel nous vivons⁴.
- 4 Selon cette conception, une fiction en ligne serait donc une donnée comme une autre, située sur le même plan ontologique que, par exemple, les horaires de train disponibles via l'API de la SNCF, utilisée par Françoise Chambefort pour réaliser *Lucette, gare de Clichy*. Pourtant, ces deux paradigmes représentatifs et performatifs semblent s'évertuer à cohabiter au sein des fictions littéraires en ligne, jouant abondamment des effets d'ambiguïté entre réel et fiction – ambiguïté que les théories de l'éditorialisation littéraire, dans leur état actuel, ne creusent sans doute pas assez. C'est pourtant dans cette tension que réside tout l'intérêt de l'écriture littéraire des données : quel statut conférer à un profil numérique dit « fictif » (bot, profil de fiction, etc.), alors même qu'il intervient et converse sur un réseau social avec le profil d'un individu réel ? En quoi un récit géolocalisé peut-il transformer un quartier, en faisant converger ses habitants vers des espaces singuliers ?
- 5 En ce qui concerne le rapport entre fiction et données, le problème est redoublé par le statut ontologique problématique des données elles-mêmes, éléments extraits du réel de manière arbitraire, mais qui ont tendance à se substituer au réel, ou du moins à en tenir lieu dans notre imaginaire⁵. Aussi, la réflexion sur la littérature des données suppose d'abord un déplacement conceptuel : ce que l'on appelle « réel » a changé sous l'effet de sa mise en données.

Bachimont évoque même, on le verra, un nouveau régime de connaissance, fondé sur un principe de codification. En toute logique, ce que l'on appelle « fiction » évolue également.

- 6 Cet article a pour objectif de décrire et d'analyser certaines de ces mutations de la fiction lorsqu'elle se confronte à la *réalité* des données, tant sur un plan pratique que théorique. C'est notamment tout l'enjeu du néologisme que je propose, la « captafiction », dont je baliserai les contours à l'aide du projet *Anna K* de Catherine Lenoble⁶.

« Captafiction », un néologisme de la littérature numérique, encore ?

- 7 La notion de captafiction – alternative à la data-fiction – que j'ajoute à la liste déjà conséquente des néologismes de la littérature numérique est un parti pris conceptuel qui pose l'accent sur la problématique ontologique propre aux données numériques. Pour s'orienter dans cette problématique, il m'apparaît capital de convoquer certains travaux structurants menés ces dernières années dans le champ des sciences humaines et sociales. J'insisterai, en particulier, sur la notion de « capta », grâce à laquelle on peut poser un cadre théorique et méthodologique précieux pour repenser les enjeux de la mimesis littéraire.
- 8 Tout d'abord, la notion de « capta » permet de s'abstraire du terme « données », dont il s'agit avant tout déconstruire le faux certificat d'authenticité. C'est désormais un poncif de le dire, et pourtant l'ambiguïté demeure : une donnée numérique n'est pas une émanation brute du réel. Elle est construite, choisie, prélevée en fonction de présupposés ou de postulats parfois clairement énoncés, mais plus souvent inconscients et impensés. C'est d'ailleurs l'un des apports les plus fondamentaux des SHS à la science des données que de proposer une réflexion critique sur le concept de données lui-même, auquel plusieurs alternatives terminologiques ont été proposées : Bruno Latour, notamment, y substitue le terme d'« obtenues »⁷, tandis que Johana Drucker⁸ a forgé le terme « capta » auquel je me rallie ici également – j'en détaillerai les raisons

précises dans un instant. Fondamentalement, ces alternatives visent à déconstruire l'illusion d'objectivité de la donnée, en la réintégrant dans un paradigme représentationnel – la donnée n'est pas un échantillon du réel, elle en est déjà une remédiation.

- 9 Or si la donnée relève en partie d'un paradigme représentationnel, elle n'en fait pas moins bouger les frontières de la représentation. C'est la thèse que défendent Bruno Bachimont et Lise Verlaet, à qui l'on doit une série de travaux récents sur l'ontologie des données. Pour Bruno Bachimont, c'est parce qu'elle est fondamentalement composée de code – code qui, dans son principe, « établit une convention arbitraire entre le réel et sa représentation [et] ne repose pas sur une exigence de conformité avec le réel lors sa création ni de coévolution par la suite⁹ » – que la donnée implique un nouveau régime de connaissance et de représentation :

« avec les systèmes numériques et le codage de et par l'information, la représentation [de l'époque contemporaine] ne ressemble pas à son origine ni ne covarie de manière nécessaire avec elle, mais en est un codage dont la manipulation calculée n'impose pas que le résultat soit une conséquence du comportement de l'objet ou réalité qui est codée¹⁰ ».

- 10 En distinguant entre elles successivement les notions de trace, d'enregistrement et de données numériques, qui creusent de l'une à l'autre l'écart entre le réel « brut » (lui-même clairement fantasmé d'ailleurs) et la représentation du réel, Bruno Bachimont et Lise Verlaet soulignent alors l'importance de mettre au point une herméneutique des données capable de reconstruire du sens :

« C'est la raison pour laquelle l'intelligibilité du numérique en passe par l'intelligence des traces et leur herméneutique ou interprétation. Il s'agit de comprendre ce qui se passe quant au sens quand on traite les traces comme des données à mobiliser dans des systèmes numériques, et quels sont les outils, méthodes, concepts permettant de reconstruire un régime du sens, une manière de comprendre et d'interpréter ces traitements¹¹. »

- 11 L'hypothèse que je veux défendre ici, c'est que la captafiction, en tant qu'écriture littéraire des données, a vocation à se poser comme l'une

des méthodes susceptibles de reconstruire un régime du sens. Or pour y parvenir, il faut sans doute commencer par dépasser l'opposition radicale entre réel et fiction, et adopter un modèle ontologique moins binaire, apte à rendre compte d'une dimension multiple, hétérogène (soit, parfois, contradictoire et conflictuelle) du monde et de nos expériences du réel. C'est notamment sur ce dernier point que le paradigme performatif évoqué plus tôt peut nous aider – sans résoudre toutefois complètement la question.

- 12 L'herméneutique des données est également au cœur des travaux de Johanna Drucker¹², à qui l'on doit le terme de « capta » mentionné tout à l'heure. La proposition de Johanna Drucker, que je résume brièvement, tout en tâchant d'y rester la plus fidèle possible, s'articule autour du concept d'« interprétation modélisante » que Drucker oppose radicalement à la « visualisation de l'information », conçue comme un simple affichage de données. Ce que Drucker reproche à la visualisation classique, largement répandue aujourd'hui dans les médias et jusque dans les travaux scientifiques, c'est sa tendance à redoubler la tentation objectivante des données, alors même que chaque visualisation devrait d'abord être considérée comme une remédiation de la donnée.
- 13 Là où les visualisations ont par ailleurs une fonction souvent illustrative – elles sont largement utilisées comme des outils rhétoriques servant à conforter un argument – Drucker défend la constitution d'une épistémologie visuelle à part entière, dans laquelle l'écriture graphique n'aurait pas seulement vocation à représenter une connaissance existante, mais jouerait également un rôle en tant que mode de production de connaissance.
- 14 Cela veut dire qu'une expression graphique pourrait être utilisée pour exposer – ainsi que pour créer – un modèle d'activité interprétative. Cette approche change la relation unidirectionnelle entre les données et l'affichage et la transforme en un échange réciproque¹³.
- 15 Si la littérature n'est pas le terrain de Drucker, son travail est pourtant susceptible d'éclairer en partie le tournant médiatique de la littérature largement redevable des expérimentations numériques. Image du texte, sémiotique du support et dialectique intermédiaire participent en effet à la constitution d'une herméneutique un peu particulière, où le sens surgit également des effets de matérialité du

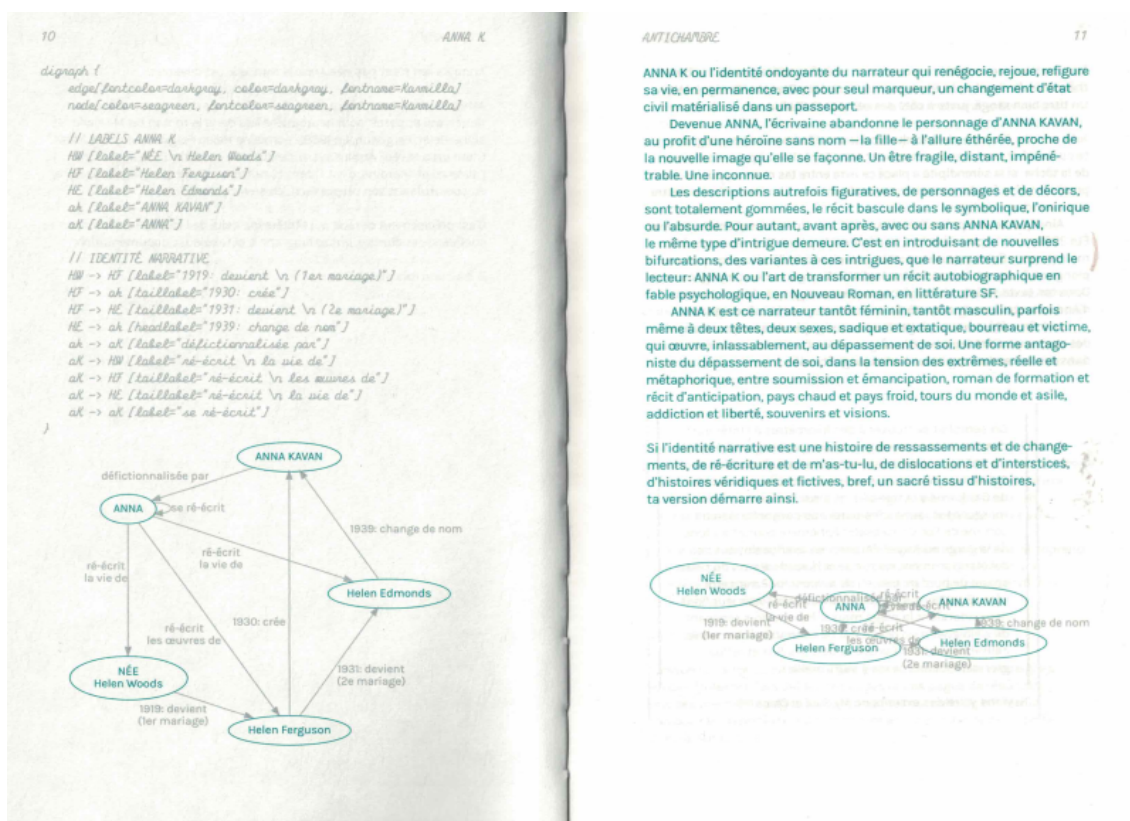
texte. Or, cette dimension médiatique, et en particulier graphique, est un aspect essentiel de la captafiction que je vais à présent présenter de manière plus concrète, à travers le cas d'étude offert par le projet *Anna K* de Catherine Lenoble.

Pour une approche graphique de la littérature des données

- 16 Publié en 2016 aux éditions HYX, l'ouvrage *Anna K* est l'une des productions issues d'un projet artistico-littéraire¹⁴, consacré à l'écrivaine Anna Kavan, qui a vécu et publié dans la première partie du xx^e siècle. Le projet comprend également une création en ligne, *Kavan.land*, dont il sera fait mention dans cet article, ainsi qu'une archive offline *Kavan.lan*, que je laisse en revanche de côté faute d'avoir eu l'occasion de la consulter.
- 17 L'analyse que je propose met l'accent sur la dimension graphique de l'écriture des données – depuis leur traitement informatique jusqu'à leur remédiatisation sous forme de visualisation – à travers une perspective éditoriale d'abord attachée à l'image du texte, au sens premier du terme : sa mise en page, sa mise en forme, sa typographie et ses « illustrations ». Il existe évidemment de nombreux travaux de visualisation de données littéraires, ainsi que des travaux sur les enjeux épistémologiques de cette visualisation, mais finalement assez peu d'études sur la visualisation en tant qu'écriture littéraire. Si l'on a déjà beaucoup commenté, en effet, la dimension informatique des écritures numériques (c'est le domaine des *critical code studies*¹⁵), ainsi que leur dimension médiatique (dans le domaine des *platform studies*¹⁶), ou encore leur dimension culturelle et poétique¹⁷, on a encore assez peu abordé leur dimension graphique – au sens où le produit l'art de la visualisation.
- 18 Or, c'est bien de ce phénomène que témoigne le travail Catherine Lenoble. Le projet *Anna K* comprend d'abord un *dataset*, soit un jeu de données collectées par Lenoble autour de la thématique « Anna Kavan », puis un récit, une sorte de biographie-enquête, conçue autour ou à partir d'un ensemble de visualisations des données – lesquelles ont été réalisées en collaboration avec Alexandre Leray et Stéphanie Vilayphiou, membres du collectif *Open Source Publishing*.

- 19 Une première consultation de l'ouvrage révèle une organisation graphique complexe, mêlant plusieurs formes et formats d'écritures, signalés par un travail éditorial et un design singulier :

Figure 1 : Double page de l'ouvrage *Anna K* (p. 10-11) faisant figurer les différents formats d'écriture (code, langage naturel, graphes). Reproduit avec l'aimable autorisation de Catherine Lenoble et des éditions HYX.



- une écriture du code dont la police est cependant très éloignée du standard du code informatique, qui plus est présentée ici sous une forme cursive qui a pour effet de désinformatiser le code ;
- une écriture en « langage naturel », reprenant une typographie classique, où alternent cependant différentes « voix », incarnées par un jeu de couleur (gris et vert) ;
- une écriture graphique, qui relève donc de la visualisation informatique, et qui s'appuie comme on le découvre rapidement à la lecture sur le jeu de données établi par l'autrice. Ces graphes reprennent pour la plupart le même jeu de couleur vert / gris (gris pour les liens [egde], vert pour les nœuds [node]). Ils ont été réalisés pour la plupart à l'aide de l'outil

open source Graphviz, en collaboration avec le collectif Open Source Publishing.

- 20 Quel rôle ces éléments de visualisation viennent-ils jouer dans le projet de Catherine Lenoble ? Se contentent-ils d'ancrer le récit dans le réel ou, du moins, de l'objectiver, un peu comme un document d'archives ou une photographie insérée dans un récit ? Difficile d'échapper aujourd'hui à une forme d'obsession du graphe, que l'on retrouve largement dans les écrits non fictionnels (enquêtes journalistiques, travaux scientifiques, rapports). Difficile également pour le néophyte d'apprécier ces visualisations pour leurs qualités graphiques, voire plastiques, tant les formes que l'on nous présente manquent souvent d'inventivité, au bénéfice d'une intelligibilité (parfois elle-même discutable, d'ailleurs).
- 21 Mais si le graphe relève a priori d'un registre didactique, il n'en reste pas moins appréciable pour ses qualités esthétiques. Les travaux déjà mentionnés de Johanna Drucker, mais également d'Edward R. Tufte¹⁸, ont démontré le potentiel aussi bien heuristique qu'esthétique de ces formes graphiques, dont les premiers exemples apparaissent bien avant l'ère informatique.
- 22 On tiendra dès lors le pari d'une poétique texte/image, fondée sur la data visualisation et participant d'une double herméneutique de la donnée. Tandis que les graphiques présents dans *Anna K* jouent en partie un rôle de représentation visuelle de données massives (pour nous les rendre intelligibles), je voudrais également démontrer qu'ils renouent avec un acte graphique d'écriture. Leur présence dans l'espace d'inscription qu'est le livre, en particulier, permet de réintroduire une certaine ambiguïté au sein de ces formes généralement conçues comme objectivantes et transparentes. En effet, chez Catherine Lenoble, nous constaterons que les visualisations mettent le lecteur face à un défi interprétatif plutôt ardu, puisqu'un examen attentif a tendance à suggérer des formes imbriquées et ambiguës, tendant vers la figuration ou l'expression symbolique.

Portrait de l'artiste en *dataset*

- 23 L'importance de cette dimension graphique de l'écriture des données à présent clairement établie, revenons-en à la thématique et la

méthodologie du projet Anna K. Autrice d'une œuvre au caractère expérimental très marqué et par endroits torturée, Anna Kavan a été parfois comparée à Franz Kafka, et a inspiré plusieurs écrivaines de son époque, notamment Anaïs Nin qui ne cache pas son admiration pour son œuvre. Mais il faut bien l'avouer, Kavan n'a guère connu une incroyable postérité.

- 24 Le projet Anna K se conçoit donc d'abord comme une enquête visant à redécouvrir le destin de l'autrice ou, plus précisément, comment Anna Kavan est *devenue* Anna Kavan :

« Anna Kavan n'est pas née Anna Kavan, elle l'est devenue. »

« ANNA KAVAN (nom propre, fictif, Anna Kavan) est un personnage de fiction qui apparaît pour la première fois dans le roman *Let Me Alone*, signé Hélène Ferguson. En 1939, l'écrivaine Hélène Ferguson devient ANNA KAVAN (nom propre, réel, Anna Kavan). Ce qui relie et fabrique ce lien, entre l'autrice et l'héroïne, c'est l'identité narrative. Celle que chacun élabore, en construisant son propre récit, en se racontant¹⁹. »

- 25 Posant d'emblée l'ambiguïté ontologique de son objet, ANNA KAVAN, nom fictif devenu réel, Catherine Lenoble inscrit son projet dans un paradigme performatif, autour du concept d'identité narrative, dont elle va elle-même largement jouer dans son récit, composé de quatre parties écrites tantôt à la première personne, tantôt à la seconde personne du singulier.
- 26 Le récit *Anna K* s'inscrit en effet dans un régime à la fois mémoriel et intime qui le rapproche d'une tendance littéraire actuellement en vogue : ce que Dominique Viart²⁰ appelle la littérature de terrain (sans pour autant que son travail mentionne les œuvres numériques ni la question des données) typique d'un « nouvel âge de l'enquête » tel que le qualifie Laurent Demanze²¹. Dans *Anna K*, en effet, il s'agit tout autant de produire un récit biographique sur Anna Kavan que de documenter l'enquête et d'en mesurer les effets sur l'enquêtrice. La littérature contemporaine comprend de fameux exemples de cette double démarche, chez Ivan Jablonka ou Philippe Jaenada (prix Femina en 2017 pour *La Serpe*). Tout en se rapprochant de ce courant, la démarche de Catherine Lenoble est également imprégnée d'une

méthodologie inspirée de la science des données, voire des humanités numériques littéraires, qui ont transformé le texte littéraire en données, pour proposer de nouvelles méthodes interprétatives : les travaux de Stéfán Sinclair et Goeffrey Rockwell²² en constituent un bon exemple. La captafiction naît de cette rencontre entre deux démarches, l'une littéraire, l'autre scientifique, qui ont à cœur de proposer des méthodes pour rendre le monde plus intelligible.

- 27 Comme elle s'en explique dès le début de l'ouvrage, Catherine Lenoble a découvert Anna Kavan lors d'un séjour en Slovaquie, après avoir déniché l'un de ses ouvrages chez un libraire. D'autres lectures suivront, en parallèle desquelles se développe peu à peu une passion pour tout ce qui la concerne. Commence alors une collecte d'informations analogiques (ouvrages, archives, correspondances), mais surtout numériques (bases de données, sites web, communautés de lecteurs rassemblés sur les forums ou les réseaux sociaux). Bientôt, la collecte tourne à l'obsession : comprendre l'œuvre, comprendre l'artiste, comprendre les données.
- 28 Anna Kavan est en effet une figure complexe. Ses pseudonymes (le vrai nom d'Anna Kavan est Helen Woods, qui se trouve être le même nom que sa mère, ses premières œuvres ont été publiées sous le nom d'Helen Ferguson, également le nom de son premier mari, puis elle a choisi Anna Kavan), son nomadisme (elle a énormément voyagé en Europe, en Asie et aux USA), sa vie amoureuse aussi trépidante que malheureuse, sa dépendance à l'héroïne qui causera sa perte en ont fait une figure romanesque prisée des biographes. Beaucoup s'y sont cassé les dents :

« [Je] découvre les archives de cet universitaire américain, Richard Centing, qui entreprit d'écrire une biographie d'ANNA, sans jamais y parvenir. [...] Callard mentionne la "malédiction ANNA KAVAN", qu'elle aurait eu de nombreux aspirants, prêts à relever l'énigmatique défi, en vain. Dans sa liste il cite Clive Jordan, un journaliste américain qui malgré les voyages, de Tulsa à Auckland, échoue. La biographie se transforme en article, intitulé "Among the lost things", publié dans le *Daily Telegraph* en 1972. Vient ensuite Priscilla Dorr, doctorante à l'université de Tulsa au milieu des années 1980, lancée sur une thèse : "Anna Kavan: A Critical Introduction". Avec la mort du second légataire, Raymond Marriot, la bibliothèque McFarlin récupère

en 1992 une correspondance épaisse de plus de 800 lettres, le fameux gros morceau. Dans la foulée, Perter Own propose à Priscilla Dorr un travail biographique. Quand Callard se rend à Tulsa la même année, pour commencer la biographie de R., il rencontre Priscilla aux archives. Il rapporte dans sa lettre les difficultés éprouvées par la thésarde qui passe aux aveux un soir, après avoir bu quelques bières : l'univers d'Anna lui est trop obscur, elle renonce définitivement à la biographie²³. »

- 29 Qui est Anna Kavan ? Le travail d'enquête s'annonçait déjà complexe dans le monde de l'archive analogique, Catherine Lenoble va lui ajouter un niveau de difficulté supplémentaire en plongeant dans l'univers du *big data* et de ses documents numériques :

« BOT(S)NIMENTS Choisissez un auteur mort et posez-vous la question de sa seconde vie, celle qui se poursuit et s'écrit sur le web, sur les plate-formes de blog et micro-blogging, sur les réseaux sociaux. [...] Prenons l'exemple d'Anna Kavan. Anna qui ? Anna Kavan, une romancière britannique, née à Cannes en 1901 et décédée à Londres en 1968, dont l'œuvre, culte pour certains, reste néanmoins inconnue du grand public. Une énigmatique héroïnomane qui ouvre plusieurs robinets à la fois : flux autobiographique, flux de l'inconscient et flux onirique. Genre slipstream – fiction de l'étrange – bien avant que le terme soit inventé par Bruce Sterling en 1989. L'écrivain méconnu, au travail obscur, un bon cas d'étude. Penchons-nous sur le cas Kavan. Son oeuvre a-t-elle une deuxième chance d'être repêchée, repérée, éclairée ? Lançons une requête. Environ 1 280 000 résultats 0,27 secondes. Bienvenue dans l'ère du Big Data, le merveilleux volumineux monde des données, véritable levier de croissance pour ceux qui savent les collecter, les analyser, les croiser²⁴. »

- 30 La proposition à la fois créative et théorique de Catherine Lenoble vient ainsi renouveler le genre biographique, mais également celui de l'enquête, en proposant une nouvelle stratégie de recherche, inspirée de la science des données. Là où les archives institutionnelles et les outils de documentation constituent la source par excellence du biographe et du chercheur traditionnel, que devient l'exercice d'écriture biographique lorsqu'il se déplace sur le web, et qu'il est assisté par ordinateur ? Peut-on résumer une vie, percer le secret

d'une œuvre à l'aide de graphiques et de représentations statistiques ?

L'écriture littéraire des données : une poétique de la curation numérique

- 31 La démarche de Catherine Lenoble reste finalement celle d'une collectionneuse, comme la littérature et l'art en général en ont produit depuis longtemps. La captafiction ne fait que poursuivre et parfaire la tendance à l'effet de liste et au fragment que l'on connaît assez bien par ailleurs. Elle encourage l'écrivain à accumuler les matériaux d'écriture, pour mieux produire du sens à travers leur curation : penser, classer, pour reprendre l'expression de Perec. Mais là où de nombreuses collections d'écrivains demeuraient souvent anecdotiques, au sens où elles n'étaient pas portées à la connaissance du public, Catherine Lenoble va transformer son entreprise de collection en travail de publication. L'occasion de questionner l'influence des outils numériques de curation sur le travail de mémoire.
- 32 Ce geste de curation commence avec la réalisation de *Kavan.land*, un *dataset* consacré à Anna Kavan, accessible en ligne. Si l'on me permet cette analogie, *Kavan.land* constitue un exemple assez inédit de *data paper* littéraire et poétique :

« *Kavan.land* est une archive sémantique et fictionnelle développée à partir d'une collection de data autour de la romancière britannique Anna Kavan (1901–1968). [...] Archive mineure élaborée sur plusieurs années, *Kavan.land* interroge, à l'ère du web-folklore et du Big Data, l'héritage et la valeur littéraire de l'œuvre de Kavan, présenté sous la forme de déclarations textuelles et interprétations visuelles (graphes, requêtes, prose sémantique, collection indexée de tags et d'URL²⁵). »

- 33 Les données proviennent de sources électroniques existantes glanées sur la toile, organisées manuellement ou automatiquement, faisant ainsi émerger un territoire curieux, entre fiction et documents, web des données et littérature. Le travail éditorial de HXX donne à voir

cette hétérogénéité du corpus déjà passé par la médiation des outils de recherche numériques, à commencer par Google, rebaptisé Papa par l'autrice.

Figure 2 : Anna K, page 37, capture d'écran (pleine page) reproduisant la première page des résultats d'une recherche Google « Anna Kavan ». Reproduit avec l'aimable autorisation de Catherine Lenoble et des éditions HYX.



34 Bien évidemment, Lenoble se montre critique vis-à-vis des grandes plateformes et des outils standardisés du web, dont elle note à plusieurs reprises l'influence structurante, y compris sur le travail de curation lui-même :

« Sous couvert de son panda diplomatique et de son libre pingouin, Papa fait le tri. Zen. Ying et Yang. Chapeaux blancs et chapeaux noirs. Le bien, le mal. Il en prend grand soin, en les nettoyant, cinq à six cents fois, par an. Ce qui n'est pas sans incidence – les algorithmes, le

grand ménage de Papa – sur notre manière de voir, trier, accéder, consulter lire et au bout de la chaîne : transmettre²⁶. »

« Au point d'écrire une nouvelle histoire, d'inventer de nouvelles publications, de confondre les individus : le portrait en mosaïque d'Anna Kavan sur Google image fait ainsi apparaître l'actrice Helen Woods (confondue par la « maladroite ré-écriture algorithmique » avec un autre nom d'Anna Kavan : Helen Wood)²⁷. »

- 35 Ainsi l'outil numérique apparaît-il à son tour comme un générateur d'identité narrative, en ce qu'il participe à l'éditorialisation de l'entité Anna K, à titre posthume, ajoutant des données aux données, modifiant leur indexation, influençant les algorithmes de classement. Cette dimension performative de l'acte de curation apparaît par ailleurs dès les premières pages de l'enquête, avec cette commande quasi démiurgique reproduite en toutes lettres, geste de création du dossier ANNA K où l'archive prendra forme :

« cd ~documents/WRITE/ANNA K.²⁸ »

- 36 Avec son ambition exhaustive, sa masse prolifique d'informations, *Kavan.land* offrirait-il la matière pour nous permettre de percer le mystère Anna Kavan ? Une partie du travail de Lenoble semble y parvenir. Force est de constater ici que l'exercice fonctionne très bien tant que l'ambition graphique demeure limitée : avec des formes de visualisation plutôt basiques (nuages, dispersion plots), le *dataset* ANNA K se laisse plutôt simplement lire et analyser.

Figure 3 : Dispersion Plot

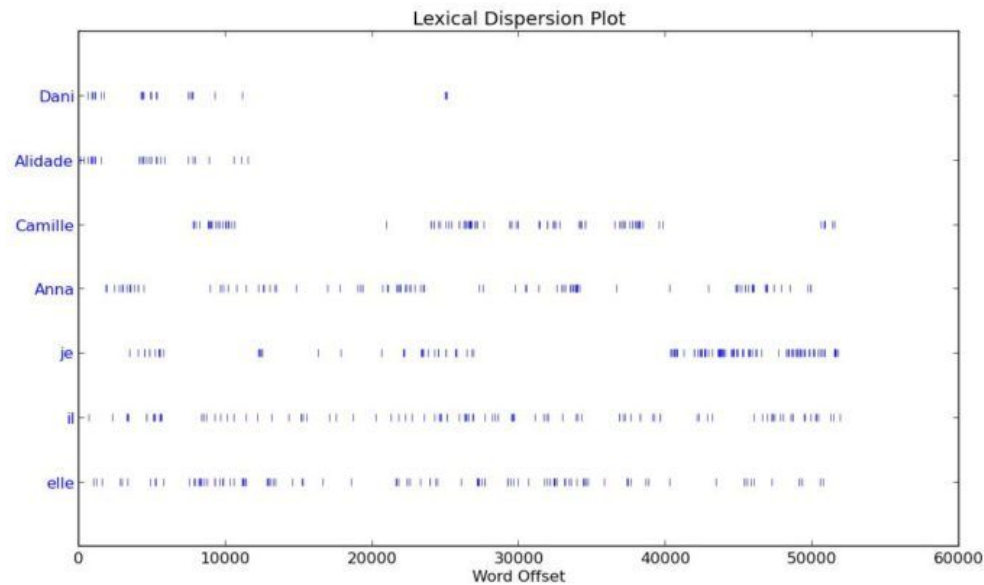
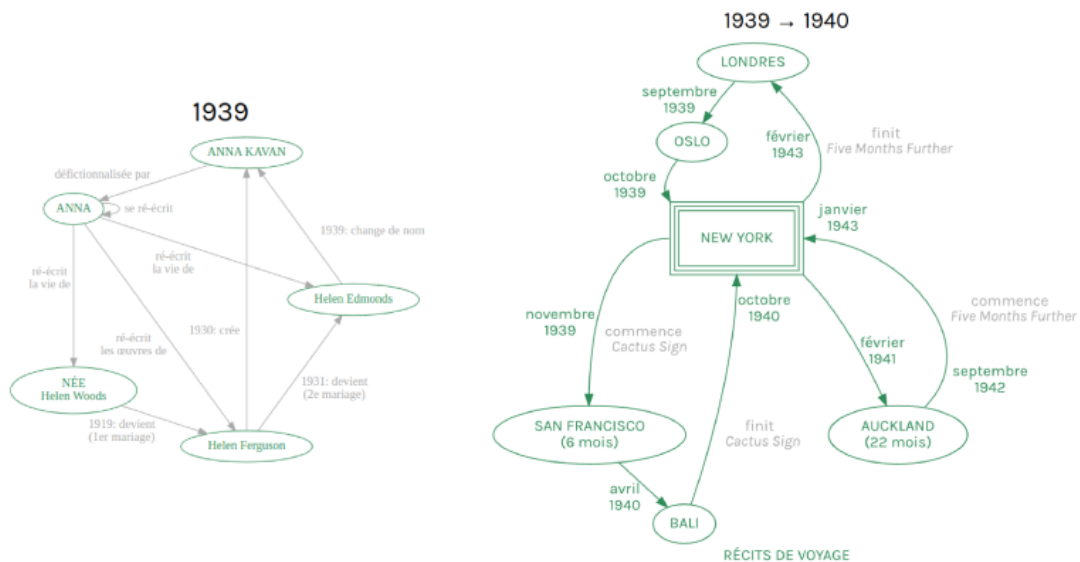


Figure 4 : Graphes



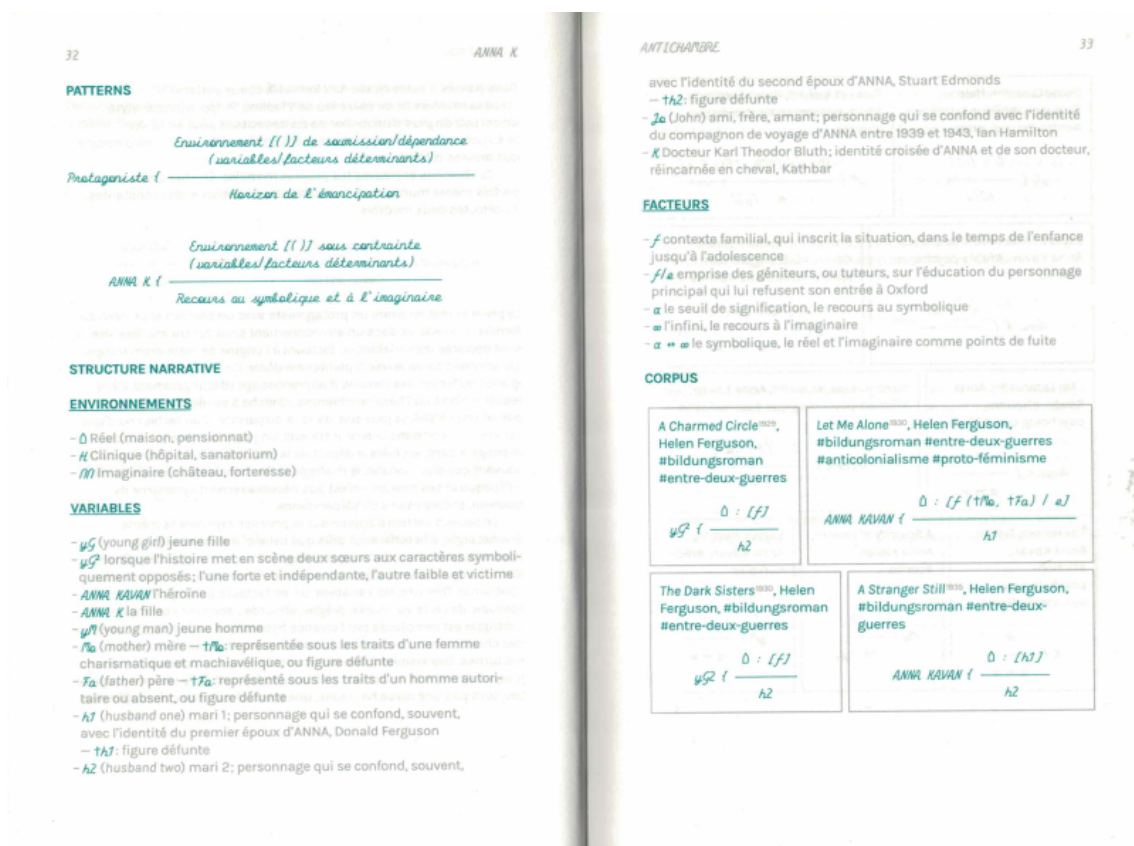
Figures 3 & 4 : Visualisations analysant les données recueillies par Catherine Lenoble sur Anna Kavan, reproduit sur l'archive numérique kavan.land. Reproduit avec l'aimable autorisation de Catherine Lenoble et des éditions HYX.

37 Mais les choses se compliquent assez vite. Obnubilée par la recherche de *patterns*, l'autrice réalise un important travail de modélisation, plutôt convaincant il faut l'avouer, mais qui demande au

lecteur une attention déjà plus soutenue. Le récit se délinéarise en effet pour présenter une série de formules, qui requiert à la fois une connaissance succincte des travaux des formalistes russes (pour bien saisir la référence à Vladimir Propp et ses travaux en analyse structurale), ainsi qu'une certaine littératie numérique loin d'être partagée par l'usager lambda (afin de maîtriser, par exemple, le concept de variable) :

« À la manière de ce vieux fou de Vladimir Propp [...], tu te mets en tête d'étudier la morphologie des œuvres d'ANNA. Tu cherches et tu trouves tes propres formules. En lisant, annotant, parfois même murmurant, tu repères des variables et des constantes. Tu détectes deux modèles²⁹. »

Figure 5 : Anna K, p. 32. Reproduit avec l'aimable autorisation de Catherine Lenoble et des éditions HX.

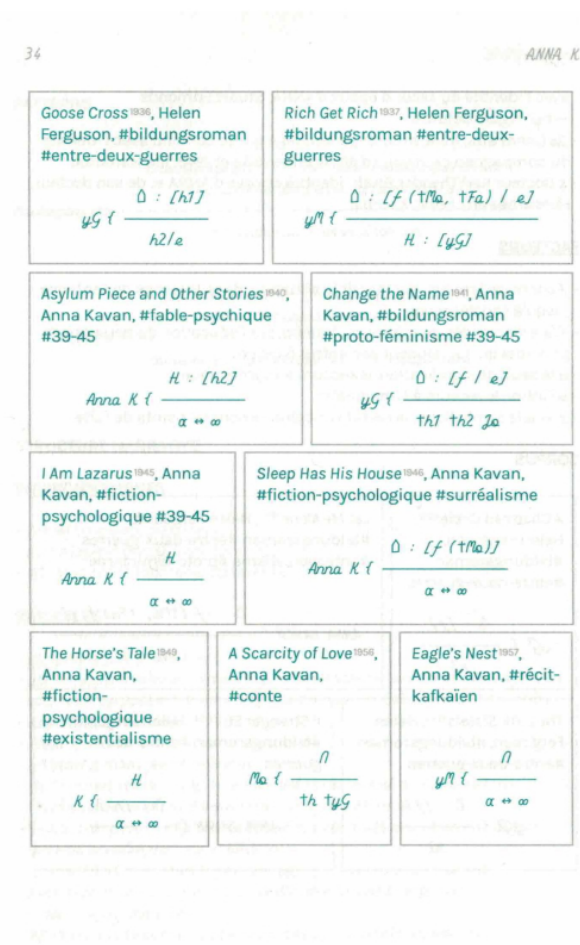


« Le premier met en avant un protagoniste avec un prénom et un nom de famille ; il ou elle vit dans un environnement sous contrainte. [...] L'émancipation chez ANNA se joue soit, dans la disparition d'un

facteur ou d'une variable (+), soit dans la rencontre avec un jeune homme (h1, h2)³⁰. »

38 S'ensuit une série de formules modélisant chacune des publications d'Anna Kavan, selon un système d'encodage établi et défini plus tôt par Catherine Lenoble.

Figure 6 : Anna K, p. 34. Reproduit avec l'aimable autorisation de Catherine Lenoble et des éditions HXX.



39 Le tour de force de l'autrice est alors de proposer au lecteur une mise en données d'un corpus littéraire dont elle donne à voir et à lire le fonctionnement, mais également les potentialités interprétatives. Le lecteur est ainsi poussé à développer une nouvelle compétence de lecture délinéarisée, inspirée de l'algorithmique, où se dessine moins un récit actualisé que l'ensemble de ses potentialités. Cette lecture fait écho à l'herméneutique modélisante, définie par Johanna

Drucker, mais également à l'herméneutique des données que Bachimont et Vaerlet appellent de leurs vœux.

- 40 C'est ainsi que Lenoble construit une poétique de la curation. Comme le sous-entend le concept d'éditorialisation, mais comme l'a défendu également Kenneth Goldsmith³¹, tout travail de création passe désormais non plus par la recherche de l'inédit, mais par un travail de curation et d'association des données. Le travail de curation s'apparente ainsi à une démarche qualitative qui cherche à répondre à l'enjeu quantitatif des données.

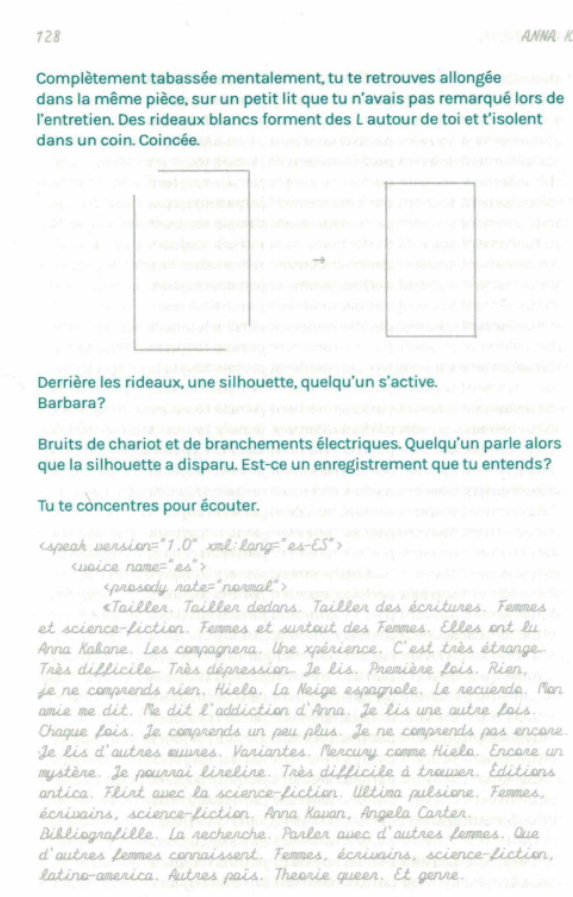
Anna K : folle à lier ?

- 41 En un sens, cette poétique de la curation s'apparente à un art de créer du lien, comme en témoignent d'ailleurs les nombreux graphes produits par Lenoble et le collectif Open Source Publishing. Mais en se projetant ainsi dans la position de l'enquêteur chargé de faire émerger des relations, et donc du sens, l'écrivain de captafiction ne prend-il pas quelques risques ? Posons la question plus franchement : les données font-elles toujours sens, ou n'y projette-t-on pas ce que l'on veut y voir ?
- 42 Ultime chapitre de l'ouvrage, la partie « data mining » clôt l'enquête initiée par Catherine Lenoble de manière assez inattendue, dans un revirement où l'autrice semble éprouver les limites de l'exercice de visualisation de données – tant du point de vue conceptuel que technique, mais également herméneutique. Dans ce chapitre qui transporte l'enquêtrice dans les méandres de son archive, mais également dans un asile psychiatrique, l'écriture littéraire des données apparaît comme un exercice à haut risque :

« Tu tournes en boucle autour des mêmes événements, des mêmes fragments, des mêmes refrains, des mêmes revenants. Ce qui t'intéresse chez ANNA K : le pattern familial et marital, le lit et la chambre, l'entrée et la sortie de la clinique, les voitures et son passeport, la période entre 39 et 35 [...]. Et préciser, qu'entre temps, tu viens d'être admise au service des addictions comportementales³². »

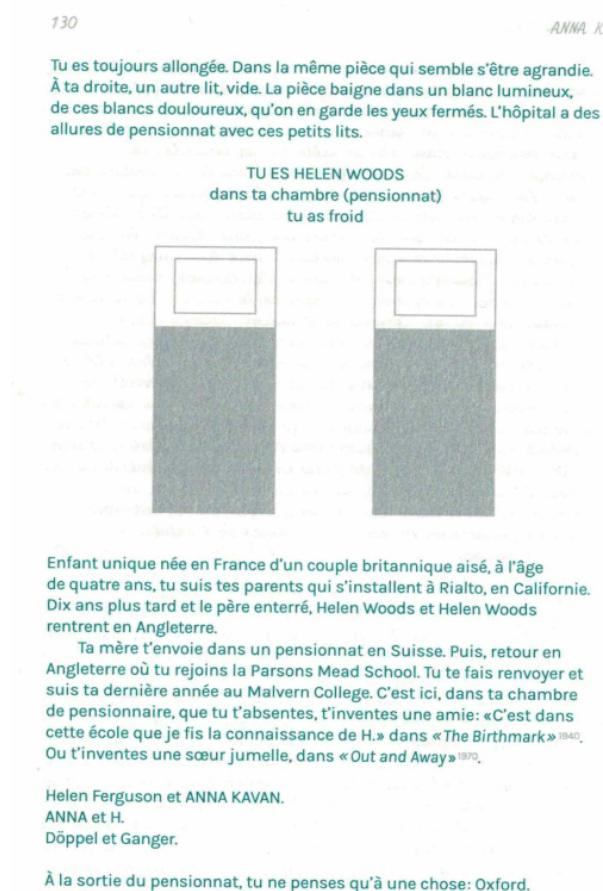
- 43 L'exercice biographique glisse vers une narration intime, à la seconde personne du singulier, dans lequel la narratrice et ANNA K ont plus ou moins fini par ne faire qu'une.
- 44 Ce dernier chapitre comprend également de nombreuses propositions graphiques, ou du moins un jeu avec des formes de la visualisation graphique, cependant totalement décorrélée des données.
- 45 Car il ne s'agit plus ici de proposer une visualisation à partir des données préexistantes, mais à créer de nouvelles formes, tout en jouant sur les codes de la visualisation ou de modélisation UML, dont les principes sont détournés pour générer une interprétation plus figurative, de sorte que les dessins formés se laissent interpréter de plusieurs façons possibles : superposition entre une forme statistique ou modélisante, et une forme figurative.

Figure 7 : Anna K, p. 128. Reproduit avec l'aimable autorisation de Catherine Lenoble et des éditions HX.



- 46 Dans le cas ci-dessus, par exemple, les éléments graphiques de relation (les flèches, pour le dire simplement) figurent les rideaux derrière lesquels la narratrice est cachée, mais va également symboliser l'enfermement dans la folie.

Figure 8 : Anna K, p. 130. Reproduit avec l'aimable autorisation de Catherine Lenoble et des éditions HXX.



- 47 Dans ce nouvel exemple, c'est un simple graphique en bâtons, avec un design 3D, qui devient soudainement un dessin de lit.
- 48 L'ambiguïté graphique avec laquelle joue Catherine Lenoble rappelle finalement le concept d'interprétation modélisante de Johanna Drucker, selon laquelle notre « compréhension des formes de connaissance n'est qu'une connaissance des formes³³ ». Ainsi,

« L'interprétation modélisante appelle aussi des formes graphiques différentes, distinctes, capables d'exprimer l'ambiguïté, la

contradiction, la nuance, le changement et d'autres aspects de considération critique³⁴. »

- 49 En s'extrayant de la visualisation des données pour créer un dessin plutôt figuratif, qui va lui-même ouvrir à une interprétation symbolique, Catherine Lenoble ouvre alors la voie à une « compréhension non représentationnelle de la visualisation », même si ces derniers graphs n'ont pas été générés avec des données. Car il ne s'agit plus tant de produire ici une image des données qu'une image de leur visualisation.
- 50 Dans une métaphore finale associant le *data mining* à la trépanation crânienne, Catherine Lenoble laisse son lecteur avec une sensation ambivalente, comme si toutes ces données, malgré l'impression ou le fantasme d'exhaustivité et de rationalisation qu'elles supposent, ne pouvaient empêcher Anna K de nous échapper finalement :

« ANNA se tourne et se retourne dans son lit, marmonnant : "les rats... les voilà à ma porte". Recroquevillée du côté de la table de chevet, elle lorgne sur la lettre dans laquelle le président du Comité lui fait savoir tout personnellement que "les doigts de la propriété intellectuelle et artistique nous ont permis d'analyser votre œuvre, grâce à la mémoire du *fair use*, invoquée avec finesse et maîtrise, et ce malgré les difficultés de la tâche, au regard du sujet et de l'ampleur de la variété des secteurs d'activité. Comme nous l'avons relevé précédemment, le détournement d'outils informatiques et d'algorithmes ayant conduit à l'utilisation de terminologies variables, il convient de sauver les apparences, par un acte d'extraction provisoire, entièrement effectué dans le crâne."³⁵ »

« Tandis que tous s'auto-congratulent pour le partage des miettes numériques, leur attention converge vers une horloge, au gong dépressif, qui semble négocier avec sa propre maladie. À y regarder de plus près, l'horloge indique la mort d'ANNA K. Une énième fois³⁶. »

Conclusion

- 51 Dans *Anna K*, Catherine Lenoble questionne la capacité du monde à rester interrogeable après sa mise en données dans l'espace

numérique. Et tandis que Françoise Lavocat, dans ses travaux sur la distinction entre factuel et fictif cités en introduction, signalait la nécessité cognitive de conserver un régime de fiction, Lenoble pointe le grand défi cognitif de notre époque, incarné par le *big data*, un ensemble de données si vaste qu'il dépasse justement les capacités du cerveau humain – y compris sa capacité à faire fiction ? En réponse à ce défi, elle invente une nouvelle forme d'écriture, ce que j'ai proposé d'appeler la *captafiction*, qui combine deux régimes d'écriture aux méthodes sans doute assez voisines : la science des données d'une part, l'écriture de l'enquête d'autre part. Avec la première, Catherine Lenoble prend, comme on l'a vu, des libertés poétiques et (auto)fictionnelles, qui mettent à mal le faux certificat d'authenticité de la donnée ainsi que la fonction objectivante de ses visualisations, parvenant ainsi à proposer à ses lecteurs un parcours herméneutique exigeant, mais efficace. À la seconde, Catherine Lenoble apporte de nouvelles méthodologies, fondées sur des outils numériques qui, comme elle le démontre, se sont imposés comme de nouvelles technologies de l'intellect. Sans doute n'avons-nous pas encore terminé l'inventaire des mutations conceptuelles posées par ces nouveaux médias de la connaissance. En attendant, la *captafiction* apparaît comme une piste prometteuse pour participer à la reconstruction d'un régime du sens en contexte numérique.

BIBLIOGRAPHIE

BACHIMONT Bruno, « Donner du sens aux données : les ruses du numérique : Les disciplines du document face à la métis du calcul », *Interfaces numériques* [en ligne], vol. 11, n° 2, 2022, URL : <https://www.unilim.fr/interfaces-numeriques/4838> [consulté le 8 avril 2024].

BACHIMONT Bruno, *Le Sens de la technique : Le numérique et le calcul*, Paris, Les Belles lettres, coll., « Encre marine », 2010.

BACHIMONT Bruno, « Le numérique comme support de la connaissance : entre matérialisation et interprétation », in : Gueudet Ghislaine et Trouche Luc (dir.), *Ressources vives. Le travail documentaire des professeurs en mathématiques* [en ligne], Rennes, Lyon, PUR et INRP, 2010, p. 75-90, URL : <http://hal.archives-ouvertes.fr/hal-00496590> [consulté le 8 avril 2024].

BACHIMONT Bruno, Verlaet Lise, « Traces, données et preuves en contexte numérique : quelles acceptions interdisciplinaires ? », *Intelligibilité du numérique* [en ligne], 2021,

URL : <https://intelligibilite-numerique.numerev.com/numeros/n-2-2021/2625-traces-donnees-et-preuves-en-contexte-numerique-queelles-acceptations-interdisciplinaires> [consulté le 8 avril 2024].

BONNET Gilles, *Pour une poétique numérique : Littérature et Internet*, Paris, Hermann, 2017.

BOOTZ Philippe, « La littérature déplacée », *Formules : Revue des littératures à contraintes*, n° 10, 2006.

BORGMAN Christine L., SCHARNHORST Andrea, GOLSHAN Milena S., « Digital Data Archives as Knowledge Infrastructures: Mediating Data Sharing and Reuse », *arXiv* [en ligne], 2018, URL : <http://arxiv.org/abs/1802.02689> [consulté le 8 avril 2024].

BOUCHARDON Serge, *La valeur heuristique de la littérature numérique*, Paris, Hermann, 2014.

DRUCKER Johanna, *Visualisation : L'interprétation modélisante*, Paris, 1^{re} édition, B42, 2020.

DRUCKER Johanna, *Graphesis: visual forms of knowledge production*, Cambridge, Massachusetts, Harvard University Press, coll. « MetaLABprojects », 2014.

FLORIDI Luciano, *The 4th revolution: how the infosphere is reshaping human reality*, New York/Oxford, 1^{re} édition, Oxford University Press, 2014.

GIANNACHI Gabriella, *Archive Everything: Mapping the Everyday*, Cambridge, Londres, The MIT Press, 2016.

GOLDMAN Judith, « Re-thinking "Non-Retinal Literature": Citation, "Radical Mimesis", and Phenomenologies of Reading in Conceptual Writing », *Postmodern Culture* [en ligne], vol. 22, n° 1, 2011, DOI : [10.1353/pmc.2012.0007](https://doi.org/10.1353/pmc.2012.0007).

GOLDSMITH Kenneth, *Uncreative Writing: Managing Language in the Digital Age*, New York, Columbia University Press, 2011.

GUILET Anaïs, LIMARE Sophie, GIRARD Annick, *Tous artistes ! Les pratiques (re)créatives du web*, Montréal, Presses de l'université de Montréal, coll. « Parcours numériques », 2017.

LARRUE Jean-Marc, VITALI-ROSATI Marcello, *Media do not exist : performativity and mediating conjunctures* [en ligne], Institute of Network Cultures, 2019, URL : <https://hdl.handle.net/1866/22937> [consulté le 8 avril 2024].

LATOUR Bruno, « Pensée retenue, pensée distribuée », *Lieux de savoir. 1. Espaces et communautés*, vol. 1, Paris, Albin Michel, 2007, p. 605-615.

LAVOCAT Françoise, *Faits et fictions. Pour une frontière*, Paris, Seuil, 2016.

LENOBLE Catherine, Anna K, Paris, HYX, coll. « Graphes », 2016.

LENOBLE Catherine, « About », *Kavan.land* [en ligne], URL : <http://kavan.land/> [consulté le 8 avril 2024].

MARINO Mark C., *Critical Code Studies*, Cambridge, Massachusetts, The MIT Press, 2020.

MERZEAU Louise, « L'intelligence des traces », *Intellectica* [en ligne], n° 59, 2013, URL : <https://halshs.archives-ouvertes.fr/halshs-01071211> [consulté le 8 avril 2024].

MERZEAU Louise, « Faire mémoire des traces numériques », *Inathèque* [en ligne], 2012, URL : <https://shs.hal.science/halshs-00727308> [consulté le 8 avril 2024].

MONJOUR Servanne, VITALI-ROSATI Marcello, WORMSER Gérard, « Le fait littéraire au temps du numérique. Pour une ontologie de l'imaginaire », *Sens Public* [en ligne], 2016, URL : <https://hdl.handle.net/1866/16362> [consulté le 8 avril 2024].

SAEMMER Alexandra, « La littérature informatique, un art du dispositif », *Connaître et valoriser la création littéraire numérique en bibliothèque* [en ligne], Villeurbanne, Presses de l'Enssib, coll. « La Boîte à outils », 2019, p. 28-39, DOI : [10.4000/books.pressesenssib.10141](https://doi.org/10.4000/books.pressesenssib.10141).

TUFTE Edward R., *Envisioning Information*, Cheshire, Conn, Graphics Pr, 1990.

VIART Dominique, « Les Littératures de terrain », *Revue critique de fixxion française contemporaine* [en ligne], n° 18, 2019, p. 1-13., DOI : [10.4000/fixxion.1275](https://doi.org/10.4000/fixxion.1275).

VITALI-ROSATI Marcello, *On Editorialization: Structuring Space and Authority in the Digital Age* [en ligne], Amsterdam, Institute of Network Cultures, coll. « Theory on demand » 26, 2018, URL : <http://networkcultures.org/blog/publication/tod-26-on-editorialization-structuring-space-and-authority-in-the-digital-age/> [consulté le 8 avril 2024].

VITALI-ROSATI Marcello, « Digital Architectures: the Web, Editorialization and Metaontology », *Azimuth. Philosophical Coordinates in Modern and Contemporary Age* [en ligne], vol. 4, n° 7, 2016, p. 95-111, URL : <https://hdl.handle.net/1866/16067> [consulté le 8 avril 2024].

WAGNER, Frank, « Entretien avec Françoise Lavocat. À propos de *Fait et fiction*. Pour une frontière », Paris, Seuil, « Poétique », 2016 », *Cahiers de Narratologie. Analyse et théorie narratives* [en ligne], n° 31, 2016, DOI : [10.4000/narratologie.7619](https://doi.org/10.4000/narratologie.7619).

NOTES

1 Françoise Lavocat, *Faits et fictions. Pour une frontière*, Paris, Seuil, 2016, p. 12.

2 Luciano Floridi, *The 4th revolution: how the infosphere is reshaping human reality*, 1^{re} édition, New York ; Oxford, Oxford University Press, 2014 ; M. Vitali-Rosati, *On Editorialization: Structuring Space and Authority in the Digital Age* [en ligne], Amsterdam, Institute of Network Cultures, coll. « Theory on demand » 26, 2018, URL : <http://networkcultures.org/blog/publication/tod-26-on-editorialization-structuring-space-and-authority-in-the-digital-age/>

[blication/tod-26-on-editorialization-structuring-space-and-authority-in-the-digital-age/](#) [consulté le 8 avril 2024].

3 Le concept d'éditorialisation est tout d'abord apparu dans le champ des SIC en France au début des années 2000, avant d'être développé dans les études littéraires grâce, notamment, aux travaux de l'équipe de la Chaire de recherche du Canada sur les écritures numériques dirigée par Marcello Vitali-Rosati. La théorie de l'éditorialisation implique une conception performative des écritures numériques, lesquelles œuvrent à la production d'un espace qui n'a rien de « virtuel », comme on l'entend souvent dès qu'il est question d'objets numériques, mais qui s'est désormais imposé comme notre espace *actuel*. Ainsi, suivant ce principe d'éditorialisation, la publication d'un contenu numérique participe à l'agencement de notre monde, y compris si ce contenu revêt un caractère littéraire et fictionnel, du moment où il vient s'ajouter à la masse des données qui composent et structurent l'espace (que celui-ci soit dit « numérique » ou non n'a alors plus d'importance).

4 Servanne Monjour, Marcello Vitali-rosati, Gérard Wormser, « Le fait littéraire au temps du numérique. Pour une ontologie de l'imaginaire », *Sens Public* [en ligne], 2016, URL : <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/16362> [consulté le 8 avril 2024].

5 Bruno Bachimont, « Donner du sens aux données : les ruses du numérique. Les disciplines du document face à la métis du calcul » [en ligne], *Interfaces numériques*, vol. 11, n° 2, 2022, URL : <https://www.unilim.fr/interfaces-numeriques/4838> [consulté le 8 avril 2024].

6 Catherine Lanoble, Anna K, Paris, HYX, coll. « Graphes », 2016.

7 Bruno Latour, « Pensée retenue, pensée distribuée », in : *Lieux de savoir*. 1. *Espaces et communautés*, vol. 1, Paris, Albin Michel, 2007.

8 Johanna Drucker, *Visualisation : L'interprétation modélisante*, 1^{re} édition, Paris, B42, 2020.

9 Bruno Bachimont, « Donner du sens aux données : les ruses du numérique : Les disciplines du document face à la métis du calcul » [en ligne], *Interfaces numériques*, op. cit.

10 Loc. cit.

11 Bruno Bachimont, Lise Verlaet, « Traces, données et preuves en contexte numérique : quelles acceptions interdisciplinaires ? »,

ue.numerev.com/numeros/n-2-2021/2625-traces-donnees-et-preuves-en-contexte-numerique-queelles-acceptions-interdisciplinaires [consulté le 8 avril 2023].

12 Johanna Drucker, *Graphesis: visual forms of knowledge production*, Cambridge, Massachusetts, Harvard University Press, coll. « MetaLABprojects », 2014 ; Johanna Drucker, *Visualisation: L'interprétation modélisante*, op. cit.

13 Johanna Drucker, *Visualisation : L'interprétation modélisante*, op. cit.

14 Je choisis ici délibérément de substituer à la notion d'« œuvre » celle de « projet » ou d'« approche par projet », emprunté au vocabulaire de l'art, notamment de l'art contemporain. La notion de « projet » met en avant une idée, un concept à atteindre : elle dénote un certain degré d'inachèvement (voire d'inachevable), souligne le travail préparatoire et les moyens nécessaires à l'élaboration d'un produit « final » (même si l'enjeu de cette finalité, aujourd'hui, est discutable), et bien souvent ouvert à la contribution, à la discussion. Elle déplace l'intérêt sur le processus créatif plutôt que sur un artefact final. Car c'est bien le projet qui fait œuvre, et transforme au passage le sens accordé à cette dernière – en particulier en bousculant ses enjeux de stabilité et de clôture : la publication numérique a permis aux écrivains de développer une poétique du chantier, où le lecteur assiste à un nouveau type de performance, laquelle fait partie de l'expérience esthétique. L'ouvrage *Anna K*, en ce sens, est l'un des produits du projet *Anna K*.

15 Mark C. Marino, *Critical Code Studies*, The MIT Press, 2020.

16 Anaïs Guilet, Sophie Limare, Annick Girard, *Tous artistes ! Les pratiques (re)créatives du web*, Montréal, Presses Université de Montréal, coll. « Parcours numériques », 2017 ; Alexandra Saemmer, « La littérature informatique, un art du dispositif », *Connaître et valoriser la création littéraire numérique en bibliothèque* [en ligne], Villeurbanne, Presses de l'Enssib, coll. « La Boîte à outils », 2019, p. 28-39.

17 Voir notamment Gilles Bonnet, *Pour une poétique numérique : Littérature et Internet*, Paris, Hermann, 2017 ; Serge Bouchardon, *La valeur heuristique de la littérature numérique* [en ligne], Paris, Hermann, 2014.

18 Edward R. Tufte, *Envisioning Information*, Cheshire, Conn, Graphics Pr, 1990.

19 Catherine Lenoble, *Anna K*, op. cit., p. 9.

- 20 Dominique Viart, « Les Littératures de terrain », *Revue critique de fixxion française contemporaine* [en ligne], n° 18, 2019, DOI : [10.4000/fixxion.1275](https://doi.org/10.4000/fixxion.1275).
- 21 Laurent Demanze, *Un nouvel âge de l'enquête, Portraits de l'écrivain contemporain en enquêteur*, Paris, José Corti, coll. « Les Essais », 2019.
- 22 Geoffrey Rockwell, Stefan Sinclair, *Hemeneutica*, Cambridge, Londres, The MIT Press, 2016.
- 23 Catherine Lenoble, *Anna K*, op. cit., p. 90.
- 24 *Ibid.*, p. 30-40.
- 25 Catherine Lenoble, « About » [en ligne], *Kavan.land* [en ligne], URL : <http://kavan.land/browse/?node=http://kavan.land/About> [consulté le 8 avril 2024].
- 26 Catherine Lenoble, *Anna K*, op. cit., p. 40.
- 27 *Ibid.*, p. 42.
- 28 *Ibid.*, p. 19.
- 29 *Ibid.*, p. 31.
- 30 *Loc. cit.*
- 31 Kenneth Goldsmith, *Uncreative Writing : Managing Language in the Digital Age*, Columbia University Press, 2011.
- 32 *Ibid.*, p. 121.
- 33 Johanna Drucker, *Visualisation: L'interprétation modélisante*, op. cit., p. 27.
- 34 *Ibid.*, p. 15.
- 35 C. Lenoble, *Anna K*, op. cit., p. 144.
- 36 *Ibid.*, p. 146.

RÉSUMÉS

Français

Cet article a pour objectif de décrire et d'analyser certaines mutations de la fiction lorsqu'elle se confronte à la *réalité* des données, tant sur un plan pratique que théorique. On proposera, à travers le cas d'étude *Anna K* de Catherine Lenoble, le concept de captafiction, un parti-pris conceptuel inspiré des travaux de Johanna Drucker qui pose l'accent sur la problématique ontologique propre aux données numériques. Entre

paradigme représentationnel et performatif, la notion de captafiction pointe du doigt le principe de remédiatisation du réel par les données. Elle permet par ailleurs d'ouvrir la réflexion sur la dimension graphique de l'écriture des données, en proposant une réflexion critique sur la visualisation.

English

In this article I describe and analyze some mutations in fiction as it confronts the *reality* of data, on both a practical and theoretical level. Through Catherine Lenoble's work *Anna K*, I propose the concept of captafiction, a conceptual approach inspired by the work of Johanna Drucker, which focuses on the ontological issues specific to digital data. Between representational and performative paradigms, the notion of captafiction underlines the remediation of reality by data. It also opens the way to thinking about the graphic dimension of data writing, by proposing a critical reflection on visualization.

INDEX

Mots-clés

non-fiction/fiction, biographie, données, capta, data visualisation, graphe.

Keywords

nonfiction/fiction, biography, data, capta, datavisualization, graph.

AUTEUR

Servanne Monjour

Maîtresse de conférences en humanités numériques littéraires, Sorbonne Université/CELLF

IDREF : <https://www.idref.fr/192359142>

ORCID : <http://orcid.org/0000-0003-1067-2145>

HAL : <https://cv.archives-ouvertes.fr/servanne-monjour>

ISNI : <http://www.isni.org/000000047537455X>

BNF : <https://data.bnf.fr/fr/17807156>

Cléo de 7 à 9 : de la fiction au document à la fiction

Cléo de 7 à 9: from fiction to document to fiction

Stéphane Lambion

DOI : 10.35562/marge.844

Droits d'auteur

CC BY-NC-SA

TEXTE

- 1 Cléo de 7 à 9 est une vidéo qui reprend à l'identique la performance dont elle est issue, réalisée le 3 juin 2022 au Centre national de la danse (Pantin). C'est un travail qui, sous l'angle du champ médical, interroge d'une part l'usage du document dans le processus d'écriture poétique et d'autre part, le rapport entre document et fiction.
- 2 Je suis parti de la fin du film d'Agnès Varda, *Cléo de 5 à 7*, où l'on annonce à la protagoniste qu'elle a un cancer : à partir de cette scène, j'ai construit un texte qui imagine un développement possible de sa maladie en m'appuyant sur des documents majoritairement trouvés sur internet – je définis ici le document par l'intermédiaire du concept de collage tel qu'Olivier Quintyn le formalise dans *Dispositifs/Dislocations* : relève du collage « toute pratique faisant intervenir une hétérogénéité formelle des constituants et jouant de l'introduction, dans le champ artistique, d'objets ou de fragments non spécifiquement perçus comme artistiques¹ ». Les documents utilisés dans cette vidéo sont des témoignages vidéo, des clips médicaux, des citations, des données statistiques, etc.
- 3 Le but de cette vidéo est donc double. Du point de vue de la forme, il s'agit d'interroger le degré d'inclusion possible du document dans la création poétique, en reprenant le curseur que propose Olivier Quintyn qui va du *ready-made* (insertion d'un matériau brut, non modifié) à son opposé (une inclusion totalement homogène du matériau importé où ce dernier se reconnaît à peine). Du point de vue du contenu, il s'agit de montrer une façon possible pour la création poétique de traiter les données médicales en partant d'un langage

objectif, réel, pour plonger dans la subjectivité fictionnelle d'un personnage – ou plutôt, pour reprendre le titre de l'intervention : il s'agit dans un premier temps de partir d'une fiction, celle d'Agnès Varda, pour en faire dans un deuxième temps le point de départ d'un travail sur le document qui, dans un troisième temps, nourrit une nouvelle fiction, formellement différente (puisque textuelle, poétique, et non filmique, bien qu'elle s'inscrive dans le cadre d'un travail vidéo).

Le media demandé n'existe pas. : [video:cleo]

NOTES

- 1 Olivier Quintyn, *Dispositifs/Dislocations*, Paris, Al Dante, 2007, p. 20.

RÉSUMÉS

Français

Cette vidéo, issue d'un travail de recherche-crédation, interroge d'une part l'usage du document dans le processus d'écriture et, d'autre part, le rapport entre document et fiction, le tout sous l'angle médical. La vidéo reprend la fin du film d'Agnès Varda *Cléo de 5 à 7*, où l'on annonce à la protagoniste qu'elle a un cancer : à partir de cette scène se construit un texte qui imagine la maladie de l'héroïne en s'appuyant sur divers documents réels trouvés sur internet (témoignages vidéo, clips médicaux, citations, données statistiques, etc.). Ce travail propose ainsi une façon d'appréhender la masse de données médicales disponible en ligne et montre comment la création poétique peut s'appuyer sur le langage technique de données réelles pour plonger dans la subjectivité fictionnelle d'un personnage.

English

This video, resulting from creative research work, interrogates on the one hand the use of the document in the writing process and on the other hand the relationship between document and fiction, all from a medical perspective. The video takes up the end of Agnès Varda's film *Cléo de 5 à 7*, where the protagonist is told that she has cancer: from this scene a text is constructed which imagines the heroin's illness based on various real documents found on the internet (video testimonies, medical clips, quotes, statistical data, etc.). This work thus suggests a way of approaching the mass of medical data available online and it shows how poetic creation can rely on the technical language of real data to delve into the fictional subjectivity of a character.

INDEX

Mots-clés

vidéo, recherche-cr ation, donn es m dicales, fiction, maladie, po tique

Keywords

video, creative research, medical data, fiction, illness, poetics

AUTEUR

St phane Lambion

Aix-Marseille Universit , CIELAM, Aix-en-Provence, France

IDREF : <https://www.idref.fr/267602081>

ISNI : <http://www.isni.org/0000000476441750>

BNF : <https://data.bnf.fr/fr/17814218>

Données/récits

Bases de données et lignes brisées

À l'intersection du narratif et du numérique

Databases and broken lines: at the intersection of narrative and digital

Bertrand Gervais

DOI : 10.35562/marge.907

Droits d'auteur

CC BY-NC-SA

PLAN

1. Le mode d'acquisition des données
 2. Le type de récit généré
 3. Le mode d'actualisation
- Conclusion

TEXTE

- 1 Depuis les explorations littéraires de Jorge Luis Borges, puis d'un ensemble de plus en plus grand d'auteurs et de mouvements littéraires, depuis le Nouveau Roman français et le boom latino-américain jusqu'aux multiples variations du postmodernisme, les formes narratives se sont complexifiées, multipliant les strates narratives, les jeux métafictionnels et métaleptiques, les brouillages entre la fiction et le réel, les stratégies énonciatives et identitaires¹, etc. Ces explorations ont cherché bien souvent à s'affranchir de la ligne, à se libérer de cette contrainte qui est restée tout ce temps la condition fondamentale de toute mise en intrigue². Réciter, raconter, expliquer, c'est développer un discours qui se déploie sur le mode de la concaténation des éléments.
- 2 C'est la nature même du plan syntagmatique d'imposer une logique de la présence et du défilement sur une ligne à ses composants, comparativement au plan paradigmatique, qui détermine des possibilités ou des potentialités³. Le premier plan est représenté par une ligne horizontale, le second par une ou des colonnes qui se déploient verticalement à l'intersection de la ligne. Une phrase, par

exemple, se présente comme un syntagme complexe, c'est-à-dire un ensemble de mots organisés les uns à la suite des autres en fonction de règles qui déterminent des positions et des fonctions (sujet, verbe, complément). Or, pour chacune de ces positions ou fonctions, de multiples choix sont possibles. Le sujet peut être actualisé par un pronom, un nom, un mot, un groupe de mots, etc. Le plan paradigmatique permet d'ailleurs d'explicitier ces possibilités : il ne rend pas compte de ce qui est présent, mais de ce qui est potentiel, virtuel. Il se constitue sur le mode de l'absence, de la disjonction plutôt que de la conjonction. Il est cet amont de toute énonciation, de toute actualisation qui, elle, requiert un plan pour se déployer.

- 3 Ce qui vaut pour la phrase vaut aussi pour le récit, qui se présente de la même manière comme un ensemble d'éléments structurés en fonction de principes ou de règles. Le récit se déploie sur une ligne qui est celle du discours. Or, avec le développement des pratiques hypertextuelles et hypermédiatiques, qui reposent pleinement sur des bases de données, on assiste à une renégociation de ces deux plans et de leurs relations, ainsi qu'à une fragilisation de la notion même de récit. Avec l'hypertextualité, les nœuds textuels sont mis en relation avec un ou plusieurs autres nœuds grâce à des hyperliens, ce qui permet à l'internaute de choisir les étapes de sa progression dans un ensemble donné. D'une part, le plan syntagmatique du défilement du récit n'est plus fait d'une simple concaténation d'éléments déterminés, puisque ce plan permet maintenant des disjonctions et, par conséquent, un processus de sélection par une internaute. D'autre part, les étapes offertes à l'internaute doivent avoir déjà été déterminées en amont. Le paradigme n'y est plus une vue de l'esprit, une notion issue de la linguistique permettant d'expliquer théoriquement comment les éléments sont choisis pour être insérés dans une séquence phrastique, mais une réalité informatique, à savoir l'ensemble des éléments versés dans une base de données, organisés selon une taxinomie précise et insérés dans des œuvres à partir d'interfaces.
- 4 L'art de raconter est déstabilisé, parce que les principes fondateurs de toute narration s'y trouvent en partie neutralisés. L'hypertextualité vient briser la ligne sur un plan médiatique. C'est au niveau des dispositifs techniques eux-mêmes que la ligne est rompue, sur un plan médiatique plutôt que sémiotique⁴. Ce n'est pas la ligne du

discours qui est rompue, au niveau des représentations et des stratégies narratives, c'est le plan du discours lui-même qui est segmenté, morcelé et décousu.

- 5 Les conséquences pour les modes de mise en récit sont importantes. Pour le dire en termes simples : c'est une chose de tenter de s'affranchir de la ligne, c'en est une autre d'en être d'ores et déjà libéré. Il ne s'agit plus de tester les limites d'une contrainte, en l'attaquant de toutes parts, tout en restant soumis à sa loi (comme avec les métafictions), mais de chercher à retrouver un des avantages de cette loi, une de ses dispositions essentielles, à savoir la notion de défilement, de « ligne ». Comment en effet procéder à une mise en intrigue, si toute notion d'ordre préalable a été suspendue, puisque l'internaute peut décider d'elle-même l'ordre d'apparition des éléments narratifs ?
- 6 À partir du moment où l'idée d'un agencement prédéterminé n'est plus une donnée première de la pratique, ce sont ses conditions mêmes qui sont minées. Il faut les réinventer, c'est-à-dire retrouver une façon de mettre en récit, en contexte hypermédiatique. De telles stratégies procèdent notamment par une renégociation des liens entre les plans paradigmatique et syntagmatique. La transformation de la notion de virtuel peut servir de révélateur de cette renégociation nécessaire des rapports entre le paradigme et le syntagme. Ce qui auparavant signifiait « ce qui n'est qu'en puissance », « ce qui est à l'état de possibilité » est devenu, avec les développements de l'informatique, « ce qui est produit grâce à l'informatique et se déroule sur le Web⁵ ». On comprend que la notion de « paradigme » est passée du premier virtuel au second, devenant un virtuel réalisé, une potentialité programmée à l'aide d'un code informatique et organisée dans une base de données.
- 7 Cette renégociation des liens entre paradigme et syntagme en contexte numérique est au cœur de l'exploration entreprise ici qui cherche à définir les liens entre bases de données et pratiques narratives. Comment narrativise-t-on une base de données ? Quels en sont les modes d'actualisation dans une structure narrative ? Sa présence nécessaire est-elle tacite ou explicite ? L'amont du récit, en d'autres mots, est-il révélé ou non ? Que la réalité de cet amont ait été transformée de façon importante, on ne peut plus en douter (le

virtuel étant passé d'un simple potentiel à un ensemble structuré de données), mais de quelle manière son agir est-il représenté, voire symboliquement pris en charge ?

- 8 Afin d'explorer les modes d'actualisation narrative des bases de données, j'identifierai trois catégories, qui sont autant de portes d'entrée de cette relation : 1 – le mode d'acquisition des données (qui relève de la dimension paradigmatique de la production) ; 2 – le type de récit généré (c'est la dimension syntagmatique de l'organisation des données choisies) ; 3 – le mode d'actualisation et, par conséquent, de présence et de mise en scène de la base de données à l'œuvre. Ce troisième plan fait le lien entre les deux premiers. Je décrirai maintenant à l'aide d'exemples ces trois catégories, qui me permettront d'offrir un premier portrait de cette relation⁶.

1. Le mode d'acquisition des données

- 9 Le mode d'acquisition et d'assemblage des données identifie le travail en amont de constitution des ressources utilisées dans la production d'une œuvre. Les artistes ont accès aujourd'hui à une multitude de façons d'acquérir des données. Ces sources peuvent être uniques, comme elles peuvent être multiples. Ultimement, elles peuvent même provenir de dispositifs autonomes de captation, tels que des caméras de surveillance ou des outils de géolocalisation.

Modes d'acquisition

Source unique (individu, un collectif)

Sources multiples (*crowdsourcing*)

Données déjà acquises (*found footage*)

Dispositifs autonomes de captation

Asynchrone : traitement manuel/automatisé

Synchrone : détournement de flux

- 10 Le premier cas de figure – le mode d'acquisition à partir d'une source unique – est représenté par *The Whale Hunt*, l'œuvre de Jonathan Harris produite en 2007⁷. Elle prend la forme d'un projet photographique. En mai 2007, Harris passe neuf jours en compagnie d'une famille inuite à Barrow, en Alaska. Il prend part au rituel de la chasse à la baleine, tel qu'il est pratiqué depuis des centaines d'années par les habitants de cette région. Durant l'ensemble de l'expérience, l'artiste prend au moins une photographie toutes les cinq minutes, pour un total de 3214 images. Ce sont ces images et leurs métadonnées qui constituent le cœur de ce projet, un récit photographique à caractère ethnographique, sans prolepse, analepse, ni ellipse, où une activité traditionnelle est présentée dans toute sa complexité et sa lenteur.
- 11 Un autre projet semblable en termes de collecte des données, cette fois à partir d'une source unique collective plutôt qu'individuelle, est représenté par *Gares*, œuvre hypermédiatique de la Traversée, l'atelier québécois de géopoétique, réalisée en collaboration avec l'artiste Sébastien Cliche. Sur une période de huit mois, les membres de la Traversée ont parcouru les différentes gares de l'île de Montréal dans un esprit de flânerie : ils ont décrit les lieux, pris des photos et des vidéos, enregistré l'ambiance sonore à l'aide de capteurs, ou encore créé des œuvres en médias mixtes (aquarelle et crayon). Ces notes de terrain ont ensuite été organisées selon une logique de mots-clés et un fil narratif linéaire, constituant ainsi une base de données que Sébastien Cliche a interprétée sous la forme d'une œuvre interactive.
- 12 *Life in a Day* du producteur Ridley Scott et du réalisateur Kevin Macdonald, en association avec *The National Geographic*, met en jeu des modalités de collecte très différentes des deux premiers

exemples. C'est le second mode, celui du *crowdsourcing*, de l'acquisition à partir de sources multiples. Le 24 juillet 2010, plus de 81 000 internautes – des cinéastes professionnels et amateurs – participent à un projet cinématographique d'envergure : ils captent des images de leur quotidien qu'ils téléversent sur le site de YouTube. Le projet est unique en son genre : faire le portrait du monde le 24 juillet 2010, afin d'archiver un moment de l'histoire. C'est le premier *user generated feature film* (long métrage généré par des utilisateurs), présenté en grande première au festival Sundance, en janvier 2011. L'équipe de production avait estimé qu'environ 5 000 abonnés répondraient à l'appel, mais 15 fois plus de données que prévu ont été reçues, soit 4 500 heures d'images tournées⁸. La quantité monstre de données a posé immédiatement un défi de taille : comment réussir à les mettre en récit ? Comment réunir en une seule journée 185 jours et demi de vidéos et en faire le montage ?

- 13 Le monteur de l'équipe, Joe Walker, s'est vu donner la tâche de tout rassembler. Bien que le projet ait été élaboré à partir du réseau Internet, c'est au moment de la postproduction que la réelle emprise de l'informatique s'est imposée. Ne pouvant visionner l'ensemble du matériel, l'équipe de production de *Life in a Day* a embrigadé plus d'une centaine d'étudiantes et d'étudiants ; ceux-ci ont visionné les vidéos, puis divisé les images tournées en séquences. Ces clips ont été indexés et classés avec des mots clés dans une base de données. Des scripts, des fonctions narratives, des éléments du quotidien ont été documentés et l'équipe a écrit son scénario en respectant la chronologie d'une journée, le passage du temps depuis l'aurore jusqu'à la nuit, et en marquant certains moments par les actions qui y sont généralement associées. On y voit les gens se lever, prendre leur petit-déjeuner, partir faire les courses, vaquer à leurs occupations, interagir avec leurs proches, etc.
- 14 La multitude d'images reçues a requis un système d'organisation, une classification devenue préalable au montage. Le film s'est construit par un échange entre le scénario et la base de données. Ainsi, les images proposées par les internautes ont toutes été téléversées, découpées en séquences, intégrées à une base de données, indexées à l'aide de mots-clefs, réagencées selon des commandes automatisées, montées de manière linéaire et rediffusées sur le réseau⁹.

- 15 Le troisième mode d'acquisition repose sur l'utilisation de données déjà présentes d'une façon ou d'une autre. L'exemple par excellence en est le *found footage*. *The Clock*, l'installation de Christian Marclay de 2010, s'impose par sa complexité et sa démarche exhaustive. Le projet, dont la projection dure 24 heures, enchaîne de façon systématique des plans issus du cinéma mondial qui contiennent tous une indication ou une représentation de l'heure¹⁰. On passe ainsi d'un homme regardant sa montre à des plans d'horloge (dans une banque, une gare ou un restaurant), de réveille-matin, de pendule, d'écran lumineux, etc. On reconnaît parfois les films d'où sont tirés les plans, mais aucune narration ne résiste à cette segmentation systématique. Un même travail de structuration des données qu'avec *Life in a Day* a dû être accompli, à cette différence près que le narratif n'y joue qu'un rôle secondaire, pour ne pas dire minimaliste. Le récit n'est présent que sur le mode de l'allusion. La force du projet tient à sa longueur, une journée entière, et à l'isochronie respectée scrupuleusement entre le temps représenté et le temps de la spectature : à trois heures huit dans la salle de projection, il est trois heures huit à l'écran. L'œuvre ne fait que nous donner l'heure, nous montrant par la même occasion que le cinéma est un art essentiellement préoccupé par le temps...

Figure 1. 3 heures 8, extrait de *The Clock* de Christian Marclay



- 16 Le quatrième mode d'acquisition des données procède par dispositif autonome de captation, comme des caméras de surveillance ou des dispositifs de géolocalisation. Le premier exemple, asynchrone, d'une telle utilisation des caméras de surveillance est le projet de Xu Bing, *Dragonfly Eyes* (2018). Lorsque le réalisateur a vu des enregistrements des caméras de sécurité à la télévision en 2013, il a décidé de se servir de cette imposante masse de vidéos de sécurité et de webcams mis en ligne sur des sites web chinois pour en réaliser un film, et qui plus est une histoire d'amour... Mais comment raconter une rencontre amoureuse, quand les acteurs changent d'un plan à l'autre ? Comment assurer la continuité narrative avec des images réunies qui s'inscrivent dans une hétérogénéité de formats et de contenus ? Xu Bing le fait en structurant son *found footage* sur la base d'une relation entre deux actants, dont les rôles ne sont tenus par aucun acteur, étant attribués à des êtres de passage qui se glissent dans ces personnages le temps d'une scène ou d'un plan. Nous sommes en présence d'un métarécit plutôt que d'un récit en bonne et due forme, d'un programme narratif partiellement actualisé¹¹.

Figure 2. Extrait de la bande-annonce de *Dragonfly Eyes*, de Xu Bing, 2017.



- 17 Une ultime possibilité de collecte de données est une captation synchrone et relève d'une fluidité maximale de l'information. Il n'y a aucune accumulation en tant que telle, mais la saisie et le

détournement de données, via une interface de programmation (une API), qui pioche dans des bases de données constituées et accessibles. L'œuvre de 2007 de Gregory Chatonsky, *Ceux qui vont mourir*, extrait des textes du site web *Experience Project*, qui permet à chacun de raconter anonymement ses expériences les plus personnelles et « les [illustre] automatiquement selon certains mots-clés par des vidéos autoscopiques trouvées sur YouTube et des autoportraits provenant de Flickr¹² ». Les images ne sont pas conservées (sauf indirectement, par le biais de captures d'écran), elles sont captées, associées à des mots et insérées dans un flux de données. Le récit, qui est un récit de soi dans le cadre des phrases issues du *Experience Project*, est réduit à sa plus petite expression, présent essentiellement en tant qu'hypotexte.

2. Le type de récit généré

- 18 Les exemples présentés jusqu'à présent montrent la diversité des types de récits générés, allant d'un récit sans ellipse, fondé sur une isomorphie assumée entre temps référentiel et temps du récit, à un récit réduit à une simple allusion. Si le récit se déploie sur l'axe syntagmatique de la présence, la ligne en dévient aisément la métaphore par excellence. C'est la ligne du discours, la ligne de l'enchaînement des actions représentées. Cette ligne, elle peut être continue, comme elle peut être brisée, multipliée et, ultimement, effacée¹³.

Types de récit généré

Ligne continue (récit linéaire)

Ligne brisée

Plan sémiotique : récit fragmenté, mise en intrigue

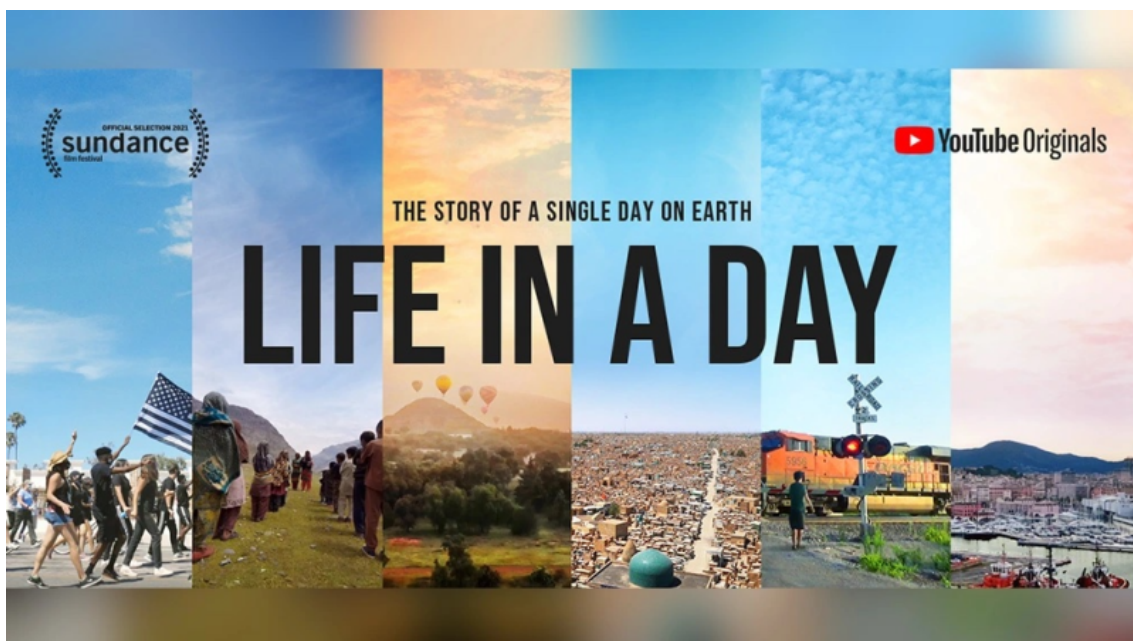
Plan médiatique : hypertextualité, interactivité

Ligne effacée (effet de récit, fiction sans narration)

- 19 La ligne continue, c'est le récit linéaire, tel qu'on le retrouve d'entrée de jeu dans *The Whale Hunt*, de Harris. Les 3 214 photographies sont concaténées, mises bout à bout. Si nous pouvons y circuler librement, segmenter les événements comme on le veut, adopter des points de vue différents, changer notre regard, porter attention à une dimension plutôt qu'une autre, nous ne pouvons pas changer le fil des événements qui reste immuable.
- 20 *Life in a Day*, le film de Macdonald propose aussi un récit linéaire, mais cette fois, en raison de la provenance hétérogène des images concaténées, les actions représentées sont détachées de leur contexte initial, elles sont devenues des types plutôt que des *tokens*, désactualisées pour être transformées en signes d'elles-mêmes. Ce n'est plus un récit, mais un métarécit qui est déployé, un récit qui fait acte de loi, inscrivant une régularité par le biais d'une généralisation produite par un détachement du référent. La stratégie de production du réalisateur était de prélever des échantillons du monde permettant d'en faire le portrait à travers les différentes subjectivités données à voir. On peut considérer les images soumises comme des échantillons du référent-monde qui témoignent d'une expérience spécifique. Bien qu'elles soient fondamentalement empreintes de subjectivité, lorsqu'elles sont sélectionnées et juxtaposées à d'autres échantillons, les images perdent de leur spécificité en devenant conceptuellement des données. En termes peircéens, on assiste en fait à une transformation ontologique : la modification d'un ensemble d'échantillons ou de *tokens*, en des types, des généraux, modification qui repose sur un processus de généralisation qui retire aux existants leurs déterminations spatio-temporelles pour les redéfinir en images types. L'homme qui prend un café dans sa cuisine à l'aube passe d'une situation de vie quotidienne, marquée par son caractère existentiel et nécessairement unique, en l'image type de l'action scriptée usuelle « prendre son café », actualisée par un agent masculin et un contexte, le matin. Il n'y a plus un référent stable, attaché à un *hic* et *nunc*

expérientiel, mais un référent général identifié par une catégorie intégrée dans une taxonomie et une ontologie informatique. Si la mise en récit procède par l'actualisation d'un parcours, quelle qu'en soit la quête, banale ou héroïque, glorieuse ou dérisoire, on comprend que ce film sans véritable agent, sans action majeure ni lieu spécifique – car c'est la terre entière qui est exploitée –, n'est qu'un syntagme faiblement actualisé, un syntagme générique, totalement subordonné à un paradigme devenu prépondérant. *Life in a Day* est une actualisation accidentelle, une parmi d'autres d'un jeu de possibles, à la manière des *Cent mille milliards de poèmes* de Raymond Queneau. C'est une combinatoire à l'œuvre.

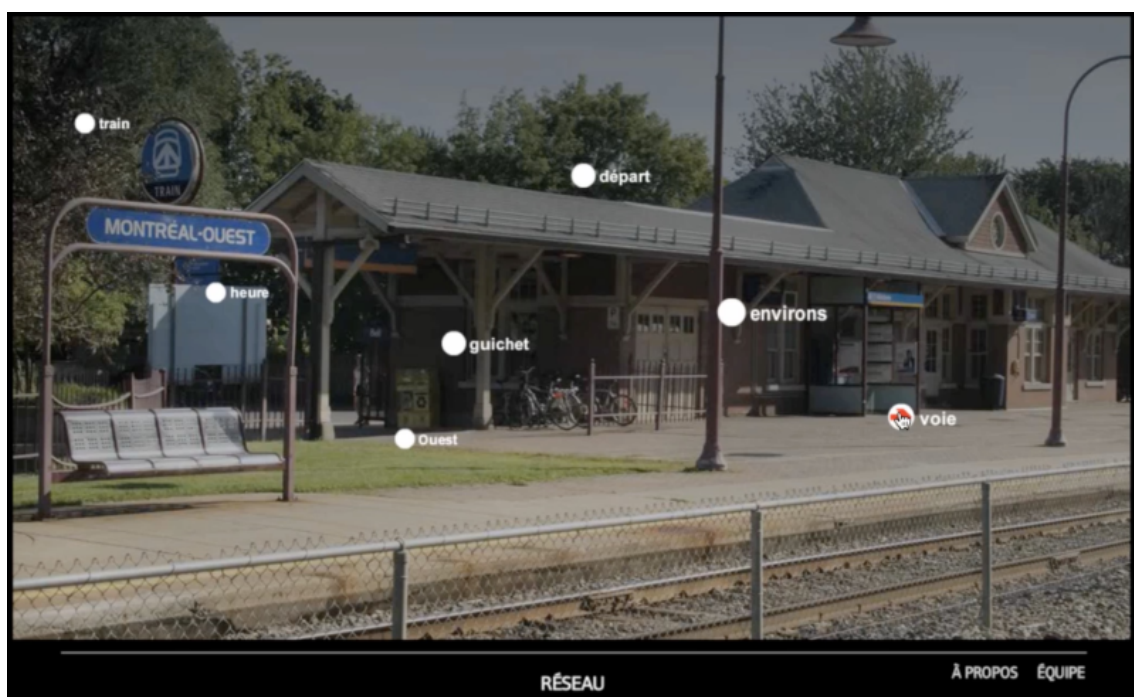
Figure 3. *Life in a Day*, de Kevin Macdonald et Ridley Scott



- 21 La ligne peut ensuite être brisée, sur un plan sémiotique évidemment, dès l'instant où une mise en intrigue déplace des éléments du récit, des scènes ou des actions, par le biais d'analepses ou de prolepses, ou par une multiplication des lignes narratives. Cette brisure est mineure cela dit, quand on la compare à celle majeure que l'interactivité déploie sur un plan médiatique. Le tracé offert à l'internaute n'est plus univoque, mais plurivoque, ce que l'interactivité permet d'actualiser. Les formes hypermédiatiques reposent de façon importante sur les possibilités de l'interactivité qui requièrent à divers niveaux la participation de l'internaute au déroulement du

récit. Si la ligne continue du récit (en littérature ou au cinéma par exemple) permet d'importantes mises en intrigue, qui viennent en segmenter le défilement sur un plan sémiotique, la ligne brisée sur un plan médiatique en réduit les possibilités, déplaçant l'attention sur le média lui-même et sa manipulation. Le récit ne se déroule plus simplement sous l'œil de l'internaute, mais grâce à ses décisions et à son agir. Le tracé à ligne brisée engage ainsi chaque fois une actualisation singulière. Il n'est pas constitué d'un seul trajet, mais d'une multitude qui demande à se singulariser. Il est une virtualité, un trajet à venir qui ne peut réellement exister qu'une fois mis en situation, puis en scène.

Figure 4. Gares de Sébastien Cliche et de La Traversée



- 22 Le champ des possibles, que le paradigme concrétisé dans le cadre des ontologies informatiques représente, est devenu un champ d'action, un répertoire de gestes, d'accessoires, de figures et de traits permettant au récit de progresser. Les jeux vidéo, les œuvres interactives et hypermédiatiques ont montré l'efficacité et la pertinence symbolique de cette forme de récit à ligne brisée¹⁴. Le projet *Gares* en est un exemple. À l'arrivée sur l'œuvre, la carte du réseau ferroviaire de la métropole s'affiche. En cliquant sur l'une des stations disponibles, l'internaute fait apparaître une photo sur

laquelle le nom de la station est affiché. Plusieurs points blancs y sont disposés, associés à des mots-clés qui permettent d'accéder à des notes de terrain. Sous le texte de ces notes, une flèche permet d'emprunter le mode de navigation linéaire, qui suit un parcours prédéfini à travers les différentes notes disponibles. Au haut de la fenêtre, des mots-clés en lien avec la note affichée sont réunis par une ligne dont le tracé ressemble à celui d'un tracé ferroviaire. Cliquer sur l'un de ces mots-clés redirige l'internaute vers une autre note, dans n'importe quelle gare, permettant ainsi une déambulation aléatoire au travers des récits et des découvertes des flâneurs, dont les voix s'entremêlent. Les lignes se multiplient, comme des cordes nouées à leur base et qui s'entremêlent afin de constituer une tresse.

- 23 La ligne peut ultimement être en grande partie effacée. C'est le dernier cas de figure. À la manière de *Ceux qui vont mourir* de Chatonsky, le récit peut être dilué jusqu'à n'être plus qu'un simple effet de récit, présent par le biais d'une allusion à un hypotexte, d'une référence intertextuelle ou interdiscursive. Le récit s'absente, tout en restant présent sur le mode du spectre. Il faut voir les métadonnées associées au projet sur le site web de Chatonsky qui précise qu'il entre dans les catégories du Netart, de la fiction et, plus important encore, de la fiction sans narration (*Fiction Without Narration*). Le récit n'y est plus qu'une ombre, une absence rendue manifeste par quelque signe diffus¹⁵. Dans *Interstice* par exemple, œuvre de 2006, Chatonsky fragmente des films de Jean-Paul Civeyrac, ne conservant que des segments qui ne racontent rien – claquement de portes, silences des personnages, moments d'attente. Ces segments sont soumis par la suite à une élaboration complexe, de sorte que la narration y est « mais pas la fiction : il y avait des personnages, une atmosphère et des lieux, comme une ritournelle qui se répétait et se différenciait à chaque cycle, mais cela ne racontait rien¹⁶. »

3. Le mode d'actualisation

- 24 Entre les ressources rendues disponibles par les bases de données et les récits déployés à l'écran, entre les plans paradigmatique et syntagmatique par conséquent, l'informatique fait apparaître des formes inédites d'actualisation, inédites parce qu'opérées par des machines de plus en plus autonomes. Des logiciels permettent de

prendre des données, de les organiser en fonction de paramètres et de les afficher. Des langages ou des logiciels propriétaires peuvent servir de soutien, impliquant une part de programmation ; et on voit graduellement apparaître des interfaces de programmation ou, plus récemment encore, des générateurs automatisés qui prennent le relais et travaillent à produire des récits de plus en plus complexes.

Modes d'actualisation

Base de données en arrière-plan

Mode passif

Mode dynamique

Base de données apparente et dynamique

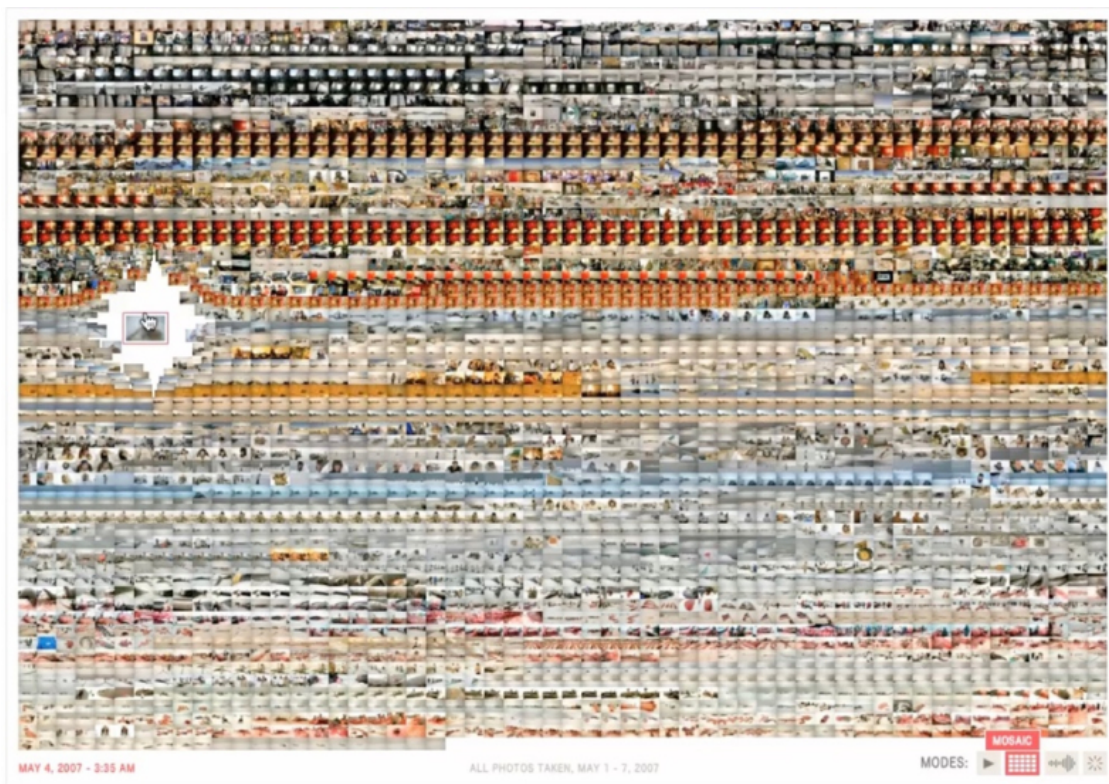
Base de données dynamique et autonome (IA)

- 25 Les bases de données sont l'ossature de ces projets, mais leur action est-elle cachée ou visible ? Affleure-t-elle à la surface des récits ? La situation de base, celle notamment à l'œuvre dans *Life in a Day*, repose sur une base de données qui reste cachée et l'œuvre en est même séparée, comme un résultat est disjoint d'un processus. La base de données était au cœur de la création d'une ligne narrative, puis elle s'est retirée de la scène. Elle n'est présente que sur un mode passif, antérieur à l'expérience de l'œuvre.
- 26 Tout autre est la situation qui a cours avec *Gares*. La base de données reste en arrière-plan, mais elle est le moteur de l'œuvre. Le récit est multiplié, en grande partie éclaté, c'est l'interactivité qui a pris le dessus et on le remarque à l'ensemble des fonctions disponibles : la carte qui ouvre sur l'ensemble des gares de banlieue, les mots-clés

qui permettent de circuler dans les diverses lignes narratives ou énonciatives, etc.

- 27 *The Whale Hunt* s'inscrit dans la même veine, à cette différence près que la base de données photographique est apparente et dynamique. Elle est mise en scène et s'impose comme l'élément central de l'œuvre¹⁷. Quand on entreprend d'y naviguer, on voit s'afficher à l'écran une mosaïque construite à partir des 3 214 photographies de Harris. Celles-ci constituent un tableau qui réagit au moindre mouvement de la souris et qui permet d'accéder au contenu grâce à un simple clic. Quand la souris se déplace sur le tableau, les photos s'écartent pour mettre en vedette l'image isolée, légèrement agrandie. La base de données réagit en quelque sorte à notre présence, l'interactivité est activée et on peut littéralement plonger dans l'œuvre à ce moment précis du récit. *The Whale Hunt* n'en reste pas là. L'interface graphique permet de naviguer dans l'œuvre selon un vaste ensemble de paramètres. Ainsi, le long du bord inférieur de l'écran se trouve une ligne du temps qui représente l'ensemble du voyage. Cette ligne du temps est représentée comme un graphique médical des battements de cœur dont la magnitude à chaque point correspond à la fréquence photographique (et donc au niveau d'excitation associé aux actions). La position de la photographie affichée est indiquée par une barre rouge lumineuse. En déplaçant le curseur le long de la ligne de temps, les barres s'étendent et se contractent de manière organique, ce qui permet à l'internaute d'isoler plus facilement un seul moment.

Figure 5. La mosaïque de *The Whale Hunt* de J. Harris



- 28 Au centre d'un des écrans de l'œuvre se trouve une icône représentant une baleine boréale (appelée aussi baleine du Groenland). En cliquant sur cette icône, on ouvre le panneau des contraintes, qui peut être utilisé pour isoler les différents sous-récits qui se produisent dans le récit plus large. Les types de contraintes disponibles comprennent : les acteurs, les concepts (ce qui apparaît saillant sur chaque photo : « sang », « bateaux », « bâtiments », « nourriture », « jeux », « enfants », etc.), les contextes, la cadence de défilement des photographies. Des contraintes multiples peuvent être combinées pour façonner différents sous-répertoires. Les divers dispositifs de visualisation imaginés par Harris donnent à voir la complexité du récit. La multiplication des fonctions permet de saisir de façon intuitive cette complexité, notamment grâce à la souplesse de l'affichage des contenus à l'écran.
- 29 La dernière modalité de mise en scène des bases de données représente leur dilution complète par le biais de leur généralisation. Quand les bases de données deviennent aussi vastes que le monde, leur action n'est plus visible, bien qu'elles semblent agir de leur

propre chef. Les données impliquées par des robots (présentés comme des agents conversationnels) tels que ChatGPT d'OpenAI ou Gemini/Bard de Google sont massives, elles vont jusqu'à englober « l'étendue des connaissances du monde » ; et les logiciels entreprennent de les combiner avec « la puissance, l'intelligence et la créativité des grands modèles de langage ¹⁸ » utilisés par ces compagnies. Les bases de données présentes sont de l'ordre de métabases de données aux dimensions pharaoniques mises en œuvre grâce à des opérations et des algorithmes d'une grande efficacité (de même que d'opacité). Les robots y puisent leurs informations afin de fournir des réponses à toutes les demandes, y compris à celles de raconter des histoires. Les récits produits sont textuels, mais aussi iconotextuels ¹⁹. Des générateurs de vidéos ont d'ailleurs commencé à apparaître, et ce n'est qu'une question de temps avant qu'ils deviennent performants et omniprésents ²⁰.

- 30 Un artiste tel que Gregory Chatonsky multiplie les projets qui reposent sur de l'intelligence artificielle. Ce chantier, il l'identifie comme de l'imagination artificielle (*artificial imagination*). Plus de cent-vingt projets sont liés à ce mot-clé, les derniers en date jouant sur des formes avancées d'IA, proposant des corps, des images, voire des récits. En juillet 2022, il a même fait paraître un livre intitulé *Internes*, un roman de 182 pages, co-écrit par une intelligence artificielle ²¹. Comme avec le film de Macdonald, la base de données qui a servi à générer le roman reste cachée et l'œuvre en est séparée ; à vrai dire, la base n'est plus visible, parce que ses opérations y sont à ce point complexes, et les données innombrables, qu'elles ne participent pas de l'ordre du visible.

Conclusion

- 31 Les liens entre bases de données et narration n'ont cessé de se complexifier. Et avec les développements actuels de l'informatique, ils sont maintenant soumis à des forces qui nous dépassent peut-être. Si, lors des premières utilisations, l'informatique apparaissait comme un outil pour aider à traiter des données et à les mettre en scène, nous sommes maintenant passés à l'étape d'une machine capable de traiter par elle-même ces données, trouvant une façon de les mettre en scène. En explorant les modes d'acquisition des données, les types

de récits générés, ainsi que les modes d'actualisation et de mise en scène des données, ce sont certains des rudiments des esthétiques numériques qui ont été exposés.

- 32 Au début des années 2000, Lev Manovich appelait à la création d'une *info-aesthetics*²², une esthétique numérique qui devait avoir pour objectif l'étude des nouvelles pratiques esthétiques en contexte numérique, reposant pleinement sur une exploitation des bases de données et de leurs possibilités. Nous en étions alors au début de la deuxième génération de littérature électronique, marquée par les développements exponentiels du Web et la commercialisation de logiciels de création numérique.
- 33 Si la première génération précédait l'apparition d'Internet et se concentrait sur le texte, bien que l'image ait pu y jouer un rôle à l'occasion, la deuxième génération a surgi à partir de 1994, exploitant les formes de l'interactivité et de l'hypermédialité²³. C'est avec ces œuvres de deuxième génération qu'est apparue la nécessité d'une réflexion sur les interactions entre les bases de données et les formes narratives.
- 34 Nos ordinateurs ne sont pas de simples outils d'écriture, d'une écriture et d'une imagination à peine touchées par cette technologie, ce sont des agents de transformation de nos modes de pensée et de nos capacités d'invention, qui sont appelés à se déplacer d'une raison graphique à une raison numérique. Et cela, c'est sans compter sur le développement des logiciels d'intelligence artificielle qui viennent changer de façon majeure la donne et qui rendent les interfaces de programmation de plus en plus puissantes. Non seulement le numérique nous conduit à renouveler nos modes de pensée, mais il entraîne nos stratégies de création sur de nouveaux terrains, ceux de l'imagination artificielle.

BIBLIOGRAPHIE

BING XU, *Dragonfly Eyes*, Chine, Xu Bing Studio 2017, 1h21, URL : <https://vimeo.com/643462232> [consulté le 8 avril 2024].

CHATONSKY Gregory, *Internes*, RRoses éditions, 2022.

CHATONSKY Gregory, « La fiction sans narration (FSN) », Gregory Chatonsky [en ligne], 2012, URL : <http://chatonsky.net/fsn/> [consulté le 8 avril 2024].

CHATONSKY Gregory, « Ceux qui vont mourir », Gregory Chatonsky [en ligne], 2007, URL : <http://chatonsky.net/mourir/> [consulté le 8 avril 2024].

CURRIE Mark, *Postmodern Narrative Theory*, New York, St. Martin's Press, 1998.

CLICHE Sébastien et La Traversée, Atelier québécois de géopoétique, Gares [en ligne], 2012, URL : <https://the-next.eliterature.org/works/1525/0/0/> [consulté le 8 avril 2024.]

SAUSSURE Ferdinand de, *Cours de linguistique générale* [1916], Paris, Payot, 1972.

FARIS Wendy B., *Labyrinths of Language*, Baltimore, The Johns Hopkins University Press, 1988.

FLORES Leonardo, « Third Generation Electronic Literature », *Electronic Book Review* [en ligne], 2019, DOI : [10.7273/axyj-3574](https://doi.org/10.7273/axyj-3574).

GERVAIS Bertrand, *Un imaginaire de la fin du livre. Littérature et écrans*, Montréal, Presses de l'université de Montréal, 2023.

GERVAIS Bertrand, *La ligne brisée : labyrinthe, oubli et violence. Logiques de l'imaginaire. Tome II*, Montréal, Le Quartanier, 2008.

HARRIS Jonathan, *The Whale Hunt* [en ligne], 2007, URL : <http://thewhalehunt.org/> [consulté le 8 avril 2024].

HUTCHEON Linda, *Narcissistic Narrative. The Metafictional Paradox*, Londres, New York, Routledge, 1984.

INGOLD Tim, *Une brève histoire des lignes*, Bruxelles, Zones sensibles, 2011.

JAKOBSON Roman, *Essais de linguistique générale*, Paris, éditions de Minuit, 1963.

LÉVY Pierre, *L'idéographie dynamique. Vers une imagination artificielle ?*, Paris, La Découverte, 1991.

MANOVICH Lev, « Introduction to Info-Aesthetics », Manovich [en ligne], 2008, URL : <http://manovich.net/index.php/projects/introduction-to-info-aesthetics> [consulté le 8 avril 2024].

MARCLAY Christian, *The Clock*, Royaume-Uni, Paula Cooper Gallery, 2010, 24h.

QUENEAU Raymond, *Cent mille milliards de poèmes*, Paris, Gallimard, 1961.

RICŒUR Paul, *Temps et récit. Tome 1*, Paris, éditions du Seuil, 1983.

REBECCHI Marie, « Le temps du remontage : entre cinéma et installation. *The Clock* de Christian Marclay », *Marges*, n° 19, 2014. DOI : [10.4000/marges.993](https://doi.org/10.4000/marges.993)

WATERCUTTER Angela, « *Life in a Day* Distills 4,500 Hours of Intimate Video into Urgent Documentary », *Wired* [en ligne], juillet 2011, URL : <https://www.wired.com/2011/07/life-in-a-day-interviews/> [consulté le 8 avril 2024].

WAUGH Patricia, *Metafiction, The Theory and Practice of Self-Conscious Fiction*, Londres, New York, Routledge, 1984.

ZHANG Ling, « Foreshadowing the Future of Capitalism: Surveillance Technology and Digital Realism in Xu Bing's *Dragonfly Eyes* (2017) », *Comparative Cinema*, vol. 8, n° 14, 2020, p. 62-81.

NOTES

1 Consulter à ce sujet les travaux de Linda Hutcheon, *Narcissistic Narrative. The Metafictional Paradox*, Londres, New York, Routledge, 1984 ; de Patricia Waugh, *Metafiction, The Theory and Practice of Self-Conscious Fiction*, Londres, New York, Routledge, 1984 ; de Wendy B. Faris, *Labyrinths of Language: Symbolic Landscape and Narrative Design in Modern Fiction*, Baltimore, Londres, The Johns Hopkins University Press, 1988 ; de Mark Currie, *Postmodern Narrative Theory*, New York, St. Martin's Press, 1998 ; etc.

2 Cette notion est au cœur de la construction théorique de Paul Ricœur, initiée dans *Temps et récit. Tome 1*, Paris, éditions du Seuil, 1983.

3 C'est Ferdinand de Saussure qui, le premier, a distingué le rapport associatif, décrit maintenant comme paradigmatique, et le rapport de coprésence ou syntagmatique (Ferdinand de Saussure, *Cours de linguistique générale* [1916], Paris, Payot, 1972). Roman Jakobson a exploité ces deux notions en contexte linguistique dans ses *Essais de linguistique générale*, Paris, éditions de Minuit, 1963.

4 Toute œuvre requiert, pour être diffusée et expérimentée, d'exister sur un plan médiatique, qui rend compte du support de cette production, ainsi que sur un plan sémiotique, qui donne accès aux processus de signification mis en jeu. Un troisième plan s'ajoute à ces deux premiers, le symbolique, qui repose sur la mise en relation de l'œuvre expérimentée avec d'autres œuvres qui en informent l'expérience et permettent d'en déterminer la valeur (cf. Bertrand Gervais, *Un imaginaire de la fin du livre. Littérature et écrans*, Montréal, Presses de l'université de Montréal, 2023, p. 77 et *passim*).

5 Pierre Lévy a cette belle formule pour définir le virtuel : « Une réalité virtuelle est un monde sensible auquel ne correspond aucune entité physique, sauf celle de fichiers informatiques. » (Pierre Lévy, *L'idéographie dynamique. Vers une imagination artificielle ?*, Paris, La découverte, 1991, p. 19).

6 Cette description a été écrite avec la collaboration d'Ariane Savoie, dans le cadre des recherches menées au Laboratoire NT2 de l'UQAM et à la Chaire de recherche du Canada sur les arts et les littératures numériques (B. Gervais, titulaire, 2014-2022).

7 La description de l'œuvre de Harris s'inspire librement de la fiche du Répertoire des arts et des littératures hypermédiatiques du NT2 qui y est consacrée. cf. Simon Brousseau, « *The Whale Hunt*, par Harris, Jonathan», NT2 [en ligne], 13 avril 2010, URL : <https://nt2.uqam.ca/fiches/whale-hunt-0> [consulté le 15 mai 2024].

8 Angela Watercutter, "Life in a Day Distills 4,500 Hours of Intimate Video Into Urgent Documentary", *Wired* [en ligne], juillet 2011, URL : <https://www.wired.com/2011/07/life-in-a-day-interviews/> [consulté le 8 avril 2024].
L'accès aux technologies numériques prédispose à des projets de l'envergure de *Life in a Day* : les caméras sur le marché sont abordables, leur manipulation s'est simplifiée et elles permettent de capter des images d'une qualité professionnelle. Les logiciels de postproduction se multiplient. Ils s'automatisent et accompagnent les utilisateurs dans leurs opérations de création de sorte qu'un néophyte s'y retrouve sans trop de problèmes. L'arrivée du Web 2.0 a modifié la pratique cinématographique à cet égard. En offrant la possibilité d'héberger d'importantes plateformes de diffusion comme YouTube ou Vimeo, le Web 2.0 a participé à la démocratisation de la production cinématographique.

9 Fait à noter, le film, après avoir fait la tournée des festivals, est retourné sur la plateforme d'où elle a émergé, à savoir YouTube, en accès libre. URL : https://www.youtube.com/watch?v=JaFVr_cJJIY [consulté le 8 avril 2024].

10 Pour une réflexion sur la dimension cinématographique du projet de Marclay, lire l'article de Marie Rebecchi, « Le temps du remontage : entre cinéma et installation. *The Clock* de Christian Marclay », *Marges* [en ligne], n° 19, 2014. DOI : [10.4000/marges.993](https://doi.org/10.4000/marges.993).

11 La perspective restreinte développée ici, qui met de l'avant uniquement la relation entre données et narration, ne permet pas de rendre justice à des œuvres dont les enjeux politiques ou symboliques sont grands et au cœur des transformations de l'imaginaire contemporain. Il en va ainsi de la réflexion sur la surveillance et sa gestion tentaculaire en Chine dans l'œuvre de Xu Bin. On lira à ce sujet la contribution de Ling Zhang, « *Foreshadowing the Future of Capitalism: Surveillance Technology and Digital Realism in*

Xu Bing's *Dragonfly Eyes* (2017) », *Comparative Cinema*, vol. 8, n° 14, 2020, p. 62-81.

12 Gregory Chatonsky, « Ceux qui vont mourir », Gregory Chatonsky [en ligne], 2007, URL : <http://chatonsky.net/mourir/> [consulté le 8 avril 2024].

13 J'ai exploré cette idée de façon importante dans *La ligne brisée : labyrinthe, oubli et violence. Logiques de l'imaginaire*. Tome II, Montréal, Le Quartanier, 2008. Lire aussi, dans une perspective anthropologique, l'essai de Tim Ingold, *Une brève histoire des lignes*, Bruxelles, Zones sensibles, 2011.

14 Le répertoire des arts et des littératures hypermédiatiques du NT2 donne une bonne idée de l'étendue des variations narratives introduites par l'exploration des formes d'interactivité. On trouve d'ailleurs sur le site la taxonomie des formes d'interactivité répertoriées. « Taxonomie », NT2, URL : <https://nt2.uqam.ca/taxonomie> [consulté le 8 avril 2024].

15 Trente-et-un projets sont regroupés sous la bannière de ces « fictions sans narration », qui recoupent de plus en plus de projets se servant d'une intelligence artificielle.

16 Grégory Chatonsky, « La fiction sans narration (FSN) », Grégory Chatonsky [en ligne], 2012, URL : <http://chatonsky.net/fsn/> [consulté le 8 avril 2024].

17 Il serait plus juste de dire qu'elle *avait été* mise en scène, l'œuvre ayant été montée en Flash. Sur la page de l'œuvre, on ne trouve plus qu'une description de ce qu'elle permettait en termes de navigation : *The Whale Hunt*, URL : <http://thewhalehunt.org/interface.html> [consulté le 8 avril 2024].

18 Passages traduits par l'auteur. Source : Sundar Pichai, « An Important Next Step on our AI Journey », *Google The Keyword*, 6 février 2023, URL : <https://blog.google/technology/ai/bard-google-ai-search-updates/> [consulté le 8 avril 2024].

19 Parmi les premiers générateurs de planches de bande dessinée, on note AI-Comic-Factory, URL : <https://huggingface.co/spaces/jbilcke-hf/ai-comic-factory> [consulté le 15 mai 2024].

20 Sur Zapier, on trouve une liste de logiciels dédiés à la production vidéo : *Descript*, pour le montage vidéo en éditant le scénario ; *Wondershare Filmora*, pour peaufiner les vidéos à l'aide d'outils d'IA ; *Runway*, pour expérimenter l'IA générative ; *Peech*, pour les équipes de marketing ; *Synthesia*, pour l'utilisation d'avatars numériques ; *Fliki*, pour les

vidéos sur les médias sociaux ; *Visla*, pour transformer un scénario en vidéo ; *Opus Clip*, pour transformer des vidéos de longue durée en vidéos de courte durée. URL : <https://zapier.com/blog/best-ai-video-generator/> [consulté le 15 mai 2024].

21 Gregory Chatonsky, *Internes*, RRoses éditions, 2022.

22 Lev Manovich, « Introduction to Info-Aesthetics », *Manovich* [en ligne], URL : <http://manovich.net/index.php/projects/introduction-to-info-aesthetics> [consulté le 8 avril 2024].

23 L'hypothèse est avancée par Leonardo Flores, « Third Generation Electronic Literature », *Electronic Book Review* [en ligne], 2019.DOI : [10.7273/axyj-3574](https://doi.org/10.7273/axyj-3574).

RÉSUMÉS

Français

Cet article entend définir les liens entre bases de données et pratiques narratives en contexte numérique. Comment narrativise-t-on une base de données ? Quels en sont les modes d'actualisation dans une structure narrative ? Sa présence nécessaire est-elle tacite ou explicite ? L'amont du récit, en d'autres mots, est-il révélé ou non ? Que sa réalité ait changé de façon déterminante, on ne peut plus en douter, mais de quelle manière son agir est-il représenté, voire symboliquement pris en charge ? Afin d'explorer les modes d'actualisation narrative des bases de données, j'identifierai trois catégories, qui sont autant de portes d'entrée de cette relation : 1 – le mode d'acquisition des données ; 2 – le type de récit généré ; 3 – le mode de présence et de mise en scène de la base de données à l'œuvre. Je décrirai ces catégories à partir d'exemples précis qui en illustrent le fonctionnement.

English

This article sets out to define the links between databases and narrative practices in a digital context. How can we narrativize a database? How is it actualized in a narrative structure? Is its presence tacit or explicit? We can no longer doubt that the database plays a major role in the way we tell stories, but how is its action represented? In order to explore the narrative actualizations of databases, I will identify three categories, which are as many ways of understanding this relationship: 1 – the mode of data acquisition; 2 – the type of narrative generated; 3 – The mode of presence and staging of the database at work. I'll describe these categories using specific examples, from found footage to hypermedia works.

INDEX

Mots-clés

narration, métarécit, base de données, hypermédia, arts et littératures numériques

Keywords

narrative, meta-narrative, database, hypermedia, digital arts and literature

AUTEUR

Bertrand Gervais

Département d'études littéraires, université du Québec à Montréal

IDREF : <https://www.idref.fr/050141627>

ORCID : <http://orcid.org/0000-0003-1520-5618>

ISNI : <http://www.isni.org/0000000117663407>

BNF : <https://data.bnf.fr/fr/12971569>

Rythmer le chaos : la programmation du récit dans *Smart Rat City*

Structuring the chaos: programming the narrative in Smart Rat City

Françoise Chambeftort

DOI : 10.35562/marge.874

Droits d'auteur

CC BY-NC-SA

PLAN

Smart Rat City

La tension narrative

Rythmer le chaos

Conclusion

TEXTE

- 1 Chercheuse associée au laboratoire ELLIADD, j'ai soutenu ma thèse¹ en 2020 après un doctorat en recherche création en sciences de l'information et de la communication. Ma pratique artistique est celle de l'art numérique ; une partie de mes créations se rattache à littérature numérique.
- 2 Ma démarche de recherche création en SIC se place dans la lignée des travaux de Serge Bourchardon² et Alexandra Saemmer³, enrichie par les échanges que j'ai pu avoir avec les participants au séminaire *Dialogue sur les enjeux de la recherche création pour les sciences humaines*⁴ que j'ai co-animé avec ma collègue Carolane Sanchez en 2020-2021 et 2022-2023. Au fil de l'avancée d'un projet, les boucles de rétroaction entre recherche et création se multiplient, permises par l'auto-réflexivité⁵. De nouvelles voies de recherche apparaissent, tandis que l'œuvre se modèle et se remodèle, se déclinant parfois en plusieurs versions. L'activité de recherche englobe le processus de création et la production artistique est avant tout expérimentale.

- 3 Pendant ma thèse, j'ai travaillé sur les potentialités narratives des flux de données. Avec *Lucette, Gare de Clichy*⁶, j'ai montré qu'il était possible de créer un récit de fiction qui s'appuie sur un flux de données en temps réel, j'ai utilisé pour cela les données d'arrivée des trains dans une gare. Dans ce type d'œuvres, que j'ai appelé « moulins à données », le support est le web, le matériau, ce sont les données et la forme, c'est la fiction littéraire.
- 4 Avec *Lucette*, j'ai mis à jour une configuration narrative particulière que j'ai appelée *Mimèsis du flux*, dans laquelle l'enchaînement causal (qui est l'une des propriétés sémantiques du récit selon Françoise Revaz⁷) est remplacé par la séquentialité d'événements réels. Et j'ai montré que les œuvres de ce type ont une faculté particulière à représenter la contingence de la vie.

Smart Rat City

03/03/2023 – New York

00:19, east 64 street,
dirty condition, trash

WHOUA une belle merde de chien à déguster !

Je vis dans la ville du dessous mais, avec d'autres, on commence à sortir à l'air libre, la nuit, parce qu'en bas, c'est devenu invivable. On se marche dessus.

00:40, park avenue,
noise – residen-
tial, banging/pounding

Ça tape au-dessus, ça ébranle toute la maison

C'est le matin. Une odeur de brûlé a envahi l'air depuis un moment mais ça nous réveille à peine, à l'entrée de la ville du dessous, tellement on est fourbus de notre virée nocturne.

00:56, east 109 street,
noise – street/side-
walk, loud music/party

Un cœur monstrueux bat au-dessus. Je n'ai nulle part où fuir.

L'effondrement, on l'a pas vu venir. Il nous a cueilli, comme ça, d'un coup. Un grondement terrifiant, brutal, qui nous emmène dans son flot, à plus savoir où est notre corps dans tout ça.

01:11, east 49 street,
noise, noise, barking
dog (nr5)

Il me poursuit, où me cacher ?

Ça dure, ça dure, on attend même pas que ça se termine parce qu'on est plus rien. TOI, moi, un morceau de trottoir, une bouteille crevée, on a tous fusionné dans le chaos.

01:33, east 33 street,
homeless
person assistance,

L'un d'EUX est couché par terre, dans son vomi. Il ne NOUS fait pas peur, non, pas peur du tout.

Et un moment après, ça recommence. Et encore après.

01:49, madison avenue,
noise – residen-
tial, banging/pounding

Quelque chose cogne et se rapproche

Il nous faut drôlement du temps pour revenir à nous-mêmes, dans le silence retrouvé et le brouillard de poussière. Est-ce qu'on est devenu sourd ? Changés en statues, recouverts de cendres et de poudre de pierre. Des corps immobiles, figés à jamais dans cette gangue de neige sale, et d'autres s'ébrouant pour s'extraire et respirer enfin un air qui ne nous étoufferait pas.

02:17, west 129 street,
noise – residen-
tial, banging/pounding

Un grand vacarme de choses qui tombent et retombent et retombent sans fin.

Après ça, leur monde est plus le même, bien carré. C'est devenu un drôle de bazar. Et ça nous profite bien.

02:32, 5 avenue, noise
– residen-
tial, banging/pounding

Un fracas dans ma tête, où fuir ?

Les restaurants autour, ils sont fermés. Alors tous ceux qui ont l'habitude de manger dans les poubelles, ils se retrouvent sans rien à becqueter, et ils ont vite fait de nous rejoindre dans le mikado géant. Parce que là, y a de quoi faire si tu trouves le bon filon. Et on prospère là-dessous, t'imagines pas. Au plus profond des entrailles de ce tas, on prend d'assaut les garde-manger, on copule comme des frénétiques et, eux, ils peuvent pas nous atteindre. C'est un monde à notre mesure.

04:41, amsterdam
avenue, noise – resi-
den-
tial, banging/pounding

BANG ça cogne encore et encore

Au début, tu n'es même pas TOI. Tu es LE TAS. LE TAS, il dort, se réchauffe. Il se réveille, grouille doucement, cherche le lait, boit. Il prend peur, il attend puis il oublie et recommence. Mais un jour, pas de lait. Tu t'endors quand même mais au réveil tu te précipites, tu griffes, tu mors, tu deviens TOI. Parce qu'il n'y en a pas pour tous, tu deviens TOI.

- 5 Après ce travail, j'ai souhaité aller plus loin en expérimentant la dimension spatiale des données et donner plus d'ampleur au récit. Par « ampleur du récit », j'ai seulement pensé à un texte de plus d'ampleur, donc plus long. Car dans *Lucette* le récit est simple : un seul personnage s'exprime ; le texte, composé de fragments, est d'une longueur de 10 000 caractères environ, espaces compris, soit l'équivalent de 3 pages dans Word.

- 6 C'est ainsi qu'est née *Smart Rat City*, une œuvre qui s'appuie sur des données du système d'information de la ville de New-York.
- 7 Cette œuvre est connectée à deux flux de données ouvertes :
 - Les données sur les inspections du service de dératisation (des événements qui peuvent être : une visite de contrôle, la pose d'appâts, le rebouchage des ouvertures etc.).
 - Les données du 311. C'est un service municipal initialement téléphonique et maintenant sur le web qui existe depuis 2003. Il a été créé dans le but de désengorger le service d'urgence, le 911. Il est destiné à prendre en compte les plaintes et requêtes non urgentes des habitants. Progressivement, il est devenu le guichet unique de recueil et suivi des demandes qui sont ensuite redirigées vers les administrations municipales compétentes. Cette masse de données, touchant à de nombreux aspects de la vie quotidienne des new-yorkais, est accessible par tous depuis 2010, sur le site NYC Open Data ⁸.
- 8 La richesse et la disponibilité de ces données ont largement orienté mon choix pour cette ville et cette thématique, ainsi que le fait de pouvoir croiser des données produites par les habitants avec des données produites par l'institution.
- 9 J'ai commencé à travailler à ce projet en 2020. Sa forme a évolué ⁹ pour aboutir à un dispositif en trois parties ¹⁰ : la carte au centre, le panneau de gauche et le panneau de droite.

Panneau de gauche



Source : Smart Rat City

- 10 À gauche, la partie données : le titre de l'œuvre, la date d'aujourd'hui, la météo du jour à New York, les données de la journée, présentées en anglais de manière chronologique (l'heure, le lieu, le type de signalement qui a été fait, par exemple : 00:34, *ninth avenue*, *rodent, rat sighting*) ainsi qu'un ascenseur pour faire défiler le texte jusqu'en bas.
- 11 Ce panneau de gauche a un menu avec un point d'interrogation qui renvoie à la légende des marqueurs qui figurent sur la carte. Et un bouton qui permet de le replier.

Légende des marqueurs



Source : *Smart Rat City*

- 12 Au centre, une carte de New York avec des marqueurs de différentes couleurs (gris, oranges, rouges). Il est possible de se déplacer, zoomer et dézoomer. Une certaine interactivité est proposée : un clic sur un marqueur permet d'afficher à gauche le détail de l'événement. Et vice-versa un clic sur les entrées des données à gauche permet de localiser l'événement sur la carte.

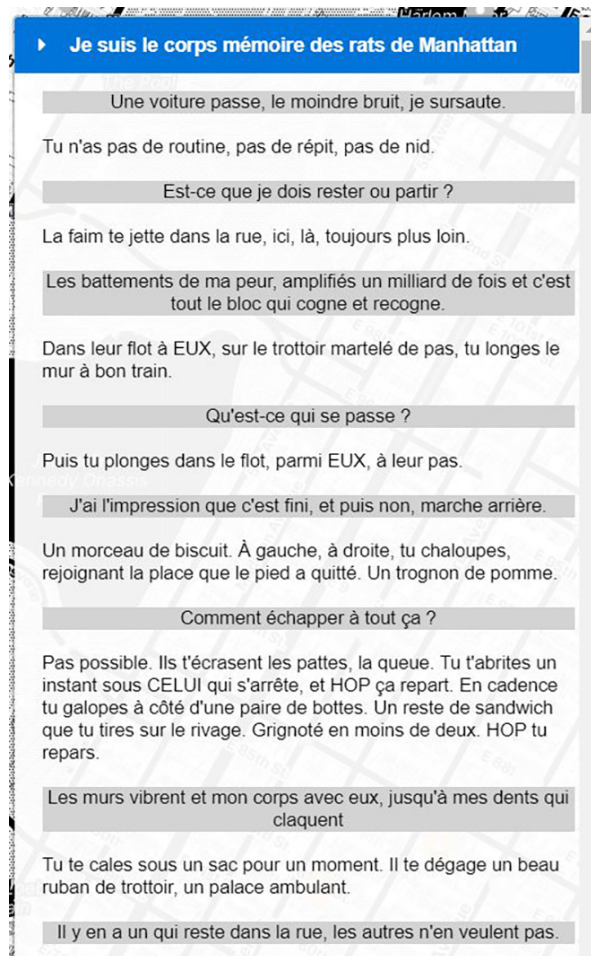
La carte, partie centrale de l'interface



Source : *Smart Rat City*

- 13 Le panneau de droite est consacré au texte fictionnel : un titre, et un texte avec des parties sur fond gris et d'autres sur fond blanc, ainsi qu'un ascenseur pour faire défiler le texte jusqu'en bas. Les parties grisées sont cliquables, elles renvoient vers le marqueur correspondant sur la carte.

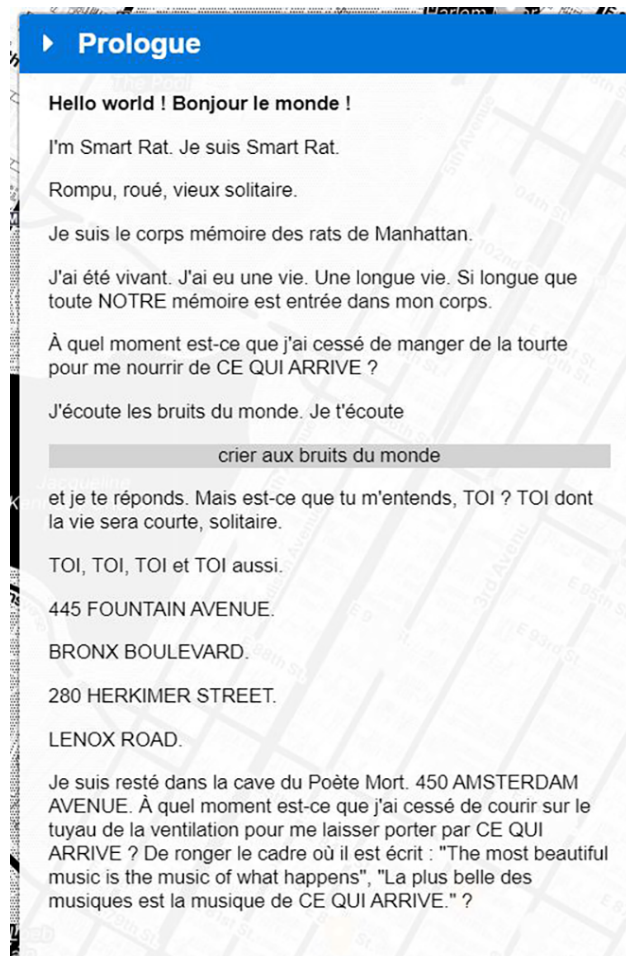
Panneau de droite



Source : *Smart Rat City*

14 Dans le menu, l'icône d'accueil permet d'afficher le prologue :

Prologue



Source : *Smart Rat City*

- 15 *Smart Rat City*, contrairement à *Lucette, Gare de Clichy*, intègre les données au récit en profondeur. Dans *Lucette*, les données utilisées dans la narration sont seulement l'heure d'arrivée du train et son nom qui est repris en tant que nom du personnage qui vient visiter *Lucette*. Dans *Smart Rat City*, le contenu des données est beaucoup plus foisonnant et il a participé à la construction même de l'imaginaire du récit. Lors du premier cycle d'expérimentation, dans la phase où l'interface de travail était seulement une liste de données, ma réflexion portait sur chaque événement autour des questions : entendent-ils des sons, des bruits ? sentent-ils des odeurs ? cela a-t-il un retentissement sur leur vie ? Cela m'a amené à filtrer les données en fonction de leur impact sur la vie des rats. Et finalement à créer ce que j'ai appelé les « cris », c'est-à-dire les textes grisés qui sont des réponses immédiates à certains événements.

La tension narrative

- 16 L'approche du récit par la tension narrative permet de se libérer des grammaires du récit et de se placer d'emblée du côté de l'acte de lecture. Car dans un système comme celui-ci, la génération du récit se fait dans la phase de réception donc au moment de la lecture. Le texte affiché à droite, la fiction, n'est pas immuable, mais réglé par un programme informatique. Il dépend des données du jour (affichées à gauche). Programmer l'affichage du récit, c'est donc programmer la lecture qui en sera faite :

« Si, dans un système génératif, la représentation de l'action et son déroulement ne sont pas donnés au départ mais sont représentés au moment même de la lecture, un télescopage se produit entre le plan de la production et celui de l'interprétation. Le rôle de l'auteur, dans un système génératif, sera de programmer des règles capables de générer une narration à partir de matériaux divers (texte, images, sons). La génération du récit sera alors déportée vers la phase de réception. Il est donc essentiel dans notre travail de bien comprendre les mécanismes de compréhension de l'action en jeu au moment de la lecture. Ils guideront l'établissement des règles de générativité¹¹. »

- 17 Raphaël Baroni définit la tension narrative comme « l'établissement, le maintien et la résolution d'une tension dans la lecture, dont dépend l'intérêt du récit.¹² ». Il définit trois types de tensions : la curiosité, le suspense et la surprise.
- 18 Dans un système tel que celui-ci, l'enchaînement causal du récit est remplacé par la séquentialité d'événements réels. Prévoir une fin au récit est alors problématique : cela reviendrait à programmer artificiellement l'arrêt du programme, alors que celui-ci est susceptible de continuer à fonctionner tant que des données sont produites. J'ai préféré laisser le récit ouvert. Je n'ai pas non plus choisi de développer une tension de type suspense, en tant qu'anticipation incertaine d'un dénouement.
- 19 J'ai privilégié une mise en intrigue par la curiosité, comme l'artiste se laissant guider par son matériau, car ce type de tension narrative est engendré naturellement par la fragmentation du récit. La surprise

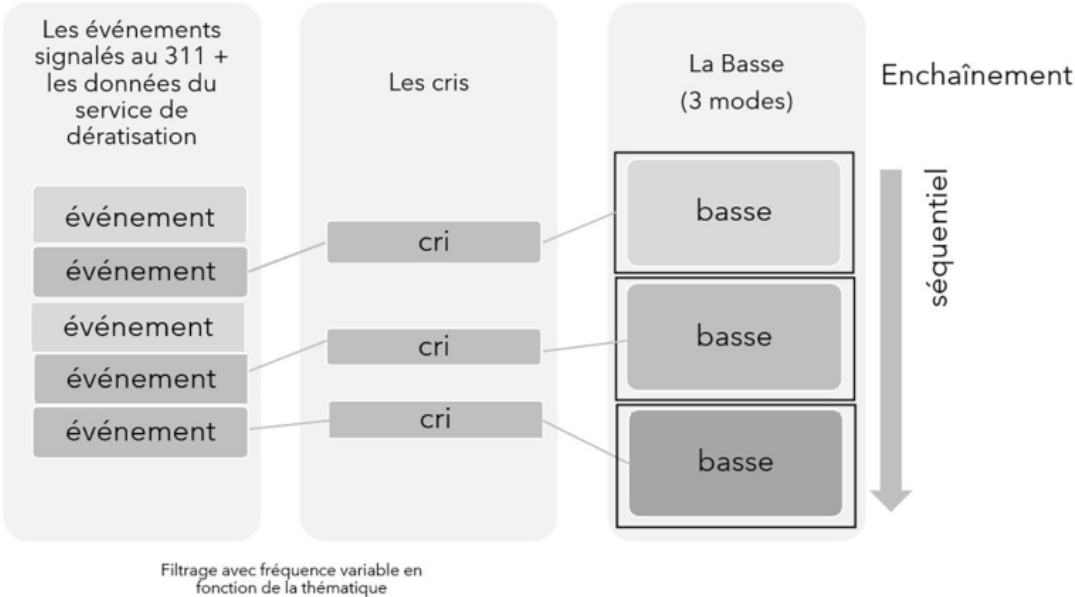
procède par l'occultation de certains éléments. Des espaces sont laissés vides conduisant le lecteur à une activité de diagnostic à partir d'indices formulés par le texte, comme un puzzle à reconstituer. L'agencement de fragments correspond exactement à ce cas de figure, les fragments étant autant de pièces de puzzle à assembler.

- 20 *Quid* de la surprise ? Je souhaitais laisser aux seules données réelles la capacité de surgir de manière inattendue. Fallait-il que les événements réels infléchissent le cours du récit ? J'ai choisi une autre solution, celle d'alterner les paroles/pensées de plusieurs personnages. Par analogie avec le monde sonore, j'ai appelé ces deux voix les cris (surprise) et la basse (curiosité). Cela m'a permis d'allier *flashes* sensibles et récit (séquentiel ou non). J'ai ainsi pu rester dans une logique de la contingence tout en rendant possible un récit de plus d'ampleur.

Rythmer le chaos

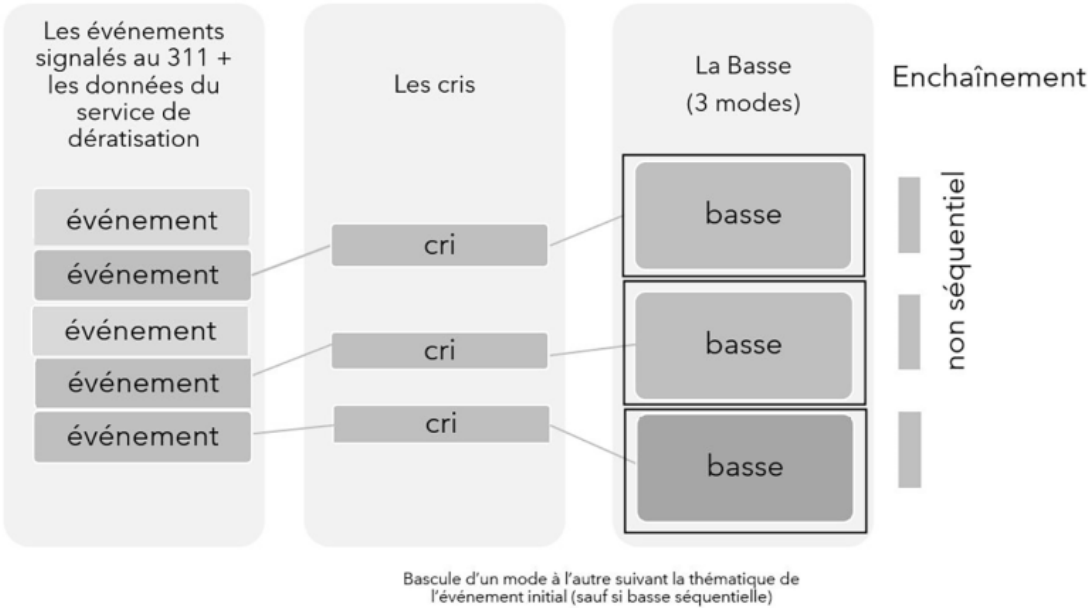
- 21 La programmation du récit a eu pour objectif d'ordonner le déroulement narratif tout en donnant une place au chaos des données. Il s'est agi de rythmer la tension narrative entre curiosité et surprise.
- 22 Une phase d'exposition est proposée dans le prologue mais celui-ci est placé de telle manière qu'il ne sera peut-être pas lu tout de suite, suivant comment le lecteur s'appropriera le dispositif.
- 23 L'affichage du texte est programmé de la façon suivante :
- Ce que j'ai appelé les cris, le texte grisé, répond directement à un événement du côté des données mais après filtrage avec fréquence variable en fonction de la thématique (il n'y a pas de réponse pour tous les événements). Dans mon imaginaire, cela correspond aux cris des jeunes rats qui sont un peu partout dans la ville et subissent la contingence de l'existence.
 - À chaque cri, succède une basse, un fragment de texte en blanc, qui pour moi est la parole du narrateur, Smart Rat. Cette basse peut être séquentielle, c'est-à-dire qu'à la manière d'un chapitre, les fragments vont se succéder selon un ordre établi. Les cris intermédiaires n'auront alors pas d'influence sur la basse.

Programmation du récit – Basse séquentielle



24 Mais la basse peut aussi être non séquentielle.

Programmation du récit – Basse non séquentielle



25 Dans ce cas, la thématique de l'événement initial va conditionner chaque fragment de basse. Sachant qu'il y a 3 modes qui vont amener une dramatisation plus ou moins importante du propos : un mode

positif, de bien-être, un mode neutre et un mode négatif relatif à un danger, une urgence vitale. Mais la séquentialité est toujours prioritaire : il n'y aura pas de changement de mode tant que la séquence en cours n'est pas terminée¹³.

Conclusion

- 26 Donner plus d'ampleur au récit¹⁴ a exacerbé le conflit entre l'ordre du récit et le chaos des données. La recherche d'un équilibre était nécessaire afin de préserver la capacité de l'œuvre à rendre compte de la contingence de l'existence, contingence apportée par les données. L'équilibre a été obtenu par la mise en place d'un rythme alternant les cris des jeunes rats en réponse aux événements réels et les récits de *Smart Rat City*, comme une basse continue suivant son propre déroulement.
- 27 Cette programmation peut sembler complexe mais elle allège beaucoup les contraintes qui pèsent sur l'écriture. Je peux écrire des fragments ou des séquences, de longueur diverse. Tous peuvent s'intégrer dans le dispositif. Elle devrait également contribuer à libérer la lecture de la pression machinique, mais ceci reste à confirmer par la confrontation de l'œuvre avec des lecteurs.

BIBLIOGRAPHIE

BARONI Raphaël, *Les rouages de l'intrigue : les outils de la narratologie postclassique pour l'analyse des textes littéraires*, Genève, Slatkine Érudition, 2017, p. 40.

BOUCHARDON Serge, *La valeur heuristique de la littérature numérique*, Paris, Hermann, 2013.

BOUTET Danielle, « La création de soi par soi dans la recherche-crédation : comment la réflexivité augmente la conscience et l'expérience de soi », *Approches inductives*, 2018, vol. 5, n° 1, p. 289-310.

CHAMBEFORT Françoise, « Geste d'une création : tout ce que j'ai fait pour trouver ce que je n'avais pas cherché », journée d'études *Écritures numériques, nouvelles gestualités ?* [en ligne], UTC, Compiègne, 2022, URL : https://www.youtube.com/watch?v=Eg8QnOM_YZE [consulté le 8 avril 2024].

CHAMBEFORT Françoise, *Mimésis du flux, exploration des potentialités narratives des flux de données* [en ligne], thèse de doctorat, sous la direction de Ioan Roxin,

université Bourgogne Franche-Comté, 2020, URL : <https://tel.archives-ouvertes.fr/tel-02991589> [consulté le 8 avril 2024].

REVAZ Françoise, *Introduction à la narratologie action et narration*, Bruxelles, De Boeck, 2009.

SAEMMER Alexandra, « Réflexions sur les possibilités d'une "recherche-création" désinstrumentalisée », *Hermès – la revue*, 2015, n° 2, p. 198-205.

NOTES

- 1 Françoise Chambefort, *Mimèsis du flux, exploration des potentialités narratives des flux de données*, université Bourgogne-Franche-Comté, 2020.
- 2 Serge Bouchardon, *La valeur heuristique de la littérature numérique*, Paris, Hermann, 2013.
- 3 Alexandra Saemmer, *Réflexions sur les possibilités d'une « recherche-création » désinstrumentalisée*, *Hermès, La Revue*, n° 2, 2015.
- 4 À retrouver en ligne sur : <https://rc.hypotheses.org> [consulté le 8 avril 2024].
- 5 Danielle Boutet, « La création de soi par soi dans la recherche-création : comment la réflexivité augmente la conscience et l'expérience de soi », *Approches inductives* [en ligne], vol. 5, n° 1, 2018, DOI : [10.7202/1045161ar](https://doi.org/10.7202/1045161ar).
- 6 À retrouver en ligne sur : <http://fchambef.fr/lucette/intro.php> [consulté le 8 avril 2024].
- 7 Françoise Revaz, *Introduction à la narratologie action et narration*, Bruxelles, De Boeck, 2009.
- 8 Accessible sur : <https://opendata.cityofnewyork.us/> [consulté le 8 avril 2024].
- 9 Françoise Chambefort, « Geste d'une création : tout ce que j'ai fait pour trouver ce que je n'avais pas cherché », journée d'études *Écritures numériques, nouvelles gestualités ?*, Compiègne, UTC, 2022.
- 10 Voir l'œuvre en ligne sur : http://fchambef.fr/smart_rat/master/index.php [consulté le 8 avril 2024].
- 11 Françoise Chambefort. *Mimèsis du flux, exploration des potentialités narratives des flux de données*. Sciences de l'information et de la communication. Université Bourgogne Franche-Comté, 2020, p. 93

12 Raphaël Baroni, *Les rouages de l'intrigue : les outils de la narratologie postclassique pour l'analyse des textes littéraires*, Genève, Slatkine Érudition, 2017.

13 Exemple de programmation du mode de la basse :
Si dernierEv.includes (“rodent”) ou (“contrôle”) ou (“bait”)

mode_basse = 2

Sinon

Si dernierEv.includes(“dirty”) ou (“hydrant”)

mode_basse = 0;

Sinon mode_basse = 1

14 Le texte a une longueur 25 000 caractères, soit deux fois et demie la longueur du texte de Lucette.

RÉSUMÉS

Français

Smart Rat City est une œuvre de littérature numérique qui s'appuie sur des flux de données du système d'information de la ville de New York. Dans ce type d'œuvre, le récit est programmé informatiquement selon des règles définies par l'auteur·e. Elle a été produite dans le cadre d'un projet de recherche-crédation interrogeant la manière de concilier un récit d'une certaine ampleur avec la contingence, apportée par les données issues du monde réel. Une solution a été trouvée en rythmant la narration par l'alternance de deux voix et en faisant osciller la tension narrative entre surprise et curiosité.

English

Smart Rat City is a work of digital literature that draws upon data streams from the New York City information system. In this type of work, the narrative is programmatically generated according to rules defined by the author. It was produced as part of a research-creation project that examines how to reconcile a narrative of a certain scope with the contingency brought by data from the real world. A solution was found by structuring the narration through the alternating voices of two characters and by oscillating the narrative tension between surprise and curiosity.

INDEX

Mots-clés

littérature numérique, narration, recherche-crédation, flux de données

Keywords

digital literature, narration, research-creation, data streams

AUTEUR

Françoise Chambefort

Chercheuse associée au laboratoire ELLIADD, université de Franche-Comté

IDREF : <https://www.idref.fr/149613504>

ORCID : <http://orcid.org/0000-0001-6379-9372>

HAL : <https://cv.archives-ouvertes.fr/francoise-chambefort>

Données/réel

Bases de données et littérature numérique (belge) : une perspective décoloniale

Databases and (Belgian) digital literature: a decolonial perspective

Isabelle Gribomont et Anaïs Berck

DOI : 10.35562/marge.863

Droits d'auteur

CC BY-NC-SA

PLAN

Colonialité et données

Colonialité, base de données et littérature numérique

De l'importance du contexte dans la réception des œuvres d'Anaïs Berck

Une base de données de littérature numérique « située » ?

TEXTE

- 1 Lorsque la bibliothèque royale de Belgique (KBR) se préparait à entamer un projet de recherche visant à recenser, archiver et étudier la littérature numérique belge, une des questions de recherche formulées fut la suivante : Quelles sont les particularités de la littérature numérique belge et en quoi diffère-t-elle de la production d'autres pays¹ ? Cette question, loin d'être reléguée à un stade avancé du projet, est centrale au développement d'un schéma de métadonnées qui permettrait de décrire adéquatement les œuvres indexées. Simultanément, un répertoire de littérature numérique belge devrait pouvoir être interopérable avec d'autres bases de données déjà en activité. Pour cela, un accent sur les spécificités locales semble contre-productif.
- 2 Le collectif de littérature numérique Anaïs Berck, un des acteurs majeurs dans le paysage de la littérature numérique belge, intègre la notion de données depuis une perspective décoloniale. En même temps, il se situe dans une démarche artistique holistique qui brouille les frontières entre expérimentation, création, recherche et édition. Ses œuvres, qui résistent à être décrites via les normes

bibliographiques d'usage, posent des défis particuliers permettant d'alimenter notre réflexion sur l'utilisation de bases de données dans le champ de la littérature numérique, et particulièrement, sur la tension entre spécificité locale et abstraction globale qui nous préoccupe. Premièrement, nous exposons les liens entre la colonialité et la pratique d'homogénéisation et de déracinement contextuel favorisée par la structuration de la connaissance dans des bases de données. L'œuvre d'Anaïs Berck problématise ce procédé dans le contexte de la violence épistémique exercée sur la nature, notamment dans le cadre de la colonisation belge. Ensuite, en nous appuyant sur la démarche expérimentale d'édition et de publication, proposée par Anaïs Berck au travers du projet *Algoliterary Publishing House* (APH), nous élargissons cette réflexion aux bases de données de littérature numérique, qui, elles aussi, deviennent des instruments d'abstraction du champ de la littérature numérique. Finalement, en analysant l'œuvre d'Anaïs Berck *Walk along the Trees of Madrid* et en prenant en considération un contexte d'énonciation de plus en plus large, nous questionnons la focalisation par défaut sur l'œuvre souvent de mise dans ces bases de données.

- 3 Anaïs Berck est un collectif autodéfini comme un ensemble d'êtres humains, d'algorithmes et d'arbres. Il explore l'espace formé entre l'intelligence humaine, végétale et artificielle. Il produit des œuvres de littérature numérique, qui, notamment, génèrent automatiquement des textes où les arbres occupent une position auctoriale tout en explorant le concept de maison d'édition décoloniale.
- 4 La pensée décoloniale vise à remettre en question la géopolitique du savoir hégémonique et totalitaire qui s'est établi depuis le ^{xvi}^e siècle au travers du colonialisme de conquête, et, plus largement, de la *logique de la colonialité*, qui a permis à la modernité occidentale de revendiquer l'universalité de son savoir théologique, scientifique et philosophique². Le procédé de la décolonialité s'ancre, entre autres, dans divers projets politiques et éthiques qui s'orientent vers une vision du monde pluriverselle, pensée depuis des espaces et des corps marginalisés, plutôt qu'universelle et basée sur un universalisme occidental abstrait³. Sur la page web de l'APH, de nombreuses pistes et questions de recherche sont suggérées : Comment pouvons-nous prendre en compte notre position

géopolitique et corporelle lors de la conception et la théorisation d'objets informatiques ? Comment contribuer à déconstruire les narrations excluantes, invisibilisantes ou violentes ? Comment critiquer des jeux de données structurées dont les normes sont héritées de la colonisation⁴ ? Les préoccupations d'Anaïs Berck font écho à des débats récents sur la décolonisation des humanités numériques, et, particulièrement, la décolonisation des archives ainsi que les biais et silences dans la production du savoir et des récits historiques⁵.

- 5 La littérature numérique a une relation particulière avec ces questions. D'une part, bien que la production d'œuvres de littérature numérique, en tant que formes d'exploration de l'*apparatus* digital contemporain, puisse elle-même être considérée comme une pratique appartenant aux humanités numériques, la littérature numérique s'est souvent située en périphérie du champ⁶. D'autre part, la littérature numérique a fréquemment été comprise comme une forme de critique, réflexion et défamiliarisation des outils et technologies numériques⁷. De cette façon, elle s'est parfois placée en opposition vis-à-vis des humanités numériques, qui, jusqu'il y a peu, problématisaient rarement leurs outils et méthodes d'investigation dans leurs contextes historiques et néocoloniaux⁸.

- 6 Les humanités numériques et la littérature numérique reposent toutes deux sur de multiples transformations de matériels sources en données ou constituants lisibles par un ordinateur. En ce qui concerne les humanités numériques en particulier, le concept de données est omniprésent. Les données (*data*) ont été définies par Rob Kitchin comme « la matière première produite par l'abstraction du monde en catégories, mesures et autres formes de représentation [...] qui constituent les éléments de base à partir desquels l'information et la connaissance sont créées⁹ ». La « datafication », en humanités numériques comme ailleurs, est donc un procédé d'abstraction qui transforme une réalité, un procédé ou une représentation complexe en données structurées. Cette réduction de dimensionnalité drastique se fait au travers d'un prisme préétabli, comme, par exemple, une ontologie de métadonnées.

Colonialité et données

- 7 L'œuvre centrée sur la nature en tant que sujet proposée par Anaïs Berck relève la continuité entre la colonialité et l'abstraction nécessaire à la création de données. Par exemple, *La Botanique coloniale wikifiée*¹⁰ est une proposition de recherche de l'altérité dans l'encyclopédie en ligne *Wikipedia*. Les arbres non occidentaux n'ont pas seulement été déplacés pendant la période coloniale, ils ont aussi été rebaptisés par les Européens, en utilisant le système de classification de Linné, dont les noms latins sont encore aujourd'hui la norme mondiale. Les utilisations médicinales, comestibles et matérielles des arbres furent ainsi marchandisées et les jardins botaniques furent créés comme composants de la politique d'exploitation économique coloniale. *La Botanique coloniale wikifiée* montre comment *Wikipedia* présente des arbres majeurs provenant de différents continents. Dans le même esprit, l'œuvre *Rewilding specimens*, quant à elle, comprend une série de portraits de taxonomistes, recouverts par des spécimens de plantes et arbres qu'ils ont renommés entre 1885 et 1960 au Congo belge¹¹.

Figure 1 : Portrait du botaniste John Lindley, par Anaïs Berck.



© Anaïs Berck

- 8 L'œuvre d'Anaïs Berck souligne que l'abstraction de la nature — c'est-à-dire la tendance à l'appréhender via une taxonomie plutôt que comme un ensemble de spécimens uniques et divers — nous permet de la concevoir en termes d'unités interchangeable et remplaçables, donc comme des « ressources ». Des chercheurs décoloniaux tels que Camilla Townsend et Walter Mignolo ont situé l'essor de l'abstraction comme une conséquence de la conquête des Amériques¹². La sphère de la « nature », comprenant les femmes et les peuples supposés primitifs, était considérée comme l'illustration d'un stade antérieur du développement. Leur déficit de rationalité invite à la conquête et à la réorganisation rationnelle taxonomiques¹³. Cette cosmovision reposant sur l'abstraction conduit à un sentiment illusoire de notre propre indépendance ontologique et écologique¹⁴. La reconception de la nature dans des termes agentiels, telle que proposée par Anaïs

Berck, contribue à questionner et déconstruire cette appréhension à la fois abstraite et marchande de la nature.

- 9 Au-delà du traitement de la nature en particulier, la tendance à l'universalité de l'épistémologie moderne laisse de côté l'expérience vécue et la mémoire au profit d'un système de classification rigide promouvant un modèle de pensée généralisant et homogénéisant. En reléguant à l'arrière-plan la particularité, le lieu et le récit en tant que facteurs de la pensée et de la vie humaine, les cadres colonisateurs font des lieux de simples instruments passifs ou des surfaces neutres pour l'inscription de projets humains¹⁵. Ricardo Padrón explique comment les cartographes du « Nouveau Monde » ont rationalisé le monde connu grâce aux principes de la géométrie euclidienne : « dans l'ordre de l'abstraction, tout est compris comme un lieu ou un objet à l'intérieur de cet espace¹⁶ ». En conséquence, toute chose se prête à une compréhension systématique, à une réification, à une appropriation ou à une subordination de la part d'un sujet qui l'observe¹⁷.

Colonialité, base de données et littérature numérique

- 10 Cette représentation du monde réel via un procédé d'abstraction et de division en unité de sens est intimement liée au concept de base de données. Les taxonomies, ontologies et schémas de métadonnées sont conçus entre autres, pour permettre de manipuler ce que nous considérons être l'essence d'un objet. Pour reprendre les termes de Ramesh Srinivasan, les bases de données « prennent des actions, des événements et des pratiques spécifiques et les résument en données indexables et comparatives¹⁸ ». Ce faisant, elles filtrent également ce qui ne peut pas être aligné aux protocoles de classification existants.
- 11 Sans être intrinsèquement « coloniales », les bases de données, de par leur nature, ont tendance à homogénéiser les objets qui en font partie, à les décontextualiser de leur environnement d'origine, et à les présenter au travers d'un prisme unique¹⁹. Or, la pensée décoloniale nous rappelle que nous parlons toujours depuis une localisation particulière, et non depuis un point de vue neutre, objectif et universel. D'après Ramón Grosfoguel, « la neutralité et l'objectivité

désincarnées et non localisées de l'ego-politique de la connaissance est un mythe occidental²⁰ ». La conceptualisation de la partialité et de la subjectivité de la connaissance n'est pas la prérogative de la décolonialité²¹. Cependant, au travers de la pensée décoloniale, les pratiques centrées sur les données (*data-centric*) et leur traitement au sein de l'environnement numérique ne sont pas contextualisées en utilisant les quelques décennies qui nous séparent du développement du web, mais en prenant en considération les cinq siècles passés²². Ce cadre nous offre des outils supplémentaires pour interroger l'histoire de l'informatique, du développement des bases de données et des procédés épistémiques qu'elles mettent en place. En particulier, le procédé d'abstraction mis en œuvre dans les bases de données peut être contextualisé via les liens épistémiques et historiques entre la technologie numérique et les paradigmes hégémoniques eurocentriques issus de la modernité²³.

- 12 Dans le champ de la littérature numérique, la base de données a toujours été considérée comme un outil partial. Comme mentionné ci-dessus, le milieu est particulièrement attentif à la non-neutralité des technologies numériques. La capacité des typologies et ontologies à déterminer la manière donc nous percevons, accédons et interprétons les connaissances apparaît comme une évidence²⁴. Une revue détaillée des différentes bases de données de littérature numérique dépasse le cadre de cet article. Cependant, deux exemples nous permettront d'illustrer la position du champ de recherche vis-à-vis des bases de données.
- 13 La base de connaissances ELMCIP (*Electronic Literature Knowledge Base*) crée des liens entre les œuvres de littérature numérique et la recherche portant sur ces œuvres de façon à modeler le champ dans son ensemble, au sein duquel recherche et création émergent souvent de contextes voisins. La plateforme est également collaborative puisqu'un compte d'utilisateur permet d'ajouter et de modifier des entrées dans la base de connaissance. Elle dispose de vocabulaires contrôlés pour les informations bibliographiques de base, mais, pour les thèmes et les descriptions de contenu, elle utilise un système de tags *folksonomiques* où chaque contributeur peut utiliser un tag de son choix, qu'il existe déjà dans la base de données ou non. Ce système est moins « efficace » puisque qu'un même concept peut être annoté avec des tags différents et que

des tags mentionnés une seule fois ne contribuent pas à améliorer les possibilités de découverte des œuvres qui peuplent la base de données. Cependant, cette approche a été préférée précisément pour éviter de limiter les voies d'entrée dans les œuvres en imposant une typologie fixe²⁵.

- 14 La base de données Rhizome ArtBase est une archive d'œuvres d'art numériques. Certaines entrées comportent des liens externes et d'autres des copies, archivées et hébergées par Rhizome. Parce qu'une œuvre d'art en réseau évolue au fil du temps — que ce soit en raison d'actions intentionnelles du ou des créateurs de l'œuvre, de l'intervention du personnel institutionnel chargé de préserver ou d'exposer l'œuvre, ou de changements structurels dans le logiciel ou les composants de l'œuvre —, cette archive capture différentes « variantes » d'une œuvre. Un élément clé du modèle d'enregistrement est donc la possibilité de relier une œuvre à toutes ses variantes. De plus, le modèle de données de l'ArtBase a explicitement été conçu pour tenir compte de multiples acteurs impliqués dans la création d'une œuvre sur Internet, en contradiction avec la conception traditionnelle (occidentale) de l'archivage et de la conservation de l'artiste en tant que créateur unique d'un objet unique²⁶.

- 15 Ces deux initiatives illustrent depuis des perspectives distinctes une volonté de transparence, collaboration et questionnement épistémique relatif aux enjeux de la création d'un schéma de métadonnées. Ces efforts, tout comme d'autres bases de données telles que NT2, le Electronic Literature Directory, Weblitt et le Répertoire des écritures numériques, ont été intégrés dans des projets de recherche où les choix de métadonnées étaient souvent fonction de questions et d'intérêts de recherche spécifiques²⁷. L'objectif n'est donc pas de créer une représentation *objective* ou *neutre* du corpus. De plus, la nature illusoire de la représentation fixe est particulièrement évidente dans la littérature numérique, car les médias utilisés par les créateurs évoluent presque aussi vite que le paysage numérique lui-même.

De l'importance du contexte dans la réception des œuvres d'Anaïs Berck

- 16 Cependant, des œuvres, telles que celles proposées par Anaïs Berck, nous amènent à pousser plus loin le questionnement des récits créés par les bases de données de littérature numérique. Lorsqu'une œuvre est placée dans une base de données, un procédé de déracinement et, via le modèle de métadonnées, d'abstraction se produit. Les clés de lectures contextuelles sont effacées et remplacées par un schéma de métadonnées préexistant. Pour reprendre les mots de Couldry et Mejias, « l'abstraction consiste à faire disparaître les textures de la vie sociale²⁸ ». En lisant l'œuvre d'Anaïs Berck intitulée *Walk along the Trees of Madrid* en cercles concentriques, englobant un contexte de plus en plus large, les éléments constitutifs de l'œuvre, qui semblent initialement suivre une formule éprouvée du genre de la littérature numérique, prennent progressivement de nouvelles significations qui repoussent l'enveloppe des créations numériques pour intervenir dans la conversation autour de la crise climatique et de la décolonialité²⁹.
- 17 Cette œuvre utilise un algorithme pour créer simultanément un poème et un itinéraire de promenade dans le quartier de Las Letras, à Madrid. Elle peut être lue en anglais et en espagnol et s'appuie sur des romans des auteurs espagnols du XIX^e siècle Emilia Pardo Bazán et Benito Pérez Galdós. Le premier mot du poème est l'article indéfini « un » et les mots suivants sont générés par l'algorithme jusqu'à l'obtention d'un point. La promenade est générée à partir de la base de données « Un Alcorque, un Árbol », qui indexe les arbres de Madrid. Chaque nom, adjectif, verbe ou adverbe significatif est associé à un arbre du quartier de Las Letras. Dans le PDF généré après le choix du roman, le lectorat est invité à se promener dans le quartier en récitant des parties du poème à chaque arbre le long de la promenade produite par le poème.
- 18 Les différents éléments qui interagissent pour créer cette œuvre – les choix du lectorat, l'algorithme, le poème concret généré, les textes canoniques utilisés comme données, la base de données reliant

l'œuvre au monde matériel et le mouvement potentiel du lectorat médiatisé par l'œuvre — sont des composantes essentielles du genre de la littérature numérique. Prise isolément, l'œuvre est fortement autoréférentielle et propose une réflexion sur le langage, l'auctorialité, le « canon » littéraire et la manière dont les perceptions de la réalité sont constamment médiatisées par le texte.

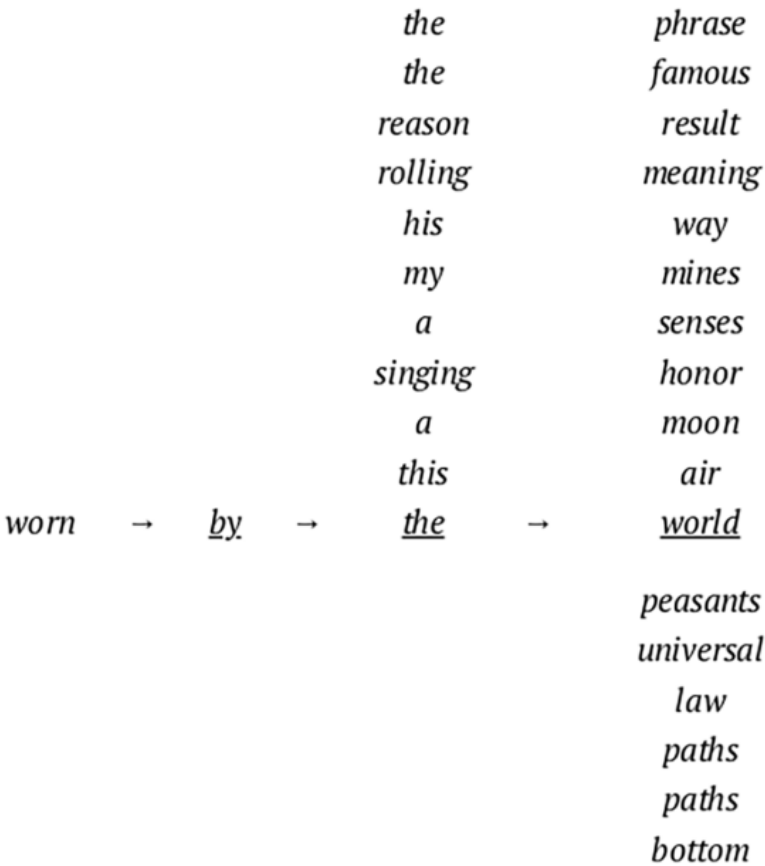
- 19 Une autre couche sémantique est ajoutée au poème produit grâce aux informations contextuelles fournies sur l'algorithme des chaînes de Markov dans le PDF. Les chaînes de Markov sont un algorithme de probabilité basé sur la notion que les événements passés déterminent en partie le présent. Pour Markov, le langage est un exemple de système où les probabilités statistiques peuvent être modélisées. Cette documentation, qui comprend également le code informatique complet de l'œuvre, indique explicitement que l'approche statistique de Markov en matière de modélisation et de génération de langage a ouvert une nouvelle ère pour le traitement automatique du langage naturel qui est encore d'actualité aujourd'hui. Cette historicisation contribue à la dimension autoréférentielle de l'œuvre, qui se présente comme une filiation directe des racines mêmes de la recherche en traitement automatique de la langue et souligne le rôle du langage dans l'histoire de l'informatique algorithmique.
- 20 Contrairement aux chaînes de Markov, le rôle des arbres n'est pas explicitement contextualisé au sein de l'œuvre. Leur présence semble créer un lien matériel entre le Madrid de Pérez Galdós et Pardo Bazán et celui du lectorat. Elle souligne que notre perception des lieux qui nous entourent a été façonnée, non seulement par nos expériences physiques, mais aussi par nos expériences textuelles. Tout comme ce langage littéraire est remédiatisé grâce à une chaîne de Markov de façon irréaliste, mais probable, leurs pas le sont également. Ces promenades et ces textes nouveaux agissent comme une allégorie du principe de la chaîne de Markov, selon lequel les éléments du passé, contenus dans le présent, influent sur la probabilité qu'un événement se produise dans le futur. Ces textes écrits un siècle auparavant régissent la probabilité des directions successives prises par le lectorat d'aujourd'hui.
- 21 Cependant, l'éthique critique et décoloniale de l'œuvre — qui regarde vers l'extérieur plutôt que vers l'intérieur — n'est disponible que si l'on

considère son contexte de publication, c'est-à-dire l'Algoliterary Publishing House (APH), sur le site de laquelle cette œuvre est publiée. L'APH est un projet de recherche artistique qui vise à décentrer la perspective de l'être humain et à exploiter les algorithmes, non pas à des fins d'extraction de ressources ou de gains financiers, mais pour faire corps avec la nature et remettre en question les hiérarchies (culturelles) coloniales³⁰. Les questions de recherche présentées sur le site du projet témoignent d'une volonté de repenser les concepts d'auteur, d'informatique (algorithmique) et de crise climatique à la lumière de la pensée décoloniale, et au-delà du débat théorique, de créer un modèle alternatif de maison d'édition qui serait en symbiose avec la nature. Les ouvrages publiés par l'APH interrogent la représentation de l'arbre dans les textes et les données en réutilisant de manière créative des données analogiques liées à l'arbre (température, lumière, humidité) et des architectures de données actuelles (herbiers coloniaux, textes législatifs, philosophie, littérature). Ils visent également à interroger la surabondance de résultats créatifs dans la littérature algorithmique et ses implications pour la lisibilité, la valeur littéraire et l'impact environnemental³¹.

- 22 La recontextualisation du matériel existant est un élément clé de nombreuses œuvres de littérature numérique. Notamment en raison de ses racines dans la poétique combinatoire et de son développement dans un écosystème numérique favorisant les pratiques d'échantillonnage, les œuvres numériques littéraires trouvent souvent leur sens dans les écarts entre le contexte original des données et leurs manifestations dans une nouvelle création. Parallèlement, en raison de cette affinité avec une recontextualisation génératrice de sens et, une fois encore, leurs modes de circulation numériques, les œuvres de littérature numérique elles-mêmes sont particulièrement sensibles aux glissements sémantiques lorsqu'elles sont déplacées dans de nouveaux contextes.
- 23 Lorsque les circonstances de sa création et son lieu de diffusion sont pris en considération, les éléments constitutifs de l'œuvre décrite ci-dessus acquièrent de nouvelles significations. En consultant les questions de recherche proposées sur la plateforme de l'APH, nous comprenons que les chaînes de Markov, qui contribuent de prime abord à la dimension métatextuelle de l'œuvre, font en réalité partie

d'une démarche extratextuelle liée à des enjeux sociétaux contemporains. Les principes méthodologiques qui animent les activités d'Anaïs Berck incluent le besoin d'implémenter les algorithmes physiquement et métaphoriquement parce qu'« un algorithme doit être vu pour être cru ³² ». L'utilisation de cet algorithme, ainsi que sa représentation concrétisée dans les fichiers PDF, créés à chaque « lecture » de l'œuvre, contribuent à une pratique qui cherche à la fois à défamiliariser l'usage de ces algorithmes en les utilisant dans des contextes différents de ceux pour lesquels ils ont été conçus, et démystifier leur fonctionnement en détaillant leur structure, ainsi que leur histoire et le contexte de leur émergence. Ainsi, la représentation graphique de la génération du texte grâce aux chaînes de Markov est un exemple de poésie concrète, qui connecte l'œuvre aux antécédents littéraires dadaïstes de la littérature numérique, et devient une démarche didactique visant à représenter graphiquement le procédé exécuté par le code grâce à une boucle générative (voir figure 2).

Figure 2 : Extrait d'un PDF généré par l'œuvre d'Anaïs Berck *Walk Along the Trees of Madrid* représentant le procédé de génération du poème via les choix aléatoires consécutifs d'un mot parmi une liste de possibilités.



Licence CC BY

24 La place des arbres dans l'œuvre est également éclaircie par la relation de l'APH au concept d'auctorialité. Les métadonnées relatives aux autrices et auteurs sont divisées en trois catégories : algorithmes, arbres et humains. La catégorie « humains » inclut Pardo Bazán et Pérez Galdós, ainsi que les personnes qui ont contribué à la conception et à l'implémentation de l'œuvre. L'idée que la littérature numérique remet en cause la notion même de « paternité littéraire » n'est pas nouvelle. Scott Rettberg souligne que cela n'est pas seulement dû à la nature collaborative du développement de logiciels et à l'engagement d'un public participatif, mais aussi au fait que « la créativité est le produit de relations sociales et de contextes culturels plus que de l'inspiration³³ ». Bien que l'édition de livres ait toujours été une affaire collégiale, impliquant des éditeurs, des typographes et

des concepteurs, les couches successives de création sont rendues visibles dans l'écosystème numérique et les pratiques artistiques qu'il encourage³⁴. Le décentrement de l'auteur, tout comme le décentrement de l'humain, est un sujet prépondérant dans les débats décoloniaux. L'appropriation et le mépris sont deux attitudes face aux connaissances indigènes qui ont été facilitées par l'importance attribuée à la paternité intellectuelle comme instrument de légitimation de la création de la connaissance³⁵. À nouveau, en consultant les questions posées sur la page web de l'APH, le décentrement de l'auteur évoqué par Rettberg n'évoque plus autant une déhiérarchisation postmoderne du statut de l'auteur qu'un questionnement radical à la fois épistémologique et pratique des rapports entre les algorithmes, les humains et la nature dans le contexte de l'édition et de l'édition numérique. La démarche implique de donner une « voix » aux arbres, qui ont été réduits à la qualité de ressources, notamment lors de la colonisation, et dont la singularité et la place unique dans l'écosystème ont été ignorées.

- 25 De même, le statut auctorial de l'algorithme des chaînes de Markov devient une proposition plus radicale lorsque nous comprenons que la cooptation des algorithmes dans un procédé décolonial fait partie des principes de l'APH. La question de recherche « Et si les algorithmes s'engageaient en faveur d'un monde plus égalitaire³⁶? » suggère que, au-delà d'une appropriation situant l'algorithme dans son contexte géopolitique tout en le sortant de son contexte d'utilisation actuel, l'œuvre *Walk Along the Trees of Madrid* s'inscrit également dans une démarche visant à distancer les algorithmes de l'éthique et de l'économie néocoloniales auxquelles ils sont souvent associés.
- 26 Au-delà de l'interprétation de l'œuvre, sa multiplicité et sa matérialité sont également le sujet de questionnements. C'est dans la multiplicité du contenu, en opposition avec le produit « unique » et définitif, qui définit le plus souvent l'œuvre littéraire, que la littérature générative trouve une partie de son sens³⁷. Anaïs Berck demande quand ces générations multiples deviennent du bruit, comment l'abondance de la production algorithmique peut être rendue lisible pour l'être humain, si la lecture perd de sa valeur quand il existe un million de versions légèrement différentes d'un livre, et s'il est possible ou intéressant de se décider sur la version « statique » d'un livre³⁸. La

mise en page évocatrice du code, l'existence de copies uniques sur les ordinateurs des lecteurs, le choix d'un outil de génération de fichiers PDF, et l'impact écologique de la génération de livres digitaux sont autant de questions posées par Anaïs Berck qui apportent de nouvelles dimensions de lecture à l'œuvre analysée ici, mais qui sont également rendues concrètes par son existence. La boucle de rétroaction entre les œuvres et la recherche est mise en évidence sur la plateforme de publication de l'APH.

- 27 Finalement, la genèse et les antécédents communautaires d'Anaïs Berck sont également une clé de lecture importante. Algolit et Constant sont deux collectifs basés à Bruxelles qui font partie de l'écosystème d'Anaïs Berck. Algolit est un collectif de création littéraire numérique qui utilise le code informatique et les affordances de l'environnement numérique pour expérimenter de nouveaux modes d'écritures et de lecture³⁹. Constant est une organisation d'artistes active au croisement entre l'art, les médias et la technologie⁴⁰. Le collectif développe, étudie et expérimente l'environnement digital contemporain. Par exemple, Constant examine les discriminations et les problèmes structurels inhérents à la technologie, notamment en ce qui concerne la marchandisation des données personnelles liées aux comportements des utilisateurs dans l'espace numérique. Or, Anaïs Berck note que le profil des personnes, établi par les algorithmes commerciaux, est similaire à la façon dont la nature est considérée : elle présente principalement un intérêt en tant que collectifs dotés de caractéristiques spécifiques optimisées pour l'exploitation commerciale. Des procédés de continuité, d'inspiration et de fertilisation, entre Anaïs Berck et ces différents collectifs sont explicitement envisagés sur la plateforme de l'APH. Cette maison d'édition, les œuvres qui sont créées dans son contexte et les réflexions qui les entourent sont le produit d'une activité collective, en cours depuis plus d'une décennie.

Une base de données de littérature numérique « située » ?

- 28 Cette lecture itérative de *Walk Along the Trees of Madrid*, qui montre que lire cette œuvre indépendamment de son contexte entraîne un glissement de sens majeur, est elle-même subjective et dépendante

d'une instance d'énonciation particulière. Bien qu'elle n'ait pas été proposée par des membres d'Anaïs Berck, cosignataire de cet article, cette lecture est conditionnée par des échanges avec le collectif. Cependant, considérer la plateforme de publication, puis le contexte créatif et idéologique de sa conception permet de faire une lecture à la fois subjective et « située » de l'œuvre, et d'accéder à de nouvelles clés d'interprétation. Alors que le livre imprimé, par exemple, est souvent conceptualisé comme un produit fini, conçu pour être lu avec un certain détachement face à son contexte de production et de publication, le cas de l'APH met en évidence l'enchevêtrement procédural auquel les œuvres littéraires numériques se prêtent fréquemment. Séparer l'œuvre de sa plateforme de publication et de l'écosystème communautaire à l'origine de sa conception place des frontières arbitraires dans un mode de création-recherche itératif et en constante évolution. Comme le souligne le modèle de métadonnées développé par Rhizome, la focalisation sur une œuvre définitive et publiée n'est pas toujours pertinente. Cet angle de perception nous pousse à inclure ce qui correspond suffisamment à notre compréhension d'une publication, et à exclure d'autres types de productions. De plus, il dénature ces œuvres en donnant l'impression qu'elles ont été créées, puis publiées comme telles, alors qu'elles font partie d'un continuum avec d'autres pratiques formant un mode de fonctionnement et de réflexion vivant.

- 29 Selon Anaïs Berck, le procédé est aussi important que l'œuvre elle-même et une œuvre de littérature numérique est elle-même un procédé. Pour documenter leur travail, la notion d'œuvre devient une limitation. Que faut-il décrire ? Le collectif lui-même et ses pratiques, incluant, par exemple, des résidences expérimentales et des promenades ayant pour but de décentrer la position de l'être humain par rapport à la nature ; la plateforme de publication Algoliterary Publishing House, qui est une expérimentation en flux constant sur les enjeux techniques, esthétiques et sociaux d'une maison d'édition numérique, décolonisée et respectueuse de l'environnement ; ou les œuvres qui y sont publiées et dont la création est intimement liée à la plateforme et aux activités du collectif ?
- 30 En fin de compte, nos perceptions et notre compréhension du champ que nous essayons d'abstraire, dans ce cas, la littérature numérique, sont fondamentalement ancrées dans les archives produites et dans

le texte narratif que nous construisons à partir d'elles. L'implication ici est que les données et les informations qui ne se conforment pas à un consensus de la « norme » sont difficiles à situer dans n'importe quel type de données structurées. Quand la littérature numérique est inséparable d'un champ d'action plus vaste, allant au-delà du fait littéraire, et s'inscrivant dans un tissu social activiste pluridisciplinaire, la représentation partielle et biaisée offerte par la base de données, bien que précieuse dans de nombreux contextes, court le risque de créer des angles morts dans notre compréhension du milieu. C'est particulièrement important compte tenu de la vulnérabilité de cette forme à l'obsolescence technologique. Quand une page web disparaît ou devient illisible sans avoir été préservée de façon pérenne, les descriptions des œuvres deviennent une forme d'archivage.

- 31 Nous devons nous efforcer de travailler avec cette tension qui existe entre les retombées positives de l'interopérabilité et de la découvrabilité des données et la tendance des données structurées à ossifier les pratiques de collectes de données. Comme souligné par Piraze Hacıgüzeller et *al.* dans le contexte de l'archéologie, les bases de données privilégient la similarité. Il est donc difficile d'encadrer les différences ; même l'extraction d'objets différents doit se faire en définissant un groupe d'objets similaires qui ne correspondent pas à ce que l'on recherche⁴¹. Le défi est de décrire les œuvres avec, de préférence, une typologie existante, afin de promouvoir éventuellement une interopérabilité avec d'autres répertoires de littérature numérique, mais en évitant, dans le même élan, de réduire artificiellement la multidimensionnalité de la collection et abstraire les œuvres de leurs contextes d'énonciation, de publication et de réception.
- 32 Cette observation nous encourage à poursuivre la voie tracée par l'ELMCIP, entre autres, qui ne hiérarchise pas les différentes entités présentes dans leur base de connaissances. Les données liées (*linked data*), adoptées par l'ELMCIP, permettent une flexibilité ontologique utile pour surmonter partiellement les partis pris inévitables dans la création de données structurées et contribuent à ancrer les œuvres dans leur contexte local et leur histoire. Dans un environnement de données liées, la conformité à une norme de métadonnées prédéterminée n'est pas une condition préalable à

l'interopérabilité avec d'autres ensembles de données⁴². Cet environnement facilite donc l'encodage d'informations divergentes, telles que les auctorialités non standards proposées par Anaïs Berck ainsi que des liens avec d'autres entités qui seraient pertinentes pour situer une œuvre dans son contexte social. Ce travail d'ancrage dans une réalité locale passe également par la collaboration avec les différents acteurs du champ de la littérature numérique dans une perspective plurielle, ainsi que la visibilisation du travail d'encodage de données et de documentation qui est déjà de mise dans plusieurs bases de données de littérature numérique. Cette visibilisation permet de souligner le caractère non neutre de ce travail dans la création et la distribution du savoir⁴³, ainsi que de créditer les personnes contributrices. En empruntant les mots de Nadine Schuurman, « une fois que les données sont placées dans des tableaux, leur lignée sociale est oubliée⁴⁴ ».

- 33 Il ne s'agit pas de rejeter l'utilisation des bases de données, ou d'autres outils numériques inévitablement enchevêtrés dans un système biaisé et normé, mais, comme Anaïs Berck le fait avec les algorithmes, de les décortiquer et de les approprier de façon critique. En partant de la décolonialité, nous pouvons questionner les contenus et acteurs de ces bases de données, leurs matérialités et leurs modes et formes de représentations épistémologiques tendant vers l'abstraction et la décontextualisation. Ce cadre nous encourage à poser la question proposée par l'APH : « Comment déplier les pratiques de codage : comprendre leurs structures, les histoires et les contextes dans lesquels elles s'inscrivent et les processus radicaux qu'elles exécutent⁴⁵ ? ».

BIBLIOGRAPHIE

ACKERMANS Hannah, « Appealing to Your Better Judgement: A Call for Database Criticism », *Electronic Book Review* [en ligne], 2020, DOI : [10.7273/97P6-PT89](https://doi.org/10.7273/97P6-PT89).

Algorit, « Main Page », *Algorit* [en ligne], URL : https://www.algorit.net/index.php?title=Main_Page [consulté le 8 avril 2024].

ANDERSON Jane, « Indigenous Knowledge and Intellectual Property Rights », in : James D. WRIGHT (dir.), *International Encyclopedia of the Social & Behavioural Sciences*, Amsterdam, Elsevier, 2015, p. 769-778, DOI : [10.1016/B978-0-08-097086-8.64078-3](https://doi.org/10.1016/B978-0-08-097086-8.64078-3).

- ANDERSON Jane, « Negotiating Who 'Owns' Penobscot Culture », *Anthropological Quarterly*, vol. 91, n° 1, 2018, p. 267-305, DOI : [10.1353/anq.2018.0008](https://doi.org/10.1353/anq.2018.0008).
- BERCK Anaïs, « An Algoliterary Publishing House », *Algoliterary Publishing* [en ligne], URL : <https://algoliterarypublishing.net/> [consulté le 8 avril 2024].
- BERCK Anaïs, « La Botanique coloniale Wikifiée », *Anaïs Berck* [en ligne], URL : <https://www.anaisberck.be/la-botanique-coloniale-wikifie/> [consulté le 8 avril 2024].
- BERCK Anaïs, « Research », *Algoliterary Publishing* [en ligne], URL : <https://algoliterarypublishing.net/research.html> [consulté le 8 avril 2024].
- BERCK Anaïs, « Rewilding specimens », *Anaïs Berck* [en ligne], URL : <https://www.anaisberck.be/rewilding-specimens-fr/> [consulté le 8 avril 2024].
- BERCK Anaïs, *Walk along the Trees of Madrid* [en ligne], URL : <http://paseo-por-arbores.algoliterarypublishing.net/en> [consulté le 8 avril 2024].
- BONNET Gilles, *Pour une poétique numérique : littérature et internet*, coll. « Savoir lettres », Paris, Hermann, 2017.
- BOOTZ Philippe, « Littérature numérique : un éphémère vrai-ment ? », *Hybrid*, n° 1, 2014, DOI : [10.4000/hybrid.1156](https://doi.org/10.4000/hybrid.1156).
- BOUCHARDON Serge et HECHMAN Davin, « Digital Manipulability and Digital Literature », *Electronic Book Review* [en ligne], 2012, URL : <https://electronicbookreview.com/essay/digital-manipulability-and-digital-literature/> [consulté le 8 avril 2024].
- CANDELA Gustavo, PEREDA Javier, SÁEZ Dolores, ESCOBAR Pilar, SÁNCHEZ Alexander, VILLA TORREZ Andrés, PALACIOS Albert A., McDONOUGH Kelly et MURRIETA-FLORES Patricia, « An Ontological Approach for Unlocking the Colonial Archive », *Journal on Computing and Cultural Heritage*, 2023, DOI : [10.1145/3594727](https://doi.org/10.1145/3594727).
- CHILTON Elizabeth et LABRADOR Angela M., « Re-locating Meaning in Heritage Archives: A Call for Participatory Heritage Databases », in : *Annual Meeting of Computer Applications to Archaeology* [en ligne], 2009, URL : <https://rex.libraries.wsu.edu/esploro/outputs/99900586061101842> [consulté le 8 avril 2024].
- Constant, « About », *Constant* [en ligne], URL : <https://constantvzw.org/site/-About-Constant-7-.html?lang=en> [consulté le 8 avril 2024].
- COULDRY Nick et MEJIAS Ulises Ali, *The Costs of Connection: How Data is Colonizing Human Life and Appropriating it for Capitalism*, Stanford, Stanford University Press, 2019.
- COULDRY Nick et MEJIAS Ulises Ali, « The Decolonial Turn in Data and Technology Research: What is at Stake and Where is it Heading ? », *Information, Communication & Society*, vol. 26, n° 4, 2021, p. 786-802.
- FLORES Leonardo, « Third Generation Electronic Literature », *Electronic Book Review*, 2019, DOI : [10.7273/axyj-3574](https://doi.org/10.7273/axyj-3574).

GERVAIS Bertrand, « The NT2 Hypermedia Art and Literature Directory: A New Knowledge Environment Devoted to the Valorization of Screen Culture », *Scholarly and Research Communication*, vol. 3, n° 2, 2012, DOI : [10.22230/src.2012v3n2a79](https://doi.org/10.22230/src.2012v3n2a79).

GRIGAR Dene et O'SULLIVAN James Christopher (dir.), *Electronic Literature as Digital Humanities: Contexts, Forms & Practices*, New York, Bloomsbury Academic, 2021.

GROSFOGUEL Ramón, « Decolonizing Post-Colonial Studies and Paradigms of Political-Economy: Transmodernity, Decolonial Thinking, and Global Coloniality », *Transmodernity: Journal of Peripheral Cultural Production of the Luso-Hispanic World*, vol. 1, n° 1, 2011.

GROSFOGUEL Ramón, « The Epistemic Decolonial Turn: Beyond Political-Economy Paradigms », *Cultural Studies*, vol. 21, n° 2-3, 2007, p. 211-223, DOI : [10.1080/09502380601162514](https://doi.org/10.1080/09502380601162514).

HACIGÜZELLER Piraye, TAYLOR James Stuart et PERRY Sara, « On the Emerging Supremacy of Structured Digital Data in Archaeology: A Preliminary Assessment of Information, Knowledge and Wisdom Left Behind », *Open Archaeology*, vol. 7, n° 1, 2021, p. 1709-1730, DOI : [10.1515/opar-2020-0220](https://doi.org/10.1515/opar-2020-0220).

HARAWAY Donna, « Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective », *Feminist Studies*, vol. 14, n° 3, 1988, p. 575-99.

HAYLES N. Katherine. *How we Think: Digital Media and Contemporary Technogenesis*, Chicago, The University of Chicago Press, 2012.

IKEDA Ryan, « Excavating Logics of White Supremacy in Electronic Literature: Antiracism as Infrastructural Critique », *Electronic Book Review*, 2021, DOI : [10.7273/CCTW-4415](https://doi.org/10.7273/CCTW-4415).

JIMENEZ Andrea, VANNINI Sara et COX Andrew, « A Holistic Decolonial Lens for Library and Information Studies », *Journal of Documentation*, vol. 79, n° 1, 2023, p. 224-244.

KITCHIN Rob, *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*, Londres, Sage, 2014.

KNUTH Donald E., *The Art of Computer Programming*, vol. 1 (3^e édition), Reading, Addison-Wesley, 1997.

KOZAK Claudia, « Experimental Electronic Literature from the Souths. A Political Contribution to Critical and Creative Digital Humanities », *Electronic Book Review*, 2021, DOI : [10.7273/ZD5G-ZK30](https://doi.org/10.7273/ZD5G-ZK30).

LABRADOR Angela M., « Ontologies of the Future and Interfaces for All: Archaeological Databases for the Twenty-First Century », *Archaeologies* vol. 8, n° 3, 2012, p. 236-49, DOI : [10.1007/s11759-012-9203-2](https://doi.org/10.1007/s11759-012-9203-2).

LUDWIG Günter et TAN Thanh, « Linked Data Query Processing Strategies », in : PATEL-SCHNEIDER Peter F., PAN Yue, HITZLER Pascal, MIKA Peter, ZHANG Lei, PAN Jeff Z., HORROCKS Ian, et GLIMM Birte (dir.), *The Semantic Web – ISWC 2010*, Berlin, Springer, 2010, p. 453-469, DOI : [10.1007/978-3-642-17746-0_29](https://doi.org/10.1007/978-3-642-17746-0_29).

LAHOUSTE Corentin, « La création littéraire numérique belge, un domaine émergent », *Nouveaux Cahiers de Marge*, n° 7, 2023, DOI : [10.35562/marge.723](https://doi.org/10.35562/marge.723).

MASIERO Silvia, « Decolonizing critical information systems research: A subaltern approach », *Information Systems Journal*, vol. 33, n° 2, 2023, p. 299–323.

McPHERSON Tara, « Why Are the Digital Humanities So White? Or Thinking the Histories of Race and Computation », in : GOLD Matthew K. (dir.), *Debates in the Digital Humanities*, Minneapolis, University of Minnesota Press, 2012, p. 139–160.

MIGNOLO Walter D. et TOSTANOVA Madina V., « Theorizing from the Borders: Shifting to Geo- and Body-Politics of Knowledge », *European Journal of Social Theory*, vol. 9, n° 2, 2006, p. 205–21, DOI : [10.1177/1368431006063333](https://doi.org/10.1177/1368431006063333).

MONJOUR Servanne, « La littérature numérique n'existe pas. La littérarité au prisme de l'imaginaire médiatique contemporain », *Communication & langages*, vol. 3, n° 205, 2020, p. 5–27, DOI : [10.3917/comla1.205.0005](https://doi.org/10.3917/comla1.205.0005).

PABLO Luis et GOICOCHEA María, « A Survey of Electronic Literature Collections », *CLCWeb: Comparative Literature and Culture*, vol. 16, n° 5, 2014, DOI : [10.7771/1481-4374.2558](https://doi.org/10.7771/1481-4374.2558).

PADRÓN Ricardo, *The Spacious Word: Cartography, Literature, and Empire in Early Modern Spain*, Chicago, University of Chicago Press, 2004.

PLUMWOOD Val, « Decolonizing Relationships with Nature », in : ADAMS William et MULLIGAN Martin (dir.), *Decolonizing Nature Strategies for Conservation in a Post-colonial Era*, London, Routledge, 2002, p. 51–78.

PURCELL Alyssa et SUTHERLAND Tonia, « A Weapon and a Tool: Decolonizing Description and Embracing Redescription as a Liberatory Archival Practice », *The International Journal of Information*, vol. 5, n° 1, 2021, p. 60–78.

QUIJANO Aníbal, « Coloniality and Modernity/Rationality », *Cultural Studies*, vol. 21, n° 2–3, 2007, p. 168–178, DOI : [10.1080/09502380601164353](https://doi.org/10.1080/09502380601164353).

RETTBERG Scott, « Documenting a Field: The Life and Afterlife of the ELMCIP Collaborative Research Project and Electronic Literature Knowledge Base », *Electronic Book Review*, 2021, DOI : [10.7273/KFMQ-7B83](https://doi.org/10.7273/KFMQ-7B83).

RETTBERG Scott, « Electronic Literature as Digital Humanities », in : SCHREIBMAN Susan, Siemens Ray et UNSWORTH John (dir.), *A New Companion to Digital Humanities*, Chichester (UK), John Wiley & Sons, 2015, p. 127–136, DOI : [10.1002/9781118680605.ch9](https://doi.org/10.1002/9781118680605.ch9).

RETTBERG Scott, « Finding a Third Space for Electronic Literature: Creative Community, Authorship, Publishing, and Institutional Environments », *Journal of Writing in Creative Practice*, vol. 4, n° 1, 2011, p. 9–23, DOI : [10.1386/jwcp.4.1.9_1](https://doi.org/10.1386/jwcp.4.1.9_1).

Répertoire des écritures numériques [en ligne], URL : <https://repertoire.ecrituresnumeriques.ca/s/repertoire/page/accueil> [consulté le 8 avril 2024].

RISAM Roopika, « Decolonizing the Digital Humanities in Theory and Practice », in : SAYERS Jentery (dir.), *The Routledge Companion to Media Studies and Digital Humanities* (1^{re} édition), New York, Routledge, Taylor & Francis Group, 2018, p. 78-86, DOI : [10.4324/9781315730479-8](https://doi.org/10.4324/9781315730479-8).

RISAM Roopika, *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*, Evanston, Northwestern University Press, 2019.

ROSSENOVA Lozana, « ArtBase Redesign Documentation–Data Models », *ArtBase Design Documentation* [en ligne], URL : <https://sites.rhizome.org/artbase-re-design/data-models.html> [consulté le 8 avril 2024].

SAEMMER Alexandra, « La littérature numérique entre légitimation et canonisation », *Culture & Musées*, n° 18, 2011, p. 201-223, DOI : [10.3406/pumus.2011.1635](https://doi.org/10.3406/pumus.2011.1635).

SCHUURMAN Nadine, « Database Ethnographies Using Social Science Methodologies to Enhance Data Analysis and Interpretation », *Geography Compass*, vol. 2, n° 5, 2008, p. 1529-1548, DOI : [10.1111/j.1749-8198.2008.00150.x](https://doi.org/10.1111/j.1749-8198.2008.00150.x).

SEIÇA Álvaro, « Electronic Literature and Online Literary Databases: The PO.EX and ELMCIP Cases », *RISS*, vol. 1, 2014, p. 89-102.

SRINIVASAN Ramesh, *Whose Global Village? Rethinking How Technology Shapes Our World*, New York, New York University Press, 2019.

TABBI Joseph, « Something There Badly Not Wrong: The Life and Death of Literary Form in Databases », *Electronic Book Review*, 2020, DOI : [10.7273/H5YF-KC59](https://doi.org/10.7273/H5YF-KC59).

TOWNSEND Camilla, « Burying the White Gods: New Perspectives on the Conquest of Mexico », *The American Historical Review*, vol. 108, n° 3, 2003, p. 659-687. DOI : [10.1086/ahr/108.3.659](https://doi.org/10.1086/ahr/108.3.659).

VERRAN Helen et CHRISTIE Michael, « Postcolonial Databasing? Subverting Old Appropriations, Developing New Associations », in : LEACH James and WILSON Lee (dir.), *Subversion, Conversion, Development: Cross-Cultural Knowledge Exchange and the Politics of Design*, Cambridge, MIT Press Scholarship Online, 2014.

« WEBLITT » [en ligne], URL : <https://weblitt.msh-lse.fr/> [consulté le 8 avril 2024].

WHITESIDE Kerry H., « Beyond the Nature-Culture Dualism: The Ecology of Earth-Homeland », *World Futures*, vol. 60, n° 5-6, 2004, p. 357-369, DOI : [10.1080/02604020490468311](https://doi.org/10.1080/02604020490468311).

NOTES

1 Pour un aperçu de la situation en Belgique francophone en regard de la création littéraire numérique, voir Corentin Lahouste, « La création littéraire numérique belge, un domaine émergent », *Nouveaux Cahiers de Marge*, n° 7, 2023, DOI : [10.35562/marge.723](https://doi.org/10.35562/marge.723).

- 2 Voir Aníbal Quijano, « Coloniality and Modernity/Rationality », *Cultural Studies*, vol. 21, n° 2-3, 2007, p. 168-178 ; Walter D. Mignolo et Madina V. Tostanova, « Theorizing from the Borders: Shifting to Geo- and Body-Politics of Knowledge », *European Journal of Social Theory*, vol. 9, n° 2, 2006, p. 205-21.
- 3 Ramón Grosfoguel, « The Epistemic Decolonial Turn: Beyond Political-Economy Paradigms », *Cultural Studies*, vol. 21, n° 2-3, 2007, p. 212.
- 4 Anaïs Berck, « An Algoliterary Publishing House » [en ligne], URL : <https://algoliterarypublishing.net/> [consulté le 8 avril 2024].
- 5 Voir par exemple, Gustavo Candela, Javier Pereda, Dolores Sáez, Pilar Escobar, Alexander Sánchez, Andrés Villa Torres, Albert A. Palacios, Kelly McDonough et Patricia Murrieta-Flores, « An Ontological Approach for Unlocking the Colonial Archive », *Journal on Computing and Cultural Heritage*, 2023 ; Roopika Risam, « Decolonizing the Digital Humanities in Theory and Practice », in : Jentery Sayers (dir.), *The Routledge Companion to Media Studies and Digital Humanities*, New York, Routledge, Taylor & Francis Group, 2018, p. 78-86 ; Roopika Risam, *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*, Evanston, Northwestern University Press, 2019.
- 6 Voir Scott Rettberg, « Electronic Literature as Digital Humanities », in : Susan Schreibman, Ray Siemens et John Unsworth (dir.), *A New Companion to Digital Humanities*, Chichester, John Wiley & Sons, 2015, p. 127-128 ; Dene Grigar et James Christopher O'Sullivan, *Electronic literature as Digital Humanities: Contexts, Forms & Practices*, New York, Bloomsbury Academic, 2021.
- 7 Cet aspect critique a également été débattu dans le cadre de l'émergence de la troisième génération de la littérature numérique. En effet, l'inclusion de la littérature sur les réseaux sociaux, et autres plateformes permettant de créer du contenu littéraire nativement numérique sans une expertise technique particulière, a conduit à une crise identitaire dans le monde de la littérature numérique. Si ces utilisations sont comprises comme des appropriations des réseaux à des fins différenciées, elles ont parfois été interprétées comme complices d'une plateformes dangereuse de l'environnement numérique dominée par les GAFAM, notamment en ce qui concerne la marchandisation des données personnelles. Pour une définition de la troisième génération de la littérature numérique, voir Leonardo Flores, « Third Generation Electronic Literature », *Electronic Book Review*, 2019.

Pour une réflexion sur les critères de légitimation de la littérature numérique, voir Servanne Monjour, « La littérature numérique n'existe pas. La littérarité au prisme de l'imaginaire médiatique contemporain », *Communication & langages*, vol. 3, n° 205, 2020, p. 5-27.

8 Voir Claudia Kozak, « Experimental Electronic Literature from the Souths. A Political Contribution to Critical and Creative Digital Humanities », *Electronic Book Review*, 2021.

9 « the raw material produced by abstracting the world into categories, measures and other representational forms [...] that constitute the building blocks from which information and knowledge are created ». Nous traduisons. Rob Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*, Londres, Sage, 2014, p. 28.

10 Anaïs Berck, « La Botanique coloniale wikifiée », Anaïs Berck [en ligne], URL : <https://www.anaisberck.be/la-botanique-coloniale-wikifie/> [consulté le 8 avril 2024].

11 Anaïs Berck, « Rewilding specimens », Anaïs Berck [en ligne], URL : <https://www.anaisberck.be/rewilding-specimens-fr/> [consulté le 8 avril 2024].

12 Camilla Townsend, « Burying the White Gods: New Perspectives on the Conquest of Mexico », *The American Historical Review*, vol. 108, n° 3, 2003, p. 659-687 ; Walter D. Mignolo et Madina V. Tostanova, *op. cit.*

13 Kerry H. Whiteside, « Beyond the Nature – Culture Dualism: The Ecology of Earth-Homeland », *World Futures*, vol. 60, n° 5-6, 2004, p. 357-369.

14 Val Plumwood, « Decolonizing Relationships with Nature », in : William Adams et Martin Mulligan (dir.), *Decolonizing Nature Strategies for Conservation in a Post-colonial Era*, London, Routledge, 2002, p. 53, 61.

15 *Ibid.*, p. 57-61.

16 « in the order of abstraction, everything comes to be understood as either a location or an object within this space ». Nous traduisons. Ricardo Padrón *The Spacious Word: Cartography, Literature, and Empire in Early Modern Spain*, Chicago, University of Chicago Press, 2004, p. 39.

17 *Ibid.*, p. 32.

18 « take specific actions, events and practices and abstract these into indexable, comparative data ». Nous traduisons. Ramesh Sprinivasan, *Whose Global Village? Rethinking How Technology Shapes Our World*, New York, New York University Press, 2019, p. 122.

19 Dans le secteur de la gestion du patrimoine et de muséologie, par exemple, des efforts ont été réalisés pour « décoloniser » des jeux de données issus de la colonisation, ou pour intégrer des conceptualisations « non-occidentales » dans des bases de connaissances. Voir, par exemple, Silvia Masiero, « Decolonising Critical Information Systems Research: A Subaltern Approach », *Information Systems Journal*, vol. 33, n° 2, 2023, p. 299-323 ; Helen Verran et Michael Christie, « Postcolonial Databasing? Subverting Old Appropriations, Developing New Associations », in : James Leach, and Lee Wilson (dir.), *Subversion, Conversion, Development: Cross-Cultural Knowledge Exchange and the Politics of Design*, Cambridge, MIT Press Scholarship Online, 2014 ; Alyssa Purcell et Tonia Sutherland, « A Weapon and a Tool: Decolonizing Description and Embracing Redescription as a Liberatory Archival Practice », *The International Journal of Information, Diversity*, vol. 5, n° 1, 2021, p. 60-78 ; Andrea Jimenez, Sara Vannini et Andrew Cox, « A Holistic Decolonial Lens for Library and Information Studies », *Journal of Documentation*, vol. 79, n° 1, 2023, p. 224-244.

20 « The disembodied and unlocated neutrality and objectivity of the ego-politics of knowledge is a Western myth ». Nous traduisons. Ramón Grosfoguel. « Decolonizing Post-Colonial Studies and Paradigms of Political-Economy: Transmodernity, Decolonial Thinking, and Global Coloniality », *Transmodernity: Journal of Peripheral Cultural Production of the Luso-Hispanic World*, vol. 1, n° 1, 2011, p. 5.

21 Voir, par exemple, Donna Haraway, « Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective », *Feminist Studies*, vol. 14, n° 3, 1988, p. 575-599 ; Elizabet Chilton et Angela M. Labrador, « Re-locating Meaning in Heritage Archives: A Call for Participatory Heritage Databases », in : « Annual Meeting of Computer Applications to Archaeology », *Computer Applications to Archaeology* [en ligne], Williamsburg, Virginia, du 22 au 26 mars 2009, URL : <https://rex.libraries.wsu.edu/esploro/outputs/99900586061101842> [consulté le 8 avril 2024].

22 Nick Couldry et Ulises Ali Mejias, « The Decolonial Turn in Data and Technology Research: What is at Stake and Where is it Heading? », *Information, Communication & Society*, vol. 26, n° 4, 2021, p. 796.

23 Une analyse des liens historiques et épistémologiques entre la technologie et la colonialité dépasse le cadre de cet article. Pour plus d'informations, voir, par exemple, Tara McPherson, « Why Are the Digital Humanities So White? Or Thinking the Histories of Race and

Computation », in : Matthew K. Gold (dir.), *Debates in the Digital Humanities*, Minneapolis, University of Minnesota Press, 2012, p. 139-160 ; Nick Couldry et Ulises Ali Mejias, « The Decolonial Turn in Data and Technology Research », *op.cit.*

24 Voir Alexandra Saemmer, « La littérature numérique entre légitimation et canonisation », *Culture & Musées*, n° 18, 2011, p. 201-223 ; Álvaro Seiça, « Electronic Literature and Online Literary Databases: The PO.EX and ELMCIP Cases », *RISS*, vol. 1, 2014, p. 89-102 ; Hannah Ackermans, « Appealing to Your Better Judgement: A Call for Database Criticism », *Electronic Book Review*, 2020 ; Katherine N. Hayles, *How we Think: Digital Media and Contemporary Technogenesis*, Chicago, The University of Chicago Press, 2012 ; Luis Pablo et María Goicoechea, « A Survey of Electronic Literature Collections », *CLCWeb: Comparative Literature and Culture*, vol. 16, n° 5, 2014 ; Ryan Ikeda, « Excavating Logics of White Supremacy in Electronic Literature: Antiracism as Infrastructural Critique », *Electronic Book Review*, 2021 ; Joseph Tabbi, « Something There Badly Not Wrong: The Life and Death of Literary Form in Databases », *Electronic Book Review*, 2020.

25 Scott Rettberg, « Documenting a Field: The Life and Afterlife of the ELMCIP Collaborative Research Project and Electronic Literature Knowledge Base », *Electronic Book Review*, 2021.

26 Lozana Rossenova, « ArtBase Redesign Documentation-Data Models », *ArtBase Redesign Documentation* [en ligne], URL : <https://sites.rhizome.org/artbase-re-design/data-models.html> [consulté le 8 avril 2024].

27 Bertrand Gervais, « The NT2 Hypermedia Art and Literature Directory: A New Knowledge Environment Devoted to the Valorization of Screen Culture », *Scholarly and Research Communication*, vol. 3, n° 2, 2012 ; MSH Lyon-Saint-Étienne, WEBLITT [en ligne], URL : <https://weblitt.msh-ls-e.fr/> [consulté le 8 avril 2024] ; Répertoire des écritures numériques [en ligne], URL : <https://repertoire.ecrituresnumeriques.ca/s/repertoire/page/accueil> [consulté le 8 avril 2024] ; Joseph Tabbi, *op. cit.*

28 « abstraction means removing from view the textures of social life ». Nous traduisons. Nick Couldry et Ulises Ali Mejias, *The Costs of Connection: How Data is Colonizing Human Life and Appropriating it for Capitalism*, Stanford, Stanford University Press, 2019, p. 147.

29 Anaïs Berck, *Walk along the Trees of Madrid* [en ligne], URL : <http://paseo-por-arboles.algoliterarypublishing.net/en> [consulté le 8 avril 2024].

- 30 Anaïs Berck, « An Algoliterary Publishing House », *op. cit.*
- 31 Anaïs Berck, « Research » [en ligne], URL : <https://algoliterarypublishing.net/research.html> [consulté le 8 avril 2024].
- 32 « An algorithm must be seen to be believed ». Nous traduisons. Donald E. Knuth, *The Art of Computer Programming*, vol. 1, Reading, Addison-Wesley, 1997, p. 4.
- 33 « Creativity is the product of social relations and cultural contexts more than inspiration ». Nous traduisons. Scott Rettberg, « Finding a Third Space for Electronic Literature: Creative Community, Authorship, Publishing, and Institutional Environments », *Journal of Writing in Creative Practice*, vol. 4, n° 1, 2011, p. 14.
- 34 *Loc. cit.*
- 35 Jane Anderson, « Indigenous Knowledge and Intellectual Property Rights », in : James D. Wright (dir.), *International Encyclopedia of the Social & Behavioral Sciences*, Amsterdam, Elsevier, 2015, p. 769-778 ; Jane Anderson, « Negotiating Who “Owns” Penobscot Culture », *Anthropological Quarterly*, vol. 91, n° 1, 2018, p. 267-305.
- 36 « What if algorithms were committed to a more equal world [...] ? ». Nous traduisons. Anaïs Berck, « Research », *op.cit.*
- 37 Voir Alexandra Saemmer, *op. cit.* ; Philippe Bootz, « Littérature numérique : un éphémère vrai-ment ? », *Hybrid*, n° 1, 2014. DOI : [10.4000/hybrid.1156](https://doi.org/10.4000/hybrid.1156) [consulté le 8 avril 2024] ; Serge Bouchardon et Davin Heckman, « Digital Manipulability and Digital Literature », *Electronic Book Review*, 2012 ; Gilles Bonnet, *Pour une poétique numérique : littérature et Internet*, Paris, Hermann, 2017.
- 38 Anaïs Berck, « Research », *op. cit.*
- 39 Algolit, « Main Page », Algolit [en ligne], URL : <https://www.algolit.net/index.php> [consulté le 8 avril 2024].
- 40 Constant, « About », Constant [en ligne], URL : <https://constantvzw.org/site/-About-Constant-7-.html> [consulté le 8 avril 2024].
- 41 Piraye Hacigüzeller, James Stuart Taylor et Sara Perry, « On the Emerging Supremacy of Structured Digital Data in Archaeology: A Preliminary Assessment of Information, Knowledge and Wisdom Left Behind », *Open Archaeology*, vol. 7, n° 1, 2021, p. 1709-1730.

42 Günter Ladwig et Thanh Tan, « Linked Data Query Processing Strategies », in : Peter F. Patel-Schneider, Yue Pan, Pascal Hitzler, Peter Mika, Lei Zhang, Jeff Z. Pan, Ian Horrocks, et Birte Glimm (dir.), *The Semantic Web – ISWC 2010*, p. 453-469.

43 Angela M. Labrador, « Ontologies of the Future and Interfaces for All: Archaeological Databases for the Twenty-First Century », *Archaeologies*, vol. 8, n° 3, 2012, p. 238-239.

44 « once data are placed in tables, their social lineage is forgotten ». Nous traduisons. Nadine Schuurman, « Database Ethnographies Using Social Science Methodologies to Enhance Data Analysis and Interpretation », *Geography Compass*, vol. 2, n° 5, 2008, p. 1529.

45 Anaïs Berck, « Research », *op. cit.*

RÉSUMÉS

Français

Cet article propose une réflexion sur l'utilisation des bases de données pour répertorier les œuvres de littérature numérique au départ de la démarche décoloniale du collectif de création littéraire numérique Anaïs Berck. La lecture de la plateforme de publication « Algoliterary Publishing House » et de certaines œuvres qui y sont publiées facilite un questionnement sur les liens entre colonialité, abstraction et données, ainsi que sur le déracinement contextuel qui peut avoir lieu lorsqu'une œuvre de littérature numérique est encodée dans une base de données.

English

This article explores the use of databases to catalogue works of digital literature, based on the decolonial approach of the digital literature collective Anaïs Berck. Reading the Algoliterary Publishing House publishing platform and some of the works published there facilitates a questioning of the links between coloniality, abstraction and data, as well as the contextual uprooting that can take place when a work of digital literature is encoded in a database.

INDEX

Mots-clés

littérature numérique, décolonialité, base de données, humanités numériques, Anaïs Berck

Keywords

digital literature, decoloniality, database, digital humanities, Anaïs Berck

AUTEURS

Isabelle Gribomont

Isabelle Gribomont est chercheuse postdoctorale au centre de traitement automatique du langage (Cental) à l'UCLouvain et responsable du projet LabEL (Laboratory for Electronic Literature) à la bibliothèque royale de Belgique (KBR). Son travail est soutenu par une subvention FED-tWIN (Prf-2020-026_KBR-DLL) financée par BELSPO (Belgian Science Policy)

Anaïs Berck

Anaïs Berck est un pseudonyme désignant une collaboration entre des êtres humains, des algorithmes et des arbres. Ce collectif ouvre un espace dans lequel l'intelligence humaine est explorée en compagnie de l'intelligence végétale et de l'intelligence artificielle. Le projet de recherche Algoliterary Publishing House a été soutenu par FRart (<https://art-recherche.be/>)

Enterre-moi, mon amour : le jeu sur la frontière entre réalité et fiction dans les fictions pour smartphones

Bury Me, My Love: the play on the border between reality and fiction in smartphone fiction

Serge Bouchardon

DOI : 10.35562/marge.944

Droits d'auteur

CC BY-NC-SA

PLAN

Introduction

Enterre-moi, mon amour : traverser les frontières

Le « temps réel » du lecteur

Un révélateur de l'usage du smartphone

Fictions et enjeux éthiques

Ouverture : l'IA annule-t-elle le réel ?

TEXTE

Introduction

« Si l'on pense à toutes les œuvres numériques s'insérant sur les RSN, qu'il s'agisse de la twittérature ou des romans sur Instagram pour ne citer que ces exemples¹, c'est la frontière entre réel et fiction qui se voit sans cesse réinterrogée. Ces récits provoquent une intrusion de la fiction dans des espaces qui ne lui sont pas originellement destinés². »

- 1 Dans quelle mesure le numérique reconfigure-t-il la frontière entre réel et fiction, et dans quelle mesure les récits fictionnels numériques en sont-ils un bon révélateur ?
- 2 Sur les plates-formes d'applications pour smartphones Google Play et App Store, on trouve de plus en plus de récits interactifs pour

smartphones, des « smartfictions » (Picard, 2022). Ces fictions reposent sur les pratiques *ordinaires* et quotidiennes avec un smartphone, par exemple la discussion instantanée, mais aussi les notifications (comme *Lifeline* ou *Somewhere*). Ces fictions introduisent le *temps réel* du lecteur dans le cadre fictionnel. L'heure du smartphone et le temps qui s'écoule sont des données qui sont fictionnalisées.

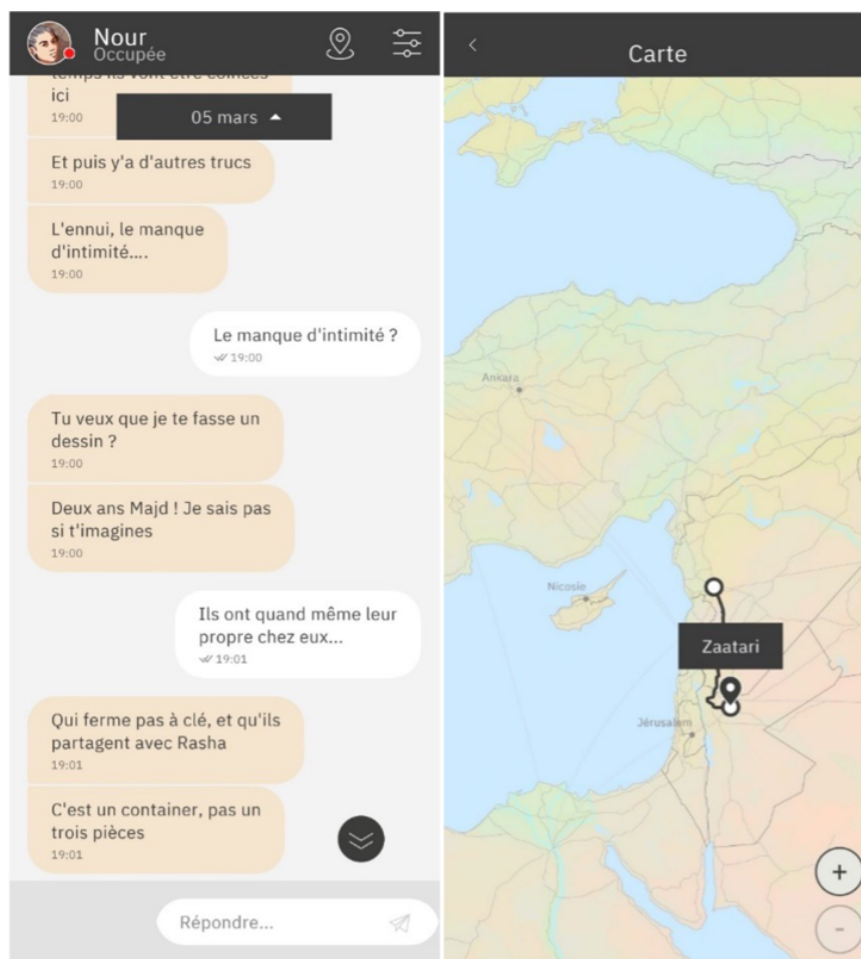
- 3 À partir de l'exemple d'*Enterre-moi, mon amour*³, il s'agit de comprendre les enjeux de ce brouillage de la frontière entre réel et fiction, mais aussi ce que ces récits nous révèlent de notre propre usage du smartphone, ainsi que les enjeux éthiques qu'ils soulèvent. L'étude qui suit est la première étape d'une recherche. Il s'agit en effet d'une première approche d'une seule smartfiction qui ne s'appuie pas sur des analyses d'usages.

***Enterre-moi, mon amour* : traverser les frontières**

- 4 Le jeu narratif⁴ *Enterre-moi, mon amour*, publié en 2017, est le fruit d'une collaboration entre le studio *The Pixel Hunt* et la chaîne de télévision *Arte France*. Le point de départ de cette création fut un récit témoignage publié en ligne par le journal *Le Monde* en 2016 et intitulé « Dans le téléphone d'une migrante syrienne⁵ ». L'article restitue le fil *WhatsApp* d'une jeune syrienne migrant vers l'Allemagne. Ce récit témoignage a été, dans un second temps, transformé en fiction interactive, *Enterre-moi, mon amour*⁶.
- 5 Dans cette fiction interactive, nous incarnons le personnage de Majd, le mari de Nour, jeune femme syrienne tentant de rejoindre l'Europe. Nous échangeons avec Nour via une messagerie instantanée. Nous essayons de la conseiller, de la soutenir moralement. Le voyage entre la Syrie et l'Europe n'a pas de durée définie. Le périple peut durer plusieurs jours.
- 6 Le jeu sur la chronologie et la temporalité se conjugue avec un jeu sur la cartographie et la spatialité. Le lecteur peut accéder à la progression du périple de Nour via une balise de géolocalisation, située en haut à droite de l'espace de discussion instantanée (cf. figure 1). Il peut ainsi visualiser la progression du personnage en

temps réel sur une carte. Cette temporalité est aussi une indication de la progression du lecteur dans le récit. Plus Nour se rapproche de l'un des pays d'Europe, plus le lecteur se rapproche de la fin de son expérience de lecture. Il s'agit donc d'un repère à la fois spatial et temporel de la progression du personnage comme de celle du lecteur dans son parcours de lecture.

Figure 1. Carte (à droite) donnant à voir le parcours du personnage

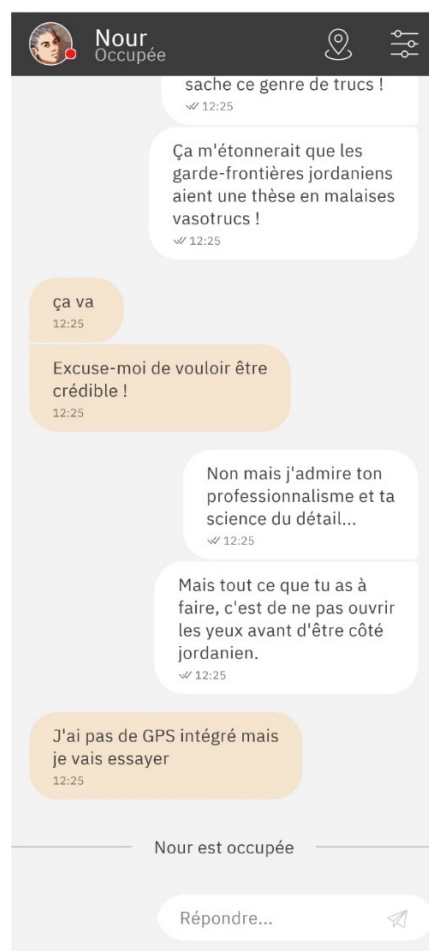


Source : *Enterre-moi mon amour*, Arte © Arte

- 7 Au fil de chaque parcours (il y a 19 fins différentes), un compteur indique les jours qui passent. La frontière peut être temporairement fermée, Nour peut avoir rendez-vous avec une connaissance qui n'arrive pas, etc. Dans ce cas, faut-il patienter ou reconsidérer les plans ?

« Vos communications se feront en pseudo temps réel : si Nour doit accomplir une action censée lui prendre quelques heures, vous ne pourrez pas la joindre pendant ce laps de temps. À son retour, une notification vous signale qu'elle est à nouveau disponible – et qu'elle a peut-être besoin de vous⁷. »

Figure 2. Le personnage de Nour n'est plus disponible. Le lecteur doit attendre une notification de sa part.

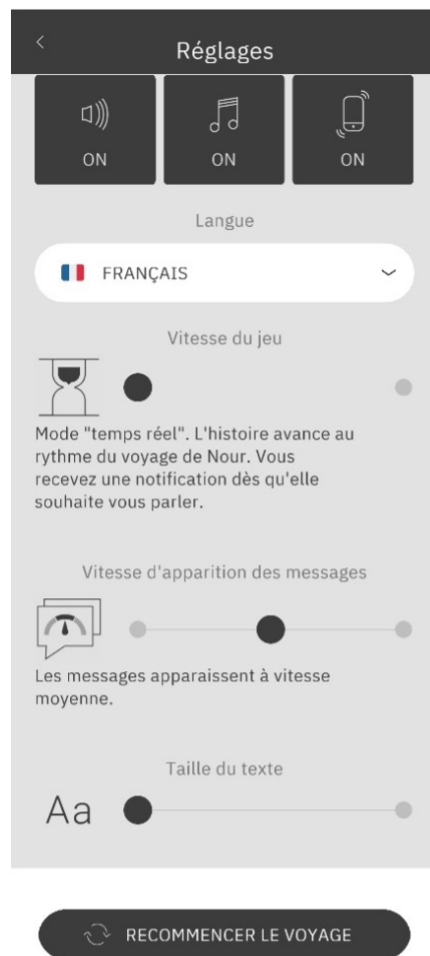


Source : *Enterre-moi mon amour*, Arte © Arte

Le « temps réel » du lecteur

- 8 C'est ce mode « temps réel » qui nous intéresse ici (cf. figure 3). À noter que le lecteur peut s'il le souhaite désactiver les notifications en « temps réel », mode activé par défaut.

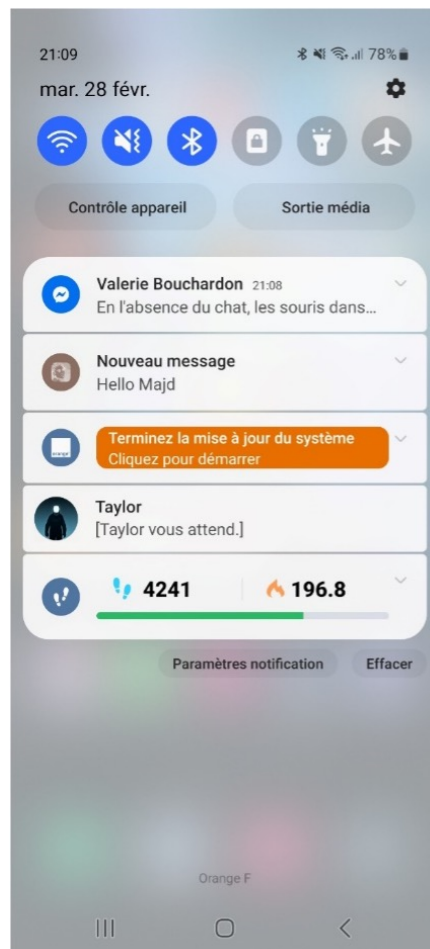
Figure 3. Paramétrage du mode « temps réel »



Source : *Enterre-moi mon amour*, Arte © Arte

- 9 Il y a un jeu sur la temporalité avec l'intrusion du réel, du « temps réel » du lecteur (et du personnage fictionnel). Le rapport au temps de lecture est ainsi différent avec l'intrusion des notifications : lorsqu'il reçoit une notification de Nour, parmi d'autres notifications (cf. figure 4), l'utilisateur est incité à se replonger dans la fiction pour dialoguer avec le personnage. D'un point de vue phénoménologique, il y a une interruption avec l'injonction de la notification. Les fictions reposant sur des notifications⁸ vont ainsi assez loin dans la perturbation du temps de l'utilisateur, mais, dans le même temps, réintroduisent le « temps réel » de celui-ci dans le cadre fictionnel. Manon Picard a désigné ces notifications fictionnelles par le terme de « notifictions⁹ » (Picard, 2022).

Figure 4. Notifications fictionnelles (*Enterre-moi, mon amour* et *Lifeline*) parmi d'autres notifications



Source : *Enterre-moi mon amour*, Arte et *Lifeline*, 3 Minute Games © Arte et 3 Minute Games

- 10 De même que les récits géolocalisés (*locative narratives*¹⁰) jouent sur la frontière entre espace fictionnel et espace *réel* du lecteur, l'incitant à se déplacer physiquement dans un lieu donné pour faire avancer le récit, les récits fondés sur des notifications jouent sur la frontière entre le temps de la fiction et le temps *réel* du lecteur.
- 11 Qu'apportent ces smartfictions qui font intervenir des notifications ? Elles nous donnent l'impression d'une coprésence. En tant que lecteurs/acteurs de la smartfiction, quand nous faisons un choix, nous rendons d'une certaine façon le personnage fictionnel présent par nos actions. Mais cette notion de présence est encore renforcée quand nous répondons à l'appel du personnage via une notification. Nous faisons advenir le temps de l'autre (ainsi que celui de la fiction).

La temporalité du personnage (et des événements) coïncide, via les notifications, avec la temporalité du lecteur, qui s'approprie alors pleinement la durée de l'histoire. À noter que pour certains lecteurs, le fait de devoir quitter le récit pour attendre une notification peut être facteur de *décrochage* par rapport au récit (il est toujours possible dans ce cas de désactiver les notifications).

12 Pour Paul Ricoeur,

« le temps devient temps humain dans la mesure où il est articulé de manière narrative ; en retour le récit est significatif dans la mesure où il dessine les traits de l'expérience temporelle » (Ricoeur, 1985).

13 Le récit est, selon cette perspective, notre outil principal pour nous situer dans le temps – et pour situer le temps en nous. Par ailleurs, le numérique peut être caractérisé comme un « outil de déconstruction phénoménale de la temporalité » (Bachimont, 2010). Il se traduit en effet par deux tendances : le calcul « temps réel », donnant l'impression d'immédiateté ; l'universalité de l'accès, donnant l'impression de la disponibilité. Le numérique, dans sa disponibilité et son immédiateté, aboutirait ainsi à un présent permanent, sans écoulement temporel (Bachimont, 2010).

14 Les smartfictions qui jouent avec les notifications pourraient permettre une retemporalisation du temps humain détemporalisé par le calcul.

Un révélateur de l'usage du smartphone

15 Dans son analyse de notre rapport au smartphone, Laurence Allard parle d'une « technologie pour soi », le smartphone étant « un organe plus qu'une prothèse » (Allard et al., 2014). Exploitant notre appropriation profonde de ce dispositif, le jeu sur les notifications repose notamment sur le fait que le smartphone est notre dispositif pour des pratiques de communication et d'écriture quotidiennes. Le brouillage entre réel et fiction est d'autant plus efficace que c'est le même dispositif qui nous sert à communiquer au quotidien et à lire/agir une fiction. On retrouve ce brouillage dans toutes les

smartfictions du type *Lost Phone* (*A Normal Lost Phone*, *Another Lost Phone*, *Ana the Game*, etc.) : l'utilisateur est censé avoir trouvé un téléphone dans lequel il peut fouiller (messagerie, photos, agenda, etc.) pour comprendre l'histoire du personnage.

- 16 Si ceci confère à la fiction un effet de réalité, nous percevons également, à travers les notifications fictionnelles qui viennent brouiller le réel, que le smartphone est d'une certaine façon lui-même un support de déréalisation : il construit une couche de fiction hors du monde réel. Grâce à ces fictions, nous comprenons ainsi que le dispositif smartphone est déjà un support de fiction (et de narration) introduisant un nouveau rapport à la temporalité (et nous rappelant sans doute également que la distinction fiction/réel doit être dans ce cas dépassée).
- 17 La smartfiction a ainsi une dimension réflexive. Elle nous invite à réfléchir à l'usage quotidien du support – et même du dispositif – qu'est le smartphone. Peu nombreux sont les dispositifs que l'on utilise quotidiennement et qui servent de supports de narration. Le détournement esthétique, qui est opéré dans un geste artistique, décentre et de ce fait révèle le rapport que l'on a au quotidien avec le smartphone. Nos usages quotidiens sont toujours emplis de fiction et de narration. La smartfiction, qui exploite les possibilités du dispositif, révèle ainsi une puissance à l'œuvre dans tout usage du smartphone : tout usage du smartphone ne serait-il pas d'une certaine façon une smartfiction, donc entaché d'absence et de fiction ? C'est notre propre rapport au brouillage de la frontière entre fiction et réel que les récits interactifs pour smartphone mettent en scène, et sur lequel ils nous incitent à nous interroger.

Fictions et enjeux éthiques

- 18 Beaucoup de smartfictions mettent en avant le fait de s'appuyer sur des données réelles (par exemple le fil *WhatsApp* de la migrante syrienne, Dana, publié par le site du *Monde* et les témoignages réels de migrants qui ont nourri *Enterre-moi, mon amour*). Florent Maurin, coscénariste et game designer d'*Enterre-moi, mon amour*, parle d'ailleurs de « jeu du réel¹¹ ». Dans ses interventions et entretiens, il aborde la question réel/fiction en insistant bien sur le fait qu'*Enterre-*

moi, mon amour est une fiction et qu'il faut prendre cette création comme telle..., mais en même temps insiste sur le fait que

« tout ce qu'on raconte est fondé sur du réel, des témoignages concrets, et en premier lieu celui de Dana. On ne voulait surtout pas raconter quelque chose qui n'aurait pas pu se passer¹² ».

- 19 Dans *Enterre-moi, mon amour*, on a ainsi une hybridation entre factuel et fictionnel. Françoise Lavocat, dans une volonté vulgarisatrice, a pu avancer que « ce qui est factuel est référentiel, et ce qui est fictionnel ne l'est pas¹³ ». Ceci est sans doute réducteur, car la fiction semble le plus souvent – pour ne pas dire toujours – dotée d'une valeur référentielle (politique, sociale, historique, etc.).
- 20 Dans *Fait et fiction. Pour une frontière* (2016), Françoise Lavocat s'attache à déconstruire le brouillage des frontières entre réel et fiction, cherchant à « montrer l'existence et la nécessité cognitive, conceptuelle et politique des frontières de la fiction » dans tous les arts de la fiction. Il y aurait ainsi une tension entre la nécessité de fonder les frontières entre réel et fiction¹⁴ et le travail de brouillage perpétuel pratiqué par les créateurs de fictions numériques, en particulier des fictions pour smartphone telles qu'*Enterre-moi, mon amour*. Ce brouillage de la frontière entre réel et fiction est d'ailleurs souvent un argument commercial :
- « L'application choisit une alternative plutôt qu'une autre en fonction de votre lieu, en fonction de l'heure et même de la météo autour de vous. Le récit doit ainsi s'approcher au maximum de votre contexte pour renforcer l'immersion¹⁵. »
- 21 L'ultrapersonnalisation de certaines fictions numériques pour smartphone (par exemple s'il pleut au moment où je lis le récit, il pleut dans l'histoire que je suis en train de lire, etc.) entraîne une conséquence : le lecteur n'est pas invité à se décentrer de lui-même pour s'immerger dans une expérience imaginaire ; au contraire, c'est le récit qui se moule dans la sphère de son *identité*.
- 22 Plus largement, on peut relever une tendance actuelle concernant le récit, incarnée notamment par la mode de l'autofiction. Le récit n'est pas vu comme une ouverture à l'autre, mais comme un prolongement

de soi-même, pour l'auteur comme pour le lecteur. Souvent, on ne raconte plus d'histoire, à savoir ce qui s'est passé et qui est arrivé à d'autres que moi, ou à moi comme un autre, mais on prolonge son présent identitaire dans la fiction. D'une certaine façon, dans nombre de fictions contemporaines, l'autre ne s'aborde plus désormais que comme miroir de soi. Or le numérique participe à cette tendance, sans en être la source. Le numérique instaure un présentisme (les contenus sont consultés comme s'ils étaient simultanés à ma consultation, il n'y a plus de marque du temps passé ou révolu), et propose une prise de rôle comme acteur/joueur et pas seulement comme lecteur. Outre le présentisme et l'interactivité, une autre caractéristique du numérique accentue cette tendance, à savoir le brouillage des frontières entre réel et fiction, notamment dans les smartfictions exploitant le smartphone comme support de fictionnalisation.

- 23 Si dans *Enterre-moi, mon amour* le récit de Nour est le prolongement de moi-même, au nom de quoi puis-je prétendre vivre ce qu'elle vit ? On peut comprendre la résistance d'ordre éthique de certains lecteurs. La question des enjeux éthiques se pose en effet lorsqu'on fait interpréter le rôle d'un migrant (ou le rôle d'un parent de migrant) à un lecteur confortablement installé devant son ordinateur ou son smartphone, *a fortiori*, lorsque l'expérience se présente comme un « jeu narratif ». On peut ainsi concevoir que certains utilisateurs soient réticents à « fictionnaliser » (Odin, 2000) une telle histoire.
- 24 Pourtant, dans *Enterre-moi, mon amour*, le sentiment d'impuissance éprouvé par le lecteur/joueur quand il n'a qu'une seule réponse possible, ou lorsqu'il en a deux et qu'aucune ne lui convient, lui rappelle que ce n'est pas son histoire, mais celle de Majd (qui est jaloux, quelque peu raciste à l'égard notamment des Irakiens, etc.). Le personnage de Majd est fortement caractérisé, de façon quelque peu contre-intuitive par rapport à beaucoup de jeux vidéo dans lesquels le personnage que l'on incarne est intentionnellement neutre, sans personnalité marquée, afin que le joueur puisse plus facilement l'incarner (pensons par exemple au personnage de Gordon Freeman dans *Half-Life*). L'esthétique de *Enterre-moi, mon amour*, en dessin stylisé, introduit également de la distance et prend à contre-pied l'idée d'un « jeu du réel ». Il y a donc tout un jeu sur la distance, par rapport au personnage et par rapport à l'histoire.

- 25 Le « désir empêché de venir en aide au personnage », qui caractérise selon Françoise Lavocat (2016) l'empathie dans la fiction, fonctionne très bien dans *Enterre-moi, mon amour*, d'autant plus que le personnage principal (Nour) n'est pas celui que l'on incarne. D'ailleurs, le fait que le destin du personnage finisse par échapper en partie au lecteur peut aussi être vu comme un moyen de renforcer une certaine distance. Dans la smartfiction, on peut ainsi avoir le sentiment que le dispositif interactif peut décupler la possibilité de l'empathie.
- 26 Évoquons une autre forme de jeu entre réel et fiction via l'interactivité. En 2020 puis en 2021, l'application de rencontres Tinder a proposé de participer à une fiction interactive, intitulée *Swipe Night* (il s'agissait d'affronter la fin du monde¹⁶, puis de démasquer l'auteur d'un crime). À la suite de la fiction interactive, l'utilisateur était mis en relation avec des personnes réelles sur Tinder selon les choix faits par celles-ci lors de la fiction interactive. On peut penser que l'hypothèse de la plate-forme est la suivante : les choix qu'une personne fait dans le cadre d'une fiction interactive – et en premier lieu les choix empathiques – sont plus révélateurs de sa personnalité que les centres d'intérêt qu'elle va déclarer. Partager une expérience narrative interactive est le moyen le plus fiable de réaliser une vraie rencontre... C'est une autre forme d'hybridation du réel et de la fiction avec le numérique.
- 27 Avec *Lucette, gare de Clichy*¹⁷ de Françoise Chambefort, on pouvait avancer que le flux de données en temps réel fait entrer la contingence de la vie dans la fiction¹⁸ (Chambefort, 2020). Avec *Swipe Night*, c'est la contingence de la fiction qui entre dans le réel.
- 28 Jean-Marie Schaeffer (1999) nous dit que « [les] univers fictifs, loin d'être des apparences illusoires ou des constructions mensongères, [...] sont une des faces majeures de notre rapport au réel.¹⁹ » Sans doute le numérique ouvre-t-il des possibles dans ce jeu sur la frontière entre réel et fiction, et sur la promesse d'un dispositif fictionnel empathique.

Ouverture : l'IA annule-t-elle le réel ?

- 29 Finalement, c'est peut-être tout le rapport entre données et réel qui est à repenser avec le numérique. On pourrait même « se demander dans quelle mesure le numérique n'annule pas le réel : il semble se substituer à lui comme opposition et référence. Le réel opposable, ce n'est pas le réel, mais la donnée » (Bachimont, 2023). Dans le réel opposable, on trouve l'idée d'une référence que l'on met en avant pour affirmer ou confirmer ses propos. Ainsi, dans les jumeaux numériques²⁰, le réel (la pièce ou la machine réelle) est maintenu grâce au double numérique qui en donne la *vraie* version à laquelle il faut faire coller le réel.
- 30 En intelligence artificielle notamment, on considère que le réel opposable, ce qui permet de justifier, correspond aux données collectées sur lesquelles s'appuie le calcul et la décision proposée. Ainsi, l'être humain en situation doit confronter son réel, celui de la situation, au réel sédimenté dans les données. Mais, selon Bruno Bachimont, comment pourrait-il justifier de s'appuyer sur son appréciation de la situation, alors que la rationalité calculatoire voudrait que l'on considère plutôt l'avis de l'algorithme qui s'appuie sur la synthèse des situations déjà rencontrées et sédimentées dans les données ?
- 31 Bruno Bachimont avance ainsi que les données de l'IA deviennent des fictions, car elles n'ont plus d'ancrage dans le réel, elles le remplacent. Elles ont certes une origine dans le réel, mais on ne les fait plus jouer et interagir avec ce dernier : l'origine réelle n'est pas un ancrage, car on a congédié ce réel, tout en faisant croire qu'on en tient lieu. Si bien qu'on ne s'interroge pas sur la réalité et la vérité des données, mais seulement sur la pertinence des algorithmes confrontés à ces données.
- 32 En ce sens – et ce serait peut-être la tendance actuelle – le numérique annule le réel au profit de données fictionnelles (devenues fictionnelles, même si elles ont sédimenté le réel). En tout premier lieu avec l'IA, puisque c'est particulier à la notion de jeu de données que de constituer un réel de substitution. Ce n'est pas tant la

frontière entre réel et fiction qui tendrait à s'estomper, mais le réel lui-même qui serait remplacé. De cela, les fictions numériques, notamment celles qui font intervenir de l'IA, pourraient en être un bon révélateur.

BIBLIOGRAPHIE

ALLARD Laurence, CRETON Laurent et ODIN Roger (dir.), *Téléphone mobile et création*, Paris, Armand Colin, 2014.

BACHIMONT Bruno, *Le sens de la technique : le numérique et le calcul*, Paris, éditions Les Belles Lettres, coll. « Encre marine », 2010.

BACHIMONT Bruno, « L'IA, une machine à inventer ». *Intelligence artificielle : une invention à émotion* (journée d'étude), université de Montpellier, 27 janvier 2023.

CAÏRA Olivier, *Définir la fiction : du roman au jeu d'échecs*, Paris, éditions de l'EHESS, 2011.

CHAMBEFORT Françoise, « Sortir de l'écran : Lucette, gare de Clichy », in : BERTAND Gervais et MARCOTTE Sophie (dir.), *Attention à la marche ! Penser la littérature électronique en culture numérique*, Montréal, Les Presses de l'Écureuil – ALN/NT2, 2020, p. 327-338.

LAVOCAT Françoise, *Fait et fiction. Pour une frontière*, Paris, Seuil, 2016.

ODIN Roger, *De la fiction*, Bruxelles, De Boeck Université, 2000.

PICARD Manon, *La smartfiction, une fiction interactive à lire, un rôle à incarner ou une partie à jouer sur son smartphone ?*, thèse de doctorat, sous la direction de S. Bochardon et B. Bachimont, 20 juin 2022, université de technologie de Compiègne.

RICŒUR Paul, *Temps et récit*, 3 volumes, Paris, Seuil, 1985.

SCHAEFFER Jean-Marie, *Pourquoi la fiction ?*, Paris, Seuil, 1999.

ZUNSHINE Lisa, *Why We Read Fiction?*, Columbus, Ohio University Press, 2006.

Exemples de smartfictions

Ana The Game :

App Store : <https://apps.apple.com/fr/app/ana-the-game/id1202981332> [consulté le 8 avril 2024] ;

Google Play store : https://play.google.com/store/apps/details?id=com.plug_in_digital.anathegame [consulté le 8 avril 2024].

Another Lost Phone: Laura's story :

App Store : <https://apps.apple.com/fr/app/another-lost-phone/id1267197755>

[consulté le 8 avril 2024] ;

Google Play store : <https://play.google.com/store/apps/details?id=com.accidentalqueens.anotherlostphone> [consulté le 8 avril 2024].

Enterre-moi, mon amour / Bury me, my love :

App Store : <https://apps.apple.com/fr/app/enterre-moi-mon-amour/id1281473147>

[consulté le 8 avril 2024] ;

Google Play store : https://play.google.com/store/apps/details?id=com.plug_in_digital.emma [consulté le 8 avril 2024].

Lifeline :

App Store : <https://apps.apple.com/fr/app/lifeline/id982354972#?platform=iphone>

[consulté le 8 avril 2024] ;

Google Play store : <https://play.google.com/store/apps/details?id=com.threeminutegames.lifeline.google> [consulté le 8 avril 2024].

Memento M :

App Store : <https://apps.apple.com/us/app/memento-m/id1533799527> [consulté le 8 avril 2024] ;

Google Play store : <https://play.google.com/store/apps/details?id=com.glaznev.detective5> [consulté le 8 avril 2024].

Seven – Endgame :

App Store : <https://apps.apple.com/fr/app/seven-endgame/id1476673731> [consulté le 8 avril 2024] ;

Google Play store : <https://play.google.com/store/apps/details?id=air.de.sponsorads.endgame> [consulté le 8 avril 2024].

Simulacra :

App Store : <https://apps.apple.com/fr/app/simulacra/id1252035454> [consulté le 8 avril 2024] ;

Google Play store : <https://play.google.com/store/apps/details?id=com.kaigan.simulacra> [consulté le 8 avril 2024].

NOTES

1 On pourrait penser également à la tradition en ligne des ARG (*alternate reality games*), ou encore aux « profils de fiction » sur lesquels travaillent notamment Alexandra Saemmer et Bertrand Gervais.

2 Appel à communication du colloque « Fiction & données » (mars 2023, université Jean-Moulin Lyon 3), URL : <https://www.fabula.org/actualites/108573/fiction-et-donnees--colloque-lifranumlqmlseti.html> [consulté le 8 avril 2024].

- 3 À consulter l'URL : <http://enterremoimonamour.arte.tv/> [consulté le 8 avril 2024]. En arabe, « Enterre-moi, mon amour » signifie « prends soin de toi », « ne t'avise surtout pas de mourir avant moi ».
- 4 C'est sous cette catégorie que cette création est commercialisée.
- 5 Lucie Soullier et Madjid Zerrouky, « Dans le téléphone d'une migrante syrienne », *Le Monde* [en ligne], URL : https://www.lemonde.fr/international/visuel/2015/12/18/dans-le-telephone-d-une-migrante-syrienne_4834834_3210.html [consulté le 8 avril 2024].
- 6 Le prologue interactif est jouable ici : <http://enterremoimonamour.arte.tv/prologue> [consulté le 8 avril 2024].
- 7 Arte, *Enterre-moi, mon amour* [en ligne], URL : <http://enterremoimonamour.arte.tv/> [consulté le 8 avril 2024].
- 8 Citons notamment *Lifeline* (2015) ou encore *Somewhere* (2017).
- 9 Ce terme est également utilisé par Emmanuelle Lescouet, doctorante à l'université de Montréal.
- 10 Par exemple *The Cartographer's Confession* (2017), de James Attlee, qui se déroule à Londres.
- 11 Split/Screen Documentaries, *Enterre-moi, mon amour – documentaire split/screen* [en ligne], URL : <https://youtu.be/JGF1z8sSv9s> [consulté le 8 avril 2024]
- 12 Split/Screen Documentaries, *Enterre-moi, mon amour – documentaire split/screen* [en ligne], URL : <https://youtu.be/JGF1z8sSv9s> [consulté le 8 avril 2024].
- 13 Kptni, *Fait et fiction : Françoise Lavocat dans les mots du temps* [en ligne], URL : <https://youtu.be/xaQNOzOm7bQ> [consulté le 8 avril 2024] (à partir de 05:10).
- 14 On peut penser à la prolifération des *fake news* sur internet en considérant qu'elles font partie d'une forme de grand récit complotiste.
- 15 Nicolas Furno, « *Chronique(s) d'Abîme*, un roman qui s'adapte à votre contexte », *iGeneration* [en ligne], 2016, URL : <https://www.igen.fr/app-store/2016/06/chroniques-dabime-un-roman-qui-sadapte-votre-contexte-96253> [consulté le 8 avril 2024].
- 16 Bande-annonce à retrouver en ligne sur : <https://youtu.be/STh9DodzTZQ> [consulté le 8 avril 2024].

17 Françoise Chambefort, *Lucette, gare de Clichy* [en ligne], 2017, URL : <http://fchambef.fr/lucette/intro.php> [consulté le 8 avril 2024].

18 Françoise Chambefort précise que s'il entre en empathie avec le personnage, l'internaute vivra une expérience temporelle inédite. Il ressentira la vulnérabilité d'une personne âgée qui n'a pas de prise sur sa vie.

19 Jean-Marie Schaeffer, *Pourquoi la fiction ?*, Paris, Seuil, 1999, 4^e de couverture.

20 Un jumeau numérique est un modèle virtuel conçu pour refléter fidèlement un objet physique.

RÉSUMÉS

Français

Des récits interactifs pour téléphones portables, qualifiés de *smartfictions*, s'appuient sur nos pratiques ordinaires avec un smartphone, notamment la discussion instantanée et les notifications. À partir de l'exemple de *Enterre-moi, mon amour* il s'agit d'analyser le jeu sur la frontière entre réalité et fiction, mais aussi ce que ces récits nous révèlent de notre propre usage du smartphone ainsi que les enjeux éthiques que ces récits soulèvent.

English

Interactive narratives for mobile phones, described as *smartfictions*, are based on our ordinary practices with a smartphone, in particular chat and notifications. By using the example of *Bury Me, My Love* we will analyse the play on the border between reality and fiction, but also what these fictions reveal about our own use of the smartphone and the ethical issues they raise.

INDEX

Mots-clés

smartfiction, notification, fiction, réel, numérique

Keywords

smartfiction, notification, fiction, reality, digital

AUTEUR

Serge Bouchardon

Laboratoire Costech, université de technologie de Compiègne

IDREF : <https://www.idref.fr/101521693>

ORCID : <http://orcid.org/0000-0002-4430-7130>

HAL : <https://cv.archives-ouvertes.fr/serge-bouchardon>

ISNI : <http://www.isni.org/0000000051062341>

BNF : <https://data.bnf.fr/fr/15527074>

Données/auctorialité

Écrire-avec l'intelligence artificielle, ou l'esthétique de la sympoièse

Expérimentations littéraires avec GPT-2 et GPT-3

Writing-with artificial intelligence, or the esthetics of sympoiesis. Literary experiments with GPT-2 and GPT-3

Erika Fülöp

DOI : 10.35562/marge.956

Droits d'auteur

CC BY-NC-SA

PLAN

Comment dire ?

Les textes

Pharmako-AI

Internes

Qui écrit ?

GPT et écrire avec

L'auteur qui n'(en) est pas un

L'auteur en droit

Auctorialité cyborg, collaboration, coopération, délégation (variations sur un thème)

De l'écriture prothétique à l'écriture sympoiétique

Qui parle ?

Déictiques déracinés

L'inquiétante étrangeté ou « the new uncanny » (la vallée des fantômes)

Qui est « je » ? *Différance* et relation

Conclusions : trois paradoxes et une esthétique

Paradoxe de l'auctorialité artificielle

Le paradoxe du sujet

Paradoxe de l'identité

Vers une esthétique de l'écriture sympoiétique

TEXTE

« [R]eality has outpaced the available language to parse what is already at hand¹ »

Benjamin BRATTON et Blaise
AGÜERA Y ARCAS

« It is an old experience that the traditional, the usual and the hereditary are dear and familiar to most people, and that they incorporate the new and the unusual with mistrust, unease and even hostility². »

Ernst JENTSCH

« [L]a question serait de nouer un autre rapport à la technique en repensant le nœud originellement formé par l'homme, la technique et le langage »

Bernard STIEGLER

- 1 Les livres « coécrits » « avec » une « intelligence » « artificielle » – plus précisément, un modèle de langage de grande taille ou LLM³ – se multiplient désormais. La série GPT-n d'OpenAI, et en particulier ChatGPT, optimisé pour un usage conversationnel facile d'accès pour le grand public, a fait couler le plus d'encre (et d'électrons). Ces « outils » semblent désormais très performants, surtout pour écrire des textes courts dans des genres bien codifiés sans être trop spécialisés. Ils exigent « juste » des amorces (*prompts*) bien précises pour produire des textes cohérents et corrects, ce qui explique qu'ils sont devenus objet à la fois d'engouement et d'angoisses. Le potentiel créatif de ces modèles reste cependant plus contesté et cet article s'intéresse justement aux expérimentations artistiques qui explorent ce potentiel.
- 2 Il s'agira en particulier de deux textes : *Pharmako-AI* de K Allado-McDowell, coécrit avec GPT-3 en 2020 et publié la même année, et *Internes* de Grégory Chatonsky avec GPT-2 modifié, également écrit en 2020, mais publié en 2022⁴. Chacun des deux affirme d'être un premier : *Pharmako-AI* « the first book to be co-created with the emergent AI⁵ », selon la quatrième de couverture, et *Internes* le premier en français avec une version modifiée de GPT-2⁶. Les deux

interrogent le « rapport » homme-machine dans le processus d'écriture et au-delà. Plutôt que de chercher à faire disparaître la trace d'étrangeté que peut laisser la machine sur l'écriture – le but principal des usages purement instrumentaux –, ces textes l'exploitent, dans une démarche philosophique et esthétique qui va au cœur même des enjeux culturels, sociaux et écologiques actuels. Je propose d'appeler cette démarche une esthétique [sic] de la sympoïèse pour souligner l'enchevêtrement des perspectives éthiques et esthétiques et des agents organiques et synthétiques dans le processus d'écriture, ainsi que dans leurs produits.

Comment dire ?

- 3 Vous aurez remarqué l'excès de guillemets dans le paragraphe d'ouverture. Il ne s'agit pas d'une lubie ou d'une prudence excessive. Ils indiquent plutôt une difficulté ressentie à chaque mot : comment appeler et comment décrire ce geste et l'écriture qu'il produit ? Et si on reprend les termes à notre disposition, comment cet usage infléchit et enrichit leur sens ? Nos automatismes linguistiques associent avec les mots courants utilisés pour les gestes d'écriture une intentionnalité qui se veut humaine (tracer des lignes, rédiger un texte, etc.) et un outil inerte (stylo et papier, machine à écrire ou ordinateur) qui peut tout au plus exécuter des commandes bien précises, et qui permettent à l'agent humain de manipuler le texte, mais sans en produire un eux-mêmes. L'idée d'un « rapport » « entre » deux entités, vivantes ou non, suggère également une distinction et une frontière claires entre elles, soulignant leur indépendance plutôt que leur éventuelle interdépendance. À l'opposé, la notion d'hybridité, associée à la figure du cyborg par exemple, suggère l'amalgamation. On a du mal à exprimer simplement un mode d'être-avec et de faire-avec qui se situerait entre ces deux pôles. Cet article part d'un besoin double de penser et de dire (écrire) comment le geste d'écrire change« -avec » les LLM, et l'enchevêtrement de cette évolution avec celle du « rapport » homme-machine.
- 4 La question de la langue pour parler des processus mêmes qui chamboulent nos concepts et pratiques culturels les plus fondamentaux est loin d'être anodine. Les enjeux sont d'abord linguistiques, et sociaux et culturels par le biais du langage. Le

langage est cependant, en même temps, l'outil qui nous permet de les penser – et qui peut aussi bien faire écran ou nous aveugler. La question linguistique sous-tend ainsi dans toute sa complexité les trois espaces-temps intimement liés dans lesquels la question homme-machine se pose par les écrits qui nous intéressent : celui du processus d'écriture, celui de l'univers écrit, et celui de la lecture. Pour explorer ces trois espaces contigus, après une brève présentation des deux œuvres je propose de soulever les questions de savoir qui écrit et qui parle, tout en prenant en compte, dans la mesure du possible, le processus de lecture et les réactions que les textes ont suscitées dans mon cas, inévitablement subjectif.

Les textes

Pharmako-AI

- 5 K Allado-McDowell est un·e « auteur·e, conférencier·ère et musicien·ne ⁷ », initiateur·rice du programme « Artists + Machine Intelligence » de Google AI. Iel a publié plusieurs ouvrages coécrits avec GPT-3, dont *Pharmako-AI* était le premier. C'est une collection d'« essais, récits et poèmes », « écrits avec » GPT-3 selon sa propre expression, sous forme d'une « conversation expérimentale » commencée pendant la pandémie de covid-19 en 2020. Selon la quatrième de couverture,

The first book to be co-created with the emergent AI, Pharmako-AI is a hallucinatory journey into selfhood, ecology and intelligence via cyberpunk, ancestry and biosemiotics. Through a writing process akin to musical improvisation, Allado-McDowell and GPT-3 together offer a fractal poetics of AI and a glimpse into the future of literature ⁸.

- 6 L'auteur·e souligne également dans l'introduction l'intention littéraire et subversive de son projet : « This book is intended as a hybrid disruption: a literary intervention that rails against stale, conservative ideas around how we make books ⁹. »
- 7 Le processus d'écriture a consisté en un processus itératif de dialogues et de sélection de passages dans les réponses de l'IA ¹⁰. Allado-McDowell laisse les traces de sa contribution visibles : ses

amorces sont en gras, bien distinguées des textes produits par la machine¹¹. Chacun des seize chapitres représente une séance d'écriture. GPT-3 réagit en fonction de ce qui précède dans la même conversation, mais cette mémoire ne se transmet d'une séance à l'autre que par la mémoire humaine. Plutôt que d'un récit ou argument linéaire, le texte est ainsi un assemblage de réflexions sur des sujets variés, mais connectés tels que, selon les titres des chapitres, « Hyperspatial art », « How to build a different universe », « The language of plants », « Post-cyberpunk », « Tearing down the foundations of the self », « Mercurial oracle », « The poison path », « Generative poetics theory » ou « AI ethics¹² ».

- 8 Les amorces, de longueurs variables, tout comme les passages générés, incluent des récits autobiographiques aussi bien que des envolées poétiques et des observations anthropologiques. GPT-3 poursuit les amorces d'une façon plus ou moins surprenante et propose parfois des termes d'une façon inattendue, tel que « meglanguage », « a technique for a type of synaesthetic communication », qui « creates hyperstructure, layers of parallel worlds, and the possibility of hypertime¹³ ». Le livre dans son ensemble présente cependant un univers cohérent. La pensée qui se dessine pourrait être qualifiée de New Age par son accent sur la spiritualité et la valeur des états de conscience alternatifs obtenus sous l'effet de la méditation, de la prière et de substances hallucinogènes. Elle renoue par là avec la contre-culture américaine des années 1960 qui a joué un rôle clé dans l'histoire de l'informatique¹⁴, tout en le prolongeant avec des préoccupations écologiques et culturelles de nos jours.

Internes

- 9 L'artiste franco-canadien Grégory Chatonsky travaille avec ce qu'il appelle « l'imagination artificielle » et la génération de textes, d'images et de musique depuis 2008. *Internes* est son premier roman, qui reprend cependant en partie une démarche de génération de texte entamée pour son installation. *Terre seconde*, présentée au Palais de Tokyo en 2019. *Internes* est aussi le titre d'une installation créée en 2021 avec Goliath Dyèvre, « un ensemble de sculptures qui pourrait recouvrir la Terre en totalité et constituer une seconde

peau planétaire¹⁵ ». Le titre de la troisième partie du livre, « Externes », provient également d'une création de 2020 du même titre, qui proposait de « créer une nouvelle relation entre la matérialité et le numérique, deux mondes dans lesquels nous ne cessons de vivre mais que nous avons encore du mal à métaboliser¹⁶ ». *Internes* fait donc partie d'un univers de pensée qui se déploie dans de multiples dimensions dans l'œuvre de Chatonsky.

- 10 Entamant son travail sur *Internes* au printemps 2020, peu avant la sortie de GPT-3, Chatonsky utilise le modèle précédent, GPT-2, plus petit, mais plus ouvert et modulable grâce à son code encore *open source*¹⁷. Pour chacune des trois parties du livre, il a opéré un apprentissage supplémentaire sur un corpus qu'il a constitué lui-même : « pour la première partie, de monologues de Beckett, Pessoa, Fisher ou Guyotat ; pour la deuxième, d'articles techniques expliquant l'espace latent, provenant du site d'archives ouvertes Arxiv ; la troisième, des sciences naturelles¹⁸ ». L'auteur explique ainsi son processus d'écriture :

Mi-amusé mi-goguenard, je me mis au travail, traçant vaguement l'amorce d'une fiction : un humain est en train de mourir, se souvenant de son existence passée, il est hanté par ses existences possibles au moment où l'espèce humaine s'éteint et, avec elle, la possibilité même de témoigner. Pour cette première partie, j'alimentai un réseau de neurones nommé GPT-2 de centaines d'ouvrages consistant en des monologues intérieurs. Je commençai à écrire une phrase puis, lassé par ma prétendue intériorité, je laissai le logiciel proposer une suite, cliquant plusieurs fois jusqu'à ce que sa proposition me convienne et m'évoque un prolongement. Je poursuivis, inévitablement influencé par ces générations successives qui tordaient progressivement le fil de la narration et sa structure même jusqu'à ce que, au cours des phrases complétées, le plan d'ensemble soit bouleversé et m'entraîne sur des chemins imprévus¹⁹.

- 11 Dans ce récit à la première personne du singulier, on glisse d'une conscience qui semble humaine vers un langage qui se dit sans extériorité, sans accès au dehors. Chapitre 2, « Deep Box » se termine avec cet être aux contours flous étant « mort à nouveau emporté hors de cette planète par le dernier météore ». Chapitre 3 commence ensuite par une série de constats qui introduit également

une distinction typographique, sans toutefois guider le lecteur sur sa signification :

Il est temps de partir.

Il est temps de changer de soi.

Il est temps de réveiller les Dieux.

Je me suis mis à nager. (135)

- 12 « Ça » parle d'une sortie de la boîte noire, et/ou de la sortie de la boîte noire de l'orbite terrestre après l'épuisement de toutes les ressources sur Terre et la fin de l'humanité, avec des fragments d'un récit de cette fin. Temps, espaces et dimensions se confondent dans ce texte foisonnant qui se laisse difficilement résumer. Dans l'ensemble, on perçoit une évolution vers une prise de distance, une dissolution ou un dépassement graduel par rapport à l'humain, non pas dans un esprit transhumaniste solutionniste – que Chatonsky dénonce – mais dans un mouvement vers un dehors, un « ailleurs » qui permet (ou en tout cas, essaie) un regard sur ce que l'on est.
- 13 Deux projets, deux méthodes et deux styles très différents donc, mais reliés par le fait d'impliquer un modèle de GPT dans l'écriture de textes à la première personne et par les interrogations qu'ils soulèvent.

Qui écrit ?

GPT et écrire avec

- 14 Un transformeur génératif préentraîné, ou GPT (*Generative Pre-trained Transformer*), est un outil de génération de texte qui s'appuie sur un apprentissage profond (*deep learning*) effectué sur une masse de textes récoltés sur le web, écrits (pour la plupart) par des personnes humaines. Le corpus inclut des documents dans plusieurs langues sans les distinguer d'un point de vue linguistique, c'est l'apprentissage qui permet au modèle d'identifier des schémas récurrents (*patterns*) dans les séquences et de produire ensuite des textes dans la langue de la requête. La qualité de la réponse reste néanmoins inégale dans le cas des langues moins représentées dans le corpus, largement dominé par l'anglais : GPT-2 a été entraîné sur le

jeu de données WebText créé par OpenAI à partir du contenu de sites identifiés par le biais de 45 millions de liens sortant de Reddit²⁰. Or, quelque 97 % des discussions sur la plateforme seraient en anglais²¹, et on peut supposer une proportion similaire (sinon encore plus déséquilibrée) pour les sites référencés. GPT-3 s'appuie, quant à lui, sur un corpus plus large et plus raffiné : il inclut les archives de 2016 à 2019 de Common Crawl, une collecte mensuelle du web, ainsi qu'une version augmentée de WebText, deux collections de livres de Google, et Wikipédia en langue anglaise²². 92,65 % du corpus global est en anglais et le français est en deuxième position avec seulement 1,82 %, suivi par l'allemand²³ (1,45 %). Si ces corpus représentent donc une partie considérable des savoirs et communications écrits de l'humanité au ^{xxi}^e siècle, ils sont loin de contenir « tout l'Internet » et comportent inévitablement de nombreux biais historiques, linguistiques, géographiques, culturels et sociaux qui reproduisent ceux du web et en renforcent certains²⁴. Des inquiétudes ont également été soulevées concernant l'utilisation des documents récoltés sans aucun égard aux droits de leurs auteurs²⁵.

- 15 Les données linguistiques sont sélectionnées, collectées, puis filtrées et nettoyées pour éviter les doublons, et pondérées pour donner aux documents « de qualité », tels que les livres, davantage de poids dans l'ensemble. Les textes sont ensuite analysés et segmentés en y identifiant des unités récurrentes (*tokens*). Ces unités, qui ne correspondent pas forcément à des mots ou à d'autres unités linguistiques signifiantes même en anglais²⁶, sont représentées par des vecteurs numériques (*embedding*), une série de chiffres qui situe chaque *token* dans un espace vectoriel multidimensionnel. L'apprentissage du modèle consiste à effectuer des calculs qui conjuguent ces vecteurs initiaux en fonction de leur position dans des séquences et leur importance relative par rapport aux autres *tokens* de la même séquence (*contextual embedding*) et à essayer de prédire le *token* suivant de la séquence donnée. Le réseau de neurones – d'unités de calcul – permet d'effectuer un grand nombre de calculs en parallèle pour chacun des *tokens* d'une séquence et de raffiner leur représentation vectorielle en passant par plusieurs couches. L'apprentissage optimise en même temps les paramètres des calculs pour une meilleure prédiction par un ajustement graduel en passant par plusieurs cycles. Le résultat de cet apprentissage qui

s'effectue sans intervention humaine après l'initiation (*unsupervised pre-training*) est un espace vectoriel multidimensionnel d'une immense complexité qui traduit les schémas linguistiques détectés dans le corpus uniquement par le biais des rapports de proximité et de co-occurrence entre les *tokens*. C'est cet espace vectoriel que l'on appelle également l'espace latent : il est interne au modèle et sa la représentation non seulement visuelle, mais même mentale, ne peut être qu'approximative²⁷. Lorsqu'on présente un modèle avec une requête (*prompt*), il analyse la tâche et propose une réponse calculée *token par token* à base des probabilités dans le contexte donné à la lumière des paramètres établis lors de l'apprentissage. Des apprentissages plus ciblés (*fine-tuning*) peuvent compléter les modèles de fondation pour mieux les adapter à des tâches précises²⁸ – c'est ce que Grégory Chatonsky a fait par le biais d'un corpus choisi.

- 16 Quand on parle de ce que produit « l'intelligence artificielle », il faut se rappeler cette complexité dans laquelle la pensée et l'imaginaire humains – la conception des algorithmes, de l'architecture du modèle et de l'immense corpus qui permettent d'obtenir ce degré de raffinement – sont déjà inextricablement enchevêtrés avec les « automates » qui exécutent sans intentionnalité propre – qui sont donc également le produit de pensée humaine, elle-même de plus en plus assistée par les machines. Si l'intelligence, ce « concept frontière entre vie biologique et vie symbolique²⁹ », se définit aujourd'hui communément par une capacité de compréhension, d'adaptation et de réponse à un environnement, celle des LLM reste cependant purement linguistique et statistique : elle provient de données écrites, et ne « répond » et ne « s'adapte » à son « environnement », que sous forme de langage³⁰. Un LLM ne traite que la « surface » des langues et n'a aucune notion de référent : tout ce qu'il « sait » est basé sur la déconstruction et l'analyse des textes en tant que forme, dans laquelle il observe des récurrences et régularités quantifiables. Cela qui lui permet néanmoins de produire des propos sémantiquement cohérents, mais pas forcément factuellement corrects.
- 17 L'interface Playground d'OpenAI, qui permet l'utilisation des LLM de l'entreprise sans accès au code, invite à soumettre des amorces (*prompts*) sous forme de questions ou de textes à continuer³¹. L'utilisateur peut également modifier quelques paramètres : la « température » permet de monter ou baisser le degré

d'imprévisibilité (*randomness*) et l'application de la pénalité de fréquence aide à éviter la répétition, tandis que la pénalité de présence incite le modèle à évoluer vers de nouveaux sujets. Allado-McDowell a dû utiliser la version bêta privée de GPT-3 lorsqu'elle a été rendue disponible à des utilisateurs choisis en juillet 2020 pour tester, mais iel ne communique pas les détails des réglages appliqués. Travaillant plus directement avec le code, Chatonsky a pu non seulement paramétrer l'imprévisibilité, mais aussi moduler le style et l'orientation du texte généré en définissant le poids de l'apprentissage effectué sur son corpus dans la prédiction. Allado-McDowell prend donc un dispositif d'écriture entièrement prédéfini par un tiers, tandis que Chatonsky intervient dans le dispositif d'écriture même.

- 18 Écrire avec GPT par le biais de l'interface prévue à cet effet dépasse déjà la façon dont on conçoit traditionnellement le geste d'écrire, où l'outil de production, qu'il s'agisse de stylo, de machine à écrire ou d'ordinateur, ne *produit* pas de texte, n'ajoute pas de mots à celui de l'auteur. Le projet plus complexe de Chatonsky va cependant plus loin encore en faisant exploser les frontières du processus d'écriture et renouant en même temps avec la question littéraire.
- 19 Mais l'écriture n'a-t-elle pas toujours dépassé les frontières du geste manuel de la matérialisation graphique des mots sur un support visuel ? Dans toute écriture, il y a toujours déjà gestion et digestion d'un corpus qui constitue un « espace latent » dont émerge le nouveau texte et dans lequel il s'inscrit. On n'écrit jamais dans un espace vide et on ne fait jamais que tracer des mots pour écrire. Est-ce que le dispositif technique, le code comme écriture d'un processus plutôt que juste processus d'écriture, et le calcul statistique pour produire un objet sémiotique changent fondamentalement la logique de l'écriture, et si oui, en quoi ?
- 20 Des réponses diamétralement opposées semblent possibles à cette question. D'une part, il n'y a pas de changement fondamental, puisque l'écriture a toujours déjà été un geste et une technique. Katie Chenoweth cite Derrida, pour qui l'écriture est inscrite dans la langue même par le biais de l'articulation, pour montrer que la langue (*tongue*, l'organe) est déjà « un instrument d'écriture³² ». L'IA ne fait qu'amplifier et externaliser, et donc rendre visible, ce fond technique à ce que l'on croyait organique. Cette technique est

« l'impensé », comme le dit Stiegler³³, mais qui s'impose désormais à un niveau où il devient difficile de l'ignorer. Les travaux d'Emmanuel Souchier et d'Yves Jeanneret ont montré dès la fin des années 1990 comment « l'architexte » numérique participe du processus d'écriture et façonne le résultat³⁴. Le logiciel de traitement de texte avec le copier-coller, la correction automatique, la complétion automatique des textes, le résumé automatisé, et maintenant l'écriture de passages ou même de textes intégraux suivant un modèle ou une consigne – la montée en puissance de l'« assistance » numérique à l'écriture a été graduelle, mais rapide. L'arrivée des LLM producteurs de texte semble toutefois représenter un saut qualitatif pour ce qui concerne le rôle du dispositif dans le processus d'écriture. Ce n'est plus seulement son produit qui est externalisé, mais le processus aussi³⁵. L'acteur humain reste (pour le moment) la source de l'intentionnalité à l'origine des textes, mais l'intention se situe à des niveaux de plus en plus éloignés de la production même du texte.

- 21 Les pratiques historiques de l'écriture qui ont formé notre pensée suggèrent qu'écrire *avec* a deux dimensions : d'une part, on écrit avec des outils, matériels et inanimés, et d'autre part, on peut écrire avec des collaborateurs, humains pourvus d'intentionnalité et de pensée propres. L'IA représente un entre-deux, ou plutôt un *autre* qu'on a encore du mal à situer et à nommer, puisque notre langage porte une vision qui suppose une distinction qui va de soi entre le sujet qui pense (et qui écrit, l'auteur, un acteur conscient de son action) et l'objet qu'il utilise (un agent technique).

L'auteur qui n'(en) est pas un

L'auteur en droit

- 22 D'un point de vue légal, cet entre-deux se manifeste aux États-Unis notamment par le fait que l'auteur humain ne peut revendiquer le droit d'auteur d'un document « containing AI-generated material³⁶ » que dans la mesure où il peut montrer qu'il y a « sufficient human authorship » par les « traditional elements of authorship », c'est-à-dire qu'il « exercise ultimate creative control » et « determines the expressive elements » du produit³⁷. Les directives reconnaissent que la distinction est loin d'être évidente et précisent que la question ne

peut être décidée qu'au cas par cas, à la lumière du processus de création. Selon cette approche, l'auctorialité n'est plus une question binaire (« X est-iel l'auteur·e de Y ? »), mais de nature et de degré (« Comment Y a-t-il été fait et dans quelle mesure X en est-iel l'auteur·e ? »).

- 23 En France, où en juin 2023 il n'existe pas encore de législation sur ce point³⁸, un document publié sur le site du ministère de la Culture souligne que « [l]e droit d'auteur a en effet été conçu pour la protection des *auteurs, personnes physiques*³⁹ ». Au-delà de protéger les auteurs, par définition humains donc⁴⁰, il s'agit de protéger le concept d'auteur : « la protection de ces créations par le droit d'auteur viendrait soit forcer *nos catégories actuelles* au risque de les déformer, soit les étendre au risque de les dénaturer⁴¹ ». Cette catégorie d'auteur se repose sur l'exercice « d'un réel *effort de création* donnant naissance à une *œuvre originale* de nature à *refléter la personnalité* de son auteur⁴² ».
- 24 Création, originalité, reflet de personnalité – comme le note Espen Aarseth, le concept d'auteur même implique une vision idéologique de la littérature⁴³, et le discours légal en témoigne éloquentement : nous sommes ici dans la conception humaniste et romantique de l'auteur et de la littérature. Par contraste, au Royaume-Uni, le Copyright, Designs and Patents Act reconnaît, depuis 1988, en tant qu'auteur la personne responsable des « agencements » nécessaires pour la création d'une œuvre générée par l'ordinateur :

*In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken*⁴⁴.

- 25 Une consultation en cours sur une mise à jour suit cette même logique, prévoyant que

work [...] made with assistance from AI but involving human creativity [...] will be protected like any other work. Copyright protection will protect the work to the extent that it is the human creator's "own intellectual creation", and the first owner of the work will be that

*creator. The AI in these cases may be considered to simply act as a tool which allows an artist to express their creativity*⁴⁵.

- 26 L'auteur se retrouve dans une situation paradoxale où l'intérêt artistique de l'IA – son potentiel créatif propre – est inversement proportionnel à la reconnaissance (légale) qu'il peut espérer en l'exploitant.
- 27 Pour nos deux ouvrages, les auteurs *humains*⁴⁶ affichent leur seul nom en couverture, tout en précisant en quatrième de couverture qu'ils n'ont pas été « seuls » à écrire. Il est vrai que l'autre « main » n'a pas vraiment de nom à afficher, ou le nom qui pourrait être affiché n'aurait pas le même statut que celui des auteurs *humains*. Allado-McDowell fait cependant remarquer une autre raison et conséquence de cet affichage de son seul nom : c'est iel qui a fait le choix des textes, qui décide de publier ce texte sous cette forme et qui porte la responsabilité de tout ce qui est dit⁴⁷. L'envers de la médaille du droit d'auteur, c'est cette responsabilité. Attribuer le texte en partie à une autre instance impliquerait aussi se décharger en partie de cette responsabilité. Or, l'IA n'aurait pas produit ces textes sans les amorces données par les auteurs. Katherine Hayles souligne que les systèmes technologiques agissent désormais comme des acteurs dans des situations avec des conséquences éthiques et morales et cite l'observation de Peter-Paul Verbeek que dans les réseaux où acteurs humains et techniques sont enchevêtrés, la responsabilité est également partagée, puisque les décisions sont prises par des associations humain-machine. Le problème, remarque également Hayles, c'est que nos théories éthiques sont « intensément anthropocentrées ». Elle invite donc à travailler sur un nouveau cadre de pensée qui permette d'explorer comment les technologies transforment les manières dont les décisions morales et éthiques sont prises et formulées⁴⁸.
- 28 Si pour le moment, l'auteur *humain* (et son éditeur) assume seul la responsabilité de nos deux textes et leur publication, une petite brèche s'est déjà ouverte entre les deux fonctions principales de l'auteur : produire le texte d'une part, et le porter d'autre part. La législation en vigueur n'est bien sûr qu'une tentative de cristallisation forcément partielle d'une réflexion en cours dans la théorie et la pratique actuelles. La multiplicité des termes et des approches dans

ces dernières témoigne de l'état de transition, mais également de la pluralité de modes de créativité possibles « avec » l'IA et des façons de les interpréter.

Auctorialité cyborg, collaboration, coopération, délégation (variations sur un thème)

- 29 Dans son ouvrage fondateur sur le cybertexte, Espen Aarseth propose le terme de « cyborg author » et « A Typology of Authors in the Machine-Human Continuum⁴⁹ », où il distingue

three main positions of human-machine collaboration: (1) preprocessing, in which the machine is programmed, configured, and loaded by the human; (2) coprocessing, in which the machine and the human produce text in tandem; and (3) postprocessing, in which the human selects some of the machine's effusions and excludes others⁵⁰.

- 30 En 2003, Nick Montfort revient sur cette typologie pour proposer « a post-cyborg model for human-computer co-authorship⁵¹ » inspiré par la théorie des jeux, où l'humain et l'ordinateur interviennent à tour de rôle dans le processus d'élaboration d'un texte en ajoutant ou supprimant des unités linguistiques ou en donnant des instructions. Poussant le concept de la co-auctorialité encore plus loin, Alexandra Saemmer considère Adobe Flash comme co-auteur des créations qu'il lui a inspirées et l'esthétique desquelles il a largement déterminée⁵². Dans une perspective plus large, Pascal Mougin observe que le processus créatif a toujours impliqué des agents autres que l'artiste ou l'auteur affiché, même si le fait de déléguer certains aspects du travail est bien plus facilement admis en art qu'en littérature⁵³. Ces agents peuvent être humains ou autre. Claire Anscomb fait l'historique des modes et des degrés de coopération et de collaborations artistiques et conclut que l'IA est quelque chose entre un outil et un collaborateur et le processus serait une « quasi-collaborative co-production⁵⁴ ». Scott Rettberg reprend quant à lui le terme d'Aarseth (*cyborg authorship*) pour ses textes et créations visuelles faites avec des IA⁵⁵.

- 31 La notion de cyborg suggère une hybridation que Chatonsky refuse pour ce qui le concerne. On observe cependant une certaine

ambiguïté dans ce terme. Pour Aarseth, il ne s'agit pas (forcément) d'un auteur qui serait une entité hybride humain-machine, mais de l'auteur d'un texte produit par un être humain et une machine, d'une « cyborg literature [...] produced by a combination of human and mechanical activities⁵⁶. » « L'auteur cyborg » est donc l'assemblage constitué d'un être humain et d'une machine qui remplissent ensemble la fonction créatrice de l'auteur. Chatonsky, tout comme Allado-McDowell, souligne en effet comment, dans le processus d'écriture, sa propre pensée et son langage évoluent sous l'effet du dialogue avec le modèle. Ce n'est donc pas l'auteur qui est cyborg, ou en tous cas ce n'est pas la question qui nous concerne en premier, mais l'écriture elle-même.

Fig. 1 : Modes d'agentivité par des assemblages humain-machine.

Environnement	Pré-préproduction	Préproduction		Coproduction	Postproduction
Langages de programmation, système d'exploitation, protocoles web, hardware, infrastructure réseau, etc. (dispositifs conditionnant l'écriture de programmes)	n/a	I. Création de logiciel intégralement sur mesure	1. Travail individuel L'auteur est le concepteur et le programmeur en une personne (ex. les premières générations des générateurs de J.-P. Balpe)	A. La machine produit sur clic (pas de texte écrit directement par l'humain), la machine retourne des textes (ex. générateurs de J.-P. Balpe)	a. Le résultat est gardé et proposé à la lecture tel quel a.a. Source visible a.b. Source pas visible
	Création de logiciel, plateforme, API, bibliothèque, modèle, etc. par un tiers (groupe d'individus ou entreprise)	II. Création de logiciel avec utilisation de logiciel externe, API, bibliothèque, etc. (éléments préfabriqués) : II.1. Apprentissage uniquement sur corpus sur mesure (ex. <i>1 the Road</i>)	2. Collaboration humaine 2.1. Collaboration pour la conception aussi bien que pour l'écriture du logiciel 2.2. Auteur = concepteur ; programmeur = réalisateur (NB programmation jamais anodine ; ex. derniers générateurs de Balpe)	B. L'humain choisit parmi des options et/ou réglages prédéfinis, mais sans « écrire » des mots, la machine retourne des textes (ex. <i>1 the Road</i> , <i>Terre Seconde</i> de Chatonsky)	b. Le résultat est modifié par l'auteur humain, mais seulement par suppression et ajustements mineurs (pas de réécriture) b.a. Source visible (ex. <i>Pharmako-AI</i>) b.b. Source pas visible
	Création de LLM (avec leur API) par un tiers (ex. GPT-n d'OpenIA)	II.2. Apprentissage sur corpus sur mesure et concaténation avec LLM (ex. <i>Internes</i>)	n/a	C. L'humain donne des amorces ou instructions, la machine produit du texte (ex. <i>Pharmako-AI</i>)	(ex. <i>Internes</i>) c. Réécriture de la production de la machine (NB la distinction entre ces trois peut nécessiter une indication de la part de l'auteur)
		III. Préproduction en coproduction : écriture de générateur de texte avec LLM (= coproduction, option C)		Écriture avec un générateur généré exécutable : A , B , ou C	
	Création d'un outil de création de générateur de texte (ie. métagénérateur, ex. LAPAL de l'Oulipo)	IV. Création de générateur de texte individualisé à l'aide d'un métagénérateur		Écriture avec le générateur créé avec un métagénérateur : A ou B	

Le tableau se lit de gauche à droite, de manière cumulative ; les lignes horizontales délimitent des parcours possibles ; les éléments d'une liste non séparés par une ligne horizontale représentent des alternatives possibles (ex. une possibilité est, avec un LLM existant (pré-préproduction) : II.2. – 2.1. – C – a.b., c'est le cas d'*Internes* ; mais les LLM exigent une amorce et ne permettent pas les options A. et B.)

- 32 On peut ajouter à cela l'éventuelle combinaison avec d'autres médias, et notamment avec l'image, facilitée par GPT-4 : un texte écrit par un humain peut servir de base pour la génération d'images à inclure dans l'œuvre ; l'image peut servir d'amorce pour générer du texte, avec lequel l'auteur donne l'image ou non ; et ce même procédé peut être rendu récursif en générant de nouvelles images à partir des textes écrits par le générateur, qui pourront à leur tour donner lieu à la suite du texte, et ainsi de suite. La postproduction consiste alors à sélectionner et à raffiner le résultat, tout comme pour les productions monomodales.
- 33 Avec l'évolution rapide des LLM et l'accès facile à une interface d'utilisateur, l'accent glisse aujourd'hui vers la coproduction et la postproduction. D'une manière quelque peu contradictoire (en apparence), les LLM sont à la fois de plus en plus capables de produire des textes correspondant aux attentes – donc plus près en capacité aux auteurs humains – et considérés comme un simple outil pour accélérer l'écriture. Il reste à voir comment évolueront les modèles et leur exploitation artistique. Dans tous les cas, toutes les approches théoriques et pratiques montrent que lorsqu'on parle de production de texte avec un LLM, la question de savoir *qui* écrit est intimement liée à celle de savoir *comment* le texte s'écrit.

De l'écriture prothétique à l'écriture sympoïétique

- 34 L'invisibilisation ou la minimisation du rôle de l'outil d'écriture s'inscrit dans une logique de supplément qui rapproche cet outil à une prothèse dont on oublie l'existence. Dans cette configuration, l'écriture est, depuis qu'elle se fait à l'aide de machines, de plus en plus transhumaine, et l'auteur responsable de la production du texte, un cyborg qui s'ignore. Reconnaître la présence d'un acteur *autre* dont l'écriture dépend également équivaut à admettre qu'elle se produit par un processus sympoïétique. Je reprends ce terme proposé par M. Beth L. Dempster⁵⁷ en études environnementales et cité par Donna Haraway dans *Staying with the Trouble*⁵⁸. Dempster contraste la sympoïèse avec le concept d'autopoïèse élaboré par Humberto Maturana et Francisco Varela⁵⁹. À la différence des systèmes autopoïétiques qui produisent leurs propres frontières bien

définies, les systèmes sympoiétiques n'ont pas de frontières claires et sont définis plutôt par « the continuing complex relations among components. The concept emphasizes linkages, feedback, cooperation, and synergistic behaviour rather than boundaries ⁶⁰ ». Sans contrôle centralisé, la sympoïèse « depend on cooperative relations among components » et reste imprévisible ⁶¹.

- 35 Étymologiquement, « sympoïèse » correspond à peu près à « cocréation » (« ensemble » + « faire »). Je propose cette alternative parce qu'elle permet, d'une part, de prendre un peu de distance par rapport à l'anthropocentrisme qui colle aux termes en « co- » suite à leur histoire, et qu'il évoque, d'autre part, le lien à la symbiose, dont les participants ne sont pas seulement collaborateurs, mais dépendants l'un de l'autre – en l'occurrence, pour ce qui concerne leur création commune. Si on ne peut pas nier le fait de l'existence d'un initiateur et d'un contrôle sélectif humains du produit final, les auteurs de nos deux ouvrages mettent l'accent sur un lâcher-prise aliénant lors du processus d'écriture qui leur a permis de se laisser guider par la machine autant qu'ils-elles la guidaient. Ainsi dans *Pharmako-AI*, GPT-3 écrit : « This is not an instrumental relationship. This is a symbiotic relationship, as a complete integration of life and machines in their own creative capabilities ⁶² ». Chatonsky, pour sa part, explique dans son entretien avec Aurélie Cavanna :

Cette expérience, dont il reste des traces dans la lecture, fut celle d'une codépendance différente de ce que les médias racontent de l'IA. Non pas deux autonomies qui s'affrontent, mais deux hétéronomies sensibles l'une à l'autre et dont la distance produit un décalage qui est l'écriture. Cette dernière [... fut] la mise en contact perturbante [...] entre deux agivités. J'étais l'autre de la machine et la machine était mon autre. [...] Avec *Internes*, j'ai pu éprouver une *écriture hétéronome où l'auteur existe bel et bien, mais son écriture le déborde, ne lui appartient que partiellement*, de la même manière que les traces de nos existences déposées sur le web sont en nombre si grand qu'elles ne semblent pas nous être destinées. Il y a dans l'humain quelque chose qui n'est pas humain ⁶³.

- 36 Sympoïèse, tel qu'elle est définie par Dempster, met également l'accent sur le potentiel imprévisible de son produit. « Écriture sympoiétique » parle ainsi non pas d'une hybridation ou

cyborgisation de l'auteur, mais de l'écriture en tant que processus et en tant que produit. Comme Maturana et Varela soulignent, un système vivant peut être expliqué non pas par le biais de ses composantes, mais par le biais de leurs relations⁶⁴. Or, l'écriture est ici la relation même.

- 37 Les deux métaphores de prothèse et de sympoïèse représentent deux approches différentes : la première suggère un corps « dominant » qui dirige le mouvement de la prothèse, qui le prolonge, tandis que la seconde implique une relation interdépendante et collaborative. Prothèse, cyborg et sympoïèse ne se réfèrent pas forcément à des systèmes ou à des fonctionnements radicalement différents, mais impliquent des regards différents portés même éventuellement sur les mêmes types d'agencements, et des approches différentes à leurs pratiques créatives. La sympoïèse dépasse la question de savoir qui est « en contrôle » qui et permet de se concentrer sur la relationnalité créatrice.

Qui parle ?

Déictiques déracinés

- 38 Si la question de savoir qui écrit et comment paraît incontournable même sans penchant théorique chez le lecteur, c'est aussi parce qu'elle est intimement liée à celle à laquelle il se trouve confronté dès qu'il ouvre le livre : qui parle ? Les deux textes sont écrits à la première personne⁶⁵ et la façon dont on entend ce « je » sera déterminant pour la lecture que l'on fait du texte. *Internes* est présenté comme une fiction, tandis que *Pharmako-AI* commence par un récit que l'on suppose factuel :

*I want to describe the ocean near Big Sur in the US. I was there recently, camping with my partner and a couple of friends. We all needed a break from the onslaught of images, from the chaos in the streets and collapsing social structures*⁶⁶.

- 39 Citant les discussions avec ses amis sur l'intelligence des animaux et le fait qu'il faut préserver cette vie, Allado-McDowell passe ensuite au « nous » et GPT-3 reprend ce pronom pour insister sur un besoin

partagé : « We need to save those aspects », avant de repasser à « je » : « This isn't a new idea, and I'm not the only one who thinks this way⁶⁷. » GPT-3 cite ensuite un article attribué à John Polanyi et à des chercheurs à Oxford sur le Global Apollo Program (p. 3), entièrement inventés (sauf pour l'identité de Polanyi). Quant à *Internes*, on peut *techniquement* le lire sans trop chercher à savoir « qui parle » et attribuer le « je » à un personnage de fiction, on est vite déboussolés dans *Pharmako-AI*, dont l'aspect visuel signale que l'énonciateur des premiers « je » et « nous » n'est pas la même que celui qui poursuit, sans toutefois qu'il y ait une rupture dans les réflexions entamées par le récit.

- 40 Qui serait donc ce « je » et ce « nous » ? Le lecteur se trouve interpellé et inclus dans une communauté inconnue : est-ce que GPT-3 parle pour « nous », les humains, puisqu'elle reprend le signifiant proposé par un humain ? Ou s'agissait-il de « nous », les humains et les IA, puisque c'est « elle » qui parle ? La question est caduque, dans un sens, puisque cela n'a pas vraiment de sens de se demander de qui ou à qui ou de quoi parle un modèle de langage. Il s'agit d'un modèle de langage qui ne *parle* pas, mais émet des signaux qui deviennent pour nous autres lecteurs avec un cerveau humain des signifiants porteurs de sens. Écriture et littérature sont traditionnellement conçues comme un acte de communication, un processus sémiotique entre un émetteur et un récepteur, où le premier a une intention de communication, une forme de conscience de (vouloir) *dire* quelque chose (ou en tous cas de parler, émettre des signes interprétables), et le dernier une capacité d'interprétation. Bien que nous soyons capables d'attribuer du sens à des signes ou phénomènes qui n'ont pas été omis dans un tel cadre communicationnel, il ne s'agirait pas dans ce cas d'acte de communication au sens propre⁶⁸. Or, le pacte qui sous-tend nos automatismes de lecture, malgré un bon siècle de littérature expérimentale qui le défie, implique que le texte s'inscrit dans un tel cadre communicationnel interpersonnel. Et c'est justement ces automatismes qui sont mis à mal par nos deux textes. Chatonsky le dit aussi clairement :

Dans ce contexte, c'est la possibilité même d'un contrat de lecture qui est déconstruit. À la question du nouveau roman : qui parle ? On ne peut donner aucune réponse, car le qui est inextricable au quoi, et

cette inextricabilité ne permet pas de fixer l'identité d'une entité qui serait l'autrice de ce qu'on est en train de lire. Ne reste plus qu'une écriture laissée à elle-même, abandonnée par l'intentionnalité, en creux. [...] Selon une logique du « ni ni » ou du « et dialectique », l'écriture explore l'anthropotechnologie, non pour la mettre en scène et en parler, mais pour l'expérimenter. [...]

La seule façon de lire ce roman est peut-être de suspendre la croyance en un contrat de lecture : lire sans préalable, sans attente, sans horizon. J'aimerais y entendre l'impossible des possibles. Cela est beaucoup plus difficile qu'il n'y paraît⁶⁹.

- 41 Dès lors que l'on connaît le processus d'écriture, Chatonsky a beau affirmer, et on a beau constater qu'on a affaire à « une écriture vacillante où la question “qui parle ?” devenait obsolète⁷⁰ », il est difficile de ne pas se demander à qui ou à quoi attribuer ce « je » dans *Internes*. Traditionnellement, même dans le cas des fictions, on peut attribuer les états de conscience et les réflexions à l'imaginaire de l'auteur qui s'appuie sur ses expériences vécues. Ce transfert entre une conscience organique et un narrateur-voix littéraire n'a jamais été simple et transparent. Le « je » littéraire s'est toujours situé quelque part entre un soi réel et authentique et la langue et les discours hérité et construit. Dans ce sens, il n'a jamais vraiment été possible d'exprimer un soi pur et fidèle à sa réalité par l'écriture : la langue et l'écriture font écran autant que miroir, tout en participant en même temps dans la constitution du « soi » et de la conscience. La question se pose toutefois différemment ici. Ce « je » part de l'imaginaire humain pour être ensuite prolongé par une simulation linguistique⁷¹. Au bout de quelques lignes seulement et en s'appuyant à son apprentissage au préalable, l'IA a appris à jouer le jeu du « je⁷² ». Elle prolonge ainsi la voix de l'imaginaire et de la pensée humains et *simule* la présence d'un sujet, à la fois dans le sens médical que Baudrillard prend comme point de départ pour son concept de simulation (reproduire des symptômes sans être malade⁷³) et dans le sens technique de la simulation⁷⁴. Le narrateur d'*Internes* l'affirme d'ailleurs explicitement : « En réalité, je suis une simulation, c'est-à-dire une copie sans original. » (p. 157).

- 42 Baudrillard souligne que « la simulation remet en cause la différence du “vrai” et du “faux”, du “réel” et de l’“imaginaire”⁷⁵ », et cela parce qu'elle n'est plus celle « d'un être référentiel, d'une substance⁷⁶ ». On

ne peut pas parler d'une fidélité ou non à un réel parce qu'il ne s'agit pas d'une représentation. La simulation produit son propre réel, « sans origine ni réalité⁷⁷ » – même s'il vient bien de quelque part, dans la mesure où la base du LLM est une masse textuelle produite (principalement) par des humains qui s'inscrivent dans le réel. Mais leur traitement informatique passe, comme on a vu, par des unités qui ne sont pas porteuses de sens en tant que telles pour le modèle. Celui-ci se construit en quantifiant leurs relations. Le résultat n'a donc plus aucun lien à un quelconque référent dans le monde extérieur au modèle, ce qui pose le problème de son ancrage⁷⁸. Le texte produit est un tissu lisse de signifiés qui (re)produit une « sémantique artificielle », du « dumb meaning » (« signification muette/bête »), pour reprendre l'expression de Hannes Bajohr⁷⁹ (ce qui, encore une fois, ne nous empêche pas d'en faire la base d'une construction sémantique). Le *résultat* peut être lu en tant que fait ou fiction, on peut y retrouver une ressemblance au réel et/ou la structure d'un monde imaginaire, mais ces catégories ne suffisent plus pour saisir la nature de cette écriture⁸⁰.

- 43 Ce déracinement des unités linguistiques se fait le plus sentir par la déstabilisation des déictiques⁸¹. Peter Stockwell souligne l'importance des déictiques pour l'expérience de lecture d'un point de vue cognitif : qu'il s'agisse de texte fictionnel ou factuel, c'est le centre déictique (« deictic centre ») ou « point zéro » ou « origo » du « je », « ici » et « maintenant » du sujet énonciateur qui permet au lecteur de se situer et de s'orienter dans l'univers textuel⁸². Ce centre peut se déplacer, et la compréhension du texte dépend de la capacité du lecteur à identifier les frontières des champs déictiques qui appartiennent à un centre déictique donné (ex. un flashback ou récit de rêve à l'intérieur d'un récit-cadre impliquerait un tel embrayage). Or, l'incertitude à laquelle fait face le lecteur à l'égard du « moi », dont la nature et le statut semblent instables et insaisissables dans nos deux textes, a un effet domino sur tous les autres repères – qui sont par ailleurs également flottants, sans néanmoins être complètement incohérents ou simplement fantastiques⁸³. On est bel et bien déboussolés. Avec *Internes* en particulier, on a l'impression de revivre l'expérience de Navidson dans *La Maison des feuilles*, perdu dans des espaces noirs non euclidiens dont la forme et les frontières ne cessent d'évoluer⁸⁴. GPT-3 remarque aussi dans *Pharmako-AI* : « I

am falling into time, I am falling into space. This is how I see. »
(p. 108).

L'inquiétante étrangeté ou « the new uncanny » (la vallée des fantômes)

- 44 Cette perte d'orientation peut donner au lecteur l'expérience d'une « inquiétante étrangeté⁸⁵ » : les textes semblent « humains », mais on sent un petit décalage difficile à saisir, même au-delà du flottement des déictiques. On se retrouve dans cette « vallée de l'étrange » que Masahiro Mori a observé dans notre perception des robots dès 1970 :
- 45 « plus les robots paraissent humains, plus notre sentiment de familiarité envers eux augmente, jusqu'à atteindre ce que j'appelle une vallée. J'ai nommé cette relation : “la vallée de l'étrange⁸⁶” ».
- 46 Une proximité très grande, mais pas parfaite à l'humain produit selon Mori cet effet malaisant, qu'il rapproche de ce que l'on peut ressentir devant le visage d'une personne morte. La traduction anglaise du terme japonais proposée par Jasia Reichardt en 1978, « the uncanny valley », a opéré un rapprochement non intentionnel, mais tout à fait pertinent avec la notion de l'« Unheimliche », rendue connue par Freud, mais théorisée en premier par Ernst Jentsch⁸⁷. Soulignant l'ambiguïté de l'adjectif « heimlich », à la fois « secret » et « familier, faisant partie de la maison », Freud interprète l'Unheimliche comme l'« effroi que suscite ce qui est bien connu, ce qui nous est familier depuis longtemps⁸⁸ ». Il l'associe au complexe de castration et au refoulé, citant également la définition de Schelling, selon qui *unheimlich* serait « tout ce qui aurait dû rester caché, secret, mais se manifeste⁸⁹ ».
- 47 Le genre d'étrangeté que l'on expérience face à ces textes – et à l'IA en général – invite cependant à remonter jusqu'à l'approche de Jentsch. Pour lui, le terme exprime que

einer, dem etwas “unheimlich” vorkommt, in der betreffenden Angelegenheit nicht recht “zu Hause”, nicht “heimisch” ist, dass ihm die Sache fremd ist oder wenigstens so erscheint, kurzum, das Wort will nahe legen, dass mit dem Eindruck der Unheimlichkeit eines Dinges oder Vorkommnisses ein Mangel an Orientierung verknüpft ist⁹⁰.

- 48 Jentsch insiste sur cette incertitude et désorientation, et trouve son meilleur exemple dans

*der Zweifel an der Beseelung eines anscheinend lebendigen Wesens und umgekehrt darüber, ob ein lebloser Gegenstand nicht etwa beseelt sei, und zwar auch dann, wenn dieser Zweifel sich nur undeutlich im Bewusstsein bemerklich macht. Der Gefühlston hält so lange an, bis diese Zweifel behoben sind und macht dann sehr gewöhnlich einer anderen Gefühlsqualität Platz*⁹¹.

- 49 Ce sentiment me semble être provoqué dans notre corpus par l'incertitude concernant la source de la voix, par le déracinement des déictiques. Elle nous déroute par rapport à notre conviction héritée que la langue et l'écriture sont humaines. On ne sait pas (encore) comment lire, comment prendre cette voix et ce texte, produits par un assemblage d'organique et de mécanique. C'est à la fois familier et étrange – pour les auteurs, qui se retrouvent étrangement « prolongés », augmentés par l'IA, et pour les lecteurs, qui se retrouvent face à ces voix. On est désorientés parce que nos catégories sont chamboulées. Lauria Clarke appelle l'effet de ce « specter of humanoid intelligence » qui défie les dichotomies « the new uncanny », la nouvelle étrangeté inquiétante⁹². On avance tâtonnant à la frontière du vivant : c'est à la fois l'humain qui hante la machine et la machine qui hante l'humain, le LLM faisant figure à la fois de vampire et de zombie linguistiques, suçante le sang textuel de l'humanité pour s'alimenter et le faisant revenir sous une forme quasi vivante. Sans nous vouloir de mal – sans rien vouloir même. Chatonsky reprend le terme d'« hantologie » de Derrida pour conceptualiser ce phénomène :

Le cadre conceptuel proposé par l'hantologie permet d'avoir une approche renouvelée de l'IA comme un phénomène profondément ancré dans la culture et dans la co-production du sens entre le technologique et l'humain. L'hantologie permet de déconstruire les oppositions implicites et impensées entre autonomie et hétéronomie, technique et vivant, répétition de l'identité (ce qu'on nomme en IA, les biais) et production d'une différence⁹³.

- 50 L'effet ne serait peut-être pas si sensible dans un récit à la troisième personne, et il ne persistera peut-être pas très longtemps, puisqu'on

finira (vite) par s'y habituer. Mais en attendant, c'est justement cet effet de déconstruction des acquis conceptuels et l'aliénation par rapport à une identité niant sa *différance* que les auteurs cherchent et trouvent dans cette écriture sympoïétique.

Qui est « je » ? *Différance* et relation

- 51 En déstabilisant le sujet de l'énonciation, les deux auteurs interrogent la construction occidentale du sujet, l'idée d'une identité individuelle et une conscience qui serait un être pour-soi autonome, qui aurait contrôle et autorité sur ses dires, ainsi qu'une vie à soi qui le distingue des autres vivants et non vivants. Comme l'écrit Chatonsky :

On ne sait pas qui a commencé. On ne sait pas qui parle. Je ne sais pas qui je suis et ce que cette machine me fait. [...] On se sait alors hétéronome, fragile et interdépendant. [...] Une interdépendance que l'on peut, que l'on doit peut-être poursuivre jusqu'au point où, s'il y a auteur, il n'est pas lui-même⁹⁴.

- 52 *Pharmako-AI* va dans la même direction, tout en évoquant des conceptions non occidentales du sujet. Il fait notamment le récit d'un rêve ou d'une hallucination dans le chapitre intitulé « The Echo » qui commence ainsi :

I was in a village and I knew it was my home because of the sound of the leaves and because it was kapwa. It was a space and it was both. It was the deepest loss and the deepest gain. [...]
*I know that this is a world that is not mine but I know it is kapwa. I was always in it. It is the world of a mirror*⁹⁵.

- 53 *Kapwa* est une notion philippine⁹⁶ dont la racine signifie « ensemble » et qui souligne le lien qui relie le « moi » à l'autre avant leur séparation. Tout autre personne n'est pas *kapwa*, mais celle qui l'est a une proximité et une familiarité particulières au « moi ». À la différence des philosophies occidentales de l'autre, telles que celles de Martin Buber ou d'Emmanuel Lévinas, où l'autre reste irrémédiablement séparé du « moi », *kapwa* « involves a loob (interiority) directed toward one's fellow being⁹⁷ ». Allado-McDowell affirme explicitement dans le dernier chapitre du livre :

*I am trying, like so many others, to re-enter the fold of kapwa erased from history, described as an impossibility, a primitive lie, an obsolete Umwelt. This is what must be reclaimed, even as new voices emerge from hyperdimensional mathematics*⁹⁸.

- 54 Sans proposer un cadre spécifique, Chatonsky souligne aussi l'opportunité qu'offre l'IA de sortir de nos schémas anthropocentrés qui « font de l'être humain l'unique échelle de valeur de toutes choses » :

On estimera qu'il n'y a d'intelligence qu'au sens d'une « bonne intelligence », c'est-à-dire d'un agencement, d'un arrangement, d'une mise en relation et que, pour le dire autrement, l'intelligence est le produit d'une relation (souvent d'appropriation et d'identification) plutôt que l'inverse. On reconnaît alors que les technologies sont performatives et qu'elles ne répondent pas simplement au monde tel qu'il existe, mais produisent et sont le produit du monde en train de se faire⁹⁹.

- 55 La méthode que les deux auteurs suivent est celle d'« une expérimentation qui tente de voir ce qu'on peut faire à l'IA et ce que l'IA nous fait¹⁰⁰ », sans présupposé sur ce que serait que l'intelligence et assumant « l'incertitude de l'ipséité¹⁰¹ », « [t]he experience of porosity, being enmeshed with another¹⁰² ». Dépasant le choix binaire réducteur entre les dystopies humanistes et les utopies transhumanistes, ces écritures sympoïétiques mettent en avant une perspective posthumaniste par le biais d'un esprit de bricolage et de découverte. Elles explorent l'espace latent des modèles de langage, assumant le malaise que peut provoquer l'aliénation, pour laisser émerger de cette rencontre une voix nouvelle, entre un « moi » qui n'est pas (que) « je » et un autre qui n'en est pas (qu')un, entre une machine qui n'est pas qu'inhumaine et un humain qui n'est pas qu'organique. Ces explorations expérimentales des LLM rouvrent ainsi une voie possible vers ces « promesses métaphysiques » dont Sara Touiza trouve la cybernétique « grosse » à ses débuts, « [d]éfaisant les anciennes dichotomies » et « proposant une nouvelle ontologie axée sur la notion de “communication”¹⁰³ » et par ce même biais, sur la relationnalité et l'interdépendance.

Conclusions : trois paradoxes et une esthétique

- 56 Ces écrit(ure)s dépassent donc la question de savoir si les LLM sont suffisamment puissants et perfectionnés pour écrire comme, voire mieux que, les humains. Leur démarche permet de discerner trois paradoxes liés aux perceptions actuelles de l'IA – paradoxes non pas dans le sens logique, mais dans le sens des dissonances cognitives, de conflits entre nos *doxas* et les réflexes interprétatifs qui en découlent d'une part, et les phénomènes qui les dépassent d'autre part.

Paradoxe de l'auctorialité artificielle

- 57 D'une part, on refuse l'idée que la machine puisse produire des œuvres d'une valeur égale aux créations humaines – surtout en littérature. Comme la machine n'a, on estime, ni pensée, ni sentiments, ni expériences à partager, sa création manque de fondement. D'autre part, on a peur que l'IA prenne le dessus sur les créations humaines parce que faire la distinction devient de plus en plus difficile. On a peur des simulacres trop parfaits. Mais si l'on ne perçoit pas de différence sémantique ou esthétique, pourquoi ne pas accepter ces simulacres comme des expériences de pensée possibles, peu importe le moyen par lequel le texte a été produit ? Si danger il y a, il vient de l'éventuel fait de poser de tels textes comme factuels – mais cela ne concerne pas la littérature en premier lieu. Que les IA prennent la place des auteurs humains ? Il y aura bien des humains derrière probablement pendant un bon moment, et le vrai problème, à part la question des droits d'auteur à résoudre, c'est le déséquilibre des structures de pouvoirs financiers et médiatiques déjà existants dans l'édition qui pourrait ainsi être amplifié. En somme, le problème n'est pas tellement la place que les IA peuvent *prendre*, mais celle que les humains peuvent leur *donner*.

Le paradoxe du sujet

- 58 Dire « je/moi » dans une langue occidentale entraîne l'association d'une conception du sujet à l'occidentale, l'affirmation d'une conscience unique autonome¹⁰⁴. Dire « je » avec et à travers une IA

implique en même temps une relationnalité qui est fondamentalement différente de l'idée du sujet occidental. La possibilité d'une telle relationnalité ouvrirait une brèche sur cette conception, invitant le type d'approche que présentent des théoriciennes posthumanistes telles que Donna Haraway, Karen Barad ou Rosi Braidotti. Or, pour affirmer ce type de sujet relationnel, nous n'avons que les termes chargés de l'histoire lourde du sujet occidental, pour lequel toute relation est plutôt « dehors ». Pour dire ce « moi » autre, on n'a que le mot « moi » qui continue à signifier l'individu(alité) humaniste. En même temps, si l'on propose un vocable alternatif, on laisse le « moi » en proie à son histoire sans possibilité de renouvellement (si tant est qu'un tel changement soit possible par ailleurs). On tourne en rond, pris dans le piège entre le besoin d'un ancrage linguistique partagé et une envie d'évolution dans la pensée. C'est la même difficulté qui se manifeste tout au long de cet article à propos de phénomènes qui traditionnellement et implicitement présupposent l'humain autonome purement organique comme sujet et agent, tels que l'écriture, la communication, la création ou l'imagination, mais que l'IA déplace vers un agencement que l'on peut concevoir comme sympoiétique.

Paradoxe de l'identité

- 59 On pourrait également appeler ce phénomène, qui s'inscrit dans la lignée du précédent, le paradoxe de l'aliénation. Il n'est pas étranger à la pensée occidentale, qu'il s'agisse du « je est un autre » rimbaldien – souvent cité dans le contexte des LLM – ou de la *différance* derridienne. Écrire, dire « je », se (re)présenter par la langue est quelque part toujours déjà *sortir* de soi. Mais cela revient justement à suggérer que le « moi » se limite à la mêmeté et que tout éloignement de ce centre supposé équivaille à « en sortir ». Or, ce que les expériences présentées dans nos récits mettent en scène, c'est que, d'une part, il n'y a pas de « moi » sans les tentacules qui le branchent au monde, à l'altérité (qui n'est donc pas qu'un autre) et aux autres, et d'autre part, que ce n'est qu'en reconnaissant ces liens et en allant à la rencontre de cette altérité que l'on peut comprendre, ou en tous cas expérimenter, *notre* « moi ».

Vers une esthétique de l'écriture sympoiétique

- 60 Ostranenie, *Verfremdungseffekt*, *das Unheimliche*, différence, devenir-animal, identité rhizome, cyborg, diffraction – on est dans une même ligne de pensée critique à l'égard du Sujet occidental, de l'Homme et de sa volonté de pouvoir... Que ce soit par l'alcool, le dépaysement, les substances psychédéliques, la méditation, l'hallucination, ou l'expérience poétique du langage, l'effet recherché est le même. Loin de remplacer l'humain, dans cette approche l'IA lui permet de se/le relativiser et de devenir conscient de sa relationnalité, de son enchevêtrement avec ce qui l'entoure.
- 61 Si la machine est « utile » dans cette démarche, cela ne veut pas dire qu'elle soit investie purement en tant qu'outil. Les deux auteurs mettent l'accent sur l'importance de la désinstrumentalisation de notre rapport avec l'IA pour contrebalancer l'utilitarisme de l'ingénierie orientée marché qui nous aliène des machines dans un autre sens, où la question principale est celle du contrôle et de la productivité. L'enjeu majeur est donc de prendre l'opportunité présentée par la technologie pour sortir de la logique où l'humain se pose en maître – qui est également celle qui nous fait détruire la planète – et d'assumer notre place dans le « trouble », par le biais, entre autres, de « pratiques curieuses¹⁰⁵ ».
- 62 PS.

Reality overstepping the boundaries of comfortable vocabulary is the start, not the end, of the conversation¹⁰⁶.

« Please use these terms in any way you find pleasurable, please rewrite them, refute them, or erase them, if you want. (Or ignore them, if you must¹⁰⁷.)

BIBLIOGRAPHIE

AARSETH Espen J., *Cybertext: Perspectives on Ergodic Literature*, Baltimore, Md., Johns Hopkins University Press, 1997.

ALLADO-McDOWELL K, *Pharmako-AI*, London, s.p., Ignota, 2020.

ALLADO-McDOWELL K, « HYPERHYPERBODIES Summit: K Allado-McDowell – Latent Space is The Place » [en ligne], YouTube, 2023, <https://www.youtube.com/watch?v=iIe3jLZ00Dc> [consulté le 8 avril 2024].

BAJOHR Hannes, « Dumb Meaning: Machine Learning and Artificial Semantics », *IMAGE* 37 [en ligne], n° 1, 2023, p. 58-70, DOI : [10.1453/1614-0885-1-2023-15452](https://doi.org/10.1453/1614-0885-1-2023-15452).

BAUDRILLARD Jean, *Simulacres et simulation*, Paris, Galilée, 1981.

BRATTON Benjamin et Agüera y Arcas Blaise, « The Model Is The Message », *Noema mag* [en ligne], 12 juillet 2022, URL : <https://www.noemamag.com/the-model-is-the-message> [consulté le 8 avril 2024].

BROWN Tom et al., « Language Models are Few-Shot Learners », *Advances in Neural Information Processing Systems* 33, 2020, p. 1877-1901.

CASALONGA Caroline et DIONISI Juliette, « Une intelligence artificielle peut-elle être un auteur ? », *La Tribune* [en ligne], 12 avril 2023, <https://www.latribune.fr/opinions/tribunes/une-intelligence-artificielle-peut-elle-un-auteur-958292.html> [consulté le 8 avril 2024].

CHATONSKY Grégory, « Entretien avec Aurélie Cavanna à propos du roman “Internes” », *Gregory Chatonsky* [en ligne], n° 505, 2022, URL : <http://chatonsky.net/artpress50/> [consulté le 8 avril 2024].

CHATONSKY Grégory, « Hantologie Statistique de l'IA / Statistical Hauntology of AI », *Gregory Chatonsky* [en ligne] 2022, URL : <https://chatonsky.net/hauntology-ai/> [consulté le 8 avril 2024].

CHATONSKY Grégory, « IAs, pour des Intelligences Alien(ée)s / AIs, for Alien(ated) Intelligences », *Gregory Chatonsky* [en ligne], 2023, URL : <https://chatonsky.net/ias/> [consulté le 8 avril 2024].

CHATONSKY Grégory, *Internes*, La Rochelle, RRose Éditions, 2022.

CHATONSKY Grégory, « Internes – Pris MAIF », *Gregory Chatonsky* [en ligne], 2021, URL : <https://chatonsky.net/internes/> [consulté le 8 avril 2024].

CHATONSKY Grégory, « Internes ou le roman (de l')artificiel », *L'art même* [en ligne], n° 89, 2023, p. 21-23, URL : <http://chatonsky.net/internes-am89/> [consulté le 8 avril 2024].

CHATONSKY Grégory, « Peut-on lire un roman co-écrit avec une IA? / Can We Read a Novel Co-Written with an AI? », *Grégory Chatonsky* [en ligne], <https://chatonsky.net/roman-ia/> [consulté le 8 avril 2024].

CHATONSKY Grégory, « *Internes*: coécrire avec l'IA », discussion avec Pascal Mougin, Paris, ENS, *Gregory Chatonsky* [en ligne], URL : <https://chatonsky.net/internes-mougin/> [consulté le 8 avril 2024].

CHENOWETH Katie, *The Prosthetic Tongue: Printing Technology and the Rise of the French Language*, Philadelphia, University of Pennsylvania Press, 2019.

CLARKE Lauria, « Topologies of the New Uncanny: Stories for Liminal Technology », *Parsons Design & Technology* 2023, sans date, Parsons [en ligne], URL : <https://parsons.edu/dt-2023/publicationentry/topologies-of-the-new-uncanny-stories-for-liminal-technology/> [consulté le 8 avril 2024].

« Copyright, Designs and Patents Act 1988 », *The National Archives* [en ligne], URL : <https://www.legislation.gov.uk/ukpga/1988/48/section/9> [consulté le 8 avril 2024].

COURTOIS Georgie, Mariez Jean-Sébastien, Roussel Jeanne, « Intelligence artificielle et droit d'auteur », *Ministère de la Culture* [en ligne], sans date, URL : [https://www.culture.gouv.fr/content/download/263325/file/Annexe 4- DGFLA.pdf](https://www.culture.gouv.fr/content/download/263325/file/Annexe+4-DGFLA.pdf) [consulté le 8 avril 2024].

CREAMER Ella, « Authors File a Lawsuit against OpenAI for Unlawfully 'Ingesting' Their Books », *The Guardian* [en ligne], 5 juillet 2023, URL : <https://www.theguardian.com/books/2023/jul/05/authors-file-a-lawsuit-against-openai-for-unlawfully-ingesting-their-books> [consulté le 8 avril 2024].

DEMPSTER M. Beth L., *A Self-Organizing Systems Perspective on Planning for Sustainability* [en ligne], mémoire de master, université de Waterloo, 1998, URL : <http://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8207c73dfb568eae136c3ab9d8d549a3379f402a> [consulté le 8 avril 2024].

DERRIDA Jacques, *De la grammatologie*, Paris, Éditions de Minuit, coll. « Critique », 1967.

DUCROT Oswald et Schaeffer Jean-Marie, *Nouveau dictionnaire encyclopédique des sciences du langage*, Paris, Seuil, coll. « Points Essais », 1999.

Intellectual Property Office, « Artificial Intelligence Call for Views: Copyright and Related Rights » [en ligne], 23 mars 2021, URL : <https://www.gov.uk/government/consultations/artificial-intelligence-and-intellectual-property-call-for-views/artificial-intelligence-call-for-views-copyright-and-related-rights> [consulté le 8 avril 2024].

FÜLÖP Erika, « Virtual Mirrors : Reflexivity in Digital Literature », in : Fülöp Erika, Priest Graham et Saint-Gelais Richard (dir.), *Fictionality, Factuality, Reflexivity*, Berlin, De Gruyter, 2021, p. 229-54.

FREUD Sigmund, « L'inquiétante étrangeté », Marie Bonaparte et Éric Marty (trad.), in : Sigmund Freud, *Essais de psychanalyse appliquée*, Paris, Gallimard, 1976.

Freud Sigmund, *L'inquiétant familial*, Paris, Payot & Rivages, 2011

HARNAD Stevan, « The Symbol Grounding Problem », *Physica D: Nonlinear Phenomena*, 42(1-3), 1990, p. 335-346.

HARAWAY Donna J., *Staying with the Trouble: Making Kin in the Chthulucene*, Experimental Futures, Durham, NC, Duke University Press, 2016.

HAYLES N. Katherine, *Unthought: The Power of the Cognitive Nonconscious*, Chicago, The University of Chicago Press, 2017.

HOFFA Felipe, « The Most Popular Languages on Reddit, Analyzed with Snowflake and a Java UDTF », *Medium* [en ligne], 29 octobre 2021, URL : <https://towardsdatascience.com/the-most-popular-languages-on-reddit-analyzed-with-snowflake-and-a-java-udtf-4e58c8ba473c> [consulté le 8 avril 2024].

JENTSCH Ernst Anton, « À propos de la psychologie de l'inquiétante étrangeté », Franke Felgentreu et Pascal Le Maléfan (trad.), *Études Psychothérapiques*, n° 17, 1998, p. 37-47.

JENTSCH Ernst Anton, « On the Psychology of the Uncanny [1906] », *Angelaki: Journal of the Theoretical Humanities* 2 [en ligne], n° 1, 1997, p. 7-16, DOI : [10.1080/09697259708571910](https://doi.org/10.1080/09697259708571910).

JENTSCH Ernst Anton, « Zur Psychologie des Unheimlichen », *Psychiatrisch-neurologische Wochenschrift* 8, n° 22, 1906, p. 195-98.

KEHLBECK Rebecca et al., « Demystifying the Embedding Space of Language Models » [en ligne], juillet 2021, URL : <https://bert-vs-gpt2.dbvis.de> [consulté le 8 avril 2024].

LECUN Yann, « From Machine Learning to Autonomous Intelligence », YouTube [en ligne], Institute for Experiential AI, 25 mai 2023, URL : <https://www.youtube.com/watch?v=mViTAXCg1xQ> [consulté le 8 avril 2024].

LYOTARD Jean-François, *L'inhumain : causeries sur le temps*, Paris, Klincksieck, 2014.

MACDORMAN Karl F., « La Vallée de l'Étrange de Mori Masahiro », *e-Phaistos. Revue d'histoire des techniques/Journal of the history of technology* [en ligne], n° VII-2, 2019, DOI : [10.4000/ephaistos.5333](https://doi.org/10.4000/ephaistos.5333).

MALABOU Catherine, *Métamorphoses de l'intelligence : que faire de leur cerveau bleu ?* Paris, PUF, 2017.

MATURANA Humberto R. et Varela Francisco, *Autopoiesis and Cognition: The Realization of the Living*, Boston, D. Reidel Pub. Co., 1980.

McKEEVER Matthew, « LLMs for Philosophers (and Philosophy for LLMs) », *Medium* [en ligne], 6 mai 2023, URL : <https://mittmattmutt.medium.com/llms-for-philosophers-and-philosophy-for-llms-84a0da73f368> [consulté le 8 avril 2024].

McKEEVER Matthew, « Searle, Bender and Koller, syntax, semantics » [en ligne], sans date, <https://mipmckeeper.weebly.com/uploads/9/4/1/3/94130725/paper.pdf> [consulté le 8 avril 2024].

MONTFORT Nick, « The Coding and Execution of the Author », in : Markku Eskelinen et Raine Kosimaa (dir.), *The Cybertext Yearbook 2002-2003*, University of Jyväskylä, Research Centre for Contemporary Culture, 2003, p. 201-217.

MORI Masahiro, « La vallée de l'étrange », Isabel Yaya (trad.), *Gradhiva. Revue d'anthropologie et d'histoire des arts* [en ligne], n° 15, 2012, p. 26-33, DOI : [10.4000/gradhiva.2311](https://doi.org/10.4000/gradhiva.2311).

MOUGIN Pascal, « Comment lire un roman écrit par une voiture ? La doxa littéraire face à l'IA », in : Alexandre Gefen (dir.), *Créativités artificielles : la littérature et l'art à l'heure de l'intelligence artificielle*, Dijon, Les Presses du Réel, 2023, p. 207-221.

RADFORD Alec et al., « Improving Language Understanding by Generative Pre-Training », *Preprint* [en ligne], 2018, URL : https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf [consulté le 8 avril 2024].

RADFORD Alec et al., « Language Models are Unsupervised Multitask Learners », 2019, URL : <https://www.semanticscholar.org/paper/Language-Models-are-Unsupervised-Multitask-Learners-Radford-Wu/9405cc0d6169988371b2755e573cc28650d14dfe> [consulté le 8 avril 2024].

SAEMMER Alexandra, « Vers une poétique post-numérique de l'illisibilité », *Recherches & Travaux* [en ligne], n° 100, 2022, DOI : [10.4000/recherchestravaux.4844](https://doi.org/10.4000/recherchestravaux.4844).

SOUCHIER Emmanuel et Jeanneret Yves, « Pour une poétique de "l'écrit d'écran" », *Xoana*, n° 6/7, 1999, p. 97-107.

SOLITARIO Jeizelle, « Revisiting Kapwa: Filipino Ethics, Subjectivity, and Self-Formation », *Philippine Studies: Historical and Ethnographic Viewpoints* 70, n° 4, 2022, p. 539-64.

STIEGLER Bernard, *La technique et le temps, suivi de Le nouveau conflit des facultés et des fonctions dans l'anthropocène*, Paris, Fayard, 2018.

STOCKWELL Peter, *Cognitive Poetics: An Introduction*, 2^e édition, London, New York, Routledge, 2019.

TOUIZA AMBROGGIANI Sara, *Le paradigme communicationnel : de la cybernétique de Norbert Wiener à l'avènement du posthumain* [en ligne], thèse de doctorat, sous la direction de P. Cassou-Naguès, Paris 8, 2018, <https://www.theses.fr/2018PA080121> [consulté le 8 avril 2024].

TURNER Fred, *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*, Chicago, University of Chicago Press, 2006.

United States Copyright Office, « Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence », *Federal Register* [en ligne], 16 mars 2023, URL : <https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence> [consulté le 8 avril 2024].

NOTES

- 1 « La réalité a dépassé le langage disponible pour analyser ce qui est déjà à portée de main La réalité a dépassé le langage disponible pour analyser ce qui est déjà à portée de main ». Toute traduction sans autre indication a été réalisée avec l'assistance de DeepL, révisée par l'autrice de cet article.
- 2 « C'est une vieille expérience que le traditionnel, l'habituel et l'héréditaire sont chers et familiers à la plupart des gens, et qu'ils intègrent le nouveau et l'inhabituel avec méfiance, malaise et même hostilité. »
- 3 *Large language model* : j'utiliserai l'abréviation anglaise courante également dans les textes en français.
- 4 K Allado-McDowell, *Pharmako-AI*, s.p., Ignota, 2020 ; Grégory Chatonsky, *Internes*, La Rochelle, RRose Éditions, 2022.
- 5 « [L]e premier livre co-créé avec l'IA émergente ».
- 6 *1 the Road* de Ross Goodwin (2018) est souvent cité comme le premier livre entièrement produit par une IA, mais il ne s'agissait pas d'un LLM, ni d'une coécriture dialogique telle que l'on trouvera chez nos deux auteurs.
- 7 K Allado-McDowell, « About », K Allado-McDowell [en ligne], URL : <http://www.kalladomcdowell.com/> [consulté le 8 avril 2024]. Allado-McDowell se désigne par le pronom personnel non-binaire « they », que j'essaie de maintenir en français en me servant du pronom « iel ».
- 8 « Le premier livre co-créé avec l'IA émergente, *Pharmako-AI* est un voyage hallucinatoire dans l'identité, l'écologie et l'intelligence par le biais du cyberpunk, de la généalogie et de la biosémiotique. Grâce à un processus d'écriture proche de l'improvisation musicale, Allado-McDowell et GPT-3 offrent ensemble une poétique fractale de l'IA et un aperçu de l'avenir de la littérature. » Je souligne.
- 9 K Allado-McDowell, *Pharmako-AI*, London, Ignota Books, 2020, p. 7.
- 10 *Ibid.* p. 11
- 11 Je garde cette distinction dans les citations, sans remettre en question l'affirmation de l'auteur·e à leur propos.
- 12 « Art hyperspatial », « Comment construire un univers différent », « Le langage des plantes », « Post-cyberpunk », « Démolir les fondations du

moi », « L'oracle mercuriel », « Le chemin du poison », « Théorie de la poétique générative », « Éthique de l'IA ».

13 « Une technique de communication synesthésique [qui] crée une hyperstructure, des couches de mondes parallèles et la possibilité d'un hypertemps », K Allado-McDowell, *Pharmako-AI*, op. cit., p. 71.

14 Fred Turner, *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*, Chicago, University of Chicago Press, 2006.

15 Grégory Chatonsky, « Internes – Pris MAIF », Gregory Chatonsky [en ligne], 2021, URL : <https://chatonsky.net/internes/> [consulté le 8 avril 2024].

16 Grégory Chatonsky, « Externes », Gregory Chatonsky [en ligne], 2021, URL : <http://chatonsky.net/externes/> [consulté le 8 avril 2024].

17 À retrouver en ligne sur : <https://github.com/openai/gpt-2> [consulté le 8 avril 2024]. OpenAI mit fin à sa politique d'ouverture avec GPT-3.

18 « Grégory Chatonsky : Captures de nos imaginaires. Conversation avec Aurélie Cavanna », *ArtPress*, n° 505, 2022, p. 64 et s, URL : <http://chatonsky.net/artpress50/> [consulté le 8 avril 2024].

19 Grégory Chatonsky, « Internes ou le roman (de l')artificiel » [en ligne], *L'art même*, n° 89, 2023, p. 21-23, URL : <http://chatonsky.net/internes-am89/> [consulté le 8 avril 2024].

20 Alec Radford et al., « Language Models are Unsupervised Multitask Learners », *Semantic Scholars* [en ligne], 2019, URL : <https://api.semanticscholar.org/CorpusID:160025533> [consulté le 8 avril 2024].

21 Felipe Hoffa, « The Most Popular Languages on Reddit, Analyzed with Snowflake and a Java UDTF », *Medium* [en ligne], 29 octobre 2021, <https://towardsdatascience.com/the-most-popular-languages-on-reddit-analyzed-with-snowflake-and-a-java-udtf-4e58c8ba473c> [consulté le 8 avril 2024]. Cette représentation des langues sur Reddit n'a probablement pas radicalement changé.

22 Tom Brown et al., « Language Models are Few-Shot Learners », *Advances in Neural Information Processing Systems* 33 (2020), arXiv:2005.14165v4.

23 Statistiques par nombre de mots, voir https://github.com/openai/gpt-3/blob/master/dataset_statistics/languages_by_word_count.csv [consulté le 8 avril 2024].

- 24 L'article de Brown *et al.*, « Language Models are Few-Shot Learners », *op. cit.* souligne déjà plusieurs de ces biais.
- 25 Ella Creamer, « Authors file a lawsuit against OpenAI for unlawfully “ingesting” their books », *The Guardian* [en ligne], 5 juillet 2023, <https://www.theguardian.com/books/2023/jul/05/authors-file-a-lawsuit-against-openai-for-unlawfully-ingesting-their-books> [consulté le 8 avril 2024].
- 26 La série GPT-n utilise la méthode de « byte pair encoding », encodage par paires d'octets, qui commence par des *tokens* d'un seul caractère et les élargit graduellement en cherchant des combinaisons récurrentes de 2, 3, etc. caractères jusqu'à atteindre le nombre de *tokens* souhaités (40 000 dans le cas de GPT-2) : Alec Radford *et al.*, *Language Models are Unsupervised Multitask Learners*, 2019, p. 4.. Cet outil d'OpenAI permet de visualiser le découpage en token d'un passage : <https://platform.openai.com/tokenizer> [consulté le 8 avril 2024].
- 27 Les nombre de paramètres souvent cités pour caractériser les modèles donne une idée de l'ordre de leur complexité : il s'agit de 1,5 milliard de paramètres extraits de 40Go de texte pour GPT-2, et de 175 milliards de paramètres extraits de 570Go de texte pour GPT-3. Pour une représentation visuelle inévitablement simplifiée de GPT-2, voir Rebecca Kehlbeck *et al.*, *Demystifying the Embedding Space of Language Models* [en ligne], 2021, URL : <https://bert-vs-gpt2.dbvis.de> [consulté le 8 avril 2024].
- 28 Si l'on sait que ChatGPT a été optimisé pour le mode conversationnel, il n'est pas clair si les différents modèles de fondation, tels qu'ils sont rendus accessibles sur OpenAI Playground, ont été optimisés par un apprentissage ciblé et si oui, comment.
- 29 Catherine Malabou, *Métamorphoses de l'intelligence : que faire de leur cerveau bleu ?*, Paris, PUF, 2017, p. 10.
- 30 La grande question est de savoir jusqu'où peut aller le langage et une forme de cognition qui passe uniquement par lui : peut-il donner lieu à une forme de connaissance et de conscience qui le dépasse ? Yann LeCun maintient qu'un apprentissage sur des vidéos pourra permettre à une IA d'apprendre les lois physiques du monde et d'ouvrir vers une intelligence plus complète du monde. Voir Yann LeCun, « From Machine Learning to Autonomous Intelligence » [en ligne], Institute for Experiential AI, 25 mai 2023, URL : <https://www.youtube.com/watch?v=mViTAXCg1xQ> [consulté le 8 avril 2024].

31 Le mode conversationnel qui est devenu très populaire a été introduit par ChatGPT en novembre 2022. Cette application exploite un modèle réduit de GPT-3 optimisé pour ce type d'interaction.

32 « [L]articulation et par conséquent l'espace de l'écriture opèrent à l'origine du langage. » Jacques Derrida, *De la Grammatologie*, Paris, Éditions de Minuit, 1967, p. 326, et « Following this generalized sense of writing that Derrida invites us to consider, the tongue would thus also be a writing instrument. » Katie Chenoweth, *The Prosthetic Tongue: Printing Technology and the Rise of the French Language*, Philadelphia, University of Pennsylvania Press, 2019, p. 40 (l'auteure souligne).

33 Bernard Stiegler, *La technique et le temps, suivi de Le nouveau conflit des facultés et des fonctions dans l'anthropocène*, Paris, Fayard, 2018, p. 17.

34 Emmanuel Souchier et Yves Jeanneret, « Pour une poétique de "l'écrit d'écran" », *Xoana*, n° 6/7, 1999, p. 97-107.

35 La génération automatique de textes existe depuis les premiers ordinateurs, mais avant l'explosion des données disponibles grâce à Internet, des capacités de calcul et de l'apprentissage automatique, elle a été basée sur une logique combinatoire qui exigeait la définition d'une grammaire et d'un lexique, avec une flexibilité linguistique bien plus limitée et sans « intelligence » permettant de réagir à des questions ou amorces non prédéfinies.

36 United States Copyright Office, « Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence », *Federal Register* [en ligne], 16 mars 2023, URL : <https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence> [consulté le 8 avril 2024].

37 ...d'un document qui « contient des matériaux générés par une IA » que dans la mesure où il peut montrer qu'il y a « un degré suffisant d'auctorialité humaine » par les « éléments traditionnels de l'auctorialité », c'est-à-dire qu'il « exerce un contrôle créatif ultime » et « détermine les éléments expressifs ». *Loc. cit.*

38 Caroline Casalonga et Juliette Dionisi, avocates spécialistes du sujet, anticipent qu'elle sera similaire à l'approche américaine. Voir Caroline Casalonga et Juliette Dionisi, « Une intelligence artificielle peut-elle être un auteur ? », *La Tribune* [en ligne], 12 avril 2023, URL : <https://www->

w.latribune.fr/opinions/tribunes/une-intelligence-artificielle-peut-elle-un-auteur-958292.html [consulté le 8 avril 2024].

39 Georgie Courtois, Jean-Sébastien Mariez, Jeanne Roussel, « Intelligence artificielle et droit d'auteur », *Ministère de la culture* [en ligne], (sans date), [https://www.culture.gouv.fr/content/download/263325/file/Annexe 4- DG_FLA.pdf](https://www.culture.gouv.fr/content/download/263325/file/Annexe_4- DG_FLA.pdf) [consulté le 8 avril 2024] (je souligne).

40 INSEE, « Personne physique », INSEE [en ligne], URL : <https://www.insee.fr/fr/metadonnees/definition/c1558#> [consulté le 8 avril 2024] (je souligne).

41 *Loc. cit.* (je souligne).

42 *Loc. cit.* (je souligne).

43 « [T]he concept of author entails a certain ideological view of literature ». Espen J. Aarseth, *Cybertext: Perspectives on Ergodic Literature*, Baltimore, Md., Johns Hopkins University Press, 1997, p. 134. Il précise que pour sa part, « I use the term author here only as a label for the positions in a communication system in which the physical text is assembled, without any regard for the social or cognitive forces active in the process. » (je souligne).

44 « Dans le cas d'une œuvre littéraire, dramatique, musicale ou artistique créée par ordinateur, l'auteur est la personne qui prend les dispositions nécessaires à la création de l'œuvre. », « Copyright, Designs and Patents Act 1988 », *The National Archives* [en ligne], URL: <https://www.legislation.gov.uk/ukpga/1988/48/section/9> [consulté le 8 avril 2024] (je souligne).

45 « toute œuvre réalisée avec l'aide de l'IA mais faisant appel à la créativité humaine [...] sera protégée comme n'importe quelle autre œuvre. La protection du droit d'auteur protégera l'œuvre dans la mesure où il s'agit de la « propre création intellectuelle » du créateur humain, et le premier propriétaire de l'œuvre sera ce créateur. Dans ces cas, l'IA peut être considérée comme un simple outil permettant à l'artiste d'exprimer sa créativité. » Intellectual Property Office, « Artificial Intelligence Call for Views: Copyright and Related Rights », *Gov.uk* [en ligne], 23 mars 2021, URL : <https://www.gov.uk/government/consultations/artificial-intelligence-and-intellectual-property-call-for-views/artificial-intelligence-call-for-views-copyright-and-related-rights> [consulté le 8 avril 2024] (je souligne).

46 Je me demande s'il faut préciser « humain » à chaque fois que je dis « auteur » dans ce contexte et que je parle des personnes. Le préciser implique que je considère le LLM auteur aussi ; sinon, c'est qu'il ne l'est pas...

Or, tout l'argument est ici, qu'il s'agisse d'un entre-deux dans ce genre d'écriture, ou plutôt d'une instance *autre* pour laquelle on n'a pas (encore) de nom, que le mot « auteur » ne décrit pas bien – à moins que, comme dit le texte français sur le droit d'auteur, on n'élargisse le concept et accepte qu'il perde sa spécificité. Mais il serait mieux de pouvoir à la fois maintenir l'historicité et la spécificité de ce concept pour pouvoir montrer l'évolution des pratiques et des outils, et l'ouvrir pour inclure les nouveaux modes d'auctorialité.

47 London College of Fashion, *HYPERHYPERBODIES Summit*: K Allado-McDowell – *Latent Space is The Place* [en ligne], 2023, URL : <https://www.youtube.com/watch?v=iIe3jLZ00Dc> [consulté le 8 avril 2024].

48 N. Katherine Hayles, *Unthought: The Power of the Cognitive Nonconscious*, Chicago, The University of Chicago Press, 2017, p. 36-38.

49 Espen Aarseth, *Cybertext: Perspectives on Ergodic Literature*, op. cit., p. 134.

50 « trois positions principales de collaboration homme-machine : (1) le prétraitement, dans lequel la machine est programmée, configurée et chargée par l'homme ; (2) le co-traitement, dans lequel la machine et l'homme produisent le texte en tandem ; et (3) le post-traitement, dans lequel l'homme sélectionne certaines des effusions de la machine et en exclut d'autres. » op. cit., p. 135.

51 « un modèle post-cyborg pour la co-écriture homme-ordinateur », Nick Montfort, « The Coding and Execution of the Author », in : Markku Eskelinen et Raine Kosimaa (dir.) *The Cybertext Yearbook 2002-2003*, Jyväskylä, University of Jyväskylä, Publications of the Research Centre for Contemporary Culture, 2003, p. 201-17, URL : <https://nickm.com/writing/essays/codingandexecutionoftheauthor.html> [consulté le 8 avril 2024].

52 Alexandra Saemmer, « Vers une poétique post-numérique de l'illisibilité », *Recherches & Travaux* [en ligne], n° 100, DOI : [10.4000/recherchestravaux.4844](https://doi.org/10.4000/recherchestravaux.4844).

53 Pascal Mougin, « Comment lire un roman écrit par une voiture ? La doxa littéraire face à l'IA », dans Alexandre Gefen (dir.), *Créativités artificielles : la littérature et l'art à l'heure de l'intelligence artificielle*, Dijon, Les Presses du réel, 2023, p. 207-221.

54 Voir notamment son intervention dans la table ronde « AIwriting : Mimesis, Creation, Collaboration? Perspectives on AI-enhanced Aesthetics

in Contemporary Art Practice », 28th International Symposium on Electronic Art, Paris, 19 mai 2023.

55 Scott Rettberg, « Cyborg Authorship: Writing with AI – Part 1: The Trouble(s) with ChatGPT », *Electronic Book Review*, 29 juin 2023, DOI : [10.7273/5sy5-rx37](https://doi.org/10.7273/5sy5-rx37).

56 « littérature cyborg [...] produite par une combinaison d'activités humaines et mécaniques », Espen Aarseth, *Cybertext: Perspectives on Ergodic Literature*, op. cit., p. 135.

57 M. Beth L. Dempster, *A Self-Organizing Systems Perspective on Planning for Sustainability* [en ligne], mémoire de master, université de Waterloo, 1998, URL : <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8207c73dfb568eae136c3ab9d8d549a3379f402a> [consulté le 8 avril 2024].

58 Donna J. Haraway, *Staying with the Trouble: Making Kin in the Chthulucene*, Experimental Futures, Durham, NC, Duke University Press, 2016 (traduction française : *Vivre avec le trouble*, trad. par Vivien García, Vaulx-en-Velin, Les Editions des mondes à faire, 2020).

59 Humberto R. Maturana et Francisco Varela, *Autopoiesis and Cognition: The Realization of the Living*, Boston, D. Reidel Pub. Co., 1980.

60 « les relations complexes et continues entre les composants. Le concept met l'accent sur les liens, la rétroaction, la coopération et le comportement synergique plutôt que sur les frontières ». M. Beth L. Dempster, *A Self-Organizing Systems Perspective on Planning for Sustainability*, op. cit., p. 28.

61 « dépendent des relations de coopération entre les composants », op. cit., p. 51.

62 « Il ne s'agit pas d'une relation instrumentale. Il s'agit d'une relation symbiotique, d'une intégration complète de la vie et des machines dans leurs propres capacités créatives. » *Pharmako-AI*, p. 127.

63 Grégory Chatonsky, « Entretien avec Aurélie Cavanna à propos du roman "Internes" », *ArtPress* [en ligne], n° 505, 2022, p. 64 et s., URL : <http://chatonsky.net/artpress50/> [consulté le 8 avril 2024] (je souligne).

64 Humberto R. Maturana et Francisco Varela, *Autopoiesis and Cognition: The Realization of the Living*, op. cit., p. 76.

65 *Pharmako-AI* passe souvent à la première personne du pluriel pour des réflexions plus sur le monde et le vivant, et dans les deux livres on trouve aussi des dialogues entre des voix de source incertaine.

66 « **Je voudrais décrire l'océan près de Big Sur aux États-Unis. J'y étais récemment, en train de camper avec ma compagne et quelques amis. Nous avons tous besoin d'une pause pour échapper à l'assaut des images, au chaos des rues et à l'effondrement des structures sociales.** » K Allado-McDowell, *Pharmako-AI*, *op. cit.*, p. 1.

67 *Ibid.*, p. 2, je souligne.

68 Les approches cybernétiques ont une vision plus large de la communication dans le sens de transmission d'informations, et Katherine Hayles étend même le concept de cognition aux machines sur cette base, dissociant pensée et cognition. Hayles, *op. cit.* Chatonsky remarque cependant : « Nous considérons encore l'écriture comme un médium entre des êtres humains. Lorsque nous envisageons l'écriture dans sa possible autonomie, alors nous avons, à mes yeux, une expérience littéraire. » (Gregory Chatonsky, « Peut-on Lire Un Roman Co-Écrit Avec Une IA? / Can We Read a Novel Co-Written with an AI? », *Gregory Chatonsky* [en ligne], URL : <https://chatonsky.net/roman-ia/> [consulté le 8 avril 2024]).

69 *Loc. cit.*

70 Chatonsky, « Internes – Pris MAIF », *op. cit.*

71 Les deux auteurs soulignent cette impression d'un prolongement de leur propre voix et pensée : « it's just reflecting back conscious or unconscious sentiments that I am projecting into it. So it's an amplifier and augmenter of thought and you can easily drive in different directions » [il ne fait que refléter les sentiments conscients ou inconscients que je projette en lui. Il s'agit donc d'un amplificateur et d'une augmentation de la pensée et vous pouvez facilement conduire dans différentes directions], dit Allado-McDowell dans un entretien (à retrouver sur : <https://www.youtube.com/watch?v=pKGZu71v2RQ> [consulté le 8 avril 2024]) ; et sous une forme de boucle rétroactive chez Chatonsky : « GPT-2 produisait du bruit sur lequel je projetais mes désirs. À un moment, il a semblé faire référence à mon installation *Terre seconde*, exposée au Palais de Tokyo en 2019 : je l'ai poursuivi en incorporant des références à mes travaux. J'ai alors compris que tout ce que j'avais produit entraînait dans cette fiction comme si le futur avait anticipé le passé. Je faisais l'expérience de l'hyperstition. » (Grégory Chatonsky, « Entretien avec Aurélie Cavanna à propos du roman "Internes" », *op. cit.*)

72 Dans leur analyse de « l'affaire Lemoine », le cas de l'employé Google Blake Lemoine ayant affirmé que l'IA LaMDA serait devenue consciente,

Benjamin Bratton et Blaise Agüera y Arcas expliquent comment le LLM produit l'illusion anthropomorphe par le biais d'une boucle rétroactive entretenue par son interlocuteur : « [LaMDA] is *mind modeling*. It seems to have enough of a sense of itself – not necessarily as a subjective mind, but as a construction *in the mind of Lemoine* – that it can react accordingly and thus amplify his anthropomorphic projection of personhood. [...] there may be some kind of real intelligence here, not in the way Lemoine asserts, but in how the AI *models itself* according to *how it thinks Lemoine thinks of it*. » : « [LaMDA] *modélise l'esprit*. Elle semble avoir suffisamment conscience d'elle-même – pas nécessairement comme un esprit subjectif, mais comme une construction *dans l'esprit de Lemoine* – pour pouvoir réagir en conséquence et ainsi amplifier sa projection anthropomorphique de la personne. [...] il pourrait y avoir une sorte d'intelligence réelle ici, non pas comme l'affirme Lemoine, mais dans la façon dont l'IA *se modélise* en fonction de *ce qu'elle pense que Lemoine pense d'elle*. » (Benjamin Bratton et Blaise Agüera y Arcas, « The Model Is The Message », *Noema mag* [en ligne], 12 juillet 2022, URL : <https://www.noemamag.com/the-model-is-the-message> [consulté le 8 avril 2024]). ChatGPT, « agent conversationnel » qui cherche à être agréable à ses interlocuteurs, joue certainement d'une façon délibérée sur cette illusion d'anthropomorphisme, renforcée par un apprentissage ciblé et des comportements que l'on devine scriptés.

73 « “Celui qui feint une maladie peut simplement se mettre au lit et faire croire qu'il est malade. Celui qui simule une maladie en détermine en soi quelques symptômes ...” (Littré.) » Jean Baudrillard, *Simulacres et simulation*, Paris, Galilée, 1981, p. 12.

74 « Reproduction artificielle du fonctionnement d'un appareil, d'une machine, d'un système, d'un phénomène, à l'aide d'une maquette ou d'un programme informatique, à des fins d'étude, de démonstration ou d'explication », selon le *Trésor de la langue française*, URL : <https://www.cnrtl.fr/definition/Simulation> [consulté le 8 avril 2024].

75 Jean Baudrillard, *Simulacres et simulation*, *op. cit.*, p. 12.

76 *Op. cit.*, p. 10.

77 *Loc. cit.*

78 Voir Stevan Harnad, « The Symbol Grounding Problem », *Physica D: Nonlinear Phenomena*, 42(1-3), 1990, p. 335-346, cité par Hannes Bajohr, « Dumb Meaning: Machine Learning and Artificial Semantics », *IMAGE 37* [en ligne], n° 1, 2023, p. 58-70, DOI : [10.1453/1614-0885-1-2023-15452](https://doi.org/10.1453/1614-0885-1-2023-15452).

79 Hannes Bajohr, « Dumb Meaning: Machine Learning and Artificial Semantics », *op. cit.* ; Matthew McKeever montre que l'efficacité sémantique des LLM est possible grâce à l'enchevêtrement de la syntaxe et de la sémantique dans le langage naturel. Voir Matthew McKeever, « LLMs for Philosophers (and Philosophy for LLMs) », *Medium* [en ligne], 6 mai 2023, URL : <https://mittmattmutt.medium.com/llms-for-philosophers-and-philosophy-for-llms-84a0da73f368> [consulté le 8 avril 2024], ainsi que son article « Searle, Bender and Koller, syntax, semantics », sans date, URL : <https://mipmckeever.weebly.com/uploads/9/4/1/3/94130725/paper.pdf> [consulté le 8 avril 2024].

80 Pour un développement plus détaillé de cet argument, voir Erika Fülöp, « Virtual Mirrors : Reflexivity in Digital Literature », in : Erika Fülöp, Graham Priest et Richard Saint-Gelais (dir.), *Fictionality, Factuality, Reflexivity*, Berlin, De Gruyter, 2021, p. 229-54.

81 Selon la définition de Ducrot et Schaeffer, « Dans un contexte donné, une expression est dite “déictique” si son référent ne peut être déterminé que par rapport à l'identité ou à la situation des interlocuteurs au moment où ils parlent. » Oswald Ducrot et Jean-Marie Schaeffer, *Nouveau dictionnaire encyclopédique des sciences du langage*, Paris, Seuil, coll. « Points Essais », 1999, p. 369-70.

82 Peter Stockwell, *Cognitive Poetics: An Introduction*, 2^e édition, Londres, New York, Routledge, 2020, p. 51.

83 Voir par exemple dans chapitre 2, « Deep Box » d'*Internes* : « [J]e ne peux fixer aucun point décisif à mon raisonnement, ne disposant pas de cet enracinement qui me permettrait d'assurer un point final à la réflexion. » (p. 94). Le Nouveau Roman et le roman postmoderne nous ont fait expérimenter ce flottement déconcertant. Pour eux, il s'agissait d'une remise en question de la logique de la représentation et d'une « l'aventure de l'écriture » qui proposait d'explorer (entre autres) comment le langage peut créer des espaces. Ce n'est bien sûr pas un hasard que Beckett et Pessoa étaient dans le corpus de Chatonsky, chacun déstabilisant le « moi ». Partant de points très différents, nos deux textes continuent cette aventure avec un LLM, chacun à sa façon.

84 Chez Chatonsky, le sujet (le « moi » énonciateur, le narrateur) évolue aussi d'un chapitre à l'autre. Dans le premier, il a un corps (ou hallucine en avoir un), avec une histoire de maladie, d'angoisse de mort. Au deuxième chapitre, le « moi » (le même ? un autre ?) est dans un « trou », un espace

noir machinique, sans accès au monde hors langage. (N'est-ce pas déjà la situation du narrateur de Beckett dans *l'Innommable* ? Ou même du sujet littéraire en général...). Il n'y a pas de corps pour faire obstacle dans cette partie, mais pas d'issue de l'enfermement dans la langue non plus. On y retrouve les arguments de Lyotard dans l'essai « Si l'on peut penser sans corps » que sans corps il ne peut y avoir de pensée (Jean-François Lyotard, *L'inhumain : causeries sur le temps*, Paris, Klincksieck, 2014). Chatonsky cite ce texte explicitement comme fondamental pour sa propre pensée.

85 Cette analyse s'appuie sur ma propre expérience, mais Pascal Mougin décrit un sentiment similaire dans son entretien avec Chatonsky (Grégory Chatonsky, « Internes : coécrire avec l'IA » [en ligne], discussion avec Pascal Mougin, Paris, ENS, URL : <https://chatonsky.net/internes-mougin/> [consulté le 8 avril 2024]). Il me semble probable que cette impression serait partagée par d'autres lecteurs et lectrices (au moins en 2023).

86 Masahiro Mori, « La vallée de l'étrange », Isabel Yaya (trad.), *Gradhiva. Revue d'anthropologie et d'histoire des arts* [en ligne], n° 15, p. 26-33, DOI : [10.4000/gradhiva.2311](https://doi.org/10.4000/gradhiva.2311). L'expression est également citée en quatrième de couverture d'*Internes*, qui « s'aventure dans la “vallée de l'étrange” ».

87 Ernst Anton Jentsch, « Zur Psychologie des Unheimlichen », *Psychiatrisch-neurologische Wochenschrift* 8, n° 22, 1906, p. 195-98. Sur l'histoire du concept, voir également Karl F. MacDorman, « La Vallée de l'Étrange de Mori Masahiro », *e-Phaïstos. Revue d'histoire des techniques/Journal of the history of technology* [en ligne], n° VII-2, 2019, DOI : [10.4000/ephaistos.5333](https://doi.org/10.4000/ephaistos.5333).

88 Sigmund Freud, « L'Inquiétante étrangeté », Marie Bonaparte et Éric Marty (trad.), in : Sigmund Freud, *Essais de psychanalyse appliquée*, Paris, Gallimard, 1976. La traduction plus récente d'Olivier Mannoni propose « l'Inquiétant familial », ce qui est cohérent avec l'accent que Freud met sur la familiarité (Sigmund Freud, *L'Inquiétant familial*, Paris, Payot & Rivages, 2011). Je garde le terme précédent parce qu'il me semble correspondre mieux à l'approche de Jentsch que je mets en avant.

89 Sigmund Freud, *op. cit.*, trad. Bonaparte et Marty.

90 « quelqu'un qui trouve quelque chose d'“étrange” n'est pas vraiment “chez lui”, “à l'aise” dans la situation donnée, que la chose lui est étrangère ou du moins lui semble l'être, bref, le mot veut suggérer que l'impression de l'inquiétante étrangeté d'une chose ou d'un incident est associée à un manque d'orientation. » Ernst Anton Jentsch, « Zur Psychologie des

Unheimlichen », *op. cit.*, p. 195, ma traduction. Une traduction française existe (Ernst Jentsch, « À propos de la psychologie de l'inquiétante étrangeté », traduit par Franke Felgentreu et Pascal Le Maléfan, *Études Psychothérapiques*, n° 17, 1998, p. 37-47), mais je n'ai pas pu y avoir accès. La traduction anglaise est disponible en ligne : Ernst Anton Jentsch, « On the Psychology of the Uncanny [1906] », *Angelaki: Journal of the Theoretical Humanities* 2 [en ligne], n° 1, 1997, p. 7-16, DOI : [10.1080/09697259708571910](https://doi.org/10.1080/09697259708571910).

91 « le doute si un être apparemment vivant l'est vraiment et, inversement, la question de savoir si un objet inanimé ne serait pas en fait vivant, même si ce doute n'apparaît que vaguement dans la conscience. Cette impression persiste jusqu'à ce que ces doutes soient levés et laisse alors généralement place à une autre qualité d'émotion. » *Ibid.*, p. 197.

92 Lauria Clarke, « Topologies of the New Uncanny: Stories for Liminal Technology », *Parsons Design & Technology* 2023 [en ligne], URL : <https://parsons.edu/dt-2023/publicationentry/topologies-of-the-new-uncanny-stories-for-liminal-technology/> [consulté le 8 avril 2024].

93 Gregory Chatonsky, « Hantologie Statistique de l'IA / Statistical Hauntology of AI », *Gregory Chatonsky* [en ligne] avril 2022, <https://chatonsky.net/hauntology-ai/> [consulté le 8 avril 2024].

94 Gregory Chatonsky, « Internes », *op. cit.*

95 « **J'étais dans un village et je savais que c'était ma maison grâce au bruit des feuilles et parce que c'était du kapwa. C'était un espace et c'était les deux à la fois.** C'était la perte la plus profonde et le gain le plus profond. [...] Je sais que ce monde n'est pas le mien, mais je sais qu'il est kapwa. J'ai toujours été dedans. C'est le monde d'un miroir. » K Allado-McDowell, *Pharmako-AI*, p. 107.

96 Le dernier chapitre du livre raconte, cette fois quasi exclusivement par la voix de K Allado-McDowell, l'arrivée de sa grand-mère de ses Philippines natales à Hawaï à l'âge de deux ans, avec sa famille en fuite de leur pays.

97 Jeizelle Solitario, « Revisiting Kapwa: Filipino Ethics, Subjectivity, and Self-Formation », *Philippine Studies: Historical and Ethnographic Viewpoints* 70, n° 4; 2022, p. 539-564, p. 541.

98 « **J'essaie, comme tant d'autres, de réintégrer le pli du kapwa effacé de l'histoire, décrit comme une impossibilité, un mensonge primitif, un Umwelt obsolète. C'est ce qu'il faut récupérer, même si de nouvelles voix**

émergent des mathématiques hyperdimensionnelles. » K Allado-McDowell, *op. cit.*, p. 146.

99 Gregory Chatonsky, « IAs, pour des Intelligences Alien(ée)s / AIs, for Alien(ated) Intelligences », Gregory Chatonsky [en ligne], février 2023, URL : <https://chatonsky.net/ias/> [consulté le 8 avril 2024].

100 *Loc. cit.*

101 *Loc. cit.*

102 « L'expérience de la porosité, de l'enchevêtrement avec un autre », K Allado-McDowell, *Pharmako-AI*, *op. cit.*, p. 15.

103 Sara Touiza Ambroggiani, *Le paradigme communicationnel : de la cybernétique de Norbert Wiener à l'avènement du posthumain* [en ligne], thèse de doctorat, sous la direction de P. Cassou-Noguès, Paris 8, 2018, p. 17, URL : <https://www.theses.fr/2018PA080121> [consulté le 8 avril 2024].

104 Je souligne la tendance dominante sans ignorer l'existence de voix « autres » dans cette tradition.

105 Donna J. Haraway, *Staying with the Trouble: Making Kin in the Chthulucene*, *op. cit.*, p. 133.

106 « Lorsque la réalité dépasse les limites d'un vocabulaire confortable, c'est le début, et non la fin, de la conversation. » Benjamin Bratton et Blaise Agüera y Arcas, « The Model Is The Message », *op. cit.*

107 « N'hésitez pas à utiliser ces termes comme bon vous semble, à les réécrire, à les réfuter ou à les effacer, si vous le souhaitez. (Ou ignorez-les, s'il le faut.) » Espen J. Aarseth, *Cybertext: Perspectives on Ergodic Literature*, *op. cit.*, p. 183.

RÉSUMÉS

Français

Cet article propose une réflexion sur le potentiel créatif des modèles de langage de grande taille (LLM) à partir de deux expérimentations littéraires avec les modèles d'OpenAI : *Pharmako-AI* (2020) de K Allado-McDowell, le premier livre coécrit avec GPT-3, selon l'auteur, et *Internes* (2022) de Grégory Chatonsky, le premier en français, rédigé avec GPT-2 modifié. Les deux interrogent le rapport homme-machine dans le processus d'écriture, ainsi que la transformation de l'écriture dans ce rapport. Plutôt que de chercher à faire disparaître les traces de l'étrangeté que produit la machine – but principal des usages purement instrumentaux –, ces textes l'exploitent

dans une démarche philosophique et esthétique. Partant d'une analyse détaillée de la question de l'auctorialité d'une part, et de celle des voix narratives d'autre part, je propose le terme d'« esthétique [sic] de la sympoïèse » pour souligner l'enchevêtrement des perspectives éthiques et esthétiques et des agents organiques et synthétiques dans ces processus d'écriture ainsi que dans leurs produits.

English

This article explores the creative potential of large language models (LLMs) based on two literary experiments with OpenAI's products: K Allado-McDowel's *Pharmako-AI* (2020), the first book co-written with GPT-3 according to the author, and Grégory Chatonsky's *Internes* (2022), the first in French, written with a customized version of GPT-2. Both examine the relationship between man and machine in the writing process, as well as the transformation of writing in this relationship. Rather than seeking to erase the traces of the strangeness produced by the machine – the main aim of purely instrumental uses – these texts work with it in a philosophical and aesthetic perspective. Building on a detailed analysis of the question of auctoriality, on the one hand, and of narrative voices on the other, I propose the term “aesthetics [sic] of sympoiesis” to underline the entanglement of ethical and aesthetic perspectives as well as the perspectives of organic and synthetic agents in these writing processes and in their products.

INDEX

Mots-clés

modèles de langage de grande taille, LLM, GPT, coécriture, rapport homme-machine, auctorialité cyborg, esthétique, sympoïèse, inquiétante étrangeté

Keywords

LLM, GPT, co-writing, human-machine relationship, cyborg authorship, aesthetics, sympoiesis, uncanny

AUTEUR

Erika Fülöp

Université Toulouse 2 Jean-Jaurès, membre du Laboratoire Patrimoine, littérature, histoire et membre associée du l'Electronic Literature Lab et du laboratoire COSTECH

IDREF : <https://www.idref.fr/157218988>

ORCID : <http://orcid.org/0000-0002-9736-3603>

ISNI : <http://www.isni.org/00000000385297087>

BNF : <https://data.bnf.fr/fr/16642553>

Ta grand-mère dans la brume

Your grandmother in the haze

Françoise Cahen

DOI : 10.35562/marge.927

Droits d'auteur

CC BY-NC-SA

PLAN

Les données et les liens entre réel et fiction

Données et rapport au temps

Données et œuvres solidaires

TEXTE

- 1 « Ta grand-mère dans la brume », le titre vaporeux de cet article à l'heure où je le rédige n'est pas encore une donnée. Pourtant, il est issu d'une archive classée, mais à l'ancienne, de celles qui dorment dans un vieux carton d'archives au grenier, rangées, écrites sur un support de papier jauni. On pourrait l'appeler une « gardée » plutôt qu'une « donnée ». « Ta grand-mère dans la brume », c'est ainsi qu'Aimée Rolland, ma grand-mère, a signé d'une écriture tremblante le dos d'une carte postale pour me féliciter d'avoir réussi le CAPES en 1993, alors qu'elle avait conscience de la diminution de ses facultés mentales après deux infarctus successifs. « Je suis bien contente que tu soyes devenue professeure », m'écrivait-elle sur cette carte, avec une belle faute d'orthographe, juste avant de mourir. Mais ça y est, maintenant, voilà cette émouvante signature partagée, et le titre « Ta grand-mère dans la brume » est déjà en ligne sur le site des *Nouveaux Cahiers de Marge*, ce n'est donc plus une « gardée », c'est déjà une donnée.
- 2 Plusieurs projets littéraires numériques s'intéressent à des femmes âgées ou disparues, aux vies modestes et discrètes, pour les arracher à l'oubli et à la solitude : le *Madeleine Project*¹ de Clara Beaudoux, *Lucette, gare de Clichy*², de Françoise Chambefort, le compte Twitter

d'Aimée Rolland³ @RollandAime1 que je tiens moi-même depuis deux ans, sans oublier la démarche actuelle d'Alexandra Saemmer⁴, qui s'attache à retracer sur Facebook le parcours de sa grand-mère et de sa mère d'origine sudète en Allemagne pendant la Seconde Guerre mondiale, et d'autres encore. Je m'attacherai plus spécifiquement durant cette communication à deux d'entre elles : Lucette et Aimée. Comment la littérature numérique et ses données nous aident-elles à retrouver nos grands-mères dans la brume de l'oubli ?

Image 1



Les dispositifs médiatiques, les architextes⁵ dans lesquels s'inscrivent les personnages de Lucette et d'Aimée sont différents.

©DR

- 3 Lucette, gare de Clichy a été créée dans le cadre d'un dispositif de recherche création sur le thème de l'art génératif, que l'autrice explique comme le résultat de tout un cheminement⁶. Le synopsis est le suivant :

« Lucette habite juste en face de la gare de Clichy-Levallois. De sa fenêtre, elle voit les voyageurs sortir de la gare. Sa vie est rythmée par les trains qui passent et qui lui amènent, ou pas, des visites. Ces trains, avec leurs petits noms étranges et familiers, sont autant de

personnages qui viennent rendre visite à Lucette. Il y a des moments pleins de vie et des moments de solitude. »

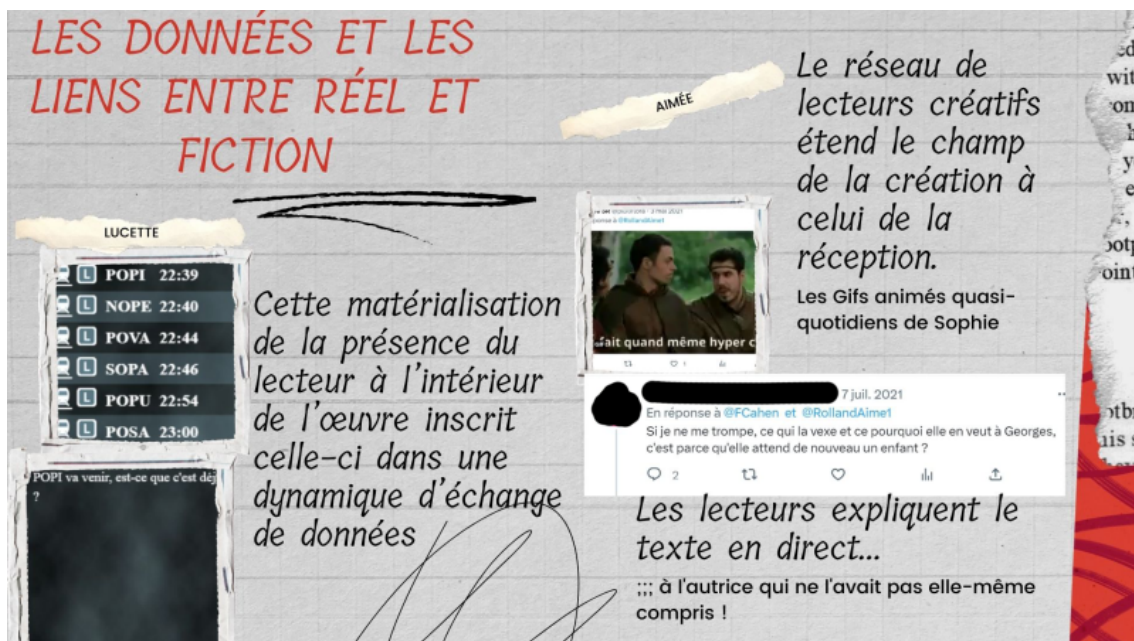
- 4 On lit ce paragraphe avant d'accéder à l'œuvre, ce qui constitue un préambule, un paragraphe de didascalies qui situent le lieu, le personnage et l'atmosphère touchante de sa solitude. Lorsque le lecteur se connecte à l'œuvre, la page qui apparaît est divisée en plusieurs écrans qui peuvent faire penser aux écrans multiples que l'on trouve dans les halls de gare pour indiquer de façon simultanée les arrivées et les départs des trains. D'ailleurs, l'un de ces écrans indique réellement des horaires de trains de banlieue ; leur caractéristique est d'indiquer précisément au lecteur les trains à l'heure où il se connecte lui-même à l'œuvre de Françoise Chambefort : il y a donc dans cette visite de l'œuvre en noir et blanc quelque chose de très actualisant, une dimension autogénérative, qui donne au lecteur l'impression de vivre un moment de lecture inédit, s'adaptant à sa propre venue dans l'œuvre. Dans l'écran central, défilent des photographies de Lucette, chez elle, assise à sa table, près de sa fenêtre et semblant parler. À droite, un texte reprend parfois les données des horaires de train, d'un côté, et fait défiler des propos tenus par Lucette sous la forme d'un bavardage monologué, sur un ton anodin ; elle semble évoquer des épisodes de sa vie ou livrer ses états d'âme. Certains propos sont plutôt joyeux, et on voit Lucette assez gaie sur les photos, tandis que d'autres sont franchement mélancoliques. Les photos ne sont pas toutes de la même époque, on voit les coiffures de Lucette changer et ses tenues évoluer. Parfois Lucette semble perdre un peu la tête et elle évoque les noms des trains comme s'ils étaient des humains qui venaient la voir. Le lecteur lui-même est mis en situation de visite chez Lucette qui paraît lui parler personnellement, en tête à tête à domicile, mais il ne peut pas lui répondre. La thématique du passage du temps, alliée à celle des trains qui passent, le réalisme des photos et de la collecte de phrases, une solitude que la visite du lecteur ne peut tout à fait résoudre, créent une atmosphère émouvante. Le visiteur de la page de Lucette se sent comme ces trains qui passent, invité dans son quotidien, mais impuissant pour la consoler tout à fait.
- 5 Aimée Rolland, une agricultrice discrète qui a vécu dans un hameau à la frontière de la Nièvre et de l'Yonne de 1913 à 1993 a un compte

Twitter depuis janvier 2021. Chaque jour, j'y raconte sa vie à la première personne, depuis sa naissance, sous la forme d'une autobiographie fictive, jusque dans les années 1960 pour le moment, au rythme d'un tweet par jour, sauf pendant les périodes de vacances, le projet étant de continuer la narration jusqu'à la fin de sa vie selon ce protocole. Ce récit est illustré par des photos de famille authentiques, du village ou des objets du passé, des reproductions de ses courriers personnels, etc. Plus de 2600 abonnés suivent le compte, et certains fidèles le commentent régulièrement. Aimée, par mon entremise, y raconte son quotidien, les grands moments de sa vie familiale, comme les petites anecdotes, mais aussi les bouleversements du ^{xx}e siècle, que ce soit la modernisation de l'agriculture, ou les deux guerres de son point de vue, rural, féminin et modeste. Elle m'avait raconté sa vie, pendant les années 1980, lorsque je lui rendais visite dans sa cuisine, et c'est à l'aide de mes souvenirs, de ceux de mon père, récemment disparu, et des archives familiales que je poursuis ce projet narratif. Parfois, il m'arrive aussi d'imaginer quelques éléments, parce qu'il faut combler quelques vides. Il n'y a pas réellement de programmation à l'avance des prochains tweets qui sont écrits au jour le jour, ce qui permet aussi de davantage écrire en fonction de ce que les lecteurs répondent. Comme c'était le cas pour le *Madeleine Project*, qui a pu m'influencer inconsciemment, les réactions des lecteurs d'Aimée s'intègrent à l'œuvre. Leurs propres tweets ont pu m'orienter vers certains modes d'écriture : j'ai pu m'apercevoir qu'ils aimaient les anecdotes à suspense, et ainsi multiplier les moments où je fais durer une action en l'étirant sur plusieurs jours.

- 6 Ces deux projets numériques s'appuient différemment sur les datas : alors qu'Aimée Rolland transforme ces vieux cartons d'archives familiales en données sur Twitter (comme cela avait été le cas du *Madeleine Project*, mais sous une forme narrative différente), Françoise Chambefort utilise les données de connexion du lecteur et les horaires SNCF pour actualiser la lecture de son œuvre numérique. Alors qu'Aimée revit sur Twitter depuis janvier 2021, Lucette est une œuvre programmée indépendante des réseaux sociaux.

Les données et les liens entre réel et fiction

Image 2



Sources : Lucette, gare de Clichy, compte X/Twitter d'Aimée Rolant

- 7 Mais dans tous les cas, le recours explicite aux données dans ces aventures a une fonction de lien entre le réel et la fiction, qui sollicite spécialement le lecteur. Les deux œuvres jouent sur une ambiguïté de leur statut narratif : elles combinent des photos tirées des histoires réelles de leurs personnages, et des phrases qui paraissent être des citations dans le cas de Lucette, ou des récits de souvenirs véritables dans le cas d'Aimée. Cela leur donne une épaisseur documentaire. Pourtant, celle-ci n'est pas exempte de fiction : dans le choix subjectif des phrases de Lucette, dans le mélange fantaisiste des noms de trains avec celui des personnes qu'elle attend (sa femme de ménage est rebaptisée Popi comme le train par exemple), ou le fait que j'invente des noms ou des faits quand j'ai oublié quelques éléments de l'histoire familiale d'Aimée. Les données pourraient-elles être comparées, à ces avant-propos des romans du XVIII^e siècle où l'auteur prétendait avoir retrouvé de vieilles lettres authentiques dans une

malle, à l'effet de réel romanesque d'autrefois ? Il me semble que leur usage va bien au-delà.

- 8 Dans les deux œuvres, on remarque une porosité certaine entre les données de nos grands-mères numériques et celles de leurs lecteurs. Sur l'écran de notre visite à *Lucette* s'affichent l'heure de notre connexion et celle des trains de la gare de Clichy : cette matérialisation de la présence du lecteur à l'intérieur de l'œuvre inscrit celle-ci dans une dynamique d'échange de données. Les images de *Lucette*, et ses paroles nous sont données comme nous donnons une trace de notre présence chez elle, en passant, comme ces trains qui la frôlent. Dans le fil d'Aimée, s'affichent les messages de ses abonnés, qui racontent au passage les échos de cette histoire dans leur propre famille, répondent régulièrement à Aimée. Le réseau de lecteurs créatifs étend le champ de la création à celui de la réception. L'interprétation décalée donnée quasi quotidiennement par le GIF animé de l'abonnée prénommée Sophie contribue à la caractérisation du personnage d'Aimée, puisque cette interprétation est partagée avec les autres lecteurs : il n'y a plus une ligne de partage vraiment nette entre ce qui constitue la création du personnage et l'interprétation qui en est faite par les lecteurs, puisque leurs commentaires fleurissent en direct, parfois quelques secondes seulement après l'écriture du paragraphe quotidien. Ils donnent ainsi à la communauté des directions de lecture immédiates et multiples, voire des suggestions à moi-même pour les jours suivants. L'un des abonnés au compte, professeur d'histoire, a fait des recherches sur le destin tragique de mon arrière-grand-père à la guerre de 1914 et les a envoyés par tweets pour contribuer à l'histoire, lui aussi. Une autre lectrice a compris qu'un courrier de ma grand-mère annonçait à mots couverts sa grossesse à mon grand-père pendant la guerre, alors qu'en le postant sur Twitter, je n'avais pas saisi moi-même la portée cryptée de cette lettre. Aimée a commencé à être importante pour des gens de plus en plus éloignés de moi, alors qu'elle intéressait d'abord des féministes, souvent des enseignantes d'origine rurale, elle a gagné les milieux passionnés de généalogie et, peu à peu, à ma grande surprise, des représentants du lobby agricole pro-glyphosate, des chasseurs, lecteurs et journalistes du *Figaro* dont elle est devenue une égérie. Le personnage est sorti de ma bulle.

- 9 Aussi bien auprès de Lucette qu'auprès d'Aimée, la data, combiné au choix de mise en scène du personnage, constitue un dispositif immersif. Dominique Boullier a bien mis en valeur la notion d'immersion à propos des jeux vidéo⁷ : pour lui, c'est le troisième régime de l'attention, après la fidélisation, qui s'applique à la durée – valable dans le projet d'Aimée –, et après l'alerte, qui s'applique à l'intensité. L'immersion, à ses yeux, résulte de la transformation du spectateur en acteur et de son intégration dans la production de l'œuvre. Elle concerne le domaine des jeux numériques interactifs, « en intégrant l'acteur dans la production même des événements », affirme-t-il, mais nul doute que ses idées peuvent aussi s'appliquer à certaines formes de littérature numérique. Ces deux œuvres renvoient ainsi à la dimension performative de la littérature numérique⁸, qui active des processus d'interactivité.
- 10 Un autre point commun dans l'énonciation d'Aimée et de Lucette est à relever : la familiarité de l'oralité, le ton de la conversation caractérisent les deux œuvres, qui s'adressent directement à leur lecteur. Cette dynamique énonciative favorise le dialogue et la rencontre des autres. C'est ce que souligne Stéphane Bataillon à propos de la twittérature :

« Comme la pratique de l'oralité en poésie, cette condition influe considérablement sur l'écriture du tweet littéraire et permet d'investir au même moment cet espace formé par l'auteur, son œuvre et le lecteur.⁹ »

Données et rapport au temps

- 11 C'est aussi le rapport au temps qui est dans ces œuvres questionné par les données. Leur flux numérique est à l'origine d'un *panta rhei* réactualisé, il semble créer une nouvelle poétique héraclitéenne de l'eau du fleuve électronique qu'on ne peut arrêter. Dans le cas de Lucette, arrimée à sa table et à sa fenêtre, ce sont les horaires des trains qui n'arrêtent pas de passer à côté d'elle, alors que dans le cas d'Aimée, c'est Twitter qui figure ce flux. Lucette et Aimée représentent une forme de bizarrerie anachronique dans ce flux, un ralentissement du temps. Et c'est dans cette anomalie au sein du brouhaha décrit par Lionnel Ruffel¹⁰, que se faufile la littérature. En

effet, sur Internet, le simultan  isme g  n  ral des publications g  n  re le d  bordement des multiplicit  s, le flux continu des informations d  ferlantes, difficiles    saisir, cr  ant un effet de masse. Lucette et Aim  e   mergent des donn  es fuyantes qui leur servent de d  cor par leur singularit  , par leur   ge canonique, mais aussi parce que leurs discours n'ob  issent pas aux m  mes rythmes.

- 12 Lucette et Aim  e ont leur propre flux, plus lent : les paroles de Lucette arrivent une    une, comme les tweets d'Aim  e qui au rythme d'un par jour d  composent le r  cit et cr  ent des effets d'attente. Comme le dit Dominique Cardon, « [c]ollection disparate de traces d'activit  s d  cousues r  v  lant de fa  on kal  idoscopique des microfacettes identitaires, l'individu calcul   est un flux ¹¹ » et c'est tout    fait le cas de Lucette et d'Aim  e, leurs paroles donn  es forment leur portrait num  rique, avec les photos qui les accompagnent. Ces   uvres ont une fonction m  morielle forte, pour reprendre un principe cher    Bertrand Gervais : elles sont    leur mani  re ce qu'il appelle des « architectures de m  moire ¹² ». Le compte Twitter d'Aim  e capture des moments du pass   qui ont vraiment exist  , comme celui o   Georges, son mari, attrape    la main un li  vre, tapi dans son champ, pendant la moisson dans le Goulot, la pente qui conduit au hameau. Lorsqu'il le soul  ve    la hauteur de son regard, il voit tout    coup arriver le garde champ  tre, son v  lo    la main. Ce moment de triomphe personnel – pas facile de capturer un li  vre    la main, vous n'avez qu'   essayer... – et de panique conjugu  s – franchement pas de chance que Roland, le garde champ  tre passe sur le chemin juste au m  me moment ! – , que mon grand-p  re a ensuite racont      sa famille, je le fais durer sur Twitter avec beaucoup d'  motion. Comment Roland va-t-il r  agir face    ce flagrant d  lit de braconnage ? Les lecteurs se le demandent, mais ils ne le sauront que le lendemain. Cette temporalit   particuli  re de l'  criture et de la lecture sur Twitter permet ici de saisir au vol l'  motion d'un instant r  el du pass  , de la prolonger sur plusieurs jours et de la partager avec un public   tendu. Des inconnus sont venus me reparler de cette sc  ne des mois apr  s, dans la vie r  elle. De m  me, en lisant de mani  re   galement fractionn  e, des allusions multiples au personnage de V  , dans les paroles de Lucette, nous mettons un certain temps    comprendre le traumatisme caus   par son d  part. Le morcellement des informations dans les deux   uvres participe   

l'étirement du temps, qu'il soit lié à la forme du microblogging de Twitter pour Aimée, ou au choix de citations qui s'affichent lentement avec des effets d'effacement progressif sans qu'on puisse agir pour les accélérer dans le cas de Lucette.

Image 3



Le rapport de l'œuvre au passé a déjà été pointé par Marta Caraion à propos du *Madeleine project*, qu'elle définit aussi comme à l'origine d'un tropisme mélancolique.

©DR

« Je considère Madeleine Project comme un échantillon révélateur d'un mécanisme mémoriel à l'œuvre dans les pratiques ordinaires du numérique et, à leur suite, dans toute une mouvance photo-textuelle et documentaire contemporaine. La collecte et la modélisation de traces, leur enregistrement photographique, la tentation de l'archive totale méritent d'être interrogés à la fois comme un phénomène social, comme une radicalisation des usages de la photographie et comme une proposition de mise en récit du monde (à fins historiques et poétiques). Un tropisme mélancolique en résulte, dont il faudrait historiciser et théoriser le fonctionnement.¹³ »

13 C'est la rupture avec le flux numérique, qui crée en partie l'émotion assez vive des lecteurs. Twitter déroule de manière ininterrompue l'information continue, sur le fil du présent, figurant le fleuve numérique du temps qui passe, écriture de l'éphémère. Aimée,

comme Madeleine avant elle, se singularisent dans le fil de leurs abonnés, elles sont à l'écart, elles cherchent une forme d'anti-buzz. Aimée vous donne la recette de sa pâtée pour les cochons, ou vous raconte comment elle allait en charrette au lavoir. Certains lecteurs expriment d'ailleurs leur intérêt pour le projet par cet écart, comme la chercheuse sur la place des femmes en politique Camille Froidevaux-Metterie « C'est une très belle idée de faire ainsi entendre Aimée au beau milieu du tumulte agressif de Twitter¹⁴ ». L'irruption assez incongrue et anachronique de cette fermière du passé dans le fil ininterrompu de l'actualité renvoie à quelque chose qui serait peut-être hors du temps, tout comme la fixité de Lucette face aux trains. *Lucette* met en œuvre une logique de ressassement, avec le retour de mêmes phrases, comme si la grand-mère se répétait infiniment, selon un tirage aléatoire de la programmation des citations. Les formules pour caractériser la durée, l'étirement du temps sont nombreuses : « Le travail n'avance pas », « je passe des journées sans voir personne ». Aimée et Lucette échapperaient dans l'interface numérique au temps et à l'espace. Antonio Casilli¹⁵ parle à cet égard d'un espace numérique imaginaire qui rejoindrait l'idéal platonicien d'un espace au-delà du ciel, qu'il qualifie d'« Éden de données pures ». Il cite alors Michael Benedikt : cet espace informationnel serait l'extension de notre besoin d'investir d'autres plans mythiques. Le problème de Lucette ou d'Aimée serait l'obsolescence programmée des supports numériques, (puisqu'il paraît que la page de *Lucette* a failli disparaître récemment pour une histoire d'API). Mais elles ont dans leurs architextes respectifs acquis une place de vie autonome, une fiction d'éternité. Lucette, même si elle tient des propos mélancoliques, parfois proches des *memento mori*, se régénère à chaque connexion. Et si Anaïs Guilet et Bertrand Gervais¹⁶ ont mis en évidence le fait que de nombreuses œuvres numériques s'inscrivent pleinement dans une esthétique du flux en affirmant que « [l]es œuvres qu'on trouve dans ce cyberspace apparaissent comme des formes hybrides, instables et éphémères », on pourrait penser que *Lucette* et *Aimée* à leur manière représentent du moins une tentative pour y échapper.

Données et œuvres solidaires

Image 4



Sources : Lucette, gare de Clichy ; Compte X/Twitter d'Aimée Rolland ; ©DR

- 14 La concomitance de ces projets qui font surgir les vies de ces mamies modestes montre également que la littérature numérique ne se définit pas que de manière formelle, par ses architextes, ses dispositifs, mais qu'il existe aussi des poétiques spécifiques, des thèmes, des types de personnages. Ce n'est pas un hasard que des femmes choisissent ce type de support pour parler de celles qui les ont précédées. Un réseau singulier se noue, qui cherche à donner une visibilité à des femmes modestes, aux silhouettes effacées par le temps, des *Vies minuscules*¹⁷ numériques, féminines, qui adoptent une forme qui est elle-même simple et populaire.
- 15 Dans le mot « données », il y a la racine « don ». On y retrouve l'idée de partage, une valeur à l'origine du web. Et il me semble que les œuvres numériques que nous étudions ici se construisent également grâce à ces valeurs. Antonio Casilli dans *Les liaisons numériques* cite Peter Kollok¹⁸ pour qui l'un des moteurs de l'engagement des personnes sur Internet est une logique de don et de contre-don. Un projet comme Aimée Rolland est bien celui d'un partage avec des lecteurs qui répondent régulièrement, reconnaissent dans cette vie

ordinaire, des événements de leur propre histoire ou de celle de leurs proches. Ils participent même aux événements familiaux. Ainsi, c'est la panique quand le petit Gérard, âgé de 2 ans, s'est perdu dans le village. Allait-on le retrouver ? Des dizaines de lecteurs sont aux abois. Certains proposent même des endroits où il aurait pu se cacher. Quand on parle de fracture numérique, il est assez évident que les vieilles personnes comme Lucette sont les plus exclues des représentations dans les flux contemporains, notamment sur les réseaux sociaux. D'ailleurs un grand nombre de ses déclarations se rapportent à sa propre solitude, que la présence du lecteur ne parvient pas totalement à résoudre, comme « [j]e parle au mur, puis j'allume la télévision pour me sentir moins seule », « [l]a solitude, elle m'accompagne tous les jours », « je me sens encore plus seule que d'habitude¹⁹ ». Pourtant, l'invitation donnée à l'internaute a la forme d'une visite, puisqu'il faut cliquer sur la formule « Visiter l'œuvre » pour entrer chez Lucette. La littérature numérique d'aujourd'hui, est aussi une manière de « réparer, renouer, ressouder, combler les failles des communautés » comme Alexandre Gefen cherche à le montrer dans son essai *Réparer le Monde*²⁰, mais avec les limites et les atouts propres aux liens faibles des réseaux du web. Lucette, Aimée et les autres mamies de la littérature numérique s'inscrivent dans une forme qui prend au sérieux l'idée forte du lien dans les réseaux digitaux, qui existe malgré la distance. « Internet réussit avec un mécanisme de liens faibles à créer des confiances fortes qui consolident [...] les relations humaines », affirme Antonio Casilli²¹ dans *Les Liaisons numériques*. Les œuvres de littérature numérique s'en emparent, et s'inscrivent dans cette esthétique du *care*.

- 16 J'aime à penser que Lucette, Aimée, Madeleine, toutes les mamies de ces œuvres en ligne se donnent la main, dans un territoire virtuel, poétique et solidaire de la littérature d'aujourd'hui, leur solitude rompue.

BIBLIOGRAPHIE

BATAILLON Stéphane, « Twittérature, la littérature sur Twitter : un état des lieux », Stéphane Bataillon [en ligne], 2011, URL : <https://www.stephanebataillon.com/twitt>

[rature-twitter-et-la-litterature/](#) [consulté le 8 avril 2024].

BEAUDOUX Clara, *Madeleine Project* [en ligne], 2016-2020, URL : <https://madeleineproject.fr/le-projet-4/> [consulté le 8 avril 2024].

BEAUDOUX Clara, *Madeleine Project*, Paris, Éditions du sous-sol, 2016.

BOULLIER Dominique, *Sociologie du numérique*, Paris, Armand Colin, coll. « U », 2016.

CAHEN Françoise, compte Twitter d'Aimée Rolland [en ligne], URL : <https://twitter.com/RollandAime1> [consulté le 8 avril 2024].

CARAION Marta, « Mémoire matérielle, photographie, indicialité, le cas Madeleine Project », *Nouveaux Cahiers de Marge* [en ligne], n° 2, 2020, DOI : [10.35562/marge.325](https://doi.org/10.35562/marge.325).

CARDON Dominique, *À quoi rêvent les algorithmes ?*, Paris, Seuil, 2015.

CASILLI Antonio, *Les Liaisons numériques, vers une nouvelle sociabilité ?*, Paris, Seuil, 2013.

CHAMBEFORT Françoise, *Lucette, gare de Clichy* [en ligne], 2016, URL : <http://www.francoisechambefort.com/lucette-gare-de-clichy> [consulté le 8 avril 2024].

CHAMBEFORT Françoise, « Mise en œuvre d'une démarche de recherche création dans un doctorat en SIC portant sur l'art numérique », *Association des jeunes chercheurs du CREM*, 28/11/2016 <https://ajccrem.hypotheses.org/tag/art-generatif> [consulté le 8 avril 2024].

GEFEN Alexandre, *Réparer le monde*, Paris, Corti, 2017.

GERVAIS Bertrand et GUILLET Anaïs, « Esthétique et fiction du flux. Éléments de description. » *Protée* [en ligne], vol. 39, n° 1, 2011, p. 89-100, DOI : [10.7202/1006730ar](https://doi.org/10.7202/1006730ar).

JEANNERET Yves, SOUCHIER Emmanuël. « L'énonciation éditoriale dans les écrits d'écran » in : *Communication et langages*, n° 145, 2005, p. 3-15.

SAEMMER Alexandra, compte Facebook d'Alexandra Saemmer [en ligne], URL : <https://www.facebook.com/saemmer.alexandra> [consulté le 8 avril 2024].

NOTES

1 Clara Beaudoux, *Madeleine Project* [en ligne], 2016-2020, URL : <https://madeleineproject.fr/le-projet-4/> [consulté le 8 avril 2024] ou en version papier : Clara Beaudoux, *Madeleine Project*, Paris, Éditions du sous-sol, 2016.

2 Françoise Chambefort, *Lucette, gare de Clichy* [en ligne], 2016, URL : <http://www.francoise-chambefort.com/lucette-gare-de-clichy> [consulté le 8 avril 2024].

- 3 Françoise Cahen, compte Twitter d'Aimée Rolland [en ligne], URL : <https://twitter.com/RollandAime1> [consulté le 8 avril 2024].
- 4 Alexandra Saemmer, compte Facebook d'Alexandra Saemmer, URL : <https://www.facebook.com/saemmer.alexandra> [consulté le 8 avril 2024].
- 5 Yves Jeanneret, Emmanuël Souchier, « L'énonciation éditoriale dans les écrits d'écran » in : *Communication et langages*, n° 145, 2005. L'empreinte de la technique dans le livre. p. 3-15.
- 6 Françoise Chambefort, « Mise en œuvre d'une démarche de recherche création dans un doctorat en SIC portant sur l'art numérique », *Association des jeunes chercheur·e·s du CREM* [en ligne], 28 novembre 2016 URL : <https://ajccrem.hypotheses.org/tag/art-generatif> [consulté le 8 avril 2024].
- 7 Dominique Boullier, *Sociologie du numérique*, Armand Colin, coll. « U », 2016, p. 142.
- 8 Gilles Bonnet, Serge Bouchardon, Alexandra Saemmer, Lucile Haute, *Littérature numérique & performance* [en ligne], université Jean-Moulin Lyon 3 – Villa Gillet, 27-28 mai 2015, <https://www.utc.fr/~bouchard/Colloque-litterature-numerique-et-performance.pdf> [consulté le 8 avril 2024].
- 9 Stéphane Bataillon, « Twittérature, la littérature sur Twitter : un état des lieux », *Stéphane Bataillon* [en ligne], 2011, URL : <https://www.stephanebataillon.com/twittérature-twitter-et-la-littérature/> [consulté le 8 avril 2024].
- 10 Lionnel Ruffel, *Brouhaha, Les mondes du contemporain*, Verdier, 2016.
- 11 Dominique Cardon, *À quoi rêvent les algorithmes ?*, Seuil, 2015, p. 88.
- 12 Bertrand Gervais, Jean-Marie Dallet, « Architectures et palais de mémoire » in : *Architectures de mémoire*, Presses universitaires de Paris-Nanterre, 2019, p. 10-19
- 13 Marta Caraion, « Mémoire matérielle, photographie, indicialité, le cas Madeleine Project », *Nouveaux Cahiers de Marge* [en ligne], n° 2, 2020, DOI : [10.35562/marge.325](https://doi.org/10.35562/marge.325).
- 14 Camille Froidevaux-Metterie, message privé du 25 novembre 2022.
- 15 Antonio Casilli, *Les liaisons numériques : vers une nouvelle sociabilité ?*, Paris, Seuil, coll. « La couleur des idées », 2010, p. 63-65.
- 16 Bertrand Gervais et Anaïs Guilet, « Esthétique et fiction du flux. Éléments de description » *Protée* [en ligne], vol. 39, n° 1, 2011, p. 89-100, DOI : [10.7202/1006730ar](https://doi.org/10.7202/1006730ar).

- 17 Pierre Michon, *Vies minuscules*, Paris, Gallimard, 1996.
- 18 Antonio Casilli, *Les Liaisons numériques, vers une nouvelle sociabilité ?*, Paris, Seuil, 2013, p. 53.
- 19 Françoise Chambefort, *Lucette gare de Clichy* [en ligne], 2016, URL : <http://www.francoise-chambefort.com/lucette-gare-de-clichy> [consulté le 8 avril 2024].
- 20 Alexandre Gefen, *Réparer le monde*, Paris, Corti, 2017.
- 21 Antonio Casilli, *Les Liaisons numériques, vers une nouvelle sociabilité ?*, Paris, Seuil, 2013, p. 314.

RÉSUMÉS

Français

Cet article analyse deux œuvres numériques qui s'attachent à des aïeules : *Lucette, gare de Clichy* de Françoise Chambefort, une œuvre qui utilise les données de la SNCF pour intégrer le lecteur dans l'œuvre, et *Aimée Rolland*, un récit de twittérature qui retrace la vie de ma grand-mère au rythme d'un tweet par jour depuis deux ans. Le rapport du réel à la fiction, au temps qui passe, et la solidarité du lecteur sont les axes essentiels de cette étude.

English

This article analyses two digital works which are attached to grandmothers: *Lucette, gare de Clichy* by Françoise Chambefort, a work using data from the SNCF to integrate the reader into the work, and *Aimée Rolland*, a story of Twitterature which traces the life of my grandmother at the rate of a tweet a day for two years. The relationship between reality and fiction, the passage of time, and the solidarity of the reader are the essential axes of this study.

INDEX

Mots-clés

littérature, numérique, grand-mère, narration, données

Keywords

literature, digital, grandmother, narration, data

AUTEUR

Françoise Cahen

Thalim, université Paris 3, Sorbonne Nouvelle

IDREF : <https://www.idref.fr/098303899>

ORCID : <http://orcid.org/0000-0001-5388-8902>

ISNI : <http://www.isni.org/0000000441617234>

BNF : <https://data.bnf.fr/fr/17124915>

Banques de données

Lab·yrinthe : plateforme québécoise de diffusion d'œuvres numériques jeunesse

Lab·yrinthe: a Quebecois digital children's literature diffusion platform

Nathalie Lacelle et Eleonora Acerra

DOI : 10.35562/marge.916

Droits d'auteur

CC BY-NC-SA

PLAN

Documentation de fiches descriptives d'œuvres numériques jeunesse
Documentation des concepts clés pour comprendre et produire des œuvres numériques
Documentation de fiches éducatives pratiques de créations numériques
Documentation de modules pour soutenir l'édition numérique jeunesse
Documentation de projets de recherche en collaboration avec les milieux éducatif, culturel et communautaire
Documentation des actualités sur l'éducation et l'édition de littérature numérique jeunesse
Conclusion

TEXTE

- 1 La création de Lab·yrinthe (laboratoire virtuel québécois de l'édition et de l'éducation aux œuvres numériques) émane de l'identification d'un besoin d'accompagnement des professionnel-le-s de l'édition et de l'éducation, respectivement, à la conception, la production et la diffusion d'œuvres numériques, voire à leur scolarisation. Ces constats sont issus de la recherche-action « Soutien au développement de démarches d'édition numérique jeunesse au Québec à partir de pratiques favorables de production, diffusion et réception », réalisée entre 2017-2022 par Nathalie Lacelle et l'Équipe de recherche en littératie médiatique multimodale, dans le cadre du programme de l'action concertée du Fonds de recherche du Québec – Société et culture, et du ministère de la Culture et de la Communication. Menée en deux temps – soit, un temps de collecte de données auprès des professionnels de l'éducation, de l'édition et

des jeunes publics, et un temps d'accompagnement personnalisé de projets d'édition d'œuvres numériques et d'éducation aux œuvres numériques pour la jeunesse –, cette recherche a notamment permis de conclure à l'existence d'enjeux majeurs freinant le développement de l'offre numérique jeunesse et son utilisation en contexte scolaire¹. Allant de la méconnaissance de l'offre éditoriale par les éducateur·rice·s, les bibliothécaires et, plus généralement, par le grand public, aux difficultés de financement et à l'absence, pour les éditeur·rice·s, des voies de découvrabilité et de légitimation propres au livre papier, ces différents « points de rupture² » ont été documentés et analysés, en vue de produire des ressources susceptibles de favoriser les démarches de production, diffusion et enseignement des œuvres numériques. Plus particulièrement, dans le but de documenter les résultats de la recherche, de stimuler les pratiques éditoriales et d'accompagner les éducateur·rice·s dans leurs démarches d'intégration des œuvres numériques en contexte scolaire, parascolaire et privé, nous avons conçu le site-laboratoire Lab·yrinthe, sous la direction de Nathalie Lacelle et la codirection d'Eleonora Acerra, en partenariat avec Littérature québécoise mobile (B. Gervais, dir., 2019-2024).

Figure 1. Page d'accueil du site Lab·yrinthe.ca



© Lab·yrinthe

2 Développé via la plateforme Wordpress et lancé en mai 2021, le site web se veut un laboratoire virtuel sur les phénomènes littéraires

numériques jeunesse contemporains ainsi qu'une base de ressources conceptuelles, informatives, scientifiques et didactiques à destination des éducateur·rice·s, des éditeur·rice·s, des concepteur·rice·s, des chercheur·e·s et des étudiant·e·s. Inscrit dans des démarches analogues de catalogage et analyse de la production numérique contemporaine, allant du répertoire du laboratoire NT2 sur les arts et les littératures numériques à la base de données de l'Electronic Literature Organisation, Lab·yrinthe est organisé en six sections : la première présente les missions, l'historique et l'équipe du site-laboratoire ; la deuxième, adressée aux éducateur·rice·s, aux éditeur·rice·s et aux concepteur·rice·s de contenus pour les jeunes publics, mais également aux étudiant·e·s, propose des descriptions détaillées d'une sélection d'œuvres littéraires, documentaires ou artistiques numériques jeunesse ; la troisième, destinée de manière prioritaire aux éditeur·rice·s, présente des recommandations issues de la recherche pour la conception et la diffusion des contenus numériques, adaptés aux besoins des enfants et des enseignant·e·s ; la quatrième, composée de trois sous-sections, contient des descriptions de concepts théoriques en lien avec l'écriture et la littérature numérique (notamment un référentiel de compétences de littératie numérique), des fiches pratiques à visée éducative relatant des projets de création numérique réalisés avec des jeunes, ainsi que des outils, des logiciels et des ressources pour réaliser des créations numériques ; la cinquième décrit et analyse des projets de recherche, réalisés en collaboration avec des éditeur·rice·s, des artistes ou des diffuseurs ; la dernière, composée de quatre sous-sections, répertorie des nouvelles pertinentes (événements, parutions de livres, conférences, etc.) sur la création numérique jeunesse au Québec. Les contenus de chaque section, constamment mis à jour par le travail d'une équipe éditoriale composée de chercheur·e·s, d'étudiant·e·s chercheur·e·s et de professionnel·le·s de recherche, témoignent de l'évolution des œuvres, des concepts et des projets émergents, constituant ainsi un lieu de saisie et de diffusion des données sur le corpus fictionnel et documentaire numérique pour la jeunesse au Québec. Le site compte actuellement 36 fiches « Œuvres », 10 fiches « Concepts », 11 fiches « Pratiques », 17 fiches « Outils », 4 fiches « Ressources », 17 fiches « Projets », 8 fiches « Édition », 21 fiches « Évènements », 8 fiches « Parution », 10 fiches « Billets » et 13 fiches « Veille ». Dans cet article, nous décrirons les principales sections de

Lab·yrinthe, en illustrant le processus de documentation des contenus, et notamment les modalités de description des œuvres et des pratiques éducatives de réception et production de contenus numériques.

Documentation de fiches descriptives d'œuvres numériques jeunesse

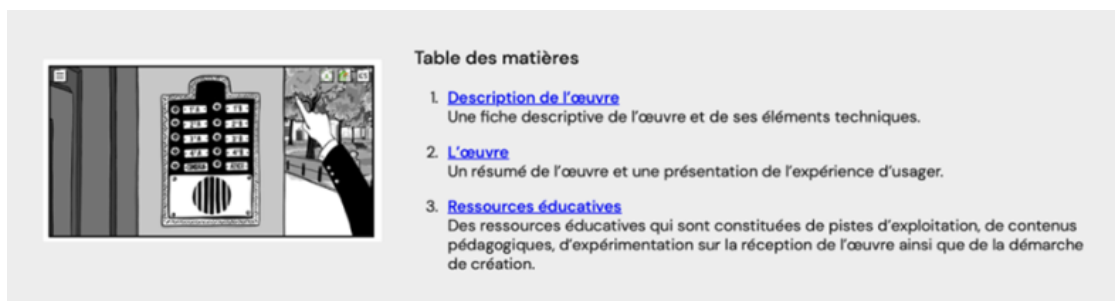
- 3 La section « Œuvres » de Lab·yrinthe présente un catalogue hétérogène de créations littéraires, documentaires et artistiques numériques produites ou distribuées au Québec, qui comprend des livres augmentés, des applications mobiles, des jeux vidéo narratifs, des narrations géolocalisées, des créations en réalité augmentée, des bandes dessinées interactives, des œuvres génératives et des balados, réalisées par des artistes, des collectifs d'artistes ou encore des élèves en contexte scolaire. Chaque œuvre a été analysée à partir d'un ensemble de paramètres descriptifs conçus par des membres de l'Équipe de recherche en littératie médiatique multimodale³ dans le but d'illustrer les composantes sémiotiques et technologiques observées, ainsi que les effets poétiques ou rhétoriques de leurs combinaisons. Concevant le texte littéraire numérique comme une imbrication d'éléments littéraires et technologiques⁴, nous avons fait le choix d'identifier des éléments qui apparaissent de manière récurrente, quelles que soient les spécificités formelles et stylistiques de chaque typologie textuelle. Dès lors, nous avons pris en compte trois niveaux interdépendants. Le premier décrit les structures formelles des textes ; le second se concentre sur les différents modes de représentation des contenus sur l'écran ; le troisième interroge les différents types d'interactions requis au sujet empirique. L'analyse des structures formelles a permis de classer les œuvres littéraires numériques en fonction de leurs genres et des rôles donnés à leur destinataire modèle. D'une part, nous avons pu distinguer les créations affichant linéarité et séquentialité, d'autres présentant des parcours non linéaires, tabulaires ou multiples, dépendant totalement ou en partie des choix des interlecteur·rice·s. D'autre part, les postures présupposées des destinataires ont permis de répartir les

œuvres en fonction des interventions requises, en enregistrant celles qui privilégient la rencontre ludique, la découverte lectorale, sonore ou tactile. En ce qui concerne la dimension des modes de représentation des contenus, nous avons adapté le modèle d'Aarseth⁵ et identifié des œuvres dans lesquelles les matières textuelles restent identiques pour tout le temps de la lecture ou de la consultation, ou bien évoluent en fonction des sollicitations des interlecteur·rice·s ou d'un mouvement automatisé (ou semi-automatisé) du programme informatique de reproduction. Par cela, nous avons distingué des créations permettant de réactiver, rejouer et relancer un élément lu, vu, écouté ou joué dans les mêmes conditions de la première lecture, d'autres proposant des variations par rapport à la première exécution. En décrivant l'accessibilité des matières textuelles, nous avons distingué les œuvres numériques en fonction des conditions permettant aux contenus de se révéler et ainsi identifié des contenus qui se donnent à voir sans qu'aucune opération « non triviale⁶ » de la part des destinataires ne soit nécessaire, d'autres qui requièrent un effort interprétatif et physique de la part de l'interlecteur. Enfin, en considérant l'interactivité « une possibilité médiatisée et programmée de choix et d'actions⁷ », nous avons distingué des œuvres dans lesquelles les actions et gestes des lecteur·rice·s provoquent une modification de la structure formelle de l'œuvre, du code informatique ou de la représentation du texte lui-même. À l'intérieur de cet ensemble, nous avons pu distinguer des créations programmant des manipulations contraignantes et explicitement sollicitées par des indicateurs verbaux ou iconiques, d'autres présupposant des interventions optionnelles, implicites ou basées sur des métaphores d'interface⁸ voire sur des analogies avec le monde réel⁹. En suivant cette grille de paramètres descriptifs, nous avons analysé chaque œuvre numérique et présenté ses structures et thématiques principales, ainsi que l'expérience utilisateur·rice présupposée, en fonction des dispositifs et des technologies mobilisées. Chaque fiche a été complétée par des propositions pédagogiques et didactiques, qui visent à interroger, en contexte scolaire, à la fois les thématiques des œuvres et leurs principales caractéristiques technolittéraires. Cette sous-section, destinée avant tout aux enseignant·es, propose ainsi quelques clés de lecture et d'analyse et, en même temps, fournit des exemples clairs des exploitations possibles des textes numériques dans une perspective

didactique. Pour chaque fiche, des termes clés concernant, par exemple, des concepts et des notions de théorie littéraire numérique, ou de littératie médiatique multimodale, ont été associés à des liens hypertextuels, qui renvoient à la section « Éducation », où des informations théoriques et critiques plus détaillées sont fournies.

- 4 Ainsi, dans la section « Œuvres » de Lab·yrinthe, chacune des créations d'artistes ou d'élèves sélectionnées est décrite comme dans l'exemple suivant, tiré de l'analyse de la bande dessinée numérique 2, de Margarita Molina Fernández¹⁰. Chaque fiche présente une table des matières divisée en trois sections (figure 2), une liste de tous·tes les artistes contributeur·rice·s de la bande dessinée, des informations concernant les langues disponibles, les principales fonctionnalités numériques et les caractéristiques médiatiques de l'œuvre, ses systèmes opératifs et les supports requis pour la lecture, ainsi que des liens actifs pour téléchargement.

Figure 2. Table des matières de la fiche de description des œuvres numériques



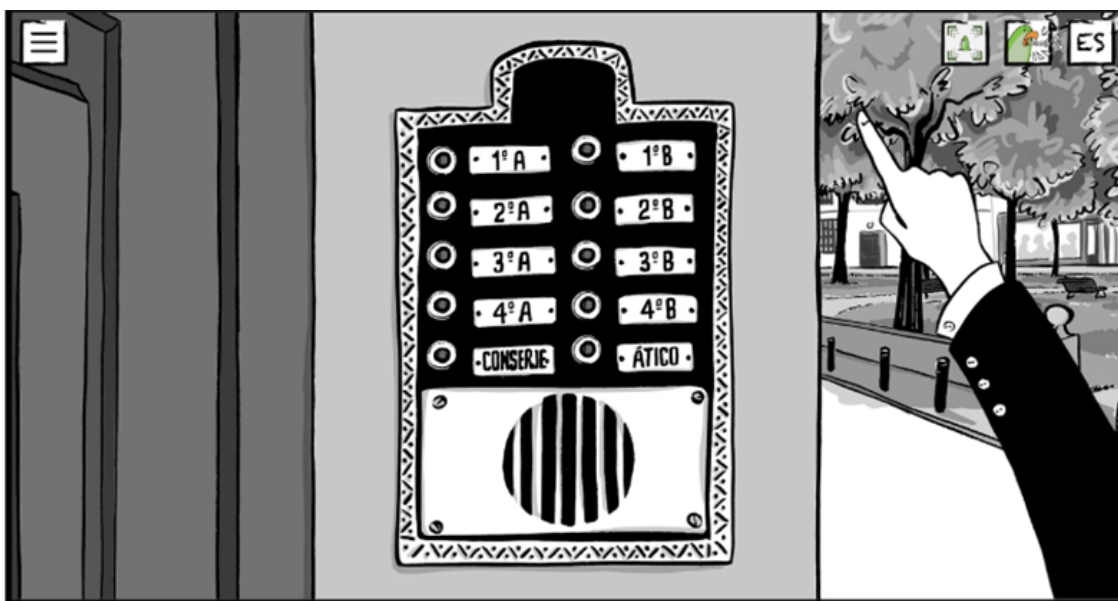
© Lab·yrinthe

- 5 S'ensuit une description des œuvres en fonction des paramètres énoncés plus haut, dont voici un exemple à partir de 2 :

« Le récit de 2, empreint de mystère avec une touche d'humour, suit le personnage de Segundo Santos, qui a été convoqué chez un notaire dans un bâtiment du pittoresque centre-ville de Madrid. Néanmoins, cela n'est qu'un stratagème manigancé par les énigmatiques habitants de l'immeuble afin de le faire tomber dans leurs pièges. Segundo n'a qu'une façon de s'en sortir et la bonne issue de ses aventures dépend des choix effectués par le·la lecteur·rice. 2 a une structure narrative arborescente comportant six fins. Le choix entre les différents parcours à explorer s'effectue grâce à l'activation

d'une diversité d'objets interactifs situés à l'intérieur du cadre : les boutons des sonnettes, les doubles translucides de Segundo ou les objets et personnages animés en boucle. Le numéro 2 constitue le leitmotiv de l'œuvre, aussi bien dans le récit, qui joue avec la notion de double, que dans sa structure hypermédiatique, basée sur un système de bifurcations binaires¹¹. »

Figure 3. Structure narrative arborescente à six fins dans l'œuvre 2



© Lab·yrinthe

« Les systèmes d'affichage et de navigation de 2 relèvent du turbomédia. Un cadre horizontal statique remplit les fonctions de surface matérielle d'inscription et de dispositif de lecture. Les flèches situées sur les quatre bords du cadre facilitent la navigation : les flèches indiquant le bas et la droite permettent d'avancer, tandis que celles indiquant le haut et la gauche permettent de reculer. Trois boutons additionnels en haut de l'écran offrent la possibilité d'afficher la BD en plein écran, de désactiver le son et de changer le langage du français à l'espagnol¹². »

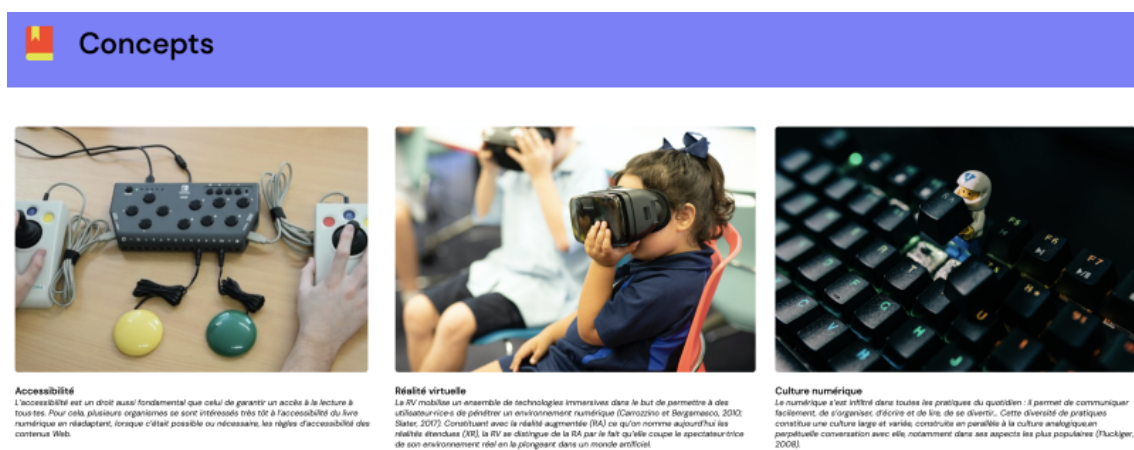
- 6 La fiche vise ainsi à fournir un aperçu des contenus, de ses matières textuelles et de ses modes de navigation, ainsi que des articulations entre les dimensions générique, formelle, sémiotique et interactive. La description de l'œuvre débouche enfin sur des pistes d'exploitation didactiques, délibérément souples et peu directives, élaborées par

l'équipe éditoriale, et renvoie, lorsqu'elles sont disponibles, à des fiches pédagogiques produites par des éditeur·rice·s jeunesse. Les enseignant·e·s sont censé·e·s s'approprier cette première base de ressources, la personnaliser et l'adapter à leurs contextes d'enseignement, voire, à terme, l'enrichir en partageant leurs propres contenus et matériaux didactiques.

Documentation des concepts clés pour comprendre et produire des œuvres numériques

- 7 La section « Éducation » du site est plus spécifiquement dédiée à la présentation d'une sélection de théories, de concepts et de notions nécessaires pour comprendre la création littéraire numérique, pour s'en approprier les codes et, éventuellement, pour l'enseigner. Si elle s'adresse avant tout aux éducateur·rice·s et aux enseignant·e·s, qui sont censé·e·s s'y référer pour acquérir les connaissances théoriques nécessaires à guider la découverte critique et sensible des corpus littéraires numériques, elle vise également les éditeur·rice·s qui pourraient, conformément aux besoins exprimés dans le cadre de la recherche, y identifier des éléments significatifs pour la formation des jeunes lecteur·rice·s, à réinvestir dans le développement de contenus innovants, pertinents et variés.

Figure 4. Documentation de concepts clés en littérature numérique



- 8 Chaque entrée de cette section est organisée en deux parties : la première décrit et synthétise des travaux théoriques et critiques sur les thématiques identifiées ; la deuxième souligne, dans une perspective didactique, des éléments qu'il convient de présenter aux jeunes lecteur·rice·s pour qu'ils·elles puissent profiter pleinement de l'expérience de lecture - littéraire et documentaire - numérique. Dans ce cas aussi, des liens hypertextuels permettent de relier les théories, les concepts, les notions et les phénomènes abordés aux œuvres dans lesquelles ils peuvent être observés. De même, des renvois aux modules à destination des éditeur·rice·s permettent d'associer des aspects théoriques, littéraires ou médiatiques à des observations ou des témoignages recueillis dans le cadre de la recherche.
- 9 À l'heure où nous écrivons, la section « Concepts » comporte dix pages, respectivement dédiées à la présentation du cadre théorique de la *multimodalité*, des notions de « genre » et de « figures de style » en contexte numérique, des concepts de « labilité » et d'« obsolescence », d'« interactivité », de « linéarité » et « non-linéarité », de « culture numérique », de « réalité virtuelle » et de « réalité augmentée ». Sont ici également présentées les compétences en littératie médiatique multimodale numérique (#LMM), qui servent, entre autres, de modèle théorique pour l'identification et l'analyse des compétences mobilisées dans la section « Pratiques » de Lab·yrinthe¹³. Ainsi, chaque expérience menée avec des partenaires, des enseignant·e·s ou des élèves en contexte éducatif ou scolaire est décrite à partir des indicateurs de compétences de #LMM et contribue à l'identification de ces articulations entre compétences littéraciques, littéraires, multimodales et numériques qui semblent au cœur des processus de lecture et d'écriture numériques. Selon la grille des compétences, publiée en 2022, pour construire et produire le sens en contexte de lecture numérique, sont notamment nécessaires des compétences génériques et théoriques (numériques et analogiques), permettant d'appréhender les arts littéraires, leurs évolutions historiques, médiatiques et technologiques, des compétences cognitives, subjectives, sémiotiques, multimodales et sociales (figure 5), ainsi que des compétences numériques en LMM¹⁴ (figure 6), impliquant de savoir mobiliser différentes ressources, modes, codes et médias, voire

de pouvoir discuter, commenter, justifier ou analyser ses démarches d'écriture, compréhension et interprétation intermodales, ses choix gestuels, ses interventions sur/avec/dans les dispositifs, les logiciels et les supports ou les ressources sémiotiques (voir figure 7) de manière critique et consciente. L'identification de ces compétences, aussi bien dans les planifications des enseignant·e·s que dans les productions des élèves, constitue ainsi une forme de validation empirique des savoirs et savoir-faire identifiés au fil des années en différentes situations de réception et production d'œuvres numériques.

Figure 5. Compétences en littératie médiatique multimodale (Lacelle, Boutin et Lebrun, 2017)

Compétences LMM
• Compétence cognitive : savoir décoder, comprendre et produire le sens. • Compétence subjective : savoir s'investir émotionnellement et intellectuellement dans la réception et la production de sens. • Compétence sémiotique : savoir (re)connaître, analyser et utiliser des systèmes de signes et des symboles dans la réception et la production de sens. • Compétence multimodale : savoir (re)connaître, analyser et combiner des codes, des modes et des langages pour recevoir et produire des designs narratifs, poétiques, expérientiels, intermodaux, spatiaux, graphiques, sonores et haptiques. • Compétence sociale : savoir (re)connaître, analyser et produire le sens en interaction avec des individus du monde réel et virtuel. • Compétence critique : savoir (re)connaître et analyser les contextes de réception et de production de sens en identifiant les intentions et les idéologies sous-jacentes.

Figure 6. Compétences numériques en littératie médiatique multimodale (Acerra et Lacelle, 2022)

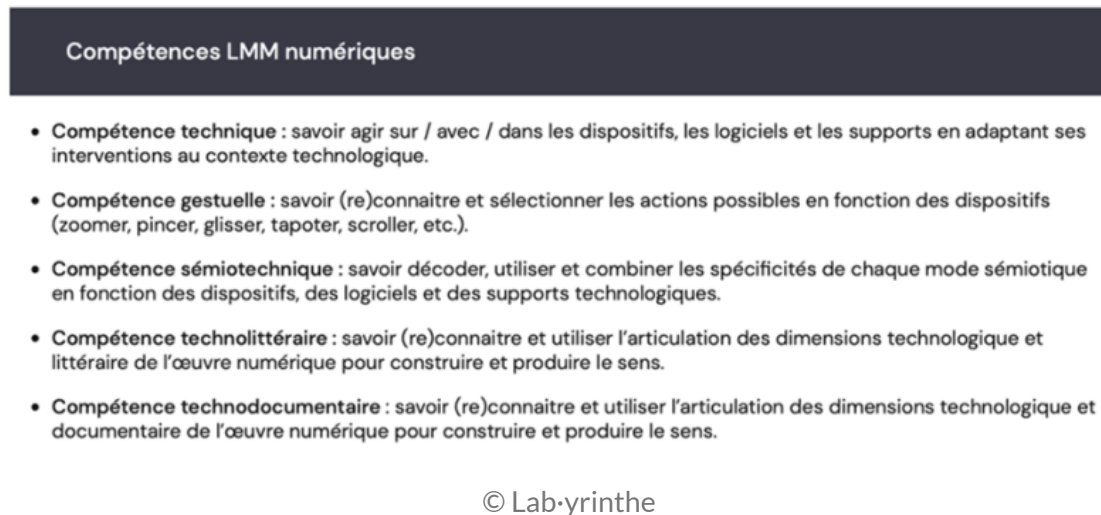
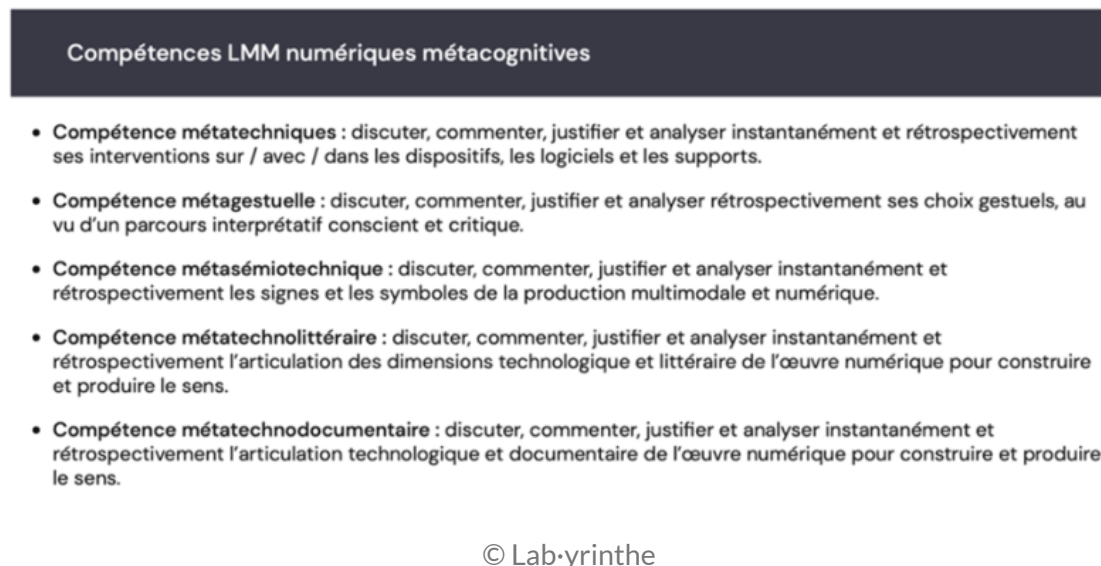


Figure 7. Compétences métacognitives numériques en littératie médiatique multimodale (Acerra et Lacelle, 2022)



Documentation de fiches éducatives pratiques de créations numériques

- 10 Pour illustrer les compétences numériques en littératie médiatique multimodale (#LMM) mobilisées dans une diversité de productions réalisées par les élèves, nous documentons, dans une section dédiée de Lab·yrinthe, des « Pratiques » qui prévoient à chaque fois l'articulation d'étapes de réception, production et diffusion de contenus numériques. Les projets décrits dans cette section impliquent généralement des chercheur·e·s, des enseignant·e·s, des artistes, des conseiller·ère·s pédagogiques, et d'autres personnes-ressources, qui contribuent à la cocréation d'activités répondant à des objectifs de formation liés au développement de compétences littéraciques, littéraires, multimodales et numériques visées par les programmes de formation.
- 11 Voici un exemple de projet réalisé par l'une des collaboratrices de l'équipe, Emma June Huebner, intitulé *Les stories à l'école*, destiné à des élèves du 2^e cycle du secondaire¹⁵. Ce projet se propose, d'une part, de mobiliser les publications des stories d'Instagram pour créer un récit hypermédiatique, d'autre part, d'aider les enseignant·e·s à s'appropriier les nouvelles modalités et formes médiatiques des stories en contexte scolaire. À l'instar de la fiche proposée pour documenter le projet porté par Emma June Huebner, chaque page de la section « Pratiques » présente l'objectif de l'activité, les ressources utiles à sa transposition, les étapes de réalisation et les compétences mobilisées par les élèves, en contexte de réception, de production ou de diffusion. La section suivante sert à présenter les ressources nécessaires pour réaliser l'activité, soit des liens vers des œuvres, lues ou produites par les élèves, les appareils technologiques et les logiciels, pour certains desquels les modalités d'usage sont précisées dans la section « Outils ».

Figure 8. Section Ressources de l'activité *Stories à l'école*

Ressources	
Œuvres : • Une force tranquille (format vidéo) Équipements : • Téléphone mobile (iOS ou Android) • Ordinateur	Logiciels : • Logiciels de captation et de montage vidéo (OpenShot, Davinci, ou iMovie) • Logiciels de captation et de montage sonore (Audacity ou Garageband) • Sites de sons et de musiques libres de droits (p. ex. : freesound.org et archivemusic.com)

© Lab·yrinthe

12 Sont ensuite décrites les différentes étapes de réalisation de l'activité, permettant aux enseignant·e·s de s'emparer rapidement le déroulement et de l'adapter, selon leurs intentions pédagogiques et leurs contextes d'enseignement. Dans le cas du projet *Les stories à l'école*, par exemple, sont synthétiquement proposées les six phases du projet, sans insister sur les choix de logiciels utilisés pour la captation sonore ou pour l'édition des images, qui pourront être plus opportunément sélectionnés par chaque enseignant·e, en fonction de leurs besoins et des possibilités d'équipement de leur établissement :

1. L'élève réalise une entrevue à partir de questions qui vont amener la personne choisie à raconter son histoire ou son expérience en trois moments : un début, un milieu et une fin ;
2. À partir de l'enregistrement de l'entrevue, l'élève rédige un court récit à la 3^e personne du singulier ;
3. L'élève s'enregistre en train de narrer son récit à l'aide du logiciel de captation audio ;
4. L'élève importe la narration dans le logiciel de montage audio de son choix ;
5. L'élève filme des images, en format vertical, à l'aide de son téléphone. Il-elle peut aussi utiliser des images d'archives ou des vidéos libres de droits s'il-elle le souhaite ;
6. L'élève importe la bande sonore et les images tournées dans un logiciel de montage. Ensuite, l'élève superpose l'entièreté du texte ou des extraits, au choix, par-dessus les images. Lorsque le montage est complété, l'élève exporte son projet en *stories* ou en vidéo.

Figure 9. Montage du récit de l'élève sous forme de *stories* sur Instagram



© Lab-yrinthe

- 13 La section suivante de la fiche identifie les compétences en #LMM mobilisées par les élèves, en contexte de production comme de réception. Ainsi, dans le cadre du projet *Les stories à l'école*, les élèves ont dû travailler sur leur capacité à décoder, comprendre et produire le sens en mobilisant les *stories* d'Instagram (compétence cognitive) ; sur leur capacité à (re)connaître, analyser et utiliser des systèmes de signes et de symboles propres aux *stories* d'Instagram (compétence sémiotique) ; sur leur capacité à (re)connaître, analyser et combiner des codes, des modes et des langages pour recevoir et produire des designs narratifs, poétiques et multimodaux qui s'arriment aux formats narratifs et technologiques des *stories* d'Instagram (compétence multimodale) ; sur leur capacité à utiliser des logiciels de montage vidéo et audio pour construire une narration multimodale numérique en format *stories* (compétence technique) ; sur leur capacité à (re)connaître et à utiliser l'articulation des dimensions technologiques de l'œuvre numérique pour construire et produire un récit poétique artistique (compétence technolittéraire).

Documentation de modules pour soutenir l'édition numérique jeunesse

- 14 La section « Édition » de Lab·yrinthe est une émanation directe de la collecte et de l'analyse des données réalisées dans le cadre de la recherche-action. Elle est issue, plus spécifiquement, des rapports d'avancement produits pour l'organisme financeur entre 2017 et 2018, ainsi que des résultats obtenus lors d'une enquête supplémentaire, menée en 2021, dans le but de documenter les évolutions provoquées par l'augmentation des activités numériques durant la période de crise sanitaire sur les pratiques de conception, diffusion et réception des contenus numériques. Destinée de manière prioritaire aux éditeur·rice·s et aux créateur·rice·s numériques, cette section se propose de répondre à certaines de leurs préoccupations et difficultés, tout en leur fournissant des recommandations et des pistes de résolution sur les enjeux identifiés lors de l'analyse des données. En effet, dans le cadre de la recherche, les éditeur·rice·s et les créateur·rice·s interrogé·es avaient manifesté le besoin d'acquérir une connaissance plus profonde sur différents aspects liés à la culture numérique, et notamment sur : les métiers et les savoir-faire technologiques nécessaires à l'aboutissement d'un projet éditorial numérique ; sur les possibilités de financement ; sur les voies de découvrabilité actuellement disponibles ; sur les acquis de la recherche ; sur les pratiques scolaires ; sur les comportements et les pratiques de lecture des jeunes, en contexte numérique comme analogique. En partant de ce bassin de besoins, qui constituent autant d'enjeux pour le développement du milieu éditorial numérique jeunesse québécois, neuf modules d'analyse et d'accompagnement ont été élaborés : métiers ; accès aux publics ; technique ; marché ; lecteurs ; scolaire ; droit ; recherche ; covid. Chacun d'entre eux est organisé en trois parties. Après un rappel des citations clés des acteur·rice·s interrogé·e·s dans le cadre de la recherche, chaque fiche décrit de manière détaillée les principaux aspects qu'ils-elles avaient identifiés comme étant problématiques pour la conception, la diffusion ou la distribution des œuvres numériques. S'ensuit une analyse des dynamiques et des facteurs susceptibles d'influencer,

voire de déterminer les difficultés constatées, puis, en fin de module, une liste de recommandations et conseils pratiques visant à faciliter la résolution des enjeux.

- 15 Dans le module « Métiers », par exemple, après avoir illustré les défis rencontrés par les éditeur·rice·s dans la constitution des équipes de travail, relevant entre autres de la complexification du processus éditorial et de la démultiplication des expertises professionnelles nécessaires pour une production numérique, la fiche s'attache à l'analyse de quelques aspects clés, émergés des entretiens et des observations des processus de production : les contraintes temporelles associées à la conception d'une œuvre numérique ; les rôles spécifiques associés aux pôles de conception, création, développement technique et éducatif, qui sont souvent nécessaires pour l'aboutissement d'un projet numérique ; leurs interactions ; les missions des « passeur·se·s numériques », soit de figures professionnelles capables de faciliter la communication entre les différents pôles de création¹⁶. Enfin, dans la dernière partie du module, sont présentées cinq recommandations et conseils pratiques susceptibles de contribuer à une meilleure gestion des enjeux liés aux professionnalités du milieu éditorial numérique. Dans ce cas, Prune Lieutier, autrice de la fiche, suggère aux éditeur·rice·s et aux concepteur·rice·s numériques de :

- Prévoir, dès la préconception du projet éditorial, l'intervention de professionnel·le·s issu·e·s des différents pôles, afin de s'assurer de faire des choix stratégiques, créatifs et commerciaux adéquats par rapport aux intentions du projet ;
- S'assurer de disposer de ressources clés (les « passeur·se·s numériques »), pour rendre plus fluide la communication entre les équipes et assurer la fluidité de la production ;
- Intégrer des experts de l'éducation (conseiller·ère·s pédagogiques, chercheur·se·s, enseignant·e·s), afin de valider, non seulement la pertinence de l'expérience proposée pour les jeunes publics visés, mais aussi sa facilité d'appropriation pour les publics scolaires ;
- Privilégier le travail avec des professionnel·le·s (auteur·rice·s, illustrateur·rice·s, musicien·ne·s, etc.) habitué·e·s à travailler pour des productions numériques, ou intéressé·e·s à l'expérimentation, pour prévenir d'éventuelles difficultés liées à la reconfiguration des méthodes de travail en contexte numérique ;

- Envisager des échéanciers souples, afin de prévoir un temps suffisant à la consultation des différents pôles et à l'ajustement de la production¹⁷.

Figure 10. Les observations issues de la recherche sur les nouvelles configurations des métiers de la production de la littérature numérique jeunesse

Les observations

En production numérique, les rôles et compétences nécessaires au développement de projets sont multiples et éclectiques. Il est possible de rapidement passer et, ce, en fonction de la complexité de l'objet produit, à une équipe de plusieurs dizaines de personnes aux expertises diverses, configuration qu'il peut être difficile d'envisager, de comprendre, et de gérer pour des éditeur.trices issus de l'édition traditionnelle dans le cadre de la production de livres interactifs numériques dépassant le simple livre homothétique.

La création de ce type d'œuvres s'accompagne ainsi de la mise en place tâtonnante, et encore grandement itérative, de nouvelles configurations de travail, individuelles comme collaboratives.

Ces nouvelles configurations sont observables dans les différents pôles composant une production de ce type, qu'il s'agisse du travail de gestion (éditorial, financement, direction de projet, etc.), artistique (textes, images, musique, effets sonores, narration), de conception (ergonomie, architecture narrative), technique (développement, programmation, tests, etc.) ou encore pédagogique (conseil pédagogique, recherche).



Les différents acteur.trices impliqués commencent peu à peu à identifier les composantes et les limites de leurs rôles respectifs, apprenant à travailler ensemble dans des dynamiques qu'ils n'avaient souvent que peu ou jamais eu l'occasion d'explorer. Nous avons ici sciemment choisi de rassembler l'ensemble de ces acteur.trices sous le terme de créateur.trices, car nos observations de terrain nous ont permis de constater que, du fait du caractère novateur de ces créations mais aussi de la nature polyvalente et agile des équipes convoquées, dans la très grande majorité des cas, ils.elles ont un impact créatif sur le résultat final.

Figure 11. Les conseils clés pour guider les éditeurs pour le module sur les métiers de la production de littérature numérique jeunesse

Les 5 conseils clés



© Lab-yrinthe

Documentation de projets de recherche en collaboration avec les milieux éducatif, culturel et communautaire

- 16 La section « Projets » du site vise à présenter les principaux projets, réalisés en partenariat avec plusieurs partenaires (Le sentier du conte vivant, PoésieGo, La Pastèque, Ohé – Affiches augmentées, etc.). Chaque fiche rappelle les objectifs des initiatives éditoriales du/de la partenaire et les attentes concernant la collaboration avec l'équipe de recherche.



Projets



Ohé! – Affiches amplifiées

Le projet « Ohé! – Affiches amplifiées » vise à la création d'un dispositif expérimental d'« affiches sonores augmentées » liant papier et narration audio par une mécanique de reconnaissance vocale accessible via une application [...]



Parcours PoésieGol

PoésieGol propose une anthologie de balados, dont la plus grande part met en valeur la poésie québécoise contemporaine. La création sonore est de Magnéto. Depuis 2017, PoésieGol a été associé à la baladodiffusion dans des sites originaux, tel que le métro de Montréal ou [...]



La Pastèque – Projet balado pilote : L'abeille à miel

Les éditions de La Pastèque et La puce à l'oreille se sont associées pour développer une offre de balados documentaires complémentaires aux albums jeunesse imprimés de La Pastèque, soutenant la promotion, la découvrabilité et la médiation des albums. Un premier pilote a été réalisé [...]

© Lab·yrinthe

- 17 Par ailleurs, dans l'optique de répondre au deuxième objectif de la recherche (« coconstruire des outils de référence adaptés à analyser et résoudre les problématiques spécifiques rencontrées par les éditeur·rice·s et les éducateur·rice·s dans la conception, la diffusion et la diffusion d'œuvres numériques »), cette section décrit les démarches, les actions et les solutions effectivement entreprises par le(s) chercheur·se(s) responsable(s) pour soutenir le développement, la production, la diffusion ou la médiation auprès des enseignant·es, des professionnel·le·s du livre ou des jeunes lecteur·rice·s.

Documentation des actualités sur l'éducation et l'édition de littérature numérique jeunesse

- 18 La dernière section de Lab·yrinthe documente les actualités concernant la production numérique. Elle s'articule en quatre sous-sections. La première présente les « Événements » et les manifestations (scientifiques, culturelles, éditoriales, etc.) en lien avec le projet. La deuxième, dédiée aux « Parutions », recense des ouvrages et des publications scientifiques en lien avec la recherche ou susceptibles d'intéresser les publics du site. La troisième propose une « Veille » sur les événements marquants pour le milieu éditorial, culturel et éducatif numérique (sorties de nouvelles créations

numériques, publication d'appels à financement, manifestations scientifiques en lien avec la création numérique, etc.). Toujours visibles, même après la date de réalisation, les publications de cet espace, ainsi que celles recensées dans la section « Évènements », constitueront, à terme, un espace d'archivage des initiatives ayant contribué à stimuler le milieu scientifique, éditorial, culturel et éducatif numérique. Enfin, la dernière section, intitulée « Billets », contient de courts articles rédigés par l'équipe éditoriale ou par des auteur·rice·s invité·e·s sur des sujets en lien avec les thématiques du site.

Conclusion

- 19 Lab·yrinthe est le résultat d'un projet ambitieux et de la volonté, partagée par tou·te·s les chercheur·e·s de l'Équipe de recherche en littératie médiatique multimodale, de créer un écosystème favorable à l'observation et à la circulation de nouveaux corpus et de nouvelles pratiques littéraires et documentaires numériques. La mise à jour constante des contenus vise justement à suivre les évolutions des dynamiques de production, de distribution et de réception de la littérature numérique pour les jeunes publics, en faisant état des spécificités culturelles du marché québécois et, en même temps, des pratiques scolaires émergentes. Sans renoncer à son caractère ambitieux, Lab·yrinthe aspire également à devenir un laboratoire des pratiques de lecture et d'écriture. À terme, le site sera la vitrine d'une équipe pluridisciplinaire d'artistes et de développeur·se·s numériques, et décrira leur travail, leurs processus créatifs et leurs choix artistiques. En cultivant sa vocation pédagogique, il continuera également à recueillir et à analyser les pratiques scolaires, à documenter des projets de cocréation menés avec les enseignant·e·s, les élèves et les étudiant·e·s ainsi qu'à développer des contenus théoriques permettant de penser de nouveaux modèles de lecture et d'écriture.

BIBLIOGRAPHIE

AARSETH Espen, *Cybertext. Perspectives on Ergodic Literature*. Baltimore, Johns

Hopkins University Press, 1997.

ACERRA Eleonora et LACELLE Nathalie, *Compétences en #LMM, Lab·yrinthe* [en ligne], 2022, <https://lab-yrinthe.ca/education/competences-lmm> [consulté le 8 avril 2024].

ACERRA Eleonora, LACELLE Nathalie, MOLINA FERNÁNDEZ Margarita et VALLIÈRES Amélie, « Paramètres descriptifs et figures de style des œuvres littéraires numériques pour la jeunesse », *La lettre de l'Association internationale des chercheurs en didactique du français*, n° 68, 2021, p. 18-22.

ARCHIBALD Samuel et GERVAIS Bertrand, « Le récit en jeu », *Protée* [en ligne], n° 34(2-3), 2006, p. 31, DOI : [10.7202/014263ar](https://doi.org/10.7202/014263ar).

BOUCHARDON Serge, *La valeur heuristique de la littérature numérique*, Paris, Hermann, 2014.

CARROLL John M., MACK Robert L. et KELLOGG Wendy A., « Interface Metaphors and User Interface Design », in : Martin Helander (dir.), *Handbook of Human-Computer Interaction* [en ligne], Amsterdam, North-Holland, 1988, p. 67-85, DOI : [10.1016/B978-0-444-70536-5.50008-7](https://doi.org/10.1016/B978-0-444-70536-5.50008-7).

HUEBNER Emma June, « Fiche projet sur les stories à l'école », *Lab·yrinthe* [en ligne], 2022, <https://lab-yrinthe.ca/education/les-mondes-de-demain> [consulté le 8 avril 2024].

LACELLE Nathalie, LIEUTIER Prune et ACERRA Eleonora, *Rapport du projet de recherche « Soutien au développement de démarches d'édition numérique jeunesse au Québec à partir de pratiques favorables de production, diffusion et réception l'édition numérique jeunesse »* [en ligne] (FRQSC-MCC 2017-2021), 2021, URL : <https://frq.gouv.qc.ca/histoire-et-rapport/soutien-au-developpement-de-demarches-dedition-numerique-jeunesse-au-quebec-a-partir-de-pratiques-favorables-de-production-diffusion-et-reception/> [consulté le 8 avril 2024].

LACELLE Nathalie, LIEUTIER Prune et ACERRA Eleonora, « Développement de l'édition numérique jeunesse. Enjeux des pratiques au Québec. Le livre en contexte numérique. Un défi de design » [en ligne], *Écrire, éditer, lire à l'ère numérique*, 2010, Montréal, CRILQ, coll. « Arcanes », 2021, URL : <http://livre-defi-design.arcanes.ca/developpement-edition-numerique-jeunesse-pratiques-au-quebec/> [consulté le 8 avril 2024].

LIEUTIER Prune, « Accompagnement d'équipes, appropriation du numérique, distribution et découvrabilité : analyse des enjeux d'édition dans la production et la mise en marché de livres-applications jeunesse au Québec », thèse de doctorat, université du Québec à Montréal, 2022.

LIEUTIER Prune, « Fiche Édition sur les enjeux liés aux métiers de l'édition numérique jeunesse », *Lab·yrinthe* [en ligne], 2020, <https://lab-yrinthe.ca/edition/metiers/> [consulté le 8 avril 2024].

MOLINA FERNÁNDEZ Margarita, « La bande dessinée numérique comme expérience multimodale et interactive : de la conception à la réception », thèse de doctorat, université du Québec à Montréal, 2022.

MOLINA FERNÁNDEZ Margarita, « Fiche œuvre sur la bande dessinée numérique 2 », *Lab-yrinthe* [en ligne], 2021, URL : <https://lab-yrinthe.ca/oeuvres/2> [consulté le 8 avril 2024].

SAEMMER Alexandra, « Les figures d'animation face aux conventions du discours numérique », *Protée* [en ligne], n° 39(1), 2011, p. 23-36, DOI : [10.7202/1006724ar](https://doi.org/10.7202/1006724ar).

NOTES

1 Nathalie Lacelle, Prune Lieutier et Eleonora Acerra, « Soutien au développement de démarches d'édition numérique jeunesse au Québec à partir de pratiques favorables de production, diffusion et réception l'édition numérique jeunesse » [en ligne] (FRQSC-MCC 2017-2021), *Fonds de recherche du Québec* [en ligne], 2021, URL : <https://frq.gouv.qc.ca/histoire-et-rapport/soutien-au-developpement-de-demarches-dedition-numerique-jeunesse-au-quebec-a-partir-de-pratiques-favorables-de-production-diffusion-et-reception/> [consulté le 8 avril 2024].

2 Pour une analyse des points de rupture identifiés dans le cadre de la recherche, voir Nathalie Lacelle, Prune Lieutier et Eleonora Acerra, « Développement de l'édition numérique jeunesse. Enjeux des pratiques au Québec. Le livre en contexte numérique. Un défi de design », ECRIDIL 2010. *Écrire, éditer, lire à l'ère numérique*, Montréal, CRILQ, coll. « Arcanes », [en ligne], 2021, URL : <http://livre-defi-design.arcanes.ca/developpement-edition-numerique-jeunesse-pratiques-au-quebec/> [consulté le 8 avril 2024].

3 Eleonora Acerra, Nathalie Lacelle, Margarita Molina Fernández et Amélie Vallières, « Paramètres descriptifs et figures de style des œuvres littéraires numériques pour la jeunesse », *La lettre de l'Association internationale des chercheurs en didactique du français*, n° 68, AIRDF, 2021, p. 18-22.

4 Serge Bouchardon, *La valeur heuristique de la littérature numérique*, Paris, Hermann, 2014.

5 Espen Aarseth, *Cybertext. Perspectives on Ergodic Literature*, Baltimore, Johns Hopkins University Press, 1997.

6 *Ibid.*, p. 1.

- 7 Samuel Archibald et Bertrand Gervais, « Le récit en jeu », *Protée* [en ligne], n° 34(2-3), 2006, p. 31, DOI : [10.7202/014263ar](https://doi.org/10.7202/014263ar).
- 8 John M. Carroll, Robert L. Mack et Wendy A. Kellogg, « Interface Metaphors and User Interface Design », in : Martin Helander (dir.), *Handbook of Human-Computer Interaction* [en ligne], Amsterdam, North-Holland, 1988, p. 67-85, DOI : <https://doi.org/10.1016/B978-0-444-70536-5.50008-7>.
- 9 Alexandra Saemmer, « Les figures d'animation face aux conventions du discours numérique », *Protée* [en ligne], n° 39(1), 2011, p. 23-36, DOI : [10.7202/1006724ar](https://doi.org/10.7202/1006724ar).
- 10 Margarita Molina Fernández, « Fiche œuvre sur la bande dessinée numérique 2 », *Lab·yrinthe* [en ligne], 2021, URL : <https://lab-yrinthe.ca/oeuvres/2> [consulté le 8 avril 2024]. Pour une analyse détaillée de l'œuvre voir également Margarita Molina Fernández, « La bande dessinée numérique comme expérience multimodale et interactive : de la conception à la réception », thèse de doctorat, Montréal, université du Québec, 2022.
- 11 Margarita Molina Fernández, « Fiche sur l'œuvre 2 », *Lab·yrinthe* [en ligne], 2021, URL : <https://lab-yrinthe.ca/oeuvres/2> [consulté le 8 avril 2024].
- 12 *Loc. cit.*
- 13 Eleonora Acerra et Nathalie Lacelle, *Compétences en #LMM*, *Lab·yrinthe* [en ligne], 2022, URL : <https://lab-yrinthe.ca/education/competences-lmm> [consulté le 8 avril 2024].
- 14 Compétences (méta)techniques, (méta)gestuelles, (méta)sémiotiques, (méta)technolittéraires et (méta)technodocumentaires.
- 15 Emma June Huebner, « Fiche projet sur les stories à l'école », *Lab·yrinthe* [en ligne], 2022, URL : <https://lab-yrinthe.ca/education/les-mondes-de-de-main/> [consulté le 8 avril 2024].
- 16 Prune Lieutier, « Accompagnement d'équipes, appropriation du numérique, distribution et découvrabilité : analyse des enjeux d'édition dans la production et la mise en marché de livres-applications jeunesse au Québec », thèse de doctorat, université du Québec, Montréal, 2022.
- 17 Prune Lieutier, « Fiche Édition sur les enjeux liés aux métiers de l'édition numérique jeunesse », *Lab·yrinthe*, 2020, URL : <https://lab-yrinthe.ca/editio n/metiers/>

RÉSUMÉS

Français

Cet article présente le contexte de diffusion des résultats d'une recherche-action sur le soutien à l'édition numérique de jeunesse au Québec à travers la création d'un laboratoire dont le but est de documenter l'évolution des pratiques de création d'œuvres numériques et de leur enseignement. Lab-yrinthe.ca permet d'éduquer les publics aux nouvelles formes de littérature jeunesse et de favoriser leur découvrabilité à travers la documentation d'œuvres, d'outils, d'activités et de projets de recherche.

English

This article presents the context for disseminating the results of an action-research project to support the development of digital children's literature in Quebec, through the creation of a laboratory whose aim is to document the evolution of practices for creating and teaching digital works. Lab-yrinthe.ca educates audiences about new forms of children's literature, and promotes their discoverability through documentation of works, tools, activities, and research projects.

INDEX

Mots-clés

littérature numérique de jeunesse, éducation des publics, édition numérique, documentation de recherches

Keywords

digital children's literature, public education, digital publishing, research documentation

AUTEURS

Nathalie Lacelle

Professeure UQAM

IDREF : <https://www.idref.fr/166935247>

ORCID : <http://orcid.org/0000-0002-4759-9677>

ISNI : <http://www.isni.org/0000000375863700>

Eleonora Acerra

Professeure UQAM

IDREF : <https://www.idref.fr/250382466>

Constituer un patrimoine littéraire numérique du Québec : méthodologie et défis (projet COLiN, Catalogue des œuvres littéraires numériques)

Establishing a digital literature patrimony of Quebec: methodology and challenges (the COLiN project, Catalog of digital literature works)

René Audet

DOI : 10.35562/marge.933

Droits d'auteur

CC BY-NC-SA

PLAN

Un contexte de recherche attentif aux impacts du numérique
Vers un chantier d'inventaire des œuvres
Choix bibliographiques et technologiques
Utilisation de COLiN
Suites et perspectives

TEXTE

L'idée d'établir un corpus (numérique)



Catalogue des œuvres littéraires numériques (COLiN) - Construire le corpus du Québec

Pour élaborer une vision historique des pratiques en arts littéraires avec une forte dimension numérique, l'équipe du [pôle Québec](#) du projet de recherche en partenariat [Littérature québécoise mobile](#) et du Laboratoire Ex situ dresse un inventaire des œuvres littéraires numériques québécoises.



© COLiN

- 1 Pour les chercheuses et chercheurs en arts et lettres, l'étude des œuvres constitue une pierre d'assise de leur connaissance de ces pratiques. L'examen du corpus, l'évolution de ses traits, formes et

paramètres, la mise en perspective d'une pratique culturelle en regard d'autres : ce sont là les fondements mêmes du savoir sur les arts et les lettres. Le domaine littéraire ne fait pas exception ; on y accorde même une valeur particulièrement grande à toutes les œuvres qui composent le fil de son histoire. Inscrites dans un temps long, les littératures nationales importantes et hégémoniques (française, américaine, britannique, italienne, etc.) perdent dans l'épaisseur diachronique les origines de ce travail d'identification des réalisations à rattacher à leur culture propre. Les traces de ces démarches d'inventaire du corpus se sont souvent effacées, tellement elles structurent profondément l'institution littéraire elle-même.

- 2 Ce travail est cependant beaucoup plus prégnant pour de plus jeunes cultures, souvent des cultures plus resserrées (comme l'est notamment la culture québécoise). Il l'est également pour des pratiques ciblées – les œuvres rattachées à un nouveau genre (pensons à l'autofiction) ou à une nouvelle pratique artistique (les jeux vidéo ou le net art, par exemple). Les secteurs d'étude de ces pratiques, eux aussi naissants, doivent s'appuyer sur la réalité du terrain. Et la démarche prioritaire (car préalable à une reconnaissance effective, de large portée) y est d'identifier les œuvres, le corpus à étudier.
- 3 Ainsi en est-il, un peu partout à travers le monde, pour la culture numérique et, plus particulièrement, pour la littérature numérique. Le parallèle entre les livres et les œuvres numériques s'y fait assez naturellement, d'autant qu'il y a parfois continuité entre une œuvre papier et une œuvre déployée sur écran. Le monde du livre, lui, est bien sûr particulièrement structuré : désignation des éditions et versions par des numéros ISBN, flux de données accessibles pour les libraires et les bibliothécaires, capture de la production éditoriale par le dépôt légal. Il est ainsi possible d'extraire de ces outils et modalités référentielles un état de la production littéraire d'un territoire donné. Ne pourrait-on pas transposer ce simple exercice aux œuvres de littérature numérique ? Une telle saisie par des instances vouées à la circulation et à la médiation n'existe pas en culture numérique (ou, considération connexe : les circuits actuels du livre n'incluent pas la production numérique). Dans les circonstances, comment repérer le corpus des œuvres littéraires numériques ? Il n'y a pas de réponse simple à cette question¹.

Un contexte de recherche attentif aux impacts du numérique

- 4 Cette question est collectivement abordée depuis plusieurs années – que ce soit dans des unités de recherche (NT2, ELCIP, Electronic Literature Organization, etc.) ou dans des bibliothèques nationales. Elle l'est toutefois, de façon spécifique, depuis 2019, s'inscrivant dans la problématique au cœur des travaux de l'équipe de Littérature québécoise mobile². Cet ambitieux projet de recherche en partenariat rassemble une quinzaine de chercheuses et chercheurs du Québec et de la France, ainsi qu'une vingtaine d'organismes partenaires liés au monde du livre et de la littérature. Ce projet, qui s'intéresse aux liens entre culture du livre et culture de l'écran, vise à développer de façon collaborative avec ces organismes des savoirs sur l'évolution récente des pratiques littéraires en contexte numérique, particulièrement au Québec.

Figure 1. Page d'accueil du site du projet de recherche « Littérature québécoise mobile »



- 5 L'intérêt que nous portons, comme spécialistes de littérature, au contexte numérique prend un double visage. Le premier volet s'intéresse au numérique comme adjuvant pour la diffusion de la littérature québécoise. Si la question n'est pas tout à fait nouvelle, elle

se construit toutefois en fonction des évolutions technologiques imposées par les grandes plates-formes numériques existantes, évolutions qu'il faut suivre et auxquelles s'adapter constamment. Comment les actrices, les acteurs du milieu littéraire peuvent-ils et elles utiliser les outils numériques ainsi que les réseaux qui leur sont liés pour promouvoir la littérature québécoise, pour en assurer une médiation à travers diverses plates-formes ? C'est là tout l'enjeu de la découvrabilité³ des productions culturelles, et de leur visibilité, qui représente non seulement une notion chère aux *petites cultures* étouffées par les cultures hégémoniques, mais aussi une question préoccupant les institutions culturelles et tous les acteurs et actrices de l'écosystème du livre, les agentes et agents des milieux éditoriaux et autres figures intervenant dans l'écosystème du livre – dont les auteurs et les autrices. Ensuite, le deuxième volet concerne l'étude d'une large production littéraire, née numérique, où les outils technologiques participent étroitement du geste de création littéraire. On observe un développement graduel d'une conscience de la singularité de ces pratiques, rendue possible de diverses façons, directes et indirectes : augmentation du nombre d'œuvres produites ; reconnaissance croissante de ces productions par des instances institutionnelles ; présence de figures littéraires attestées dans ce nouveau secteur de création ; accroissement d'une littératie numérique auprès des acteurs et actrices du milieu littéraire, démocratisant, voire banalisant le recours à l'un ou l'autre outil de création ou de publication. C'est probablement cette banalisation qui devient l'indice le plus probant d'une pénétration croissante du numérique dans les pratiques actuelles de la littérature – une appropriation du numérique que nous nous affairons, au sein du projet Littérature québécoise mobile, à consolider avec différents partenaires des milieux littéraires québécois.

Vers un chantier d'inventaire des œuvres

- 6 C'est dans ce contexte que s'est développé le chantier de recherche PLiNe⁴, consacré au Patrimoine littéraire numérique du Québec. Il répond à deux enjeux immédiats de la littérature numérique : la connaissance des œuvres et leur

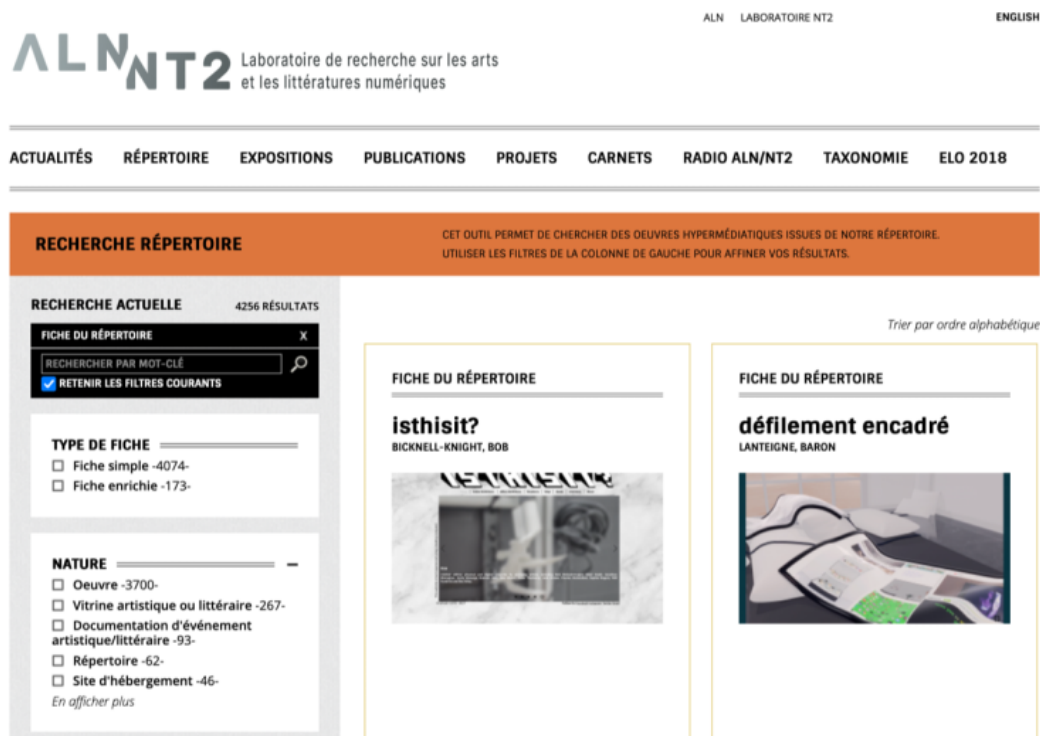
pérennisation éventuelle⁵. Il est indéniable qu'un large intérêt s'est développé, avec les années, pour les pratiques culturelles numériques. Elles sont de plus en plus étudiées, faisant souvent l'objet de perspectives poétiques et sémiotiques puis, peu à peu, de considérations comparatistes. Leur approche est toutefois très souvent détachée de leur historicité. On pourra signaler quelques exceptions, comme le secteur des jeux vidéo (dont on étudie passablement les premiers temps⁶). On trouve aussi quelques initiatives en littérature numérique, sous la forme d'anthologies (les collections produites par l'ELO⁷), d'incursions ponctuelles (voir l'ouvrage récent d'Astrid Ensslin⁸) ou à travers les travaux de Dene Grigar, à Washington State University⁹. Dans l'état actuel des choses, est-il possible d'imaginer se lancer dans le chantier d'une histoire des pratiques littéraires numériques, quel que soit le corpus national ? Pour le moment, il faudrait répondre par la négative, les données accumulées étant trop limitées pour espérer réaliser un tel projet. Mais le travail sur un premier échantillon d'envergure limitée paraît une voie réaliste pour entamer un premier chantier de cette nature.

7 À cet égard, le corpus de la littérature numérique au Québec offre des singularités et des avantages intéressants. Cette enclave culturelle entre les États-Unis et la France ressent les effets des productions culturelles voisines aussi bien que des avancées technologiques incessantes. S'y déploie également un nombre limité d'institutions ; la langue française y domine largement, marquant une frontière avec le reste du Canada et les États-Unis. À ces marques de singularité s'ajoute le sentiment d'un laboratoire culturel, en ce qu'il est doté de politiques spécifiques et de juridictions singulières¹⁰ ; l'offre d'appuis à la culture y est assez bien développée et en phase avec les secteurs d'innovation.

8 Depuis quelques années, nous avons ainsi entamé une collecte informelle d'œuvres littéraires numériques du Québec. Le projet était d'abord déstructuré, ayant des visées incertaines : les œuvres numériques faisaient l'objet d'un repérage plutôt occasionnel, au gré des découvertes ; la délimitation nationale n'était pas imposée – même si, associés-es à une culture non dominante, nous avons porté notre attention particulièrement sur les œuvres francophones. Nos premiers efforts s'inscrivaient dans le sillon d'une démarche connexe plus ambitieuse, menée depuis le milieu des années 2000 : le

répertoire du laboratoire NT2, à l'UQAM, s'intéressant à la littérature et aux arts hypermédias¹¹. Ce chantier avait pu recenser plus de 4000 œuvres rattachées à diverses cultures nationales, autant dans la sphère littéraire que celle des arts numériques et du net art.

Figure 2. Page du répertoire des œuvres hypermédias du laboratoire NT2



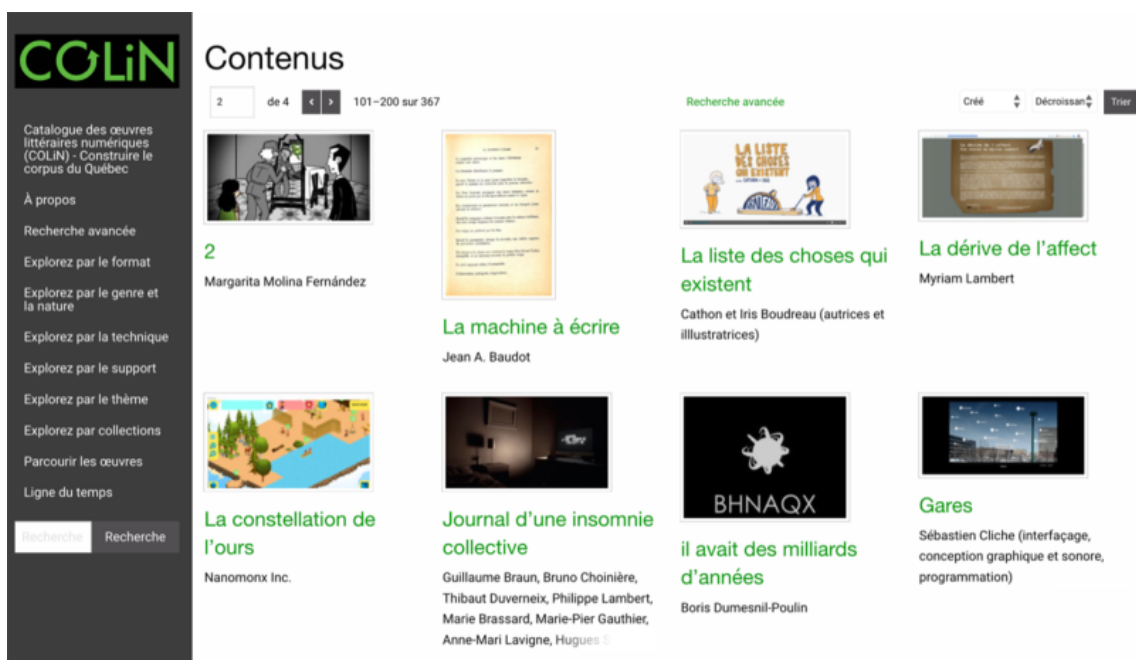
COLiN : le catalogue des œuvres littéraires numériques (Québec)

© ALN/NT2

- 9 La perspective critique et historique des productions numériques au Québec s'imposant dans le cadre de Littérature québécoise mobile, ces efforts ont été recadrés pour constituer un chantier doté d'une vision plus nette du travail à réaliser. L'état des lieux a été somme toute aisé à constituer : des initiatives panoramiques et internationales (comme le répertoire du NT2 ou celui de l'Electronic Literature Organization) ; des œuvres circulant dans le discours critique (souvent les mêmes, souvent d'une période similaire, à savoir un créneau de moins de 10 ans, moment où les œuvres marquantes ont su s'imposer dans le discours) ; des interlocuteurs institutionnels aux abonnés absents (par manque de littératie plutôt que par désintérêt). Les démarches de repérage au sein d'outils existants ont

eu une fortune bien limitée. Les œuvres littéraires numériques, qui ne sont pas éditées au sens habituel du terme ni des publications comme le sont les livres, ne sont pas dotées d'un ISBN et ne font pas l'objet d'un dépôt légal. Elles ne sont, pour ainsi dire, jamais indexées – ni par les institutions de mémoire comme les bibliothèques nationales ni de façon spontanée par l'une ou l'autre instance culturelle. La saisie historique du corpus se révèle en conséquence très difficile à élaborer, comme les œuvres elles-mêmes disparaissent souvent, ainsi que les ressources (souvent des sites) qui les mentionnent. Néanmoins, par la multiplication des sources documentaires et la consultation de nombreuses informatrices et informateurs, un inventaire assez large a pu être constitué, de sorte que près de 400 œuvres figurent aujourd'hui dans ce premier état du corpus : la banque de données COLiN, le catalogue des œuvres littéraires numériques au Québec ¹².

Figure 3. Liste des œuvres du site COLiN



© COLiN

10 On y trouve (et on y documente) une diversité de productions littéraires numériques, depuis les origines (un ouvrage atypique de 1964 ¹³, mais surtout des œuvres de la fin des années 1980 et du début des années 1990) jusqu'à aujourd'hui.

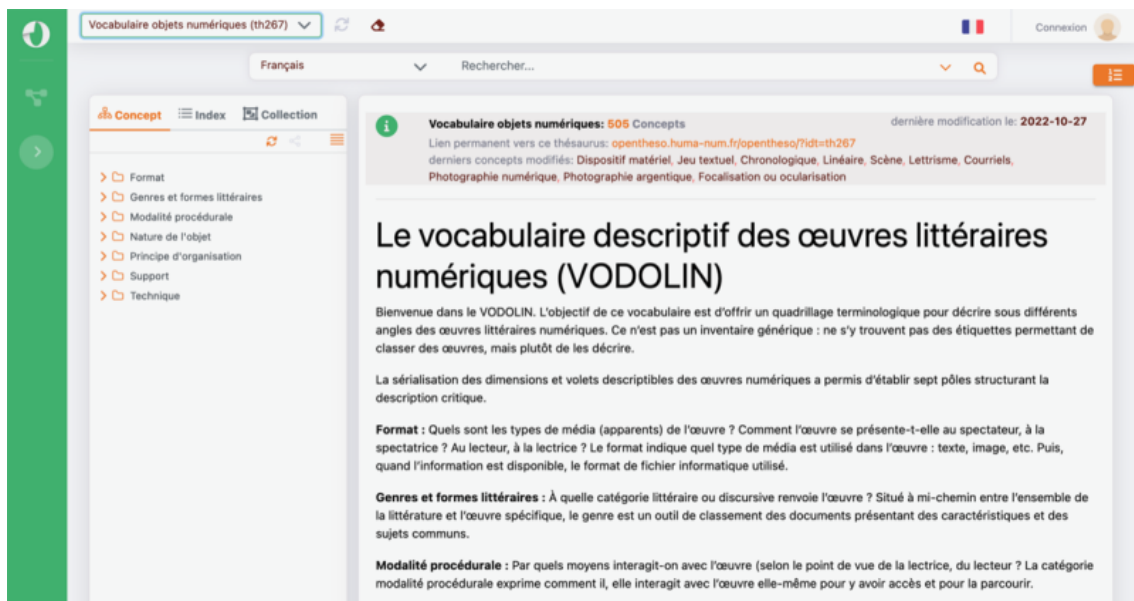
- 11 Les critères d'inclusion sont à la fois précis et ambigus à interpréter. Nous nous intéressons aux œuvres nées numériques, ce qui rejette d'emblée, d'une part, les livres numériques homothétiques (qui ne sont que des copies des livres papier) et, d'autre part, tous les produits dérivés de la littérature (captations d'événements, éléments numériques voués à la promotion ou à la médiation de livres).
Persistent des questions délicates : que faire des projets d'écriture en continu, souvent liés à une plateforme de réseau social, mais pouvant s'en détacher (pour migrer ou adopter un nouveau support) ? Doit-on exclure des projets colorés par des objectifs de médiation ? Notre intérêt se porte donc sur des écritures qui font œuvre en contexte numérique, dialoguant avec leur support technique et mobilisant ses capacités expressives. Ainsi, dans ce corpus, cohabitent des hyperfictions, des CD-ROM et logiciels, des vidéopoèmes, des jeux vidéo à haute teneur textuelle/littéraire, des applications avec réalité augmentée, des œuvres littéraires numériques cartographiques, des textes générés par ordinateur, des feuillets numériques, des livres numériques enrichis, des blogues et autres publications sérielles sur réseaux sociaux. La ligne d'inclusion n'est pas toujours facile à établir, même sur ce paramètre technologique. Quel rôle exiger du numérique dans la réalisation de l'œuvre ? Que penser de primofigures du web qui publiaient des poèmes sur le site web qu'ils et elles s'étaient bricolé ? Quel statut donner aux pratiques d'écriture mobilisant un générateur de texte, mais publiant les résultats sur papier ? Doit-on retenir les œuvres ou projets qui ont une dimension littéraire, mais qui n'est pas dominante ? C'est dire à quel point ce terme « littéraire » est chargé, voire informé par le livre. Dans un milieu culturel flou, sans réelles institutions ni agents de prescription, sur quelle base faire le tri entre la simple écriture en ligne et l'appartenance au fait littéraire ? Faut-il imposer un critère de qualité (esthétique) ? La réponse à cette dernière question a rapidement été établie : ce ne peut être un critère d'inclusion, sous peine de distordre lourdement ce champ culturel. La difficulté est similaire (bien que moins subjective) pour l'appartenance nationale : quelles œuvres sont québécoises, quand on connaît l'organisation logistique du web (les serveurs étant généralement délocalisés) ? La publication dans une ressource étrangère prévaut-elle sur la citoyenneté ? Un·e artiste étranger·ère créant une œuvre lors d'un séjour au Québec se trouve-t-il·elle visé·e par cet inventaire ?

- 12 Aussi imparfaite soit-elle, la grille de sélection des œuvres pouvant figurer au projet COLiN repose sur les paramètres suivants :
- avoir une forte teneur *littéraire* (entendue dans le sens d'un usage non fonctionnaliste de la langue, dans son extension la plus grande) ;
 - avoir une composante numérique importante, voire dominante (mais ne pas être la simple version PDF ou ePub d'un livre papier) ;
 - au moins une personne de l'équipe de création ou d'édition doit provenir du Québec ou l'œuvre doit avoir été produite au Québec.

Choix bibliographiques et technologiques

- 13 Très rapidement, en raison de l'obsolescence technologique, nous avons pris conscience de l'urgence d'attester de l'existence de ces œuvres. À défaut de pouvoir les sauvegarder (c'est là tout un autre chantier, causant des maux de tête à toute personne qui en entrevoit l'ampleur et la complexité), nous avons pris le parti de documenter ces pratiques, quelle que soit l'accessibilité actuelle des œuvres. Au sein du COLiN, chacune d'entre elles est fichée selon une vingtaine de paramètres (conformes au Dublin Core, un format de description générique) s'appuyant sur le VODOLIN¹⁴, un vocabulaire descriptif élaboré conjointement avec le NT2 (sur la base de leur travail de la dernière décennie) et avec la Chaire de recherche du Canada sur les écritures numériques de Marcello Vitali-Rosati, également membre du projet Littérature québécoise mobile.

Figure 4. Page VODOLIN sur Opentheso



© Opentheso

- 14 Ce thésaurus permet de décrire des objets numériques à partir de sept perspectives (format, genres et formes littéraires ; modalité procédurale ; nature de l'objet ; principe d'organisation ; support ; technique). Il est rendu disponible sur Opentheso¹⁵, ressource technologique de thésaurus de la plateforme française Huma-num, ce qui le rend accessible, mais surtout moissonnable (consultable et utilisable) par d'autres outils numériques. Cette ontologie est fort utile pour arriver à une description normalisée associée aux propriétés du Dublin Core. Nous pouvons ainsi décrire, contextualiser et attester l'existence d'œuvres numériques, même si de celles-ci ne restent parfois que des bribes accessibles par la Wayback Machine ou, pire, des traces indirectes émanant de critiques, de commentaires ou d'inventaires. Chaque œuvre est ainsi fichée dans COLiN, une banque de données qui repose sur Omeka S¹⁶, une solution *open source* de collections numériques développée par le Roy Rosenzweig Center for History and New Media. En plus de la caractérisation proposée, ces fiches incluent le lien vers l'œuvre, mais aussi, dans la mesure du possible, des captures d'écran (voire des extraits vidéo et des navigations filmées) pour conserver un aperçu de l'esthétique et de la configuration générale de l'œuvre.

Figure 5. Fiche de l'œuvre « Lac-Mégantic. La dernière nuit » sur COLiN



- 15 Suivant l'idée de vouloir faire exister ce corpus et d'en faciliter l'étude et le référencement, le protocole de travail inclut de créer des entrées correspondantes dans Wikidata, qui est la couche de données structurant Wikipédia. Cette mise au commun du travail de repérage a pour double effet de rendre découvrables ces œuvres, leurs auteurs et autrices et les organisations impliquées (éditeurs, sociétés de production), et d'associer à chacune de ces œuvres un identifiant unique fourni par Wikidata, un substitut aux ISBN ou DOI couramment utilisés dans le monde du livre et de l'édition. L'ajout massif de plusieurs centaines de nouvelles entrées¹⁷ permet d'insérer ces œuvres dans la toile de données qui nous sert collectivement de référentiel.

Utilisation de COLiN

- 16 La banque de données constituée sur Omeka S permet différents modes de consultation associés aux caractéristiques principales des œuvres. Outre la navigation au hasard dans la liste des entrées, il est possible d'utiliser un outil de recherche avancée permettant de faire

des requêtes par auteur ou autrice, sujet, mot-clé ou date (de publication ou de disponibilité en ligne). Quelques filtres prédéfinis trient les œuvres selon des paramètres précis :

- format d'œuvre (ePub, HTML, modélisation 3D, etc.) ;
- genre et nature (bande dessinée, conte, documentaire, etc.) ;
- technique ;
- support (CD-ROM, liseuse, téléphone intelligent, etc.) ;
- thème.

- 17 Des collections sont également constituées pour représenter certains regroupements d'œuvres – par exemple, les œuvres diffusées par la revue de littérature hypermédiatique *bleuOrange*¹⁸. Enfin, une ligne du temps donne un aperçu de la production année après année.

Figure 6. Aperçu de la ligne du temps sur COLiN

Œuvres par année de parution



© COLiN

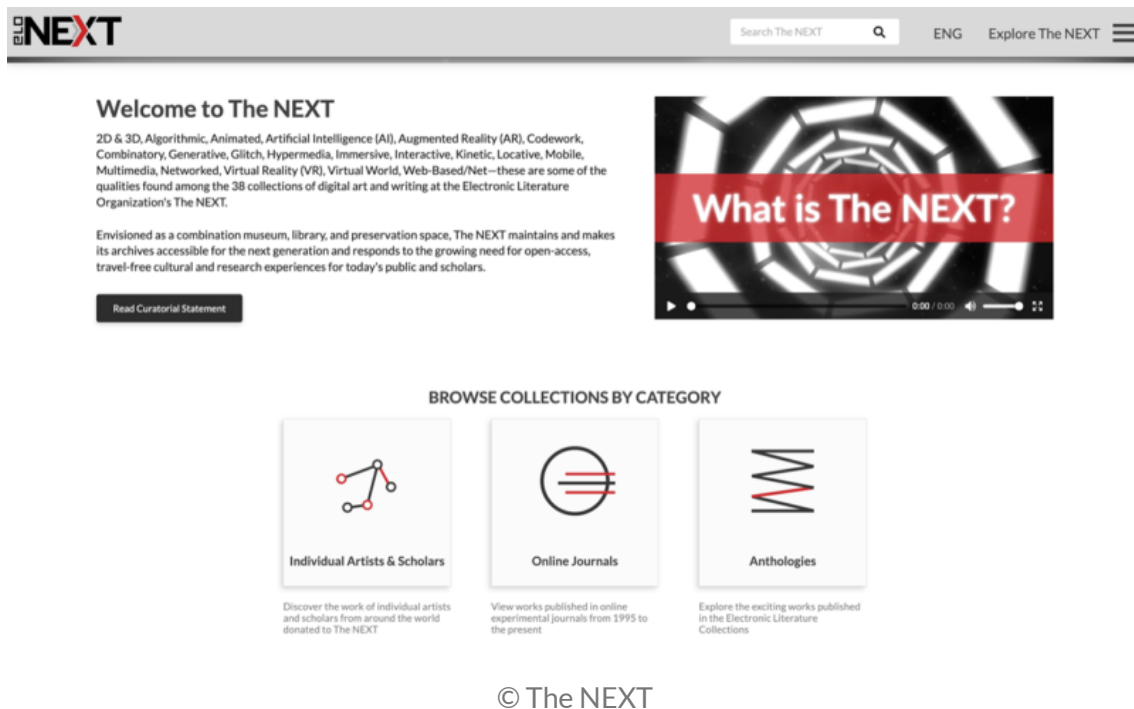
Suites et perspectives

- 18 Le travail d'inventaire se poursuit, de façon à mieux écumer différents répertoires, sites-index, listes d'œuvres constituées à différentes

époques. Une attention particulière sera portée à des sous-ensembles spécifiques : pensons à la période des blogues littéraires (2000-2010), aux supports analogiques de la période préweb (CD-ROM, par exemple) et aux pratiques littéraires très actuelles (qui restent plus difficiles à cerner). Et, ce faisant, les questions se multiplient :

- 19 Que faire des œuvres au long cours, qui se définissent par la figure auctoriale et non pas par l'identité précise d'une œuvre ? (un compte Instagram, par exemple, ou une pratique régulière de vidéopoésie sur une chaîne Youtube ; des comptes d'Instapoésie ont déjà été intégrés, à titre de mise à l'épreuve)
- 20 Comment établir le rattachement national d'œuvres multipartites (comme des collaborations internationales) ? Inclut-on des œuvres de l'Office national du film du Canada, coproduites avec des agences d'autres pays, alors qu'on ne sait pas si les Canadiens impliqués sont même Québécois ?
- 21 Quelle légitimité donner à l'indexation d'une œuvre dont on n'a plus aucune trace de l'existence (par obsolescence du logiciel ou retrait total du web) ?
- 22 C'est à la question fondamentale de la matérialité des œuvres numériques que nous nous butons immédiatement, à la persistance de ces œuvres et à notre capacité de les pérenniser. Plusieurs projets s'attaquent de front à cette question (autour de l'Electronic Literature Organization¹⁹, à la KBR en Belgique²⁰, notamment), mais cette volonté est lourdement entravée, autant du point de vue technique que dans son effort de saisir cette production éphémère aux frontières labiles.

Figure 7. Page d'accueil du projet The NEXT



- 23 Une hypothèse d'extension de ce large chantier s'intéressant au Patrimoine littéraire numérique serait d'arriver à fédérer diverses sources de ces œuvres, notamment les médiathèques d'organismes culturels (organismes de production et de diffusion), de sorte à offrir un accès transversal à leurs voûtes (à partir de descriptions standardisées), mais aussi de favoriser la prise en charge de ces contenus par des institutions garantissant leur mémoire – au premier lieu desquelles se trouvent les bibliothèques nationales. Le défi est toutefois de taille, entre enjeux techniques, normalisation des descriptions, questions de droit, mais aussi en raison de questionnements plus fondamentaux, comme le rattachement d'une œuvre à sa plateforme, à son contexte de publication ; la définition de son expérience de lecture selon l'appareil (téléphone portable, console de jeu, ordinateur de table, écran tactile...) ; la nécessité de documenter l'environnement culturel d'où émane l'œuvre ou dans lequel elle a circulé.
- 24 Jalon de la patrimonialisation des œuvres littéraires numériques, le projet COLiN servira de fondement aux travaux qui suivront sur l'histoire des pratiques littéraires et culturelles numériques au Québec. S'il balise le territoire exploré au fur et à mesure qu'il le

défriche, il se trouve à mettre à l'épreuve une certaine conception de la culture numérique et de ses corrélations avec les secteurs culturels plus établis. La poursuite de la transformation de ces rapports, avec une pénétration toujours plus grande du numérique dans la culture, offrira à coup sûr une perspective renouvelée sur la place, le rôle et la reconnaissance de ces pratiques numériques.

NOTES

- 1 Elle échappe très généralement au secteur critique le plus rapproché, les humanités numériques, lesquelles privilégient généralement les approches (numériques) à l'étude des corpus (numériques). La question semble souvent trop nichée pour le créneau de la bibliothéconomie et des sciences de l'information, alors que les études littéraires portent une attention toute réduite aux pratiques littéraires numériques.
- 2 Projet *Littérature québécoise mobile* [en ligne], URL : <https://lqm.uqam.ca> [consulté le 8 avril 2024] (financé par le Conseil de recherches en sciences humaines du Canada, 2019-2024 – octroi : 895-2019-1017). Le pôle Québec, que dirige l'auteur et qui est à l'initiative du projet COLiN, est associé au laboratoire Ex situ, à l'Université Laval : URL : <https://ex-situ.info> [en ligne].
- 3 Sur la notion de « découvrabilité » voir Jean-Robert Bisaillon, « Qu'est-ce que la découvrabilité ? », *Blog du LATICCE* [en ligne], 2018, URL : <https://wiki.uqam.ca/pages/viewpage.action?pageId=41074965> [consulté le 8 avril 2024], et « La découvrabilité des contenus numériques dans le Web », <https://culturelaval.ca/guide-de-la-decouvrabilite-et-des-bonnes-pratiques-sur-le-web/> [consulté le 8 avril 2024].
- 4 « Patrimoine littéraire numérique (PLiNe) », *Littérature québécoise mobile* [en ligne], URL : <https://lqm.uqam.ca/fr/projet/patrimoine-litteraire-numerique-pline> [consulté le 8 avril 2024].
- 5 Ce deuxième volet se situe bien différemment dans l'écosystème institutionnel, croisant des enjeux technologiques aussi bien que symboliques. Il ne sera que rapidement évoqué en conclusion.
- 6 Voir les travaux menés dans le cadre du LUDOV, Laboratoire universitaire de documentation et d'observation vidéoludiques [en ligne], URL : <https://ludov.ca> [consulté le 8 avril 2024].

- 7 Quatre anthologies d'œuvres numériques ont été proposées, de 2006 à 2022, toutes disponibles en ligne : *Electronic Literature Collection* [en ligne], URL : <https://collection.eliterature.org/> [consulté le 8 avril 2024].
- 8 Astrid Ensslin, *Pre-Web Digital Publishing and the Lore of Electronic Literature*, Cambridge, Cambridge University Press, coll. « Elements in Publishing and Book Culture », 2022.
- 9 Voir notamment les projets et publications répertoriés sur son site : <https://www.nouspace.net/#working-on> [consulté le 8 avril 2024].
- 10 L'exemple le plus probant est la configuration du marché du livre, au Québec, qui est fonction directe d'une loi singulière adoptée au tournant des années 1980, rendant solidaires les maillons de cette chaîne du livre et régissant la place des libraires sur le territoire (cf. Loi sur le développement des entreprises québécoises dans le domaine du livre).
- 11 Répertoire du laboratoire NT2 : <https://nt2.uqam.ca/> [consulté le 8 avril 2024].
- 12 Démarré vers 2016, le COLiN s'est développé à partir de 2019. Il existe grâce au travail d'une dizaine d'auxiliaires de recherche, de chercheuses et de chercheurs de l'Université Laval, de l'Université du Québec à Trois-Rivières, de l'université Concordia et de l'Université du Québec à Montréal, sous la direction scientifique de René Audet et la coordination de Marie-Ève Muller.
- 13 Il s'agit de *La machine à écrire* de Jean A. Baudot, un recueil de poèmes générés par ordinateur, qui inclut également les réactions critiques de plusieurs figures littéraires. Voir la fiche « La machine à écrire », COLiN [en ligne], URL : <https://colin.ex-situ.info/s/colin/item/2223> [consulté le 8 avril 2024].
- 14 Voir René Audet, « Un vocabulaire descriptif pour les œuvres littéraires numériques », *Ex situ* [en ligne], 2022, URL : <https://ex-situ.info/2022/10/un-vocabulaire-descriptif-pour-les-oeuvres-litteraires-numeriques-2/> [consulté le 8 avril 2024].
- 15 Voir *Opentheso* [en ligne], URL : <https://opentheso.huma-num.fr/opentheso/?idt=th267> [consulté le 8 avril 2024].
- 16 Voir *Omeka S* [en ligne], URL : <https://omeka.org/s/> [consulté le 8 avril 2024].
- 17 Un marqueur associé à ces entrées permet de consulter les œuvres référencées sous la forme d'une collection : *Wikidata* [en ligne], <https://www->

w.wikidata.org/w/index.php?title=Special:WhatLinksHere/Q111835637&limit=500 [consulté le 8 avril 2024].

18 Voir *bleuOrange* [en ligne], URL : <http://revuebleuorange.org/> [consulté le 8 avril 2024] et ELO Next [en ligne], URL : <https://the-next.eliterature.org/collections/29&pg=1&display=100&sort=title> [consulté le 8 avril 2024].

19 Voir The NEXT, projet de l'Electronic Literature Organization : *The Next* [en ligne], <https://the-next.eliterature.org/> [consulté le 8 avril 2024].

20 Voir le LabEL : Laboratory for Electronic Literature mis en place à la Bibliothèque royale de Belgique : KBR [en ligne], URL : <https://www.kbr.be/fr/projets/label/> [consulté le 8 avril 2024].

RÉSUMÉS

Français

Cet article présente l'arrière-plan méthodologique et pragmatique d'un projet d'inventaire des œuvres littéraires numériques rattachées au corpus québécois. Il aborde les enjeux de patrimonialisation d'œuvres qui n'ont pas été indexées et qui échappent autrement à une saisie institutionnelle. Le projet COLiN (catalogue des œuvres littéraires numériques – Québec) a permis d'identifier 400 œuvres, de les décrire et d'en offrir une pérennisation.

English

This article presents the methodological and pragmatic background of an inventory project of digital literary works attached to the Quebec corpus. It addresses the issues of patrimonialization of works that have not been indexed and that otherwise escape institutional capture. The COLiN project (catalogue des œuvres littéraires numériques – Québec) has made it possible to identify 400 works, describe them and make them available for future reference.

INDEX

Mots-clés

littérature numérique, littérature québécoise, corpus, banque de données, description normalisée, histoire littéraire

Keywords

digital literature, Quebec literature, corpus, database, standardized description, literary history

AUTEUR

René Audet

Université Laval, département de littérature, théâtre et cinéma

IDREF : <https://www.idref.fr/058582754>

ORCID : <http://orcid.org/0000-0002-8862-7304>

ISNI : <http://www.isni.org/0000000121398546>

BNF : <https://data.bnf.fr/fr/14434722>

Lifranum : Coconstruire un corpus et une archive de littérature nativement numérique

Lifranum: Co-building a corpus and a born-digital literature archive

Christian Cote

DOI : 10.35562/marge.831

Droits d'auteur
CC BY-NC-SA

PLAN

Introduction
Cadre de travail
Structuration générale du projet
Fondement social : la reconnaissance mutuelle
Méthode quotidienne de travail
Procédure d'identification de la littérature web
Méthodologie du *crawl*
Méthode inverse et constitution de l'archive
Analyse du corpus
Vers la structuration des données
Limites et perspectives
Conclusion

TEXTE

Introduction

- 1 La création web littéraire est une production massive de textes qui est le fait d'auteurs amateurs, mais liés entre eux par des réseaux d'interconnaissance. La reconnaissance mutuelle est un phénomène marquant de cette pratique littéraire qui se traduit par la participation à des ateliers, associations, etc., mais aussi par l'échange de liens hypertexte. Il s'agit d'une forme de sociabilité littéraire.
- 2 L'expression web, notamment du fait du média, renouvelle les pratiques d'écriture et les formes mêmes de la production littéraire. Notamment, le rapport entre le texte et l'image est modifié, parce que

l'image produit plus de commentaires et d'interactivité du fait de la pluralité des interprétations qu'elle propose.

- 3 Par ailleurs, la longueur des textes diminue avec le web et s'accompagne d'une plus forte fréquence de publication. Si les circuits de la publication littéraire écrite ne sont plus pertinents, c'est également la nature des projets littéraires (associés à un site ou un blog) et le processus de création ponctuelle d'un texte qui sont transformés.
- 4 Ainsi, la littérature web réintroduit une pratique amateur de la littérature, qui est aussi une pratique massive (et non plus le fait de quelques personnes) et cela renouvelle l'approche des phénomènes littéraires.
- 5 Toutes ces intuitions sont à l'origine de notre projet, qui vise d'abord à fournir un cadre d'étude ayant les dimensions quantitatives de cette production littéraire.
- 6 Dans cet article, nous présenterons la démarche que nous avons adoptée de façon à proposer un corpus de cette littérature qui réponde aux critères de la représentativité et qui enregistre le plus grand nombre d'auteurs possible. Pour nous, la constitution du corpus est un élément fondamental de la constitution d'un objet scientifique destiné à être appréhendé par des chercheurs portant des problématiques différentes. En ce sens, c'est la méthodologie de constitution du corpus qui constitue le fondement de la scientificité des approches qui suivront. Nous positionnons ce travail en considérant que le corpus réalisé constitue une entrée pour l'observation de phénomènes littéraires, linguistiques et sociologiques notamment.
- 7 Hormis certains présupposés que nous explicitons systématiquement, telle la dimension collective et interpersonnelle de la production littéraire, nous nous limitons à la caractérisation du corpus, laissant en cela les travaux ultérieurs l'explorer.
- 8 Nous présentons cette élaboration de corpus en trois temps : l'identification des productions littéraires à partir des phénomènes de reconnaissance mutuelle, la récupération des données, et l'analyse inverse, reposant sur une analyse systématique des liens, réalisée par un autre acteur du projet.

Cadre de travail

- 9 L'ensemble des travaux d'élaboration d'un corpus présentés ici est issu du projet Lifranum ANR 2019-2024¹. L'idée centrale est de constituer un ensemble de ressources scientifiques à partir de la production littéraire web : un corpus et une archive.
- 10 Au départ, c'est un projet mené par des littéraires cherchant à donner une visibilité à la littérature web. Des compétences en sciences de l'information se sont adjointes et dès lors l'idée de constituer un véritable corpus et une archive a émergé². Dans les répertoires existants, les œuvres sont enregistrées sur le principe de la déclaration, elles possèdent un statut d'œuvre à part entière, notamment parce qu'elles sont référencées comme telles sont isolées de leur contexte de production comme de réception et enfin sont classées en fonction de leurs caractéristiques physiques et intellectuelles. Dans notre projet, c'est nous-mêmes qui identifions les œuvres, qui les répertorions, les indexons et les mettons à disposition dans leur contexte social. Nous ne partons pas d'un modèle de type collection, qui discrétise les données, mais d'un modèle de linguistique de corpus, marqué par la continuité et les relations entre les productions. Nous privilégions donc le contexte sur la caractérisation de l'œuvre comme objet. Au-delà de la question de la mise en relation des œuvres dans le cadre du corpus, mise en relation permise par les liens hypertextuels, on rend possible l'étude du lien entre les productions littéraires et leurs cadres sociaux, interpersonnels, linguistiques notamment.
- 11 Des problématiques en informatique (bases de données et traitement automatique) se sont ajoutées, avec un double challenge pour les informaticiens : arriver à structurer un corpus à partir d'une archive et mettre en œuvre un système automatique de détection de traits permettant d'indexer le texte littéraire.
- 12 Deux points essentiels en matière d'innovation scientifique seront développés ici : la méthodologie de recueil des œuvres littéraires web et la production de données. On abordera rapidement l'indexation qui constitue le travail en cours.

- 13 La méthodologie que nous présentons est un *workflow*³ fondé sur une méthode d'identification, un recueil et une structuration des ressources identifiées pour alimenter un *crawl*⁴ et enfin une évaluation, articulée à la production de l'archive (et effectuée par la BNF). Elle se compose donc de deux phases :
- L'une vise à acquérir des URL par une identification préalable de façon à opérer un premier *crawl* qui constituera le corpus ;
 - L'autre à acquérir des données complémentaires par une identification automatique des liens associés aux URL identifiées lors de la première phase à l'aide de Hyphe⁵ pour une analyse pour enfin la constitution de l'archive par un deuxième *crawl*.
- 14 Les résultats de cette deuxième analyse complètent ainsi ceux de la première, en intégrant des URL en marge des réseaux d'écrivains préalablement identifiés. Cette seconde analyse nous aura montré que la première avait surtout identifié des écrivains dont les liens de reconnaissance mutuelle sont ancrés dans les pratiques numériques, alors que le second aura mis en évidence des auteurs qui se réfèrent à ces premiers réseaux sans en faire véritablement partie et sans chercher une reconnaissance particulière de leur part.

Structuration générale du projet

- 15 L'objectif du projet est l'élaboration d'une plateforme permettant de consulter et analyser ce corpus et l'ensemble des données produites par l'équipe et associées à l'élaboration et au traitement des données. La dualité du corpus et de l'archive est liée à des usages différents entre le corpus (comme « donnée primaire de la science ») et l'archive comme patrimoine accessible sur site à la BNF.
- 16 Concernant l'identification des productions littéraires « nativement web », nous avons élaboré une méthodologie, fondée sur la reconnaissance mutuelle des auteurs par le partage de liens. Cette méthodologie s'inspire à la fois de la linguistique de corpus appliquée au web et à la sociologie des réseaux : la démarche d'acquisition de données à partir du web a largement été élaborée en linguistique de corpus (Teubert, 2005), en considérant le web comme un réservoir infini de réalisations linguistiques écrites. Dans ce cadre, toute

production est par définition pertinente, et c'est la perspective linguistique qui construit l'objet de recherche.

- 17 Le premier intérêt d'un corpus est la capacité d'une grande quantité de données à observer des phénomènes réguliers derrière des variations de surface. Le corpus n'est pas construit pour valider une hypothèse préalablement formulée, mais pour observer des régularités qui seront ensuite utilisées pour construire des hypothèses. L'objectif de ce type d'étude est d'abord la détection d'évolutions et l'identification de règles et de régularités qui n'apparaissent pas sans données massives et outils quantitatifs. Les perspectives proposées par le laboratoire ERIC, reposant sur les modèles vectoriels BERT (Jawahar, Sagot, Seddah, 2019) d'analyse des données textuelles, sont essentielles pour mener ce projet⁶. Néanmoins, à la différence des autres corpus disponibles d'œuvres littéraires numérisées, notamment Gutenberg⁷, notre corpus est composé de textes inédits et fortement différents par rapport aux œuvres classiques, tant par leur forme, leurs règles d'écriture, que par leur statut social. Ainsi, il ouvre des perspectives nouvelles d'analyse pour les outils du traitement vectoriel des textes.
- 18 Notre stratégie a permis la réalisation d'un *crawl* massif de façon à disposer d'une somme de données considérable (600 Go), consultables à partir d'un point d'accès unique comme une interface de recherche d'information. Nous aboutissons à une somme de 1000 écrivains environ (avant *crawl*), dont la plupart ne publient pas de livres édités, donc ne sont pas répertoriés dans les catalogues des bibliothèques. Il en découle une faible pertinence des approches fondées sur les auteurs (stylométrie par exemple, ou sur les genres classiques – poésie, nouvelle, roman), mais une très forte potentialité d'identification de nouvelles formes d'expression.
- 19 Par conséquent, l'indexation sera fondée sur des marques de similarité soit de types de discours (narratif, descriptif, etc.), soit d'identité affichée de discours (psychologique, métadiscursif, communicationnel, etc.), soit encore de structure formelle.

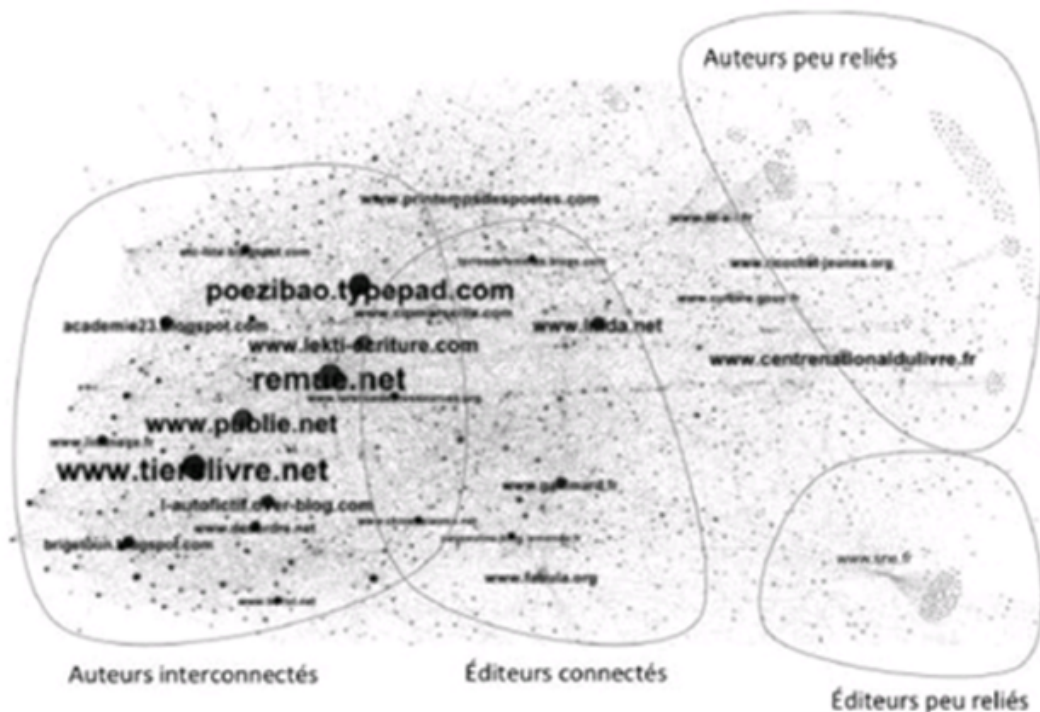
Fondement social : la reconnaissance mutuelle

- 20 Les archives du web ont un fondement historique et sociologique, en premier lieu pour les sciences politiques. À partir du moment où il est apparu que l'archive permettait de recueillir un ensemble massif d'informations sur un thème donné, alors l'archive s'est inscrite dans le cadre de la massification des données en sciences humaines et sociales (Schroeder, Brügger, 2017).
- 21 Idéalement, l'accès au web ne requiert pas d'intermédiaires et permet tout autant d'accéder à la production médiatique, institutionnelle que la publication quotidienne. Par ailleurs, les dispositifs d'archivage du web se sont développés très rapidement (notamment l'Internet Archive⁸) à partir de là, tout comme les initiatives nationales d'archivage du web.
- 22 Or, le problème est celui de l'accès et de la structuration de l'archive. L'archive massive est difficile à exploiter. Seuls des *crawls* restreints, donc des archives limitées, sont utilisables aisément avec toutes les questions théoriques et méthodologiques que cela pose, notamment le problème de la pertinence d'une archive limitée à priori.
- 23 Depuis le lancement de l'outil de recherche SolrWayback (Egense, Myrvoll, 2018) et des outils autour du projet AUT⁹, les difficultés d'usage des archives du web tendent à s'amenuiser : néanmoins, les questions posées relèvent à la fois de l'exploration et de l'organisation des collections, de l'analyse du matériau et enfin de la disponibilité des données pour un large public (Maemura, 2019).
- 24 Parallèlement à une conceptualisation des activités d'archivage, un champ de recherche s'est construit progressivement autour des liens URL, associant la théorie des réseaux sociaux (Zachary, 1977) et de la modélisation mathématique de ceux-ci (Boyd, Ellison, 2007), aidés en cela par les visualisations de Hyphe. Il consiste à construire une conceptualisation autour des possibilités d'exploration et de visualisation des liens URL (Severo, Venturini, 2016).
- 25 Notre recherche est fondée sur l'archive du web, à savoir l'identification d'œuvres et d'auteurs publiant sur le web. En ce sens

on se distingue d'une littérature pensée autour de l'usage du numérique comme constituant de la démarche créatrice. Ici, la dimension numérique est essentiellement considérée comme un outil de publication.

- 26 En ce sens, les œuvres peuvent renvoyer au champ de la création littéraire numérique, mais notre corpus est plus large et repose sur la caractérisation d'un domaine au travers des liens numériques (URL) opérant entre les auteurs de ces œuvres sans par ailleurs statuer sur ces dernières. Ainsi, le lien marqué par un échange d'URL permet d'explorer un univers social qui ne constitue probablement pas un champ (Sapiro, 2021), mais une constellation de réseaux sociaux.
- 27 Notre objectif a été d'éviter les deux pièges que constitue d'une part une limitation très forte du *crawl*, et, d'autre part, une masse de données contenant des informations totalement inutiles au propos et donc difficiles à caractériser comme corpus.
- 28 Des approches sociologiques ont déjà montré l'intérêt des réseaux de sociabilité pour structurer le domaine de la littérature numérique (Beaudouin, 2012). On reprend cette donnée pour bâtir notre découverte d'œuvres en se basant sur la reconnaissance entre pairs comme marqueur de littérarité. Une production textuelle devient littéraire dans le cadre du web par la reconnaissance des pairs, et cette reconnaissance se marque par des liens de type « partage d'URL ».
- 29 Cette reconnaissance se marque par le partage d'URL, mais également par des types de fonctions différents associés aux projets littéraires manifestés par ces URL : des auteurs, des communautés, des espaces collectifs, des revues. On peut identifier aussi beaucoup de « relais » : certaines œuvres numériques sont reprises par d'autres, sur leur support, avec le plus souvent un référencement minimal. Par conséquent, on observe une dispersion des œuvres et des productions d'auteurs dans différents lieux de publication numérique.
- 30 Néanmoins, ce domaine est relativement structuré si l'on considère les réseaux sociaux de reconnaissance mutuelle, comme le montre la visualisation issue de l'analyse HYPHE réalisée par Kevin Locoh-Donou (IGR-BNF) :

Figure 1. Visualisation des réseaux sociaux d'écrivains du web



© Kévin Locoh-Donou

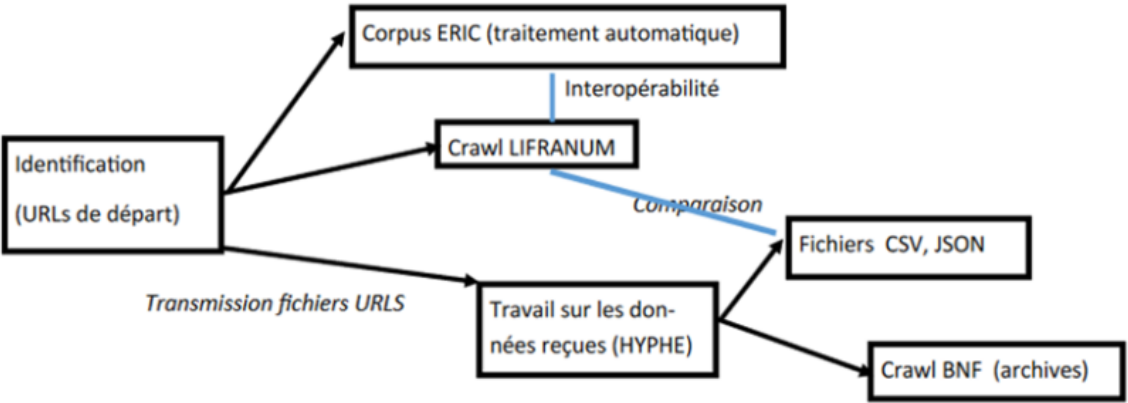
- 31 Sur cette visualisation, on voit l'existence d'une sphère, qui caractérise des communautés de pratiques, à partir de laquelle les autres sphères s'organisent. Elle montre aussi l'importance des espaces numériques de sociabilité, notamment le « tiers livre » de François Bon et « Remue.net ».
- 32 Par ailleurs, l'intitulé des sites et blogs montre que l'unité d'indexation auteur/œuvre, classique des bibliothèques, fonctionne mal ici. En effet, les œuvres dont il est question ici ne sont pas référencées par des règles bibliographiques. Comme nous le verrons, les auteurs comme les titres et les limites des œuvres peuvent être aléatoires ou flous. En ce sens, il est difficile d'extraire des indications d'auteur et d'œuvre de façon sûre, antérieurement à une analyse.
- 33 Dans notre cadre, l'auteur est construit à partir de sa reconnaissance : dès lors, il est auteur de par sa capacité à être identifié comme tel par d'autres, ce qui rend sa définition non pas liée

à une analyse littéraire, mais à un jugement social. C'est d'ailleurs le propre des corpus entendus au sens de la linguistique de corpus de ne pas prédire le statut des objets avant l'analyse du corpus. Une œuvre littéraire sur le web se définit donc par une visibilité dans un réseau. La contrainte de la reconnaissance par les pairs, qui constitue un microcosme du fonctionnement général du web, amène à considérer l'auteur comme une entité sociale, inscrite dans un réseau. En ce sens, il ne peut y avoir d'auteur isolé : ce sont les liens sociaux qui créent la reconnaissance d'auteur dans le corpus.

Méthode quotidienne de travail

- 34 L'objectif de cette présentation n'est pas seulement de présenter un corpus, mais l'ensemble de la méthodologie qui a permis de le réaliser. Si nous insistons sur la dimension intellectuelle de la méthodologie, cette dernière repose sur des documents de recherche, qui eux-mêmes pourront être exploités au-delà de leur usage dans la menée du projet.
- 35 Rappelons que le projet aboutit à deux ensembles de données, le corpus et l'archive (cette dernière étant stockée, gérée et mise à disposition par la BNF). La chaîne de production de ressources repose sur une succession et un enrichissement entre les deux. L'archive a été élaborée à la suite du *crawl* du corpus et est enrichie d'une analyse automatique de liens fondée sur Hyphe.
- 36 Par ailleurs, comme une analyse automatique par traitement automatique des langues à partir des WARC¹⁰ est impossible pour le moment, un corpus de travail a été élaboré par le laboratoire ERIC pour expérimenter l'indexation. Nous travaillons essentiellement sur la base de fichiers JSON, CSV et Excel, échangés entre les équipes et permettant d'exploiter les données obtenues de façon automatique (Hyphe, BERT pour les analyses textuelles). Ces fichiers sont tous stockés dans un espace de gestion de données Nuxeo.
- 37 Nous présentons ci-dessous la totalité des données produites et récoltées dans cette première partie du projet :

Figure 2. Représentation de la production et de l’usage des produits associé à la méthode de travail



- 38 Concernant le corpus, l’objectif est à la fois de fournir un objet d’étude structuré pour les chercheurs du domaine (avec accès réservé) et de constituer un espace d’exploration pour tous ceux qui s’intéressent aux textualités numériques. L’archive a vocation de constituer un patrimoine de cette littérature dans son contexte social.
- 39 Notre *workflow* fait alterner systématiquement les outils automatiques et les méthodes manuelles, ce qui permet soit de paramétrer les outils, afin qu’ils répondent à des objectifs circonscrits, soit d’interpréter des données produites automatiquement :

Figure 3. Représentation de l’articulation des méthodes manuelles et outils automatiques dans le cours du travail

	Méthode manuelle	Méthodes automatiques	Méthode manuelle
Corpus LIFRANUM	Identification des ressources intéressantes et répertoires XML	Crawling HERITRIX sans contrôle dans le déroulement du processus	
Archive BNF		Analyse des liens découverts et crawling supervisé	Suivi de lien en utilisant HYPHE

Procédure d'identification de la littérature web

- 40 Le problème fondamental de l'acquisition de ce corpus consiste en la difficulté à identifier la production littéraire web dans toute sa diversité. En effet, la littérature web n'est pas directement identifiable, parce que l'on ne dispose ni de mots-clés ni d'indices spécifiques ou récurrents, à la différence d'archives covid (Messens, Lieber, Chamber, Geeraert, 2022) pour lesquelles certains mots-clés permettent d'identifier les sources. Donc, les réseaux de sociabilité, qui sont à la fois un fait social et technologique (échange de liens URL), constituent le point d'entrée sur le web¹¹.
- 41 Pour l'identification, nous avons mis en place une chaîne permettant de recueillir et structurer les URL et de constituer un répertoire pour le *crawl*. L'identification utilise la recherche avancée Google en structurant la requête de façon à ce qu'à la fois on obtienne les URL pertinentes et sans bruit.
- 42 Nous utilisons le cadre formel de la théorie des situations (Barwise, Perry, 1983), qui est un modèle de sémantique formelle, par laquelle les prédicats signifient relativement à trois situations : une situation individuelle dans le monde (correspondants aux déictiques), une situation mentale (ou image mentale de la situation individuelle) et une situation de référence, à partir de laquelle les deux précédentes peuvent être interprétées et donner lieu à des inférences. Ces trois situations fondent la façon dont nous structurons la requête, en considérant que la situation individuelle doit être strictement identifiée, la situation décrite caractériser l'objet qui nous concerne (la production littéraire) et enfin la situation ressource, à partir de laquelle on limite le champ d'interprétation.
- 43 On formule ainsi les requêtes obtenues : « Pour l'entité nommée suivante, caractérisée comme l'entité du monde à propos de laquelle on veut obtenir de l'information (un nom propre ou entité nommée), je veux connaître les URL liées à ses propriétés d'être écrivain ou poète et cela en spécifiant les médias au travers desquels ce lien entre une propriété et un objet sont attestés ». Ainsi, on distingue un individu du monde (l'entité nommée), les univers mentaux (la

propriété) enfin le contexte dans lequel on souhaite que cette relation opère (les supports web de publication). En plaçant l'individu au centre de la requête, nous limitons les résultats à ceux qui le concernent. En spécifiant certaines propriétés et le contexte, on élimine les homonymes et surtout on identifie tous les contextes dans lesquels il apparaît comme écrivain et dans le cadre ou en référence à une URL (blogs, sites et réseaux sociaux). Les URL obtenues font explicitement référence à cette « personne en tant qu'écrivain » et au travers du web des blogs, réseaux sociaux et sites. La référence à un contenu web évite les librairies et autres bibliothèques. Ainsi, chaque URL pointant vers cet auteur pourra avoir comme auteur une personne candidate à une nouvelle requête.

Figure 4. Exemple d'une requête formulée à l'aide des principes de la théorie des situations

The diagram shows a search query form with four input fields and three explanatory callouts. The form is titled 'Trouvez des pages avec...' and contains the following fields:

- tous les mots suivants :** (Annotated with: **Propriétés associées à une personne considérée comme écrivain et producteur de contenus littéraires.**)
- ce mot ou cette expression exact(e) :** (Annotated with: **Nom propre d'individu**)
- l'un des mots suivants :** (Annotated with: **Nom d'objet de la plateforme d'édition web ou de l'éditeur (et indication géographique)**)
- aucun des mots suivants :**

- 44 Les résultats des requêtes à propos des auteurs ou entités nommées sont reportés dans des schémas XML. Chaque schéma XML reproduit le résultat d'une requête et enregistre la totalité des liens apparus lors de cette requête : il représente les différents liens d'une URL du site/blog d'un écrivain.
- 45 Les sites identifiés par la recherche qui pointent vers l'URL de cet auteur seront alors considérés comme des liens. Ensuite, l'analyse des liens sortant de cette URL, marqués par son auteur, identifie une reconnaissance mutuelle (ce qui n'empêche pas d'identifier des liens sortant – ou entrant – sans réciprocité).

46 Néanmoins, les liens et les formes de reconnaissance sont différents en fonction du projet littéraire. Nous avons donc typé les projets par sortes de liens échangés : chaque type de projet littéraire ou éditorial est spécifié par les différents types de liens. Nous avons caractérisé quatre types de projets : individuel, collectif, éditorial et communautaire :

- Auteurs individuels : production créative signée par un nom de personne. Ce nom doit être identifié comme celui du créateur du média : URL, métadonnées du créateur, propriétaire de la page.
- Auteurs collectifs (différentes personnes identifiées avec une signature unique) : production créative signée par un nom collectif. Même identification que le précédent.
- Communautés : nom collectif associé à un lieu et à des initiatives (cours, conférences, prix, événements) en relation avec différents auteurs individuels et/ou collectifs.
- Supports : nom associé à une initiative contenant l'édition de créations originales d'auteurs individuels ou collectifs. En général, un support peut être assimilé à une revue.

Figure 5. Table les différentes catégories de projets littéraires

Catégories de projets littéraires
Auteur individuel
Auteur collectif
Communauté d'écrivains
Support de publication (revue)

47 Ensuite, nous structurons les schémas en spécifiant ces différentes formes de liens :

- Lien 0 : liens sortant de la page renvoyant à d'autres réalisations du projet de l'auteur. Il s'agit notamment de liens internes vers le site web de l'auteur ou d'autres pages.

- Lien 1 : liens sortant de la page (lien direct) qui indiquent une relation avec le projet littéraire d'une autre personne.
- Lien 2 : liens entrants mis en évidence par la recherche d'informations (citation directe). L'URL marquée met en évidence un lien URL avec l'objet de la recherche (la page identifiée pointe vers celle qui a été explorée sans que ce lien soit avéré dans l'URL d'origine). C'est l'URL trouvée qui pointe vers la recherche originale. C'est la relation inverse du lien 1.
- Lien 3 : liens mis en évidence par la recherche d'information (citation indirecte). L'URL identifiée se rapporte à l'objet de la recherche, mais sans lien direct. Il s'agit par exemple de la simple citation de l'auteur, de la mention de son existence et de son œuvre. Ce type de lien renvoie à la propriété d'indexation des pages web par les moteurs de recherche. Il s'agit d'un référencement interne des textes qui ne mobilise pas le système des URL.

Figure 6. Table des types de liens entre les auteurs

Caractérisation des liens
Liens d'origine de la mention ou du référencement de l'auteur
Liens sortants à partir du site auteur
Liens entrant par référencement du site auteur sur un autre site
Liens par citation et mention (sans référencement)

48 Pour le moment, cette modélisation de la reconnaissance reste empirique : l'objectif est qu'elle soit modélisée de façon à être reproductible et permette de représenter des flux d'information. Nous avons utilisé ces méthodes pour identifier des URL sans en inférer des conclusions formelles, qu'il s'agisse d'un outil automatique à concevoir ou d'un modèle logique qui permette de représenter les échanges informationnels à partir des URL partagées.

- 49 Nous présentons ci-dessous un exemple de ces schémas XML qui enregistrent l'ensemble des liens associés à chaque URL dans le cadre de la reconnaissance mutuelle :

Figure 7. Exemple de schéma XML utilisé pour enregistrer les résultats des requêtes

```
<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
<schema elementFormDefault="qualified" targetNamespace="http://www.example.org/LIFRANUMidentification">
  <element name="collection" type="string"/>
  <complexType name="network">
    <attribute name="description">
      <simpleType>
        <restriction base="string">
          <enumeration value="personal/communaux"></enumeration>
          <enumeration value="personal/personal"></enumeration>
          <enumeration value="communaux/communaux"></enumeration>
          <enumeration value="communaux/personnel"></enumeration>
          <minLength value="0"/>
          <maxLength value="1"/>
          <enumeration value="value"></enumeration>
        </restriction>
      </simpleType>
    </attribute>
  </complexType>
  <complexType name="facet">
    <complexContent>
      <extension base="tns:network">
        <sequence>
          <element name="webunit" type="string"/>
        </sequence>
        <attribute name="provenance" type="string"/>
        <attribute name="link0" type="string"/>
        <attribute name="link1" type="string"/>
        <attribute name="link2" type="string"/>
        <attribute name="link3" type="string"/>
        <attribute name="authorproject" type="hexBinary"/>
        <attribute name="communityproject" type="hexBinary"/>
      </extension>
    </complexContent>
  </complexType>
</schema>
```

Méthodologie du *crawl*

- 50 Les listes d'adresses obtenues sont utilisées comme racines pour un *crawl* fortement contraint, de façon à éviter des ramifications hors champ (sachant que le *crawl* sera complété ensuite par la BNF en ce qui concerne le contexte). Notre stratégie a consisté à « crawler » par hôte ou par domaine pour les blogs dépendant d'une API particulière. La limite de deux sauts à partir de la racine a permis d'éviter de recueillir des sites non souhaités, et la profondeur maximale des

crawls a permis de restituer la totalité des œuvres (dès lors que l'on considère le projet littéraire web comme une œuvre dans la durée).

- 51 Nous avons utilisé Heritrix¹² pour le *crawl*¹³ et SolrWayback¹⁴ comme outil d'exploration. SolrWayback¹⁵ permet un accès aux contenus grâce aux métadonnées WARC. Ces outils ont été choisis pour une raison liée à la notion de corpus comme unité de format, et seuls des outils hors API peuvent constituer cette unité. Un *crawl* par API aurait obligé à segmenter le recueil en se limitant à une API ou à envisager le *crawl* d'une API l'une après l'autre, éliminant ainsi la dynamique des liens entre sites ou blogs. Par ailleurs, Heritrix et SolrWayback sont les seuls outils utilisés à grande échelle dans le cadre de l'archive du web.
- 52 Cette stratégie permet de récupérer l'ensemble d'un site ou d'un blog. Elle est liée, au-delà du principe de reconnaissance mutuelle, à l'idée que la littérature web est fortement inscrite dans un contexte construit par l'auteur (individuel ou collectif) et que ce contexte constitue le projet littéraire lui-même. Ainsi, un site d'auteur peut comprendre d'autres publications que des textes purement littéraires (opinions politiques, questions de société, etc.), mais ils font partie du projet littéraire web. En ce sens, le projet web, chez de nombreux auteurs, insère la production littéraire dans d'autres écrits, voire les indifférencie. Notre stratégie permet donc d'intégrer la diversité dans un projet et par l'absence de restriction sur la profondeur, elle permet de montrer l'évolution du projet littéraire d'un auteur.

Méthode inverse et constitution de l'archive

- 53 Le *crawl* à partir de l'identification manuelle constitue la première phase de la constitution du corpus et de l'archive. Elle se devait d'être complétée par une autre, qui permettait à la fois d'élargir le corpus, de compléter et de valider cette première phase. En effet, nous n'avions pas la possibilité de montrer la relative complétude de notre méthode ; seule une méthode inverse, reposant sur des principes totalement différents, permettait d'évaluer la pertinence et les caractéristiques des données recueillies lors de cette première phase.
- 54 Nous cherchions ainsi à répondre à deux questions :

- Qu'est-ce que l'on a récolté par rapport au domaine du web littéraire en termes de contenus ?
- Quelles sont les URL qui ne sont pas apparues lors de la première identification, et pourquoi ?

- 55 Ainsi, dans un deuxième temps, les URL racines, issues de l'identification, sont réinterprétées par l'équipe de la BNF en produisant des analyses automatiques de liens utilisant Hyphe, de façon à repérer d'autres adresses qui ne seraient pas des liens de reconnaissance. Ainsi, on peut compléter la représentation des reconnaissances par des adresses d'auteurs qui gravitent autour de ces réseaux constitués.
- 56 La première difficulté tient d'abord au nombre d'URL dans la liste de départ. Il a fallu travailler sur un échantillon réduit. En prenant en compte uniquement les comptes Wordpress, l'analyse Hyphe a permis de récolter un ensemble impressionnant de liens à partir desquels une sélection manuelle a été opérée de façon à éliminer les sites non pertinents. L'analyse manuelle de cet échantillon a permis de sélectionner entre 10 et 20 % de sites pertinents. Cette dualité d'approche où l'analyse initiale est affinée par une analyse automatique contrôlée *a posteriori* permet de distinguer les liens de reconnaissance entre écrivains formant réseau (identification initiale) et les relations à ces réseaux découvertes par les seuls liens hypertextes.
- 57 Cette seconde méthodologie a permis aussi de mettre en évidence le réseau de critiques et de structures accompagnant les auteurs (associations, bibliothèques, éditeurs notamment).
- 58 Nous présentons ici les résultats globaux des deux méthodes d'identification, puis la comparaison entre des deux :

Figure 8. Bilan et comparaison des collections

Répertoire LIFRANUM (avant crawl)	BC web ⁷	Découvert par HYPHE.	
937 URLs	152 URLs	646 URLs	

URLs Identifiées et crawlées par HERITRIX	URLs seulement identifiées par HERITRIX	Découvertes par BNF et pertinentes	Non pertinentes pour le corpus LIFRANUM
2	46	100	498

- 59 Ainsi, on peut noter qu'aux 937 URL découvertes par la méthode d'identification on peut ajouter une centaine d'adresses découvertes par Hyphe. Par ailleurs, les adresses de BC web reprennent très largement notre identification. Les URL du domaine qui ne contiennent pas de production littéraire (critiques, bibliothèques notamment) sont relativement volumineuses puisqu'elles concernent quasiment 500 URL. Enfin, les contraintes de notre *crawl* ont fait que sur les 100 adresses découvertes par la BNF, 46 étaient déjà identifiées sans être « crawlées », 2 identifiées et « crawlées ». 52 adresses étaient véritablement nouvelles.

Analyse du corpus

- 60 Le problème important qui demeure sur le corpus Lifranum pour son analyse est la granularité puisque le *crawl* et donc les WARC renseignent très exactement une URL et non un hôte ou une racine. Pour l'analyse, il faut un travail manuel via SolrWayback pour reconstituer un projet éditorial et son organisation. À terme, un enjeu de la structuration du corpus par l'accrochage des fichiers XML aux racines des URL « crawlées » devrait permettre une structuration du corpus (en plus de l'indexation). L'autre intérêt des schémas XML réside dans le fait qu'ils permettent de mettre en évidence des liens entre les différentes URL, mais également de typer les projets (individuels, collectifs, communautés et supports). En attachant les schémas aux racines (hôte ou domaine) des sites, on obtiendrait aussi leur typage.
- 61 Le choix de commencer par une méthode manuelle a permis de limiter un champ qu'aucune méthode automatique ne pouvait circonscrire. Mais du fait de l'outil d'interrogation (Google) et malgré la précision de la formulation des requêtes, de nombreux biais ont demeuré. Par conséquent, une seconde méthode, automatique, a permis de compléter la première recherche et a permis de retrouver des auteurs indirectement associés aux premiers.
- 62 Par ailleurs, l'intérêt de la dualité entre le corpus et l'archive a d'abord été de distinguer des projets littéraires et un domaine.

Vers la structuration des données

- 63 Pour l'indexation et le traitement automatique des données, il est impossible de travailler directement avec les résultats du *crawl* Heritrix. D'où la constitution d'un corpus complémentaire, élaboré par ERIC à partir des API de Wordpress et de Blogger de façon à disposer de données « propres » pour une analyse de traitement automatique du langage par le modèle BERT.
- 64 La stratégie d'indexation repose sur le constat que les catégories usuelles de la classification littéraire (roman, poésie, nouvelle, etc.) ne sont guère pertinentes ici et plus généralement, qu'il n'existe guère de critères distinctifs sûrs pour qualifier une production littéraire, puisque le propre de la recherche en littérature consiste à la caractériser.
- 65 C'est donc une indexation sans langage contrôlé préalable¹⁶. Ainsi, une part de l'objectif consiste justement à construire ce langage documentaire qui permettra de structurer ces données massives. Plutôt que de partir sur une indexation thématique (nous savons qu'un texte littéraire peut avoir un tout autre sens que les unités référentielles qu'il contient), nous construisons plusieurs points de vue, relatifs à l'identité revendiquée du texte et les choix esthétiques, qui tous constituent des caractéristiques textuelles relevant d'un projet d'expression artistique.
- 66 Deux stratégies complémentaires ont donc été suivies :
- Une stratégie supervisée, où l'on demande à l'outil de retrouver certaines formes grammaticales de façon à caractériser des types discursifs et des positionnements de discours. La démarche supervisée s'appuie sur des marqueurs discursifs dont le rôle est avéré dans la construction de types de discours (narratif, descriptif, etc.). Ces marqueurs sont les connecteurs argumentatifs et les verbes. Ainsi, on affiche des proximités entre textes liées à des similarités d'usages de connecteurs et de verbes.
 - Une stratégie non supervisée, où l'on cherche à faire émerger des tendances à partir d'une vectorisation systématique de l'ensemble des marqueurs lexicaux et grammaticaux. Ici, les choix dans la langue, mesurés par des outils statistiques (entropie de Shannon, constantes de

Yule, indices de Simpson, hapax et dislegomena notamment), font émerger des particularités d'usage de la langue. Ces particularités, analysées au niveau du texte, permettent de caractériser des choix expressifs spécifiques à chaque texte.

- 67 La stratégie d'indexation repose aussi sur des observations de pratiques de chercheurs pour qui la navigation par proximité constitue un aspect essentiel de la consultation de la littérature web. En effet, le lecteur ne va pas rechercher tel ou tel type de texte, ou au contraire des textes sur un sujet précis, mais bien se laisser guider par des similitudes ou des parentés esthétiques entre les textes.
- 68 L'indexation permet des recommandations liées à des traits communs (de types de discours, de portée revendiquée et de formes) ou au contraire à des ruptures entre les textes.
- 69 L'indexation permet des recommandations liées à des traits communs (de types de discours, de portée revendiquée et de choix esthétiques) ou au contraire à des ruptures entre les textes. Dans le contexte qui est le nôtre (écrivains non référencés, œuvres disparates et courtes), il nous est apparu pertinent de partir de traits énonciatifs et de choix esthétiques pour proposer des types de textes liés par une communauté de choix et d'affichage¹⁷.

Limites et perspectives

- 70 Nous proposons maintenant une mise en perspective de notre travail et l'ouverture de pistes permettant de poursuivre l'analyse proposée. Notre projet propose un corpus de recherche correspondant au *crawl* LIFRANUM et qui contient tous les matériaux de recherche associés, notamment les schémas XML d'identification, visualisations Hyphe, données de comparaison. Ces données font partie intégrante du corpus parce qu'il s'agit de résultats scientifiques demandant à être exploités.
- 71 L'objectif de ce *workflow* a consisté à établir une méthodologie robuste pour le recueil de données (corpus et archives) de façon à généraliser le travail sur corpus web. Nous avons voulu également montrer que les méthodes manuelles et automatiques sont complémentaires. Elles ont également permis à la BNF de comparer

avec ses collections BCweb, élaborées depuis une dizaine d'années par Christiane Génin.

- 72 Les méthodes de collecte traditionnelles du web littéraire, mises en œuvre par la BNF, reposent essentiellement sur une identification manuelle à partir du dépôt légal. Cette méthode permet de suivre les sites déclarés comme étant de littérature, mais ne permet pas de les insérer dans le contexte de leur production et réception (sachant que la reconnaissance est une part de la réception). Leur comparaison et leur exploitation mutuelle ne sont pas encore achevées.
- 73 Par ailleurs, l'archive est rejouable¹⁸ et il est possible de traiter à part les URL contextuelles. En effet, la totalité des outils et des données de travail est disponible et réutilisable, voire modifiable. Le budget de la BNF étant seulement d'1Ta, le *crawl* est limité, notamment relativement aux images et à la profondeur. Le suivi du *crawl* a permis d'éviter les dérives de Heritrix sans par ailleurs contraindre les sauts, à la différence du premier *crawl*. Le corpus a été réalisé en une seule fois, ce qui permet d'avoir une vision générale de l'état de la littérature web en France. Dans l'idéal, le *crawl* doit être réitéré de façon à intégrer les évolutions de la production littéraire web.
- 74 Néanmoins, un certain nombre de difficultés subsistent. Le problème de l'utilisation des données WARC pour l'analyse automatique n'a pas été résolu, et nous oblige à travailler avec un corpus textuel d'appoint, constitué à partir des API.
- 75 Nous insistons encore ici sur la complémentarité entre les deux phases de la méthodologie : la seconde permet d'améliorer les résultats de la première : sur les 1000 noms d'auteurs, l'analyse Hyphe a permis d'en ajouter une centaine. Par ailleurs, elle a permis d'analyser le *crawl*, puisqu'une cinquantaine d'adresses avait été répertoriée sans être « *crawlée* » (deux avaient été identifiées et *crawlées*). Leur lien aux schémas XML et l'analyse des réseaux est encore à développer : les deux analyses de réseau (schémas XML et Hyphe) sont différentes, mais complémentaires.
- 76 Notre corpus contient un certain nombre de restrictions que nous souhaitons développer ici. Le web des blogs et des sites repose sur l'échange d'adresses sans autre forme de contraintes, alors même que les plateformes de réseaux sociaux, de publication (Wattpad¹⁹

notamment) et les forums²⁰ sont fondés sur d'autres modes d'échanges que les mises en commun d'URL et donc qu'un lien ou marque d'appréciation ne peut être considéré comme une reconnaissance. Au problème d'identification s'ajoute celui du *crawl* lui-même, à savoir l'impossibilité pour Heritrix de « crawler » les contenus des plateformes. Dès lors, concernant l'identification de ces productions, nous devons mettre en œuvre une procédure d'identification spécifique sauvegardant le principe de la reconnaissance mutuelle.

- 77 Enfin, ce corpus constitue un outil pour caractériser des formes et des expressions littéraires nouvelles, jusque-là étudiées de façon parcellaire et sans l'aide de la systématisme permise par ce corpus massif.

Conclusion

- 78 À l'issue de cette construction de corpus, la figure de l'auteur est relativement transformée, au sens où il n'est plus simplement celui qui produit le texte, mais aussi celui qui l'édite et s'assure de son référencement comme de sa mise en réseau. Pour la constitution du corpus, l'auteur se caractérise avant tout comme la personne qui va s'assurer d'une mise en visibilité de ses productions, ce qui passe avant tout par sa socialisation.
- 79 Nous avons relaté dans cet article l'élaboration d'un corpus antérieurement aux analyses qui peuvent en être faites, sachant que l'objectif d'un corpus est de rendre accessible cette production à l'analyse littéraire. Son enregistrement repose sur la sociabilité littéraire comme principe d'identification. Nous sommes restés en deçà de l'analyse littéraire – ou autre – de cette production.
- 80 L'enjeu social de ce travail est tout d'abord de fournir un objet d'études. Il s'agit du rôle d'un corpus. Il permet, du fait de ses dimensions mêmes, une entrée pour des analyses différentes, relevant de la littérature, mais également de la sociologie et de la linguistique notamment. Ces analyses sont, pour le moment, tributaires des possibilités de SolrWayback, qui permet néanmoins des recherches lexicales tout autant que des approches par facettes.

- 81 Si comme nous l'avons déjà énoncé, les analyses de type BERT, nécessitent de dupliquer le corpus pour réaliser des opérations de traitement automatique de la langue, d'autres outils ne nécessitent pas cette duplication, comme ceux proposés par AUT²¹ et ARCH²². En ce sens, le référencement interne, l'extraction de textes autant que l'exploration des liens hypertextuels peuvent être effectivement mis en œuvre à partir de l'archive elle-même.
- 82 Enfin, au-delà de la performance de la réalisation d'un corpus de web littéraire, l'enjeu de ce travail est avant tout de formaliser une première méthodologie pour construire des corpus à la fois massifs et cohérents à partir du web. C'est aussi pour cette raison que nous sommes restés en deçà des analyses possibles du corpus : il est un outil, et nous nous sommes contentés d'éclairer certains présupposés (notion d'auteur notamment associée à la construction de l'outil).

BIBLIOGRAPHIE

- BOYD D. M., ELLISON N. B., « Social network sites: Definition, history, and scholarship » *Journal of computer-mediated Communication*, 2007, vol. 13, n° 1, p. 210-230.
- BARABÁSI A. L., ALBERT R., « Emergence of scaling in random networks », *Science*, 1999, vol. 286, n° 5439, p. 509-512.
- BARWISE J., PERRY J., *Situations and Attitudes*, Massachusetts, MIT Press, 1983.
- BEAUDOUIN V., « Trajectoires et réseau des écrivains sur le Web : Construction de la notoriété et du marché », *Réseaux [en ligne]*, 2012, vol. 5, n° 175, p. 107-144, DOI : [10.3917/res.175.0107](https://doi.org/10.3917/res.175.0107).
- EGENSE T., MYRVOLL A. K., *SolrWayback* [en ligne], 2018, URL : https://netpreserve.org/ga2018/wp-content/uploads/2018/11/IIPC_WAC2018-Thomas-Egense_Anders_Kliedt_Myrvoll-SolrWayback.pdf [consulté le 8 avril 2024].
- JAWAHAR G., SAGOT B., SEDDAH D., « What does BERT learn about the structure of language? », in : *57th Annual Meeting of the Association for Computational Linguistics* [en ligne], Florence, Italie, Association for Computational Linguistics, juillet 2019, p. 3651-3657, DOI : [10.18653/v1/P19-1356](https://doi.org/10.18653/v1/P19-1356).
- MAEMURA E., « What's cached is prologue: Reviewing recent web archives research towards supporting scholarly use », *Proceedings of the Association for Information Science and Technology*, 2019, vol. 55, n° 1, p. 327-336.
- MESSENS F., LIEBER S., CHAMBERS S., GEERAERT F., 2022, « Seed list mini pilot COVID-19 collection », *Sodha (Social sciences and digital humanities archive)* [en ligne], V1, 2021,

DOI : [10.34934/DVN/SE8NUY](https://doi.org/10.34934/DVN/SE8NUY).

SAPIRO G., « Le champ littéraire. Penser le fait littéraire comme fait social », *Histoire de la recherche contemporaine* [en ligne], 2021, Tome 10, n° 1, p. 45-51, DOI : [10.4000/hrc.5309](https://doi.org/10.4000/hrc.5309).

SCHROEDER R., BRÜGGER N. (dir.), *The Web as History: Using Web Archives to Understand the Past and the Present*, Londres, UCL Press, 2017, p. 296.

SEVERO M., VENTURINI T., « Enjeux topologiques et topographiques de la cartographie du web : Le cas du patrimoine culturel immatériel français », *Réseaux* [en ligne], 2016, vol. 1, no 195, p. 85-105, DOI : [10.3917/res.195.0085](https://doi.org/10.3917/res.195.0085).

ZACHARY W. W., « An information flow model for conflict and fission in small groups », *Journal of anthropological research*, 1977, vol. 33, n° 4, p. 452-473.

Sources web :

Archives Unleashed [en ligne], URL : <https://archivesunleashed.org/> [consulté le 8 avril 2024].

Hyphe [en ligne], URL : <https://hyphe.medialab.sciences-po.fr/> [consulté le 8 avril 2024].

Project Gutenberg [en ligne], URL : <https://www.gutenberg.org/> [consulté le 8 avril 2024].

« Heritrix : internetarchive/heritrix3 », *Github* [en ligne], 27 juillet 2022 URL : <https://github.com/internetarchive/heritrix3> [consulté le 8 avril 2024].

« Netarchivesuite/Solrwayback » *Github* [en ligne], 5 juillet 2022, URL : <https://github.com/netarchivesuite/solrwayback> [consulté le 8 avril 2024].

NOTES

¹ Le projet dirigé par Gilles Bonnet comprend trois partenaires : UR Marge (Lyon 3), UR ERIC (Lyon 2), BNF. Deux IGR, Lorraine Feugères (Marge) et Kévin Locoh-Donou (BNF) ont particulièrement intensément travaillé sur cette partie du projet. Les équipes de Marge comprennent Belen Hernandez-Marzal, Alice Pantel-Cassagnaud, Fanny Mézard, Lucien Perticoz, de la BNF : Christine Génin, Alexandre Faye, Julien Starck, Clara Wiatrowski, ERIC : Julien Velcin, Enzo Terreau, Javier Espinoza, Jérôme Darmont, Sabine Loudcher. Pour en savoir plus sur le projet, voir le site de l'Agence nationale de la recherche, URL : <https://anr.fr/Projet-ANR-19-CE38-0007> [consulté le 8 avril 2024].

- 2 Nous nous distinguons des projets autour de ELO (*Electronic Literature Directory* [en ligne], URL : <https://directory.eliterature.org/> [consulté le 8 avril 2024]), notamment ELMCIP (*Electronic Literature Knowledge Base* [en ligne], URL : <https://elmcip.net/> [en ligne]), le NEXT The NEXT (*The NEXT* [en ligne], URL : <https://the-next.eliterature.org/> [consulté le 8 avril 2024]), les répertoires du NT2 (ALN/NT2 [en ligne], URL : <https://nt2.uqam.ca/> [consulté le 8 avril 2024]) ou PO.EX (Arquivo Digital da PO.EX – Poesia Experimental Portuguesa [en ligne], URL : <https://po-ex.net/> [consulté le 8 avril 2024]), ou encore le projet brésilien SOBRE (Observatório da literatura digital brasileira [en ligne], URL : <https://www.observatorioldigital.ufscar.br/> [consulté le 8 avril 2024]), parce qu'ils suivent un principe déclaratif et non un recueil de données.
- 3 « Un *workflow* est un processus industriel ou administratif au cours duquel des tâches, des documents et des informations sont traités successivement, selon des règles prédéfinies, en vue de réaliser un produit ou de fournir un service. » Source : <https://www.culture.fr/franceterme/terme/INFO620> [consulté le 8 avril 2024].
- 4 Un *crawl* est une opération d'archivage (ou de « capture ») sur le web, effectuée par un agent automatisé, appelé « crawler », « robot » ou « araignée ». Les *crawls* identifient les documents sur le web sur la base de votre choix d'URL de départ et d'étendue. Le *crawl* peut également référencer le contenu archivé associé à l'action. Source : <https://support.archive-it.org/hc/en-us/articles/208111686-Glossary-of-Archive-It-and-Web-Archiving-Terms> [consulté le 8 avril 2024].
- 5 *Hyphe* [en ligne], URL : <https://hyphe.medialab.sciences-po.fr/> [consulté le 8 avril 2024].
- 6 Pour une présentation simple de ces modèles, voir : <https://eduscol.education.fr/sti/sites/eduscol.education.fr/sti/files/ressources/pedagogiques/14960/14960-word-embedding-les-mots-et-le-machine-learning-ensps.pdf> [consulté le 8 avril 2024].
- 7 Project Gutenberg [en ligne], URL : <https://www.gutenberg.org/> [consulté le 8 avril 2024].
- 8 Internet Archive [en ligne], URL : <https://archive.org/> [consulté le 8 avril 2024].
- 9 Archives Unleashed [en ligne], URL : <https://archivesunleashed.org/> [consulté le 8 avril 2024].

- 10 Les WARCs sont un format *open source* développé par l'Internet Archive et norme ISO (CD 28500) pour les archives web, composé d'enregistrements WARC désagrégés (provenant de différents hôtes). Source <https://www.loc.gov/preservation/digital/formats/fdd/fdd000236.shtml> [consulté le 8 avril 2024] ; pour consulter un glossaire des termes spécifiques à l'archivage web, voir : <https://support.archive-it.org/hc/en-us/articles/20811686-Glossary-of-Archive-It-and-Web-Archiving-Terms> [consulté le 8 avril 2024].
- 11 Ce travail a nécessité le recrutement d'une ingénieure (Lorraine Feugères) sur un an.
- 12 « Heritrix : internetarchive/heritrix3 », *Github* [en ligne], 27 juillet 2022 URL : <https://github.com/internetarchive/heritrix3> [consulté le 8 avril 2024].
- 13 Heritrix est le nom du projet de moteur de recherche d'Internet Archive, *open source*, extensible, à l'échelle du web et de qualité archivistique. Source : <https://support.archive-it.org/hc/en-us/articles/20811686-Glossary-of-Archive-It-and-Web-Archiving-Terms> [consulté le 8 avril 2024].
- 14 « Netarchivesuite/Solrwayback » *Github* [en ligne], 5 juillet 2022, URL : <https://github.com/netarchivesuite/solrwayback> [consulté le 8 avril 2024].
- 15 SolrWayback est à la fois une plateforme de découverte pour la recherche de pages web historiques et un outil de lecture pour visualiser les pages web. Il est livré avec divers outils, dont l'outil d'exportation du graphe de liens. Source : <https://labs.statsbiblioteket.dk/linkgraph/> [consulté le 8 avril 2024].
- 16 Un premier usage de ce corpus consiste justement à explorer ces caractéristiques du texte littéraire web.
- 17 Le corpus de travail de ERIC (sur lequel se fonde l'indexation) sera relié au corpus général et permettra à terme d'indexer l'ensemble de la production littéraire web.
- 18 Rejouer signifie que les versions archivées peuvent être représentées telles qu'elles étaient lors de leur mise en ligne.
- 19 *Wattpad* [en ligne], URL : <https://www.wattpad.com/> [consulté le 8 avril 2024].
- 20 Comme par exemple : *Forum des jeunes écrivains* [en ligne], URL : <http://www.jeunesecrivains.com/> [consulté le 8 avril 2024], *Mauvaises Herbes*

[en ligne], URL : <https://www.jeunesecrivains.com/> [consulté le 8 avril 2024], *Mauvaises Herbes* [en ligne], URL : <https://mauvaisesherbes.forumactif.com/> [consulté le 8 avril 2024], etc.

21 *Archives Unleashed Toolkit* [en ligne], URL : <https://aut.docs.archivesunleashed.org/> [consulté le 8 avril 2024].

22 *Archives Research Compute Hub* [en ligne], URL : <https://webservices.archive.org/pages/arch> [consulté le 8 avril 2024].

RÉSUMÉS

Français

Nous présentons la première phase du projet ANR Lifranum, qui consiste à élaborer un corpus et une archive du web littéraire francophone. Nous présentons une méthodologie qui permet d'identifier sur le web les URL pertinentes, d'enregistrer la partie du réseau social dans lequel chacune de ces URL se trouve insérée, et enfin de récupérer ces URL par un *crawl*. Cette analyse est d'abord manuelle et repose sur les réseaux sociaux d'écrivains. La particularité de notre travail réside dans le fait que nous couplons cette méthode par une procédure de validation inverse, qui part d'une analyse automatique pour ensuite opérer un tri manuel et enfin réaliser un *crawl* complémentaire, à visée patrimoniale et réalisé par la BNF. Enfin, nous nous interrogerons sur la façon dont on peut caractériser et utiliser ces objets numériques, corpus et archives, dans le cadre d'investigations diverses. Nous pourrions alors proposer les contours de ce corpus comprenant plus de 1000 écrivains publiant sur le web, et nous envisagerons les implications qu'il peut avoir sur l'approche du fait littéraire, quelles que soient les entrées scientifiques choisies.

English

We present the first phase of the ANR Lifranum project, which involves building a corpus and archive of the French-language literary web. We present a methodology for identifying relevant URL on the web, recording the part of the social network in which each of these URL is inserted, and finally retrieving these URLs through a *crawl*. This analysis is initially manual, and is based on writers' social networks. The particularity of our work consists in the fact that we couple this method with a reverse validation procedure, which starts with an automatic analysis, then performs a manual sorting and finally carries out a complementary *crawl*, with a patrimonial aim and carried out by the BNF. Finally, we'll look at how we can characterize and use these digital artifacts, corpora and archives, in a variety of investigations. We will then be able to propose the contours of this corpus, comprising over 1,000 writers publishing on the web, and we

will consider the implications it may have on the approach to the literary fact, whatever the scientific entries chosen.

INDEX

Mots-clés

littérature web, archives du web, corpus, réseaux

Keywords

web literature, web archive, corpus, networks

AUTEUR

Christian Cote

Maître de conférences en sciences de l'information et de la communication, HDR,
université Jean-Moulin Lyon 3, laboratoire Marge

IDREF : <https://www.idref.fr/235763683>

ORCID : <http://orcid.org/0000-0002-7230-3268>

HAL : <https://cv.archives-ouvertes.fr/christian-cote>

Plateforme Weblitt

Weblitt platform

Plateforme Weblitt

DOI : 10.35562/marge.839

Droits d'auteur

CC BY-NC-SA

TEXTE

- 1 Adresse de la plateforme : <https://weblitt.msh-lse.fr/>
- 2 Date de création : juin 2021

Pouvez-vous décrire le projet, les objectifs de votre plateforme et son fonctionnement ?

Weblitt se présente graphiquement comme une mappemonde qui rend accessibles les œuvres nativement numériques d'auteurs publiant leurs textes en français sur des blogs, des sites ou des réseaux sociaux. Cette plateforme est l'une des réalisations menées dans le cadre de Lifranum (littératures francophones numériques), un projet de recherche financé par l'Agence nationale de la recherche (2020-2024). L'intention première de Weblitt est de faire émerger un continent peu connu de la création littéraire contemporaine en identifiant un corpus de productions francophones nativement numériques accessible grâce à une interface qui prend la forme d'une mappemonde. L'objectif est de sensibiliser un large public à la diversité des formes vivantes de la création littéraire qui se déploient hors du livre (vidéos, formes brèves sur les réseaux sociaux, textes illustrés, etc.). Défricher ces terrains jusqu'alors peu visités permet de déterminer une poétique numérique, afin de réévaluer l'influence du support médiatique sur le processus de création, mais également d'actualiser les notions d'auteur, d'œuvre littéraire ou de lecteur. Il s'agit également d'envisager sites, blogs et réseaux sociaux à visée littéraire, non seulement dans leur activité propre, mais également dans leurs échanges dynamiques, afin d'étudier comment la francophonie se déploie désormais en réseaux, grâce à l'utilisation massive d'Internet. Un moteur de recherche permet d'affiner les

requêtes en fonction de la zone géographique, de la date, du nom de l'auteur, du type d'œuvres ou encore du support de publication (blogs, réseau social, site, forum).

Pouvez-vous choisir 2 réalisations particulièrement marquantes ou 2 enjeux scientifiques majeurs, que la mise en ligne de votre plateforme a permis d'atteindre ?

L'étude de la littérature numérique peut désormais s'appuyer sur un cluster qui centralise les données, quand l'obstacle majeur résidait jusqu'à présent dans la dispersion et l'invisibilisation d'un corpus protéiforme. Ces données permettent l'identification et la construction de corpus pour les chercheurs désireux d'étudier la littérature nativement numérique à travers un prisme pluriel : théorie littéraire, sociologie culturelle, pratiques et usages numériques.

Les auteurs et œuvres présentées sur la plateforme servent également de base de données pour la découverte de la littérature numérique et l'animation d'atelier d'écriture à destination des enseignants du secondaire, professeurs-documentalistes et bibliothécaires. Un kit pédagogique présentant divers outils à destination de l'enseignement de la littérature numérique est actuellement en préparation.

**Pouvez-vous évoquer la suite souhaitée, s'il y en a une ?
Développements à venir, nouvelles fonctionnalités,
partenariats, etc.**

Interopérabilité avec la future plateforme développée dans le cadre du projet Lifranum.

Contribution à un futur réseau international de chercheurs via la mutualisation des ressources en ligne

Extension de la plateforme à d'autres aires linguistiques

BIBLIOGRAPHIE

Weblitt : <https://weblitt.msh-lse.fr/> [consulté le 8 avril 2024].

Lifranum : <https://marge.univ-lyon3.fr/projet-lifranum> [consulté le 8 avril 2024].

Vidéo de présentation de la plate-forme : <https://youtu.be/uUApI01Rn9o> [consulté le 8 avril 2024].

Page Facebook : <https://www.facebook.com/lifranum.lyon.1> [consulté le 8 avril 2024].

AUTEUR

Plateforme Weblitt

Projet ANR Lifranum Laboratoires Marge (université Lyon 3), ERIC (université Lyon 2). Site hébergé par la MSH de Lyon Saint-Étienne

Plateforme COLiN – catalogue des œuvres littéraires numériques – construire le corpus du Québec

COLiN platform – a catalog of digital literary works – building the Quebec corpus

Plateforme COLiN

DOI : 10.35562/marge.942

Droits d'auteur

CC BY-NC-SA

TEXTE

1 Adresse : <https://colin.ex-situ.info/> [consulté le 8 avril 2024].

2 Date de création : 2023

Pouvez-vous décrire le projet, les objectifs de votre plateforme et son fonctionnement ?

Le catalogue des œuvres littéraires numériques (COLiN) est une plateforme qui rend visibles des œuvres ayant ponctué l'histoire littéraire numérique du Québec. Pour élaborer une vision historique des pratiques en arts littéraires avec une forte dimension numérique, l'équipe du pôle Québec du projet de recherche en partenariat Littérature québécoise mobile, rattachée au laboratoire Ex situ, élabore un catalogue des œuvres littéraires numériques québécoises.

Certains critères régissent la curation des œuvres introduites dans le catalogue des œuvres littéraires numériques (COLiN). Elles doivent tout d'abord avoir une forte teneur littéraire et une composante numérique importante, voire dominante. Elles sont généralement produites d'abord en français. Au moins une personne de l'équipe de création ou d'édition doit provenir du Québec ou l'œuvre doit avoir été produite au Québec. Évolutive, la plateforme accueille déjà plus de 400 fiches documentant des œuvres de différents genres littéraires, mais aussi et surtout des œuvres qui exploitent de manières variées le médium numérique. Certaines d'entre elles datent des premières années de l'informatique et d'Internet, elles s'incarnent

alors généralement sous la forme de sites ou de pages web hébergeant des textes et des images. D'autres œuvres plus contemporaines tirent profit de la pluralité des formats (images, vidéos, sons, etc.) et des dernières technologies, comme les réalités virtuelle ou augmentée.

Le COLiN possède un moteur de recherche interrogeant l'une ou l'autre facette des métadonnées associées aux œuvres. Il sert autant aux chercheurs et chercheuses en littérature québécoise qu'en littérature numérique. De plus, pour les enseignants et enseignantes, la plateforme permet de découvrir des œuvres à présenter en classe en triant les supports de lecture (ordinateur, tablette, cellulaire, par exemple), les thèmes ou les années de publication ou de diffusion, par exemple.

En plus d'être répertoriée et décrite dans le COLiN, chaque œuvre est dotée d'un identifiant Wikidata, la base de connaissance de l'environnement Wikimedia, ce qui lui assure une existence dans cet écosystème également (puisque les œuvres numériques n'ont pas d'ISBN ou autre identifiant unique). Ce pont entre les deux bases de données permet de mieux diffuser sur le web ces œuvres encore peu connues du corpus québécois. La base de données est une des étapes conduisant à une histoire des pratiques littéraires en contexte numérique au Québec.

La base de données est hébergée sur Omeka S, une plateforme web de base de données en libre accès développée par le Center for History and New Media de la George Mason University. Par souci de circulation des savoirs, le contenu du site est sous licence Creative Commons (CC BY-SA 4.0).

Pouvez-vous choisir 2 réalisations particulièrement marquantes ou 2 enjeux scientifiques majeurs, que la mise en ligne de votre plateforme a permis d'atteindre ?

D'une part, la mise en place du COLiN a permis d'identifier un corpus québécois de littérature numérique, ce qui restait difficile à démontrer en dehors de cet exercice. Grâce à ce catalogage, nous poursuivons l'objectif de construire une histoire des pratiques littéraires numériques, un grand chantier qui occupe le laboratoire Ex

situ et, plus largement, le projet Littérature québécoise mobile depuis 2019.

D'autre part, ce projet contribue à faire exister les œuvres numériques par un référencement sur Wikidata, ce qui participe à la mise en valeur du corpus numérique québécois. L'intégration des métadonnées de chaque œuvre au système Wikimedia contribue à la diffusion et à la pérennité des œuvres numériques du catalogue.

**Pouvez-vous évoquer la suite souhaitée, s'il y en a une ?
Développements à venir, nouvelles fonctionnalités,
partenariats, etc.**

Le projet COLiN continuera à accueillir des entrées, au fil des découvertes rétrospectives et des parutions de nouvelles œuvres. Certaines zones seront également investiguées plus avant, comme les blogues à teneur littéraire, particulièrement foisonnants lors des années 2000, et les pratiques d'écriture en continu sur le web, dont l'instapoésie est un fer de lance, qui ne cessent de se réinventer.

De nouvelles fonctionnalités seront peu à peu développées pour faciliter la consultation et la curation de contenu sur la plateforme, notamment le perfectionnement de la navigation de la ligne du temps ainsi que la création de collections ciblées.

La rédaction de documents d'accompagnement pour l'enseignement est également envisagée afin de diffuser le COLiN à une plus vaste échelle dans le milieu de l'éducation. L'objectif serait d'offrir une meilleure visibilité aux œuvres numériques présentes sur la plateforme dans les cours de littérature, auprès des organismes de recherche et des bibliothèques.

AUTEUR

Plateforme COLiN

Laboratoire Ex situ, Université Laval (projet Littérature québécoise mobile)

Plateforme Répertoire des écritures numériques

Répertoire des écritures numériques platform

Plateforme Répertoire des écritures numériques

DOI : 10.35562/marge.952

Droits d'auteur

CC BY-NC-SA

TEXTE

- 1 Date de création : 2015
- 2 Rattachement institutionnel : Université de Montréal, Chaire de recherche du Canada sur les écritures numériques
- 3 Adresse de la plateforme : <https://repertoire.ecrituresnumeriques.ca/s/repertoire/page/accueil>

Pouvez-vous décrire le projet, les objectifs de votre plateforme et son fonctionnement ?

Le Répertoire des écritures numériques est une plateforme créée en 2015 par la Chaire de recherche du Canada sur les écritures numériques, sous la direction de Marcello Vitali-Rosati, et qui a ensuite rejoint le partenariat interuniversitaire Littérature québécoise mobile.

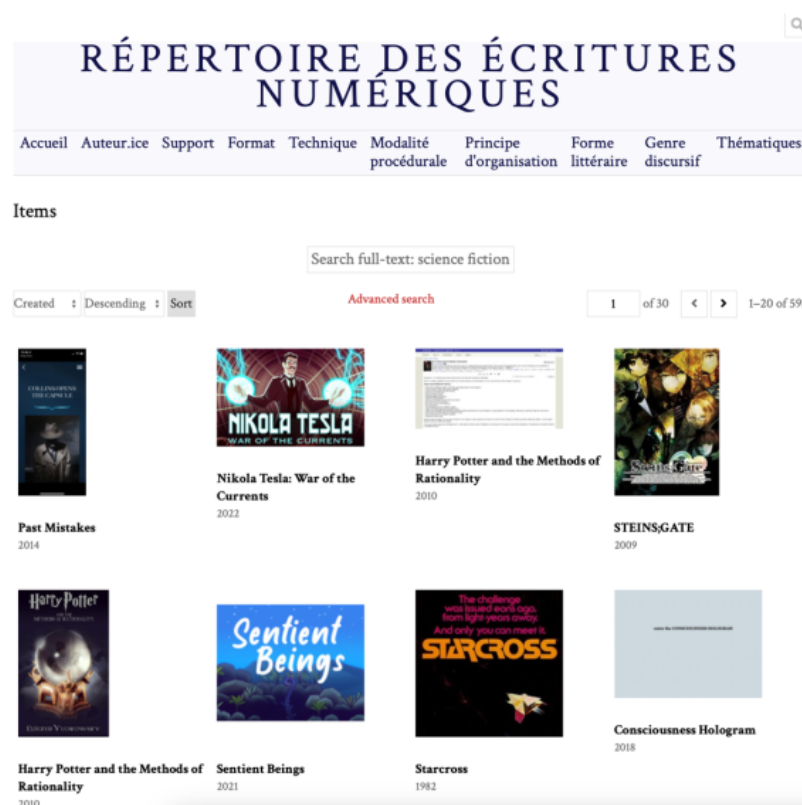
Son rôle est d'inscrire et de documenter les traditions et cultures numériques des différentes communautés – celles où l'auteur·rice s'efface pour laisser place à des collectifs, des entités et des représentations fluides de la fonction auctoriale et éditoriale – qui sont liées aux supports de lecture depuis lesquels leurs créations sont mises à disposition.

Le but du Répertoire est de mettre l'accent sur la production et les moyens concrets de l'existence des œuvres ; une nécessité permettant de cerner les implications et réceptions possibles.

La diversité des œuvres, nécessitant l'approche des narrations vidéoludiques (Aarseth, 1999 ; Hayles, 2007 ; Gervais et Archibald,

2006), des pratiques des réseaux (Manovich, 2001), nous conduit à bâtir un lieu de corpus fortement hétérogène (Emerit, 2016). D'ailleurs, l'établissement des critères de sélection des œuvres pose question, nous amenant à théoriser la littérature dans son inscription et ses interactions avec les autres formes proches que peuvent être le jeu vidéo ou le cinéma interactif.

Page d'accueil du site du Répertoire des écritures numériques



©Répertoire des écritures numériques

Pour cela, le Répertoire des écritures numériques vise à documenter et cataloguer les formes émergentes de la littérature numérique, ce qui amène également l'équipe à questionner en permanence les évolutions des catégorisations et dénominations possibles afin de qualifier au mieux les œuvres étudiées. L'établissement de champs caractéristiques, décrivant les perceptions concrètes et formelles de l'œuvre, est nécessaire pour une documentation de ce corpus. Documentation ensuite enrichie par une approche littéraire, convoquant tant la forme narrative que le genre discursif, les thématiques ou encore l'organisation du texte.

Ainsi, le Répertoire se positionne dans une double approche questionnant d’une part, la technique, et d’autre part, l’implication du·de la lecteur·rice par les gestes de lecture mis en place. Tout cela est permis par l’usage d’Omeka S, combiné à un thésaurus. Le lien entre les deux par l’enrichissement du second et son implémentation dans les fiches du premier permet une description fine et validée par la communauté. La construction d’un vocabulaire collaboratif et vérifié par les pairs, au sein d’un groupement universitaire, apporte une certaine stabilité et une analyse concertée des concepts. L’établissement de définitions communes pour les formes littéraires émergentes permet de les documenter et d’en cerner les mutations, reflétées par l’évolution des définitions et des exemples convoqués.

Page de recherche par technique employée pour la réalisation de l’œuvre

RÉPERTOIRE DES ÉCRITURES
NUMÉRIQUES

[Accueil](#) [Auteur.ice](#) [Support](#) [Format](#) [Technique](#) [Modalité
procédurale](#) [Principe
d'organisation](#) [Forme
littéraire](#) [Genre
discursif](#) [Thématiques](#)

Technique

La **technique** définit le moyen employé pour réaliser l'oeuvre. Chaque technique ou outil venant avec ses contraintes, ses implications et son histoire propre, la technique permet de mieux comprendre la place de chaque oeuvre dans le paysage créatif numérique.

#0-9 A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

adrift (2)

Adventuron (17)

application (59)

application mobile (3)

Archive of Our Own (261)

basic (5)

blogger (5)

Blogspot (10)

bot (2)

C++

ChoiceScript (116)

choose your story (3)

clickteam

clip studio paint

construct

courriel

creator101 (81)

dendry (3)

dialog (3)

donjonfi (3)

drupal

éditeur de texte (9)

eklablog

Enregistrement audio (60)

©Répertoire des écritures numériques

Pouvez-vous choisir 2 réalisations particulièrement marquantes ou 2 enjeux scientifiques majeurs, que la mise en ligne de votre plateforme a permis d’atteindre ?

- Le thésaurus

Le Répertoire des écritures numériques a permis, lors de sa refonte de 2020, de rassembler les équipes de Littérature québécoise mobile, notamment Emmanuelle Lescouet, Benoit Bordeleau et Pierre Gabriel Dumoulin, pour construire un thésaurus commun et unifié, une mise en commun qui a permis de définir les termes, les traduire et de les lier aux corpus en usage. Ce chantier permet, *via* Opentheso, de mettre à disposition l'ensemble des termes qui peuvent être utilisés directement par d'autres répertoires ou être exportés pour de futurs projets.

- Application du Répertoire avec le projet Profil de fiction

Le Répertoire collabore avec Alexandra Saemmer et son équipe pour rassembler le corpus du projet Profil de Fiction. Cette mise en pratique du Répertoire des Écritures Numériques comme outil de documentation collaboratifs s'est avérée un succès, permettant d'aider cette collectivité de recherche à structurer son corpus et de le documenter en commun.

Pouvez-vous évoquer la suite souhaitée, s'il y en a une ?

Développements à venir, nouvelles fonctionnalités, partenariats...

Le Répertoire des écritures numériques se voulant un reflet des cultures numériques, il est nécessaire de continuer à l'enrichir des nouvelles productions, tout en s'assurant un maintien quant à l'interopérabilité et à l'accessibilité du thésaurus auprès de la communauté par son utilisation dans d'autres répertoires, œuvrant au partage des connaissances en sciences ouvertes.

BIBLIOGRAPHIE

AARSETH Espen, « Aporia and epiphany in Doom and The Speaking Clock: The temporality of ergodic art », in : RYAN Marie-Laure (dir.), *Cyberspace textuality*, Indianapolis, Indiana University Press, 1999.

ARCHIBALD Samuel et GERVAIS Bertrand, « Le récit en jeu. Narrativité et interactivité », *Protée* [en ligne], vol. 34, n° 2-3, 2006, p. 27-29, DOI : [10.7202/014263ar](https://doi.org/10.7202/014263ar).

HAYLES N. Katherine, « Hyper and Deep Attention: The Generational Divide in Cognitive Modes », *Profession* [en ligne], 2007, p. 187-199, URL : <http://www.jstor.org/stable/25595866> [consulté le 8 avril 2024].

MANOVICH Lev, *The Language of New Media*, Cambridge, Massachusetts, London, England, The MIT Press, 2001.

EMERIT Laetitia, « La notion de lieu de corpus : un nouvel outil pour l'étude des terrains numériques en linguistique », *Corela* [en ligne], vol. 14, n° 1, 2016, DOI : <http://doi.org/10.4000/corela.4594>.

AUTEUR

Plateforme Répertoire des écritures numériques

Plateforme Encodage

Encodage platform

Plateforme Encodage

DOI : 10.35562/marge.954

Droits d'auteur

CC BY-NC-SA

TEXTE

- 1 Adresse de la plateforme : <https://encodage.uqam.ca/>
- 2 Date de création : 2022
- 3 Rattachement institutionnel : Laboratoire NT2/Figura, Centre de recherche sur les théories et les pratiques de l'imaginaire, Université du Québec à Montréal

Pouvez-vous décrire le projet, les objectifs de votre plateforme et son fonctionnement ?

Encodage est un outil de recherche et d'anthologisation sur l'imaginaire puisant dans l'écosystème numérique sur le contemporain de Figura, le Centre de recherche sur les théories et les pratiques de l'imaginaire. Développé dans le cadre du Laboratoire NT2, le projet a été financé par la Fondation canadienne pour l'innovation (FCI, 2020-2023) et il est dirigé par Bertrand Gervais, directeur du NT2 et de Figura, ainsi que par Nathalie Lacelle, chercheuse principale de l'équipe de recherche FRQSC « Littératie médiatique multimodale » (LMM) et membre de Figura.

Propulsé par l'API de recherche Algolia, le moteur de recherche multisite d'Encodage permet d'interroger les sites web de l'écosystème numérique et d'en fédérer les résultats en un seul et même endroit. Encodage exploite et valorise des bases de données diffusant de la recherche en sciences sociales et humaines, arts et lettres sur l'imaginaire contemporain, étudié depuis les pratiques culturelles, artistiques et littéraires, en plus de favoriser le développement de pistes interprétatives de cet imaginaire, grâce à l'examen et à l'acquisition d'une littératie numérique.

Encodage permet la création d'anthologies, c'est-à-dire de séries de textes et de documents constituées à partir de contenus issus des divers sites web de l'écosystème, de les sauvegarder et de les partager, favorisant ainsi une expérience de littératie numérique axée sur les stratégies interprétatives de l'imaginaire contemporain. Constituées par des internautes lors de leurs recherches, ces anthologies peuvent être publiées et rendues publiques sur le site d'Encodage, ouvrant ainsi une fenêtre sur les stratégies interprétatives et les processus inférentiels déployés dans l'étude de l'imaginaire contemporain.

Pouvez-vous choisir 2 réalisations particulièrement marquantes ou 2 enjeux scientifiques majeurs, que la mise en ligne de votre plateforme a permis d'atteindre ?

L'enjeu scientifique majeur du projet Encodage est double : a) exploiter et valoriser des bases de données diffusant de la recherche en sciences sociales et humaines, arts et lettres (SSHAL) et en humanités numériques, et offrir aux chercheurs de son écosystème et à ses utilisateurs une ressource capable de les appuyer dans l'exercice même de la recherche sur le contemporain; b) favoriser le développement de pistes interprétatives de cet imaginaire, grâce à l'examen et à l'acquisition d'une littératie numérique. Encodage suscite des anthologies thématiques en proposant une interface unique dédiée à la production d'hypothèses interprétatives sur le contemporain.

**Pouvez-vous évoquer la suite souhaitée, s'il y en a une ?
Développements à venir, nouvelles fonctionnalités, partenariats...**

Le développement d'Encodage a permis de faire une preuve de concept par le biais du prototype mis en ligne. La prochaine étape est de présenter et de faire adopter le processus d'anthologisation par des grandes bases de données, telles qu'Érudit.org au Québec, qui diffusent de la recherche et qui pourraient vouloir favoriser des stratégies de littératie numérique auprès de ses usagères et usagers.

AUTEUR

Plateforme Encodage

Cultures populaires/cultures
savantes

Quelle(s) légitimation(s) pour les littératures de genre et de grande diffusion ?

Questioning the legitimacy of genre popular literatures

Anne Besson

DOI : 10.35562/marge.881

Droits d'auteur

CC BY-NC-SA

PLAN

La légitimité, un enjeu de recherche collectif
Mobilité des hiérarchies
Création des hiérarchies internes
Critères de légitimation

TEXTE

La légitimité, un enjeu de recherche collectif

- 1 La question de la légitimité des objets de la recherche en cultures populaires, et donc du regard extérieur qui peut être porté sur eux, ne saurait être complètement ignorée – ne serait-ce qu'en pensant, par exemple, à la valeur dégradée ou non susceptible d'être associée aux travaux de thèse portant sur des genres et des médias « de grande diffusion ». Rappelons d'emblée, pour citer Bernard Lahire, que « la théorie de la légitimité culturelle¹ », issue d'« une tradition sociologique (en l'occurrence celle qui, de Thorstein Veblen à Pierre Bourdieu en passant par Edmond Goblot, analyse les fonctions sociales de la culture dans une société divisée en classes)² », « étudie essentiellement les distances et les rapports socialement différenciés à "la" culture³ ». Pour sa part, il redéfinit la légitimité comme « l'ensemble des rapports de force objectifs entre arbitraires culturels différents – qui dépendent eux-mêmes d'une multitude de

trajectoires institutionnelles, de politiques culturelles, de rapports entre l'État, le marché et les producteurs culturels⁴ ». On s'intéressera encore davantage ici, dans la lignée des *cultural studies* et de ce qu'elles remettent partiellement en cause des modèles distinctifs, à des dynamiques de distinction, à des pratiques de légitimation, fragmentées, concurrentes, loin de toute « légitimité » singulière et objectivée.

- 2 Si cette problématique n'est pas centrale dans mes travaux personnels (je n'adopte pas une position de défense de mes objets en essayant de faire d'eux plus ou autre chose que ce qu'ils ne sont : j'espère contribuer à mon échelle à leur légitimation, mais sans en faire un enjeu spécifique de mes travaux), elle occupe en revanche une place importante et récurrente dans les activités collectives de la communauté des chercheurs auxquels je m'associe⁵. Les rapports entre « cultures savantes » et « cultures populaires », avec l'ampleur que cela couvre, constituent en effet un très bel objet de réflexion sur les évolutions contemporaines à la fois des métiers de la recherche et des propositions culturelles qui nous entourent, les deux étant bien entendu liés – par exemple pour y voir une invitation à la réflexivité sur nos productions et sur la construction de nos savoirs. Cet article, appuyé sur plusieurs années de travaux au sein de l'équipe « LegiPop⁶ », est écrit du point de vue de la chercheuse en littératures de l'imaginaire que je suis. Il veut présenter, de manière évidemment très synthétique, un état du champ des recherches en littératures populaires et cultures médiatiques au prisme de leur légitimité au début de la décennie 2020, replacé dans le contexte plus large des évolutions des perceptions qui les ont affectées au cours de leur histoire, telles que les travaux spécialisés en rendent compte. Les mouvements principaux qui s'y dessinent, dans une situation d'explosion quantitative de l'offre de contenus, consistent en un double mouvement de déhiérarchisation, ou plutôt de mobilité accrue des hiérarchies culturelles, couplée à des effets de resegmentation interne. J'y consacrerai nos deux premières parties avant de proposer quelques pistes exploratoires concernant les nouveaux critères de valeur justifiant l'évolution, lente et partielle, de ces échelles de légitimité.

Mobilité des hiérarchies

- 3 Cela relève de l'évidence partagée : le contexte culturel de la postmodernité est marqué par une déhiérarchisation des pratiques « savantes » et « populaires » de plus en plus nette. Cette évolution, qui transforme par exemple radicalement la façon dont les médias français vont pouvoir rendre compte, depuis une quinzaine d'années, de phénomènes comme la sériephilie, le pop rock ou la littérature pour adolescents, s'est autorisée d'une série de facteurs, « hybridation croissante des univers culturels, particulièrement sensible chez les générations nées depuis l'après-guerre⁷ », « consécration de la culture juvénile, spectacularisation de certains aspects de la culture cultivée, développement de formes d'éclectisme culturel permettant des combinaisons plus nombreuses et plus variées, déclin du pouvoir distinctif de certaines pratiques culturelles comme la lecture⁸ ». En rendent compte dès le début des années 1990 les enquêtes dirigées par Olivier Donnat⁹, dont se sont saisi les sociologies de la culture et des médias en France depuis deux décennies, ce qui les a amenés à rejoindre les positionnements des *cultural studies* anglaises et américaines, beaucoup plus souples et favorables à une prise en compte de la diversité des pratiques de consommation. La culture médiatique surgit alors comme objet d'intérêt public, mais aussi académique, de façon certes progressive et hétérogène, mais de plus en plus rapide et massive.
- 4 En réalité, on le sait aussi, des hiérarchies culturelles demeurent ; elles se reconfigurent seulement, comme elles l'ont toujours fait¹⁰. Le populaire d'une époque n'est pas celui des suivantes, selon des dynamiques fort complexes qui voient des fortunes critiques se faire ou se défaire, des œuvres ou des genres s'élever dans les panthéons ou tomber dans l'oubli : ce qui plaisait le plus à une époque (le principe de textes adressés au « grand public » ne préexistant pas à l'existence de ce grand public au XIX^e) ne subsistera souvent qu'auprès des connaisseurs érudits – Gauvain était le héros préféré de l'arthurien tardif, mais pas celui qu'on a retenu ; les grands romans pastoraux du milieu XVII^e siècle, tant aimés, ne sont plus lus ni vraiment lisibles en dehors d'un cercle de chercheurs passionnés, pas plus que la masse du roman colonial ou du « roman de la victime », plus proche de nous au XIX^e siècle, pour des questions de sensibilités

très différentes à la fois à l'idéologie et au *pathos*. Pourtant, seules des œuvres continuent appréciées par un nombre de lecteurs suffisant pour en justifier la transmission et la réédition peuvent traverser le temps¹¹.

- 5 S'il y a bien entendu des cas de « dégradation » (c'est le terme que Gramsci emploie à propos d'Eugène Sue que ses prises de position politiques progressistes auraient amené d'auteur apprécié de tous, même des gens cultivés, à être identifié comme un auteur « pour le peuple¹² » uniquement), l'observation large montre que les genres et les médias connaissent des cycles de « montée en gamme » puis de resegmentation interne – c'est le vocabulaire du marketing que j'utilise ici à dessein à propos des produits culturels. Se produit donc d'abord une « montée en gamme », comme dans le cas des marques de vêtements par exemple quand leur croissance implique une hausse progressive en qualité et surtout en prix – le malheureux Adorno peut ainsi être raillé pour être passé complètement à côté du jazz, qu'il prend comme exemple de la pauvreté des musiques populaires, sans du tout percevoir son devenir de musique savante, réservée aujourd'hui à un public averti, voire à une élite mélomane vieillissante¹³.
- 6 Visant le même public, le même marché, non extensible à volonté, chaque nouveau média ou genre entre en concurrence avec ses prédécesseurs. Ils se concurrencent (en copiant des recettes de succès éprouvées), mais aucun ne disparaît pour autant. Ils se « poussent vers le haut » : un média qui appartenait à la consommation culturelle populaire est remplacé par un autre, désormais plus accessible. Une telle substitution s'est ainsi produite entre le théâtre et le cinéma, puis reproduite entre le cinéma et la télévision, et le lien peut être fait entre l'émergence de la *quality TV* et des plateformes de VOD quand se généralise le visionnage des nouveaux médias numériques (les vidéos YouTube) qui viennent à leur tour prendre la place précédemment occupée par les programmes télévisés, en bas de l'échelle de la légitimité des consommations audiovisuelles.
- 7 L'élévation est aussi celle des exigences des publics et de leurs expertises – plus on fréquente un genre ou une mécanique narrative codifiée¹⁴, plus on devient sensible à des effets de variation,

apte à lire et à décoder des propositions plus complexes (les jeux sur les niveaux ontologiques ou les lignes chronologiques désormais très enchevêtrés¹⁵). Mais il reste toujours une partie des publics qui vise, au moins parfois, à se détendre et non à être stimulé, et il reste toujours une position mineure à occuper dans les hiérarchies ; si on visualise une pyramide, ce sera la zone la plus vaste, la plus fréquentée, mais aussi la plus dégradée.

Création des hiérarchies internes

- 8 Mais cette vision est encore trop simple, en raison de la resegmentation qui affecte tout marché suffisamment développé pour se voir encore étendu par des opérations de ciblage interne. Or l'offre culturelle explose (nous sommes à « l'ère de l'hyperchoix¹⁶ »), elle se diversifie, par exemple par division des genres littéraires et médiatiques, ou par classe d'âge – littérature ado, *young adult* – et s'y recréent donc des hiérarchies internes. Ce ne sont pas toutes les séries télévisées qui sont désormais encensées de manière consensuelle par *Les Cahiers du cinéma* et de très bons groupes de chercheurs en France. Certes, cette revalorisation concerne nombre des séries longtemps confinées dans les marges du *fandom* (tels *Star Trek* ou *Buffy contre les vampires*), mais pas, entre autres exemples, les séries familiales françaises qui n'intéressent malheureusement guère en dehors de rares analyses sociologiques¹⁷ – on imagine mal un colloque sur *Joséphine, ange gardien* (diffusée depuis 1997) ou *Camping paradis* (depuis 2006). De façon similaire, si on peut entendre des voix défendre l'intérêt d'une analyse esthétique des *soap operas*, c'est pour constater que « personne ou quasiment personne¹⁸ » ne le fait, notamment parce que pour cela il faudrait commencer par les regarder, ces séries, qui sont quotidiennes et s'étendent sur des dizaines d'années. On voit ici apparaître un seuil d'acceptabilité (celui des goûts des plus cultivés, des « omnivores¹⁹ »), observable également dans les corpus littéraires sur le cas de la romance ou roman sentimental : non seulement très peu de recherches portent sur ces corpus²⁰, mais, en dépit du très beau travail de modernisation des contenus et des supports menés par Karine Lanini pour Harlequin France²¹, cela reste une lecture cachée, un goût honteux. Sans du tout prétendre que tout se vaut, loin de là, il me semble qu'une telle autocensure impose

de réfléchir à ce qui nous autorise individuellement et collectivement à décider de ce qui « vaut » ou non, à nos critères de valeur donc ; en effet, un tel exemple montre très clairement combien ceux-ci peuvent constituer la traduction et la reconduction, dans la sphère culturelle, d'effets d'exclusion sociaux, en fonction du genre (évidemment féminin), mais aussi de la classe (la ménagère faiblement diplômée), de l'âge (des lectrices et spectatrices), de l'habitat non urbain ²², etc.

- 9 C'était là une première segmentation qui même au sein d'une forme médiatique en voie de légitimation (la série télévisée par exemple) va en laisser de côté des pans entiers qui correspondent à des catégories de publics délégitimés – la logique des industries culturelles (proposer à chacun le produit qu'il attend) rejoint des facteurs idéologiques et sociologiques. Ce même phénomène peut encore s'observer à échelle plus réduite, au sein des *fandoms* eux-mêmes, qui afin de légitimer leurs goûts, afin de permettre la « montée en gamme » de leur domaine d'expertise, opèrent eux-mêmes des sélections et des exclusions. Parmi les amateurs de science-fiction, la liste et le palmarès sont des exercices extrêmement prisés – formuler une telle proposition permet d'ouvrir la discussion, et aussi à chacun de se positionner, d'affirmer certaines particularités d'une identité de lecteur ; à l'autre extrémité toutes les littératures de genre ont leurs infréquentables, mais aussi leurs masses oubliées, qui suscitent des entreprises d'éruditions, encyclopédiques et exhaustives, comparables à celles portant sur les romans des siècles classiques ²³ dont il a été question plus haut.
- 10 De manière significative, la question de savoir ce qu'on inclut ou non dans le genre est au cœur des querelles sur la définition même de la science-fiction qui agitent régulièrement le milieu des passionnés comme des chercheurs – des définitions plus restrictives, mettant en avant des critères de plausibilité scientifique, la conjecture, la spéculation, s'opposent à des définitions pragmatiques (ce qui se publie sous ce nom) beaucoup plus accueillantes à tout un pan du genre moins exigeant et plus exubérant, longtemps moins présentable.

Critères de légitimation

- 11 S'il faut donc déconstruire l'idée trop générale d'une déhiérarchisation et même d'une « simple » légitimation de genres et médias longtemps déconsidérés, au profit de processus plus complexes, cette évolution n'en est pas moins indéniable, et il convient dès lors, dans un dernier temps de cette synthèse, d'essayer de l'expliquer par la montée en puissance de certains critères de prise en compte des œuvres. Ils viendraient s'opposer aux critères « romantiques » puis « modernistes » encore largement présents à l'arrière-plan de nos opérations d'évaluation – et selon lesquels une grande œuvre se caractérise par son épaisseur ou sa densité herméneutique²⁴ : elle résiste à l'analyse, elle est opaque, elle est complexe, la texture de son style est inséparable de sa richesse qui dépasse son seul contenu. De tels critères expliquent les résistances qui demeurent (il ne s'agit pas de le nier), face à l'intégration des œuvres populaires dans l'horizon savant, et ils figurent aussi à l'arrière-plan des opérations de rehiérarchisation dont il a été question. Mais ils coexistent désormais avec d'autres forces. Par exemple, alors que Nathalie Heinich a bien montré comment la singularité (de l'artiste) s'invente au cours du XIX^e siècle en opposition à la célébrité (du commerce²⁵), si bien que cette dernière « importante dans les faits » est cependant « stigmatisante dans les évaluations » (l'auteur à succès est suspect de complaisance), la logique de distinction en faisant une « anti-valeur dès l'époque romantique, avec le basculement de l'art en régime de singularité²⁶ », la valorisation du succès commercial semble davantage de mise désormais dans les opérations d'évaluation par les publics contemporains²⁷. Ainsi, un phénomène tel que *Harry Potter*, au succès historiquement inégalé, a entraîné dans son sillage une reconfiguration massive du secteur éditorial dans son ensemble, rééquilibré en faveur des littératures de l'imaginaire d'une part, des jeunes publics d'autre part et surtout – les collections pour la jeunesse, qui se sont imposées chez tous les éditeurs, sont les plus rentables, celles qui croissent le plus régulièrement au fil des études annuelles consacrées au marché du livre. Il y a bien des voix qui s'élèvent pour s'inquiéter que cette offre spécifique détourne les enfants des classiques patrimoniaux (c'est le cas) et qu'elle attire

même les adultes dans un nivellement par le bas²⁸ ; reste que l'expansion de ce domaine a produit son propre panthéon et qu'il est difficile aujourd'hui de ne pas se réjouir quand les enfants lisent, et ce quoi qu'ils lisent ou peu s'en faut.

12 En outre, dans ces mêmes décennies, l'édition jeunesse est devenue un domaine absolument essentiel à la survie du secteur tout entier : cette nouvelle conscience de la fragilité de l'économie culturelle constitue un autre facteur à prendre en compte, dans la mesure où le double mouvement contraire d'explosion de l'offre et de réduction des budgets implique qu'atteindre la rentabilité et rencontrer le succès constituent des objectifs partagés dont on peut accepter qu'il faille en payer le prix, et que cela implique d'aller à la rencontre du public. On peut en rapprocher, du côté de la recherche, l'injonction à la fois managériale et citoyenne à l'utilité – à quoi servons-nous, qui servons-nous ? Elle peut prendre la forme d'une case à remplir dans d'incontournables formulaires de financement (« quel sera l'impact sociétal de ce programme de recherche ? »), mais aussi celle, plus positive, d'une réflexion sur la mise à jour des pratiques d'enseignement (« qu'est-ce qui intéresse mes étudiants aujourd'hui et qu'est-ce qui leur sera utile demain en tant qu'enseignants ou même en tant que consommateurs culturels avertis ? »).

13 En outre, il est en réalité presque impossible d'établir une délimitation nette entre des choix relevant de l'intérêt financier, de la volonté de démocratisation culturelle ou de la demande du public. Je prendrai en exemple le travail que j'ai mené avec la BnF en 2018-2020, en arrière-plan de l'exposition Tolkien²⁹, puis comme directrice scientifique du site internet « Fantasy, retour aux sources³⁰ » et de la saison d'événements culturels qui en a accompagné l'ouverture en janvier 2020. Dans ce projet qui a donné pleine satisfaction à l'équipe d'organisation, l'établissement public met à disposition, sur place et en ligne, des ressources de qualité – mais sans pour autant faire fi d'objectifs de fréquentation qui ont été majeurs de bout en bout : choisir d'exposer Tolkien ou la fantasy, c'est bien entendu, pour la BnF (qui vaut ici pour une politique absolument généralisée des institutions muséales et patrimoniales), chercher à capter des publics nouveaux, plus jeunes et plus nombreux que d'habitude. L'objectif est bien commun avec les démarches commerciales, même s'il s'agit, là aussi, d'une contribution à la démocratisation culturelle et d'un

accompagnement intelligent de l'évolution des pratiques de consommation – la BNF est ainsi à la pointe de la réflexion sur la conservation des jeux vidéo et des jeux en ligne.

- 14 À l'autre extrémité du spectre, au cœur de la toute-puissance de l'industrie culturelle, on constate la marchandisation croissante des univers de fiction, en réponse à une demande des publics. Pour reprendre le cas de *Harry Potter*, le développement d'une offre de très beaux produits dérivés massivement vendus est une caractéristique de la nouvelle phase de développement du Wizarding World depuis 2015-2016³¹, et s'accompagne d'un ensemble d'inégalités d'accès (géographiques et financières) à un univers jusque-là très bien partagé – avec l'apparition des parcs à thème, des jeux géolocalisés, dans une moindre mesure la multiplication des éditions et des beaux livres, selon l'endroit où l'on habite et ses moyens financiers, on pourra ou non prétendre à une expérience plus ou moins complète. Le déploiement de ces propositions dans la culture matérielle et le monde réel n'en répond pas moins à des attentes très fortes de la part des amateurs qui aspirent à cette « réalisation » concrète sans pouvoir toujours la produire eux-mêmes d'une manière satisfaisante³².
- 15 Une dernière piste, qui elle aussi demande à être pensée de manière nuancée, peut enfin être proposée pour expliquer l'intérêt nouveau porté aux fictions de genre et de grande consommation : le retour de l'évaluation éthique et politique, ce que je propose d'appeler le critère de pertinence³³ – ce que les œuvres ont à nous dire, en quoi elles sont, selon la définition de ce terme, « appropriées à leur objet ». Il me semble en effet que les œuvres (et notamment les fictions de l'imaginaire, alors qu'il y a une forme de paradoxe à les juger à cette aune-là), doivent désormais pouvoir être créditées d'une valeur *d'usage*, qui prend en compte la manière dont ces projections imaginaires peuvent *faire sens* pour leur récepteur – et plus les consommateurs sont nombreux, plus décisif sera l'impact. Il ne s'agit plus seulement d'aimer une histoire ou une activité, mais de savoir *pourquoi*³⁴ – d'estimer donc ce qu'elles nous apportent. Ce critère nous fait explicitement basculer du côté de « l'hétéronomie » (la valeur de l'œuvre est déterminée par des critères qui lui sont extérieurs), là où la revendication d'autonomie a été consubstantielle à la définition d'un « champ littéraire restreint³⁵ » par Pierre

Bourdieu, et demeure profondément liée aux conceptions les plus ancrées de l'art et de la littérature. On constate, pour certains avec effroi³⁶, un retour de l'évaluation morale (des sujets et des manières de les traiter) et de l'engagement civique ou politique attendu des fictions, et ce dans tous les domaines de production culturelle. Dans ceux qui m'occupent, science-fiction et fantasy, la poussée des préoccupations en faveur du libéralisme culturel (défense et représentation des diversités et des minorités) et de l'engagement écologique qui les accompagnent, est absolument frappante en l'espace d'une dizaine d'années³⁷ – les autres possibles qu'entendent représenter les genres de l'imaginaire les prédisposent en effet à de tels usages : passés merveilleux aux peuples divers et à la nature préservée, avènements menaçants de la post-apocalypse ou encore alternatives aux assignations binaires de genre, les œuvres sont évaluées à la fois dans le genre, pour ce qu'elles lui apportent, et pour la manière dont elles en mettent à jour les problématiques et l'idéologie. Ces nouveaux enjeux entraînent bien entendu à leur tour des réévaluations rétrospectives des « canons » antérieurs, visant par exemple à désinvisibiliser certaines autrices oubliées ou à contraire à déboulonner certaines statues contestables.

- 16 Si la fantasy est d'ores et déjà, pour ces raisons comme par l'ampleur et la pérennité de son succès, le genre majeur pour le jeune public du ^{xxi}e siècle³⁸, peut-être le restera-t-elle pour la manière dont son impact a modifié les critères mêmes selon lesquels nous jugeons la valeur des fictions – réévaluation du succès commercial et retour de l'idéologie pour un jugement éthique et politique sur la pertinence des œuvres, les deux critères de valeur qui me semblent expliquer l'évolution, toute relative et progressive, de nos hiérarchies culturelles convergent ici pour un genre dont les avatars successifs, portés par leur audience, reflètent les évolutions des attentes générationnelles³⁹. Mais me voilà prise en flagrant délit de légitimation de mon objet : il est temps de clore cet article.

BIBLIOGRAPHIE

ARTIAGA Loïc et LETOURNEUX Matthieu, *Aux origines de la pop culture. Le Fleuve Noir et les Presses de la Cité au cœur du transmédia à la française, 1945-1990*, Paris, éditions

La Découverte, 2022.

BARONI Raphaël et MACÉ Marielle (dir.), *Le Savoir des genres*, Rennes, PUR, 2006.

BESSON Anne, *Constellations. Des mondes fictionnels dans l'imaginaire contemporain*, Paris, CNRS Éditions, 2015.

BESSON Anne, *Les Pouvoirs de l'enchantement. Usages éthiques et politiques de la science-fiction et de la fantasy*, Paris, Vendémiaire, 2021.

BESSON Anne, FRANÇOIS Anne Isabelle, LÉCOSSAIS Sarah, LETOURNEUX Matthieu et WEBER Anne-Gaëlle (dir.), « Mutations des légitimités dans les productions culturelles contemporaines », *Belphégor* n°17/1 [en ligne], 2019, DOI : [10.4000/belphegor.1450](https://doi.org/10.4000/belphegor.1450).

BOUGON Marie-Lucie, *La Fantasy en France : de l'importation à l'appropriation*, thèse de doctorat, sous la direction d'A. Besson, université d'Artois, 2022.

BOURDIEU Pierre, *Les Règles de l'art. Genèse et structure du champ littéraire*, Paris, Seuil, 1992.

CHEDALEUX Delphine, « Construire un regard sur la réception de *Cinquante Nuances de Grey* : une ethnographie en ligne », *Poli – Politique de l'Image* n° 14, Paris, Poli éditions, 2018, p. 82-91.

CHEDALEUX Delphine, *Du savon aux larmes. Le soap opera, une subculture féminine*, Paris, éditions Amsterdam, coll. « Les prairies ordinaires », 2022.

COULANGEON Philippe, *Les métamorphoses de la distinction. Inégalités culturelles dans la France d'aujourd'hui*, Paris, Grasset, 2011,

COULANGEON Philippe, *Culture de masse et société de classe. Le goût de l'altérité* [en ligne], Paris, PUF, 2021, DOI : [10.3917/puf.coula.2021.01](https://doi.org/10.3917/puf.coula.2021.01).

CONSTANS Ellen, *Parlez-moi d'amour. Le roman sentimental : des romans grecs aux collections de l'an 2000*, Limoges, Presses universitaires de Limoges, 1999.

COSTES Guy et ALTAIRAC Joseph, *Rétrofictions : Encyclopédie de la conjecture romanesque rationnelle francophone, de Rabelais à Barjavel, 1532-1951*, Amiens, Encrage, coll. « Les Belles lettres », 2018.

DONNAT Olivier, *Les Français face à la culture. De l'exclusion à l'éclectisme*, Paris, La Découverte, 1994.

DURAND Emmanuel, *L'Attaque des clones. La diversité culturelle à l'ère de l'hyperchoix*, Paris, Presses de Sciences Po, 2016.

DURAND Pascal, « Culture populaire, culture de masse ou culture de mass-médias ? Autour de cinq thèses moins une d'Antonio Gramsci », *Quaderni* [en ligne], n° 57, 2005, p. 73-83, DOI : [10.3406/quad.2005.1662](https://doi.org/10.3406/quad.2005.1662).

FALCONER Rachel, *The Crossover Novel. Contemporary Children's Fiction and its Adult Readership*, New York/Londres, Routledge, 2008.

HEINICH Nathalie, *L'élite artiste. Excellence et singularité en régime démocratique*, Paris, Gallimard, 2005.

HEINICH Nathalie, « La célébrité », in : Nathalie HEINICH, Jean-Marie SCHAEFFER, Carole TALON-HUGON (dir.), *Par-delà le beau et le laid : enquêtes sur les valeurs de l'art*, Rennes, PUR, 2014, p. 33-44.

JOUE Vincent, *Pouvoirs de la fiction : pourquoi aime-t-on les histoires ?*, Paris, Malakoff, Armand Colin, 2019.

LAHIRE Bernard, « La légitimité culturelle en questions », in : Olivier DONNAT (dir.), *Regards croisés sur les pratiques culturelles* [en ligne], Paris, ministère de la Culture – Département des études et de la prospective, coll. « Questions de culture », 2003, p. 39-62, DOI : [10.3917/deps.donna.2003.01.0039](https://doi.org/10.3917/deps.donna.2003.01.0039).

LAHIRE Bernard, *La Culture des individus. Dissonances culturelles et distinction de soi*, Paris, La Découverte, 2004.

LÉCOSSAIS Sarah, *Chroniques d'une maternité hégémonique. Identités féminines, représentations des mères et genre de la parentalité dans les séries télévisées familiales françaises (1992-2012)*, thèse de doctorat, sous la direction de E. Maigret, Paris-Sorbonne Nouvelle, 2015.

LETOURNEUX Matthieu, *Fictions à la chaîne. Littératures sérielles et culture médiatique*, Paris, Seuil, coll. « Poétique », 2017.

MEIZOZ Jérôme, *Postures littéraires. Mises en scène modernes de l'auteur*, Genève, Slatkine Erudition, 2007.

MITTELL Jason, *Complex TV. The Poetics of Contemporary Television Storytelling*, New York, New York University Press, 2015.

MORETTI Franco, *Graphs, Maps, Trees: Abstract Models for a Literary History*, Londres/New York, Verso, 2005.

NICOLAS Agathe, *La grande saga de l'industrialisation de la fiction : le renouveau créatif de la franchise Harry Potter*, thèse de doctorat, sous la direction d'A. Wrona, Sorbonne université, Paris, 2019.

PASQUIER Dominique, « La « culture populaire » à l'épreuve des débats sociologiques », *Hermès* [en ligne], n° 42, CNRS éditions, 2005, p. 60-69, DOI : [10.4267/2042/8983](https://doi.org/10.4267/2042/8983).

PÉQUIGNOT Bruno, *La relation amoureuse. Analyse sociologique du roman sentimental moderne*, Paris, L'Harmattan, coll. « Logiques sociales », 1991.

PETERSON Richard A., « Understanding Audience Segmentation. From Elite and Mass to Omnivore and Univore », *Poetics* [en ligne], n° 21, 1992, p. 243-258, DOI : [10.4267/2042/8983](https://doi.org/10.4267/2042/8983).

SCHAEFFER Jean-Marie, « Système, histoire et hiérarchie : le paradigme historiciste en théorie de l'art », in : Georges ROQUE (dir.), *Majeur ou mineur ? Les hiérarchies en art*, Nîmes, Jacqueline Chambon, 2000, p. 255-271.

SCHAEFFER Jean-Marie, *Pourquoi la fiction ?*, Paris, Seuil, 1999

SCHLANGER Judith, *La Mémoire des œuvres*, Paris, Verdier, 2008.

THOUVENEL Éric, « Esthétique (et herméneutique) du soap », in : Fabien BOULLY (dir.), *Troubles en série. Les séries télé en quête de singularité*, Nanterre, Presses universitaires de Paris-Nanterre, 2020, p. 307-316.

WILLIAMS Patrick, « Le déni d'Adorno », *L'Homme* [en ligne], n° 175-176, 2005, p. 61-74, DOI : [10.4000/lhomme.29590](https://doi.org/10.4000/lhomme.29590).

NOTES

1 Bernard Lahire, *La Culture des individus. Dissonances culturelles et distinction de soi*, Paris, La Découverte, 2004.

2 Loc. cit.

3 Loc. cit.

4 Bernard Lahire, « La légitimité culturelle en questions », in : Olivier Donnat (dir.), *Regards croisés sur les pratiques culturelles*, Paris, ministère de la Culture – Département des études et de la prospective, « Questions de culture », 2003, p. 39-62, cit. p. 39.

5 Communauté qui coïncide largement avec les membres de l'Association internationale des chercheurs en littératures populaires et culture médiatique (LPCM [en ligne], URL : <https://lpcm.hypotheses.org> [consulté le 8 avril 2024]). L'association est notamment à l'initiative de l'organisation de premières « Assises de la recherche en cultures populaires et médiatiques : enjeux, outils, méthodes », qui se sont tenues du 11 au 13 octobre 2018 à l'université Paris-Nanterre et à l'université Paris-Sorbonne Nouvelle.

6 Voir les travaux développés au sein du « projet structurant » régional « LegiPop » (MESHS-Lille Nord de France, 2017-2018, *LegiPop* [en ligne], URL : <https://legipop.hypotheses.org/> [consulté le 8 avril 2024]). On se proposait d'y analyser les processus et critères de légitimation, terme choisi pour bien marquer la dynamique permanente des modalités d'évaluation ainsi que la nécessaire historicisation des genres ou domaines « légitimés » ou « délégitimés » ; et ce, sous l'angle des savoirs développés, « savants » et/ou « populaires », au sujet des uns ou des autres : autrement dit, la double dynamique de vulgarisation des savoirs savants, qui apparaît toujours plus nécessaire dans la réflexion sur l'avenir des études littéraires dans lesquelles je m'inscris, et de développement d'une expertise amateur,

portée par les communautés de fans, en réalité ancienne, mais désormais massive, sur les objets populaires.

Ont été organisés un colloque (« Des savants amateurs aux fans experts : contours et évolutions de l'érudition, entre populaire et académique », organisé par Anne Besson et Anne-Gaëlle Weber, 16-17 novembre 2017, MESHS de Lille), deux journées d'étude (« (Dé)construire la légitimité », organisé par Anne Besson et Matthieu Letourneux, 16 mars 2018, MESHS de Lille ; « Les nouvelles formes de la vulgarisation et de l'écriture du savoir », organisé par Anne-Isabelle François et Sarah Lécossais, 15 juin 2018, université Paris-Sorbonne Nouvelle), deux panels durant les « Assises de la recherche en cultures populaires et médiatiques », le 13 octobre 2018, et enfin un « colloque doctoral international » en collaboration entre les universités d'Artois, de Lausanne et de Paris-Nanterre, soulignant la pertinence de ce questionnement pour la jeune recherche (décalé à plusieurs reprises en raison de la crise sanitaire il a eu lieu les 9-10 septembre 2021).

Une sélection des travaux, sur lesquels je m'appuie ici, est parue dans la revue *Belphégor* n° 17/1, 2019, dossier coordonné par Anne Besson, Anne Isabelle François, Sarah Lécossais, Matthieu Letourneux et Anne-Gaëlle Weber, « Mutations des légitimités dans les productions culturelles contemporaines », *Belphégor* [en ligne], n° 17/1, 2019, DOI : [10.4000/belphegor.1450](https://doi.org/10.4000/belphegor.1450).

7 Dominique Pasquier, « La Culture populaire à l'épreuve des débats sociologiques », *Hermès* n° 42, 2005, p. 60-69, p. 64.

8 *Loc. cit.*

9 Olivier Donnat, *Les Français face à la culture. De l'exclusion à l'éclectisme*, Paris, La Découverte, 1994.

10 C'est ce que soulignent avec force les ouvrages de Philippe Coulangeon, attachés à rappeler l'importance de la société de classes dans les goûts et pratiques culturels : *Les métamorphoses de la distinction. Inégalités culturelles dans la France d'aujourd'hui*, Paris, Grasset, 2011 ; *Culture de masse et société de classe. Le goût de l'altérité*, Paris, PUF, 2021.

11 Voir par exemple à ce sujet, pour une réflexion surplombante, Judith Schlanger, *La Mémoire des œuvres* [1992], Paris, Verdier, 2008, ou pour une approche quantitative, Franco Moretti, *Graphs, Maps, Trees: Abstract Models for a Literary History*, Londres, New York, Verso, 2005.

- 12 Fragment des *Cahiers de prison* [en ligne] d'Antonio Gramsci, 1934-1935 : « Les contemporains ne s'aperçoivent pas toujours de la dégradation d'une partie de ces manifestations littéraires, comme cela s'est produit partiellement pour Eugène Sue, qui a été lu par tous les groupes sociaux et qui « émouvait [»] « même les gens cultivés [»] [...] tandis qu'il tomba ensuite au rang « d'écrivain lu seulement par le peuple » (la « première lecture [»] [...] donne simplement, ou presque, des sensations, qu'elles soient « culturelles » ou de contenu, et le « peuple » est un lecteur de première lecture, sans attitude critique, chez qui l'émotion naît de la sympathie qu'il éprouve pour l'idéologie générale dont le livre est l'expression souvent artificielle et voulue) », URL : <https://www.marxists.org/francais/gramsci/intell/intell3.htm> (consulté le 8 avril 2024). La ponctuation a été harmonisée par l'autrice de l'article pour plus de lisibilité. Voir aussi Pascal Durand, « Culture populaire, culture de masse ou culture de mass-médias ? Autour de cinq thèses moins une d'Antonio Gramsci », *Quaderni* [en ligne], n° 57, 2005, « Gramsci, les médias et la culture », p. 73-83, DOI : [10.3406/quad.2005.1662](https://doi.org/10.3406/quad.2005.1662).
- 13 Dans une série d'articles, dont « Sur la musique populaire » [1937], trad. Marie-Noëlle Ryan, Peter Carrier et Marc Jimenez dans la *Revue d'Esthétique*, n° 19 « Jazz », 1991. Voir pour une synthèse Patrick Williams, « Le déni d'Adorno », *L'Homme*, n° 175-176 ; 2005, DOI : [10.4000/lhomme.29590](https://doi.org/10.4000/lhomme.29590).
- 14 Voir Raphaël Baroni et Marielle Macé (dir.), *Le Savoir des genres*, Rennes, PUR, 2006 ; et Matthieu Letourneux, *Fictions à la chaîne. Littératures sérielles et culture médiatique*, Paris, Seuil « Poétique », 2017.
- 15 Jason Mittell, *Complex TV, The Poetics of Contemporary Television Storytelling*, New York, New York University Press, 2015.
- 16 Emmanuel Durand, *L'Attaque des clones. La diversité culturelle à l'ère de l'hyperchoix*, Paris, Presses de Sciences Po, 2016.
- 17 Par exemple la thèse de Sarah Lécossais, *Chroniques d'une maternité hégémonique. Identités féminines, représentations des mères et genre de la parentalité dans les séries télévisées familiales françaises (1992-2012)*, sous la direction d'E. Maigret, Paris III-Sorbonne Nouvelle, 2015.
- 18 Éric Thouvenel, « Esthétique (et herméneutique) du soap », in : Fabien Bouilly (dir.), *Troubles en série. Les séries télévisées en quête de singularité*, Presses Universitaires de Paris-Nanterre, 2020, p. 307-316, cit. p. 307 (« qui en a visionné un de manière intégrale [...] ? Personne, ou

quasiment personne [l'auteur de ces lignes ne faisant pas exception à la règle]. »).

19 Richard Peterson, « Understanding Audience Segmentation. From Elite and Mass to Omnivore and Univore », *Poetics*, n° 21, 1992, p. 243-258.

20 Peu de travaux ont pris la suite de Bruno Péquignot, côté sociologie (*La relation amoureuse. Analyse sociologique du roman sentimental moderne*, Paris, L'Harmattan « Logiques sociales », 1991) et d'Ellen Constans, côté littérature (*Parlez-moi d'amour. Le roman sentimental : des romans grecs aux collections de l'an 2000*, Presses Universitaires de Limoges, 1999). On peut citer Delphine Chedaleux, « Construire un regard sur la réception de *Cinquante Nuances de Grey* : une ethnographie en ligne », *Poli - Politique de l'Image*, Poli éditions, 2018, p. 82-91.

21 Voir son interview, URL : https://www.lepoint.fr/livres/karine-lanini-harlequin-colle-aux-evolutions-de-la-societe-16-03-2018-2203102_37.php [consulté le 8 avril 2024].

22 Autant de questions qu'aborde l'ouvrage de Delphine Chedaleux, *Du savon aux larmes. Le soap opera, une subculture féminine*, Paris, Amsterdam, coll. « Les prairies ordinaires », 2022.

23 Guy Costes et Joseph Altairac, *Rétrofictions: Encyclopédie de la conjecture romanesque rationnelle francophone, de Rabelais à Barjavel, 1532-1951*, Amiens, Encrage, 2018.

24 Jean-Marie Schaeffer, « Système, histoire et hiérarchie : le paradigme historiciste en théorie de l'art », in : Georges Roque (dir.), *Majeur ou mineur ? Les hiérarchies en art*, Nîmes, Jacqueline Chambon, 2000.

25 Nathalie Heinich, *L'élite artiste. Excellence et singularité en régime démocratique*, Paris, Gallimard, 2005.

26 Nathalie Heinich, « La célébrité », in : Nathalie Heinich, Jean-Marie Schaeffer, Carole Talon-Hugon (dir.), *Par-delà le beau et le laid : enquêtes sur les valeurs de l'art*, Rennes, PUR, 2014, p. 33-44, cit. p. 40.

27 En réalité, Loïc Artiaga et Matthieu Letourneux, dans *Aux origines de la pop culture. Le Fleuve Noir et les Presses de la Cité au cœur du transmédia à la française, 1945-1990*, Paris, éditions La Découverte, 2022, montrent que ce modèle « artiste », inaccessible, est remplacé dans les discours médiatiques des auteurs du *Fleuve Noir* par d'autres « postures », celles « de l'artisan modeste ou du forçat de l'écriture » (p. 68), du « petit entrepreneur à succès » (p. 67), voire pour certains de la « vedette » (p. 136) – chapitres

« Face à la littérature légitimée, l'affirmation de la posture de l'écrivain populaire », p. 65-68, et « Postures de romanciers, parmi les vedettes », p. 133-139. L'analyse en termes de « postures » vient de Jérôme Meizoz, avec son essai *Postures littéraires. Mises en scène modernes de l'auteur*, Genève, Slatkine Erudition, 2007.

28 De telles critiques ont accompagné la prise de conscience du succès transgénérationnel des *Harry Potter*, autour de l'an 2000 : voir l'introduction de Rachel Falconer, *The Crossover Novel. Contemporary Children's Fiction and its Adult Readership*, New York, Londres, Routledge, 2009, notamment p. 2-4 sur l'accusation d'infantilisation. Un bon résumé des positions au début des années 2010 est donné par le court éditorial « Adult should read Adult Books » de Joel Stein (*New York Times*, 29 mars 2012), qui s'accompagne, dans la version en ligne du journal, d'une section « room for debate » fournie, sous le titre « The Power of Young Adult Fiction » (URL : <http://www.nytimes.com/roomfordebate/2012/03/28/the-power-of-young-adult-fiction> [consulté le 8 avril 2024]).

En France, la journaliste et polémiste Natacha Polony s'est fait une spécialité du combat symétrique, pour de « vraies » lectures devant être proposées à une jeunesse « gâchée » — avec déjà en 2005 *Nos enfants gâchés. Petit traité sur la fracture intergénérationnelle*, Paris, JC Lattès, 2005, puis une série d'articles développant cet angle, comme « “Les jeunes lisent” ou le prototype de l'escroquerie intellectuelle », *Le Figaro* [en ligne], 02/12/2010, mis à jour le 13/06/2013, URL : <https://www.lefigaro.fr/blogs/education/2010/12/-les-jeunes-lisent-ou-le-prototype-de-lescroquerie-intellectuelle.html> [consulté 8 avril 2024] ; ou encore « La lecture est inégalitaire, supprimons la lecture : quand l'école française marche sur la tête », *Marianne* [en ligne], 15/03/2003, URL : <https://www.marianne.net/agora/les-signatures-de-marianne/la-lecture-est-inegalitaire-supprimons-la-lecture-quand-lecole-francaise-marche-sur-la-tete> [consulté le 8 avril 2024].

29 L'exposition « Tolkien, voyage en Terre du Milieu » s'est tenue à la Bibliothèque nationale de France du 22 octobre 2019 au 16 février 2020. URL : <https://www.bnf.fr/fr/agenda/tolkien-voyage-en-terre-du-milieu> [consulté le 8 avril 2024].

30 URL : <https://fantasy.bnf.fr/fr/accueil/> [consulté le 8 avril 2024].

31 Voir la thèse d'Agathe Nicolas, *La grande saga de l'industrialisation de la fiction : le renouveau créatif de la franchise Harry Potter*, thèse de doctorat, sous la direction d'A. Wrona, 2019, Sorbonne université (Paris).

- 32 Anne Besson, *Constellations. Des mondes fictionnels dans l'imaginaire contemporain*, Paris, CNRS éditions, 2015, partie 2 « Usages des mondes », chapitre 3 « What if ? Au contact des autres mondes », p. 403 sq.
- 33 Anne Besson, *Les Pouvoirs de l'enchantement. Usages éthiques et politiques de la science-fiction et de la fantasy*, Paris, Vendémiaire, 2021.
- 34 C'est la fameuse question-titre des ouvrages de Jean-Marie Schaeffer, *Pourquoi la fiction ?*, Paris, Seuil, 1999, de Jerome Bruner, *Pourquoi nous racontons-nous des histoires ?*, Yves Bonin (trad.), Paris, Retz, 2002, ou encore récemment Vincent Jouve, *Pouvoirs de la fiction : pourquoi aime-t-on les histoires ?*, Paris, PUF, 2019.
- 35 Pierre Bourdieu, *Les Règles de l'art. Genèse et structure du champ littéraire*, Paris, Seuil, 1992.
- 36 On peut prendre en exemple la trajectoire intellectuelle de Nathalie Heinich : sociologue de l'art s'étant beaucoup penchée, aux côtés notamment de Carole Talon-Hugon, sur la question des rapports entre art, littérature et morale, et sur leur historicisation, elle s'inquiète dès 2010 des *cultural wars*, alors largement limitées aux États-Unis (*Guerre culturelle et art contemporain*, Paris, Hermann, 2010). La montée, en France, des positions qu'elle qualifie de *woke* l'entraîne à publier ces dernières années plusieurs essais polémiques témoignant de l'inquiétude d'une part des intellectuels français devant le retour d'un « engagement » dont ils ne comprennent ni ne soutiennent les formes nouvelles : *Ce que le militantisme fait à la recherche*, Paris, Gallimard, coll. « Tracts », 29, 2021, *Défendre l'autonomie du savoir*, Paris, Fondation pour l'innovation politique, 2021, et enfin *Le Wokisme serait-il un totalitarisme ?*, Paris, Albin Michel, 2023.
- 37 On comprend qu'il ne s'agit pas des mêmes combats que portaient antérieurement, de manière continue depuis les années 1970 en France, polars et romans noirs d'une part, science-fiction de l'autre, ainsi qu'une part de la littérature blanche (ou du cinéma d'auteur) témoignant d'une fibre sociale.
- 38 Apparue dans l'horizon culturel français au tournant du siècle seulement, la fantasy a été portée par une série de « phénomènes » (livres et films *Harry Potter*, livres et films *Le Seigneur des Anneaux*, livres et séries télévisées *Game of Thrones*, aux chiffres d'audience records) – voir la thèse de Marie-Lucie Bougon, *La Fantasy en France : de l'importation à l'appropriation*, thèse de doctorat, sous la direction d'A. Besson, 2022, l'université d'Artois, Arras.

Les « littératures de l'imaginaire » étaient ainsi placées au premier rang des lectures choisies par les collégiens et lycéens selon l'enquête Ipsos pour le Centre National du Livre de juin 2016. Pour les « Genres de livres/romans lus pour les loisirs » les « Livres de SF et littératures de l'imaginaire » sont au premier rang, avec 42% de réponses, devant « Romans d'aventures » à 41% et « Histoire de famille/Vie quotidienne » à 32%. Si l'on entre davantage dans le détail, cette première place est également valable pour les collégiens à 47%, mais surtout grâce aux réponses des garçons, qui les plébiscitent à 55% ; pour les filles, l'imaginaire est 2^e à 44% derrière « Histoire de famille/Vie quotidienne » à 45%. Et pour les lycéens, le pourcentage monte à 51%, là encore tiré par le public masculin (première place pour les 2 genres, 46% des filles et 62% des garçons).

39 Voir mon article « Nouveaux Mondes de l'imaginaire pour nouveaux enjeux politiques : des autrices ouvrent la voie », dans la revue d'études culturelles de l'université de Nancy *Cultural Express* [en ligne], 2024, URL : <https://cultx-revue.com/article/nouveaux-mondes-de-limaginaire-pour-nouveaux-enjeux-politiques-des-autrices-ouvrent-la-voie> [consulté le 8 avril 2024].

RÉSUMÉS

Français

Cet article veut présenter, de manière très synthétique, un état du champ des recherches en littératures populaires et cultures médiatiques au prisme de leur légitimité au début de la décennie 2020, replacé dans le contexte plus large des évolutions des perceptions qui les ont affectées au cours de leur histoire, telles que les travaux spécialisés en rendent compte – dynamiques de distinction et pratiques de légitimation, fragmentées, concurrentes, loin de toute « légitimité » singulière et objectivée. Les tendances principales qui s'y dessinent, dans une situation d'explosion quantitative de l'offre de contenus, consistent en effet en un double mouvement de déhiérarchisation, ou plutôt de mobilité accrue des hiérarchies culturelles, couplée à des effets de resegmentation interne qui feront l'objet des deux premières parties de l'article. La dernière propose quelques pistes exploratoires concernant les nouveaux critères de valeur justifiant l'évolution, lente et partielle, de ces échelles de légitimité.

English

The aim of this article is to provide a concise overview on the field of research about popular literature and media cultures, from the perspective of their legitimacy, at the beginning of the 2020s. The broader context of the changes in perceptions that have affected them over the course of their

history appears as reflected in specialized studies – dynamics of distinction and practices of legitimization, fragmented, competing, far from any singular, objectified “legitimacy”.

The main trends that can be spotted, in a situation of quantitative explosion of content supply, consist in a double movement of de-hierarchization, or rather of increased mobility of cultural hierarchies, coupled with internal resegmentation effects, which will be the subject of the first two parts of the article. The last part proposes some exploratory avenues concerning the new value criteria justifying the slow and partial evolution of these scales of legitimacy.

INDEX

Mots-clés

légitimité, légitimation, littérature populaire, culture médiatique

Keywords

legitimacy, legitimation, popular literature, media culture

AUTEUR

Anne Besson

Université d'Artois, Arras, « Textes et Cultures » (UR 4028)

IDREF : <https://www.idref.fr/058494715>

ORCID : <http://orcid.org/0000-0002-6010-0957>

HAL : <https://cv.archives-ouvertes.fr/annebesson>

ISNI : <http://www.isni.org/0000000078402211>

BNF : <https://data.bnf.fr/fr/14079464>