



Dans une base de données lointaine, très lointaine : les humanités numériques pour explorer les univers étendus

Emmanuelle Lescouet, Université de Montréal, CA, emmanuelle.lescouet@umontreal.ca

Pierre Gabriel Dumoulin, Université du Québec à Montréal, CA, dumoulin.pierre_gabriel@uqam.ca

Amélie Vallières, Université du Québec à Montréal, CA, vallieres.amelie@uqam.ca

La popularisation du numérique et la facilité d'accès de nombreux outils de création ont permis aux lecteur-riche-s, joueur-euse-s et spectateur-riche-s de participer à l'établissement d'univers étendus. Ces individus sont tout autant des artistes confirmé-e-s, rémunéré-e-s pour ce travail, que des amateur-riche-s, plus ou moins expert-e-s dans leur domaine et dans l'œuvre à laquelle iels participent. Comment toutes ces productions s'articulent-elles pour former un univers cohérent? Quels outils et méthodes pouvons-nous mobiliser pour les comprendre? La diversité des propositions narratives, de la fanfiction textuelle aux fangames, en passant par des vidéos, reprises sonores, œuvres graphiques, etc., pose la question d'une multimodalité nécessaire à la compréhension d'une fiction. La diversité des supports mobilisés par les univers étendus (romans, films, séries, jeux vidéo, jeux de rôles, fanfictions, visual novels, séries audios parodiques ou non, pièces de théâtre, etc.) permet d'appuyer notre étude sur un vaste corpus que nous pourrions étudier quantitativement autant que qualitativement. Dans cet article, nous faisons une étude exploratoire entre la production des fans et la production officielle d'œuvres littéraires dans l'univers étendu de Star Wars. Nos objectifs sont de développer une méthodologie d'analyse issue des humanités numériques s'intéressant à l'étude des littératures de l'imaginaire, pour documenter la représentativité et l'inclusivité dans ces œuvres. Nous procédons d'abord par la contextualisation des études en littératures de l'imaginaire et au processus de documentation de ces littératures. Ensuite, nous définissons notre méthodologie d'analyse, du moissonnage de données dans les API au traitement des données dans OmekaS. Enfin, nous concluons cet article par une discussion sur les résultats obtenus et les pistes de recherche vers lesquelles ceux-ci nous encouragent.

The increased interest in digital technology and the ease of access to numerous creative tools enable readers, gamers, and spectators to take part in the creation of extended universes. These individuals are just as much established artists—paid for their work—as they are amateur artists,



more or less experts in their field or in the work in which they are involved. How does all of this work fit together in a coherent universe? What tools and methods can we use to understand it? The diversity of narrative proposals, from textual fanfiction to fangames, videos, audios, graphic works, etc., raises the question of whether multimodality is useful to understand fiction. Since extended universes may be mobilized in a wide range of media (novels, films, series, video games, role-playing games, fanfiction, visual novels, audio series, plays, etc.), we are able to conduct qualitative and quantitative research on our corpus. In this article, we conduct an exploratory study of the relationship between fan production and official production of literary works in the extended universe of Star Wars. Our objectives are to develop an analytical methodology, drawn from works in the digital humanities interested in the literature of the imagination, to document representativeness and inclusivity in the works studied. We begin our article by contextualizing studies of these forms of literature and the process of documenting them. Next, we define our analytical methodology, from data harvesting in APIs to data processing in OmekaS. Finally, we conclude with a discussion of the results obtained and the research avenues they yield.

Introduction

La popularisation du numérique et la facilité d'accès grandissante aux outils de création numériques ont permis aux lecteur·rice·s, joueur·euse·s et spectateur·rice·s de participer à l'établissement d'univers étendus (Wuyckens 2021) qui dépassent les créations originales. Sans l'intervention de communautés de créateur·rice·s reliées par des intérêts communs (Jenkins 2006), l'expansion de ces univers serait impossible. Ces individus sont autant des artistes confirmé·e·s ou rémunéré·e·s pour ce travail, que des amateur·rice·s plus ou moins expert·e·s dans leur domaine et dans l'œuvre à laquelle iels participent. Comment toutes ces productions s'articulent-elles pour former un univers cohérent? Quels outils et méthodes pouvons-nous mobiliser pour les comprendre? La découvrabilité de ces œuvres, entre aires linguistiques et entre communautés, est problématique. Chaque ensemble peine à se connecter aux autres, au Canada cela est flagrant entre les productions anglophones et francophone par exemple. La production diverse et extrêmement dense (nous parlons de la publication de plusieurs dizaines de milliers d'œuvres plus bas) empêche une vision d'ensemble ainsi que des stratégies de diffusion unitaires, ces ensembles immenses se subdivisent en de nombreux ensembles liés aux parcours de lecture de chaque individu et de ses pairs immédiats. D'ailleurs, le cloisonnement relativement hermétique des communautés de fans se réverbère également dans les communautés de spécialistes, alors que la rencontre entre les milieux scientifiques et professionnels serait nécessaire pour saisir les corpus de façon globale. En brossant le portrait d'univers étendus ainsi que des acteurs qui les alimentent, notre projet espère présenter les liens qui se tissent entre les communautés de même que les impacts que celles-ci peuvent avoir sur les univers étendus auxquels elles participent. Ce projet espère établir une continuité entre tou·te·s ces acteur·rice·s et brosser le portrait d'univers narratifs étendus, comme c'est le cas dans la série Star Wars.

La diversité des propositions narratives, de la fanfiction textuelle aux *fangames*, en passant par des vidéos, des reprises sonores, des œuvres graphiques, etc., pose la question de la multimodalité nécessaire à la compréhension d'une fiction. La diversité des supports mobilisés par les univers étendus (romans, films, séries, jeux vidéo, jeux de rôles, fanfictions, *visual novels*, séries audios parodiques ou non, pièces de théâtre ...) permet d'appuyer notre étude sur un vaste corpus que nous pourrions étudier quantitativement autant que qualitativement. De plus, une réflexion à la croisée des productions officielles et de celles des fans permettent en mettre en relief certaines différences, dans les thèmes abordés que dans les représentations genrées et identitaires. Pour cerner de tels corpus mixtes, il est impératif de développer une méthode d'analyse commune à toutes ces œuvres et d'être en mesure de les aborder comme un ensemble

de données. Ce groupement pose notre premier enjeu de recherche : que documenter et comment le documenter ? Si ces corpus peuvent exister individuellement dans leurs disciplines respectives, leur mise en commun encourage une réflexion sur les pratiques de typologisation des données de corpus hétérogènes en ce que les descripteurs, tant techniques que conceptuels, ne trouvent pas toujours d'échos entre eux.

Dans cet article, nous proposons d'ouvrir la réflexion en s'intéressant, d'une part, aux études en littératures de l'imaginaire et sur les représentations queer, et, d'autre part, en détaillant notre cadre méthodologique de même que notre processus d'analyse, lequel a permis la mise en place du projet Répertoire. Enfin, nous proposons en fin d'article l'analyse des données récoltées en janvier 2023 sur l'univers étendu de Star Wars, analyse qui nous permettra de mettre en relief l'importance des méthodologies de recherche numériques pour comprendre les univers étendus.

1. Documenter les littératures de l'imaginaire

En littératures de l'imaginaire, les univers étendus décrivent des fictions se déployant sur plusieurs supports, avec plusieurs sources officielles (plusieurs créateur-riche-s, par exemple). Dans ces univers, différentes histoires sont racontées dans des mondes seconds constants, qu'elles soient dans des lieux divers ou à des époques distinctes. Afin de saisir les nuances de telles œuvres, d'autant plus dans une perspective quantitative qui permet d'avoir une vision globale d'un corpus de quelques milliers d'œuvres, il importe d'établir certains points de repères pour situer les pratiques d'écriture, d'édition et de lecture des œuvres qui alimentent les univers étendus dans lesquels les fans, professionnel-le-s rémunéré-e-s comme bénévoles enthousiastes, participent.

1.1 Diversité des corpus, thématiques et médias

Les littératures de l'imaginaire ont su, depuis les années 1980, se multiplier sur divers formats et supports (Besson 2015). La diversité des productions de l'imaginaire (fantasy, fantastique et science-fiction) en fait un champ d'études à la fois transdisciplinaire et intermédial. Cette diversité tend à diviser les recherches menées en fonction de leurs caractéristiques techniques : certaines vont se concentrer sur les littératures, d'autres sur les formes ludiques, d'autres encore sur les formes numériques ou sur celles audiovisuelles. Non seulement il est nécessaire de délimiter un espace interdisciplinaire commun afin que ces études puissent dialoguer entre elles, mais en plus, il faut définir les paramètres qui permettront aux différentes communautés de spécialistes de partager et corréler leurs données. Le besoin de prendre en compte la diversité de ces corpus est à l'origine du projet Répertoire, fondé en janvier 2023. (Le projet est

financé de juin 2023 à juin 2024 par la Chaire de recherche du Canada sur les écritures numériques. Le carnet de recherche Hypothèses dédié au projet offre un suivi au mois du chantier [Repertorium 2023]).

Les éléments constants du genre décrits par le mégatexte rejoignent les artefacts communs de la culture numérique. Ces ensembles convergent dans une mythographie collective, se superposant souvent dans la culture populaire, permettant d'étudier la perméabilité de ces concepts. À cet effet, Astrid Ensslin souligne l'importance et les effets des technologies sur les œuvres littéraires numériques et vidéoludiques, notamment en soulevant les compétences, certes, mais aussi l'ouverture des lecteur-ric-e-s à de nouveaux modes interprétatifs interactifs (Ensslin 2014, 6). L'impératif multimodal des univers étendus déplace la focale de cette recherche littéraire : notre corpus hybride nécessite des particularités de nomenclature, de description et de terminologie. Cette construction, au cœur des études sur les cultures de l'imaginaire, nous force à adopter une position de consensus éclairé à même de satisfaire les communautés de recherche tout en autorisant une interopérabilité avec les milieux professionnels, tant du côté des éditeurs que des libraires.

L'intime imbrication du mégatexte fictionnel (Bréan 2012 ; Bréan 2022) et de ce que nous pourrions appeler un mégatexte technologique (Ensslin 2014) mène à des habitudes construites sur des dictionnaires de gestes consistants (Masure 2017 ; Lescouet 2022). Si l'érudition de genre induite par la constitution des encyclopédies fictionnelles rassemblées au sein du mégatexte fictionnel se retrouve dans la maîtrise des usages — des littératies — numériques et médiatiques, elle se spécialise pour certains univers étendus, comme pour certaines franchises ludiques. Ainsi si une amateur-ice de science-fiction saura reconnaître un vaisseau spatial, un-e fan de Star Wars aura des compétences plus précises pour reconnaître certains « modèles » ; de même les joueur-euse-s de jeux de tir auront un certain nombre d'automatismes et d'habitudes quant au maniement des armes à leurs disposition, mais des joueur-euse-s de *Star Wars: Battlefront* (Pandemic Studios 2004) seront en mesure de saisir les contrôles d'un nouvel opus bien plus rapidement.

Aussi, les pratiques de rédaction des fanfictions (Barnabé 2014) sur des plateformes dédiées, qu'il s'agisse de blogues, de Wattpad ou encore Fanfiction.net, pousse les communautés à se regrouper dans des lieux distincts. Le but de cette recherche est d'interroger les adaptations formelles qui sont faites au sein des univers étendus pour prendre en compte les particularités des différents supports, mais aussi de comprendre comment ces changements de plateforme nourrissent le *world building* (Saint-Gelais 1999 ; Bréan 2012) en l'étendant. Nous interrogerons enfin leur impact sur l'immersion : la diversité des supports, et donc des expériences possibles au sein

d'un même univers construit, renforce-t-elle sa crédibilité? Ou au contraire, cette extension dissout-elle le focus et donc la suspension d'incrédulité demandée?

1.2 Participation à la construction de l'imaginaire

Loin d'être un phénomène nouveau, la production de fan est étudiée depuis plus de 40 ans. Dès 1975, Lichtenberg, Marshak et Winston décrivaient, dans leur chapitre « Do-It-Yourself Star Trek – the Fan Fiction », le même genre d'histoire que les fans d'aujourd'hui publient sur des sites web comme Fanfiction.net et Archive of Our Own (Lichtenberg, Marshak, et Winston 1975). De plus en plus de personnes participent à une forme d'activité organisée pour/par des fans : lire ou écrire des fanfictions, créer des fanarts, faire des films ou des vidéos de fan, organiser des wikis, collectionner de la marchandise, faire du cosplay, etc. (Guo 2020). Comme le souligne Coppa, « fandom itself is moving fast from subculture to culture » (Coppa 2017, 1). Sans faire le récit des études des fans, nous ne pouvons passer sous silence l'importance des communautés de fans dans la création d'univers étendus.

En particulier, la migration des fanfictions, soit des récits fictionnels écrits par des fans et qui s'inspirent d'œuvres préexistantes, sur Internet a considérablement accru la portée et la visibilité du phénomène. Alors qu'à une époque, les fans devaient se retrouver dans des conventions pour échanger des fanzines et discuter avec les membres de leur communauté plutôt restreinte (Jenkins 2006), il est maintenant possible pour tout un chacun de pleinement participer. En effet, le support numérique a facilité la production des textes et a permis leur distribution à plus grande échelle, mais a aussi entraîné d'importantes modifications dans l'organisation de l'activité fanique (Barnabé 2014). Maintenant, il est possible pour les auteur-ric-e-s de publier progressivement leur texte, de le faire relire par un-e ou des bêta-lecteur-ric-e-s (*beta readers*), de le modifier rétroactivement et d'y ajouter des notes (*author's notes*, A/N) au début ou à la fin pour transmettre des messages directs au/à la lecteur-ric-e, qu'il s'agisse de décharge (*disclaimer*), de traumatisme (*trigger warning*) ou simplement pour donner des indications au regard de l'œuvre (genre de fanfiction, couples [pairings], headcanons, etc.). Du côté des lecteur-ric-e-s, les diverses plateformes permettent dorénavant de créer des listes de favoris, de s'abonner à des fanfictions ou des auteur-ric-e-s, en plus de laisser des commentaires et des évaluations (*reviews*) à toutes les étapes de la création d'une fanfiction, ce qui peut en influencer l'écriture. De plus, la multiplication des textes disponibles a transformé les modalités de leur consommation, si bien qu'une taxonomie s'est naturellement développée autour des œuvres. Les lecteur-ric-e-s et auteur-ric-e-s ont ainsi développé une méthode de classification qui permet l'utilisation de filtres (p. ex. genres, personnages et relations) et de tri (p. ex. dernière date de

publication, texte complété) pour parvenir à naviguer le vaste ensemble de fanfictions désormais disponibles.

Alors que la fanfiction a longtemps été définie en opposition à la fiction professionnelle, il faut plutôt comprendre que le caractère de « fan » est plus important que l'aspect légal ou monétaire de l'activité créative. Selon Coppa, « most fanfiction writers write for love rather than for money, and to their own specifications rather than the market's » (Coppa 2017, 3). De la même manière que les contes, fables et légendes sont réappropriées et racontées autrement, les mythes contemporains (Jenkins 2006) des littératures de l'imaginaire se voient réécrits et transformés par leurs fans. La production fanique n'est pas simplement la continuation ou l'interprétation d'une histoire, mais bien une nouvelle histoire qui se déroule dans l'œuvre originale, laquelle existe dans, pour et sous l'influence d'une communauté donnée. Cela va au-delà de l'idée qu'un fandom est une audience ou un marché pour la production. Un fandom (mot anglais composé de fan [pour fanatic] suivi du suffixe dom [pour domain]) est la sous-culture propre à une base de fans, c'est-à-dire tout ce qui touche au domaine de prédilection d'un groupe de personnes et qui est organisé ou créé par ces mêmes personnes. Un fandom peut se développer autour de n'importe quel domaine d'intérêt ou d'activité humaine : dans le cadre de notre projet, il s'agit des textes de la littérature de l'imaginaire. Il faut comprendre que les œuvres de fanfiction sont créées selon des conventions littéraires, des attentes et des désirs de cette communauté (Coppa 2017), et que c'est en son sein et pour celle-ci que des pratiques, des genres, des critiques sont élaborés. D'ailleurs, comme l'évoque Hellekson (Hellekson 2009, 115), les productions de fans « are incorporated into a multivocal dialogue that creates a metatext, the continual composition of which creates a community [...]. When the fan work is proffered, it is taken into the metatext. The individuality of that piece is lost; it becomes part of something greater ». C'est pourquoi il nous paraît primordial de considérer l'apport des fans dans le *word building* des littératures de l'imaginaire.

2. Construire des lieux d'étude de corpus pour comprendre les tendances d'un ensemble large

Les humanités numériques, tout comme l'analyse linguistique de corpus, offrent les assises théoriques et méthodologiques nécessaires pour penser de larges corpus et entamer un dialogue avec les données qu'ils contiennent (Hoover, Culpeper, et O'Halloran 2014). En constituant des bases de données sur les littératures de l'imaginaire, il est possible d'identifier des pratiques d'écriture ou de lecture, des thématiques ou des enjeux récurrents, ou encore de commenter l'histoire, en temps réel, de l'évolution de ces littératures. Le portrait global rendu possible par l'analyse de

ces données permet la production de connaissances situées, dans l'espace et le temps, sur la consommation des œuvres. Une telle approche devient d'autant plus pertinente lorsqu'il s'agit d'interroger la représentativité, l'inclusivité et les enjeux socioculturels des minorités (Liu 2020).

Notre approche fait écho aux travaux sur la lecture distante (*distant reading*), à la fois comme piste d'analyse et méthodologie. Le développement et le perfectionnement des méthodes de collecte et de traitement de données encouragent un renouvellement constant des méthodes d'analyse, tant qualitatives que quantitatives, et favorisent l'interdisciplinarité, notamment en brouillant les frontières entre les disciplines (Underwood 2016, 531). Les échanges entre les différents modes et milieux de recherche favorisent l'émergence de nouvelles pistes interprétatives — ou de nouvelles lectures — et invitent les chercheur·se·s à élargir l'horizon de leurs recherches. Notamment, l'indexation de données tirées de larges corpus pour en tirer des modèles d'analyse permet la découverte de faits saillants autrement difficiles à percevoir. En ce sens, l'accessibilité à des données fiables et structurées devient un enjeu significatif, notamment lorsque ces données sont issues de milieux marginaux ou minoritaires ; procéder à une « lecture distante » d'un corpus permet de formuler de nouvelles inférences (Underwood 2016, 530). Si un rapport plus distant et détaché des données peut révéler de nouvelles relations, il demeure néanmoins important, surtout dans un milieu littéraire, de conserver une approche critique des modèles développés, puisque ceux-ci ne peuvent prétendre à l'exhaustivité des formes de représentations.

Dans l'introduction de leur ouvrage collectif *Corpus Linguistics and Translation Tools for Digital Humanities* (Maci et Sala 2022), les auteur·rice·s soulignent que les humanités numériques ne doivent pas se limiter à l'utilisation de ressources informatiques, suivant l'influence des études de corpus linguistiques, pour procéder à leurs analyses ; plutôt, elles invitent la discipline à encourager l'intervention humaine (Maci et Sala 2022, 3), à aller au-delà des modèles et des données pour mieux prendre en compte leurs complexités, notamment en ce qui a trait aux enjeux éthiques et représentationnels. La standardisation des modèles pourrait masquer les spécificités des minorités en se concentrant seulement sur les principales tendances observées par une analyse de données distante (voir Spivak 2005, 107–108). Travailler des ensembles de données, lesquelles peuvent inclure des références de genre, de culture ou d'identité, revêt un caractère expérimental et exploratoire, en ce que ces données exigent une analyse qualitative plus fine que ce qu'une méthodologie de lecture distante pourrait produire, surtout dans des contextes où la langue s'éloigne des modèles standardisés (par exemple avec les néo-pronoms et l'écriture inclusive). Suivant la proposition de Matthew Jockers pour qui l'expérimentation est un processus de va-et-vient entre les

observations et les conclusions (Jockers 2013, 6), nous désirons expérimenter, de façon abductive, avec les données et méthodes. Ce faisant, nous cherchons à recontextualiser nos données, mais aussi à les documenter, afin de saisir la circularité des inférences et de nuancer l'expérience interprétative (Dumoulin 2024).

Avec notre projet, nous désirons développer une base de données structurées et standardisées qui permet de procéder à une analyse macroscopique (sur les sujets ou les thèmes, par exemple) des œuvres tout en offrant la possibilité d'effectuer des requêtes sur différents enjeux, dans un désir de procéder à une description adéquate et représentative des œuvres et de leur contenu. En ce sens, il ne s'agit pas de développer un modèle d'analyse nécessairement réutilisable, comme cela peut être le cas avec différentes modélisations de données (voir Houston 2023), mais plutôt de concevoir un modèle d'analyse qui repose sur un rapport abductif aux données observées et traitées, avec un retour constant à la base de données et à sa classification interne, autant sur les données contenues dans les œuvres que leurs métadonnées.

2.1 Vers des lieux de corpus pour les littératures contemporaines

Les études de corpus constituent une des principales approches en études littéraires. Elles discutent des thèmes, des œuvres ou encore des auteur·rice·s et mettent en relation plusieurs œuvres dans un exercice d'analyse croisée, mais reposent souvent sur une méthodologie d'analyse linguistique des textes, plutôt que sur les métadonnées du texte ou sur les bouleversements socioculturels qui marquent la nature même des œuvres. Le développement des outils numériques depuis la seconde moitié du XX^e siècle renouvelle autant la méthodologie que la façon d'analyser les données contenues dans les œuvres.

En tant que discipline, les humanités numériques mobilisent de nouvelles techniques et méthodologies, rendues possibles par le caractère computationnel des œuvres et outils, pour encourager de nouvelles pistes de recherche (Mohany 2018, 371). Elles encouragent la mise en place de méthodes mixtes pour analyser les corpus qui ne sont pas strictement quantitatifs (Hoover, Culpeper, et O'Halloran 2014, 4). En se développant et se perfectionnant, les méthodes d'analyse de corpus propres au numérique favorisent l'adéquation entre les méthodes qualitatives et quantitatives, puisqu'elles permettent de récolter un large échantillon de données, autant sur l'œuvre que dans l'œuvre, pour rendre compte des spécificités des textes. L'interprétation subjective du texte, son analyse linguistique et littéraire, se voit suppléer par des données explicitées au moyen d'interfaces qui rendent accessible le contenu des bases de données. Il faut néanmoins tenir compte, comme l'encourage Alan Liu, qu'un corpus donne plus que des données sur des œuvres : « Properly edited, a corpus does not just

present a collection of materials but makes visible the editorial practices by which it was created, selected, balanced, corrected, and transformed to become a representation of sociocultural diversity» (Liu 2020, 141). Dès lors, autant la collecte en elle-même que ses paramètres revêtent une subjectivité interprétative que toute analyse de corpus se doit de reconnaître.

Dans la constitution, puis l'analyse, de notre base de données sur les univers étendus, nous avons décidé de nous éloigner d'une analyse linguistique de corpus pour poser la question des caractéristiques qui structurent la création de ces univers. Cela passe par la collecte de métadonnées sur l'œuvre, mais aussi par le croisement de certaines données intrinsèques aux œuvres, notamment en ce qui a trait aux enjeux de représentation qui sont à l'origine de notre projet de recherche. L'analyse statistique de ces données offre une idée d'ensemble, comme nous le verrons plus loin, sur les pratiques d'écriture dans les univers étendus, dont celui de Star Wars. Nous joignons ainsi des pratiques de la littérature et des humanités numériques, dans un objectif de mieux prendre en compte la réalité, complexe et mouvante, de cette forme de littérature.

2.2 Usage des bases de données et des humanités numériques

L'avantage pour des corpus larges et divers est évident : si nous ne pouvons, seul-es, expérimenter et documenter sur des fiches dans la base de données toutes les œuvres narratives des littératures de l'imaginaire, ou ne serait-ce qu'un segment d'entre elles, les outils numériques peuvent nous aider à gagner un temps précieux. Automatiser cette documentation, au moins pour une partie significative, est une étape importante du développement de cette recherche.

Les corpus larges et divers offrent des perspectives avantageuses, autant pour les chercheur-euse-s que les professionnel-le-s, voire même pour les fans. Toutefois, un des principaux enjeux est, à juste titre, d'expérimenter et documenter l'ensemble des œuvres et de les déposer dans un tout cohérent et structuré : pour développer une base de données qui recoupe toutes les œuvres narratives des littératures de l'imaginaire, ou ne serait-ce qu'un segment d'entre elles, il est impératif de mobiliser les outils et pratiques numériques pour accélérer la constitution de la base données et la rendre accessibles. Automatiser cette documentation, au moins pour une partie significative, est une étape importante du développement de notre recherche.

D'abord, il est possible de procéder à une collecte de données plus efficaces sur les littératures de l'imaginaire à l'aide de ressources numériques, notamment en mobilisant des listes de publications des libraires (comme Catalyst au Québec) ou encore en recueillant les métadonnées des sites de publications ou d'autopublications (comme Fanfiction.net ou Wattpad). Pour cette première phase, nous nous sommes

concentrées sur Archive of Our Own (AO3), dont les données sont publiques et facilement accessibles. Plusieurs raisons nous ont amené-e-s à retenir ce répertoire de productions de fan : son statut *open source*, sa mission d'archivage ainsi que ses principes de tolérance et l'absence de censure qui le rend unique, et finalement la reconnaissance littéraire qu'on lui donne et qui lui a valu un Hugo Award en 2019. (Le Hugo Award est un prix littéraire américain décerné chaque année aux meilleures œuvres de science-fiction et de fantasy de l'année écoulée. Le nombre de catégories récompensées a évolué au fil du temps, à mesure que les littératures de l'imaginaire se sont étendues à de nouveaux médias tels que le cinéma et la télévision.) Le fonctionnement d'AO3 s'appuie sur des bénévoles (principalement des femmes) qui non seulement fournissent, pour certain-e-s, les fanfictions, mais, pour les mêmes ou pour d'autres, veillent et améliorent les systèmes de balises et de classement des diverses œuvres stockées (des fanfictions, des œuvres d'art, des dessins, des enregistrements, des *podcasts*, etc.). Ajoutons que AO3 est reconnu comme étant la plateforme qui accepte la plus grande diversité de formats de productions faniques (des fanfictions, des œuvres d'art, des dessins, des enregistrements, des *songfics*, des listes, etc.), ce qui nous paraissait plus approprié, considérant la dissémination sur Internet des corpus produits par les fans.

De plus, les données collectées peuvent aussi être manipulées plus facilement, notamment à l'aide de l'arsenal d'outils statistiques disponibles : il est alors possible de poser des questions aux données et d'avoir des réponses chiffrées rapidement. Si cette sortie est limitée, ne donnant qu'une réponse factuelle, sans contexte ou signification, elle peut permettre de rebondir sur une assomption, confirmer une intuition ou simplement pointer des faits invisibles ou difficilement visibles autrement. Ici, nous allons nous intéresser, dans ce premier temps du projet, à des questions simples de dates et de quantité chiffrées, pour ensuite pouvoir élargir l'usage de la base à des questions des milieux professionnels, notamment en éducation.

Avant de manipuler des données, il est nécessaire de choisir quelles données nous voulons. Pour ce projet, nous avons fait le choix de documenter des ensembles cohérents par blocs. Nous entendons par là un ensemble d'œuvres regroupées autour d'une thématique (un genre, un sujet, etc.) ou d'une présence auctoriale forte. Dans le cadre de cet article, nous avons décidé de nous concentrer sur un seul de nos corpus, soit celui des œuvres liées à l'univers de Star Wars, bien que d'autres choix auraient été possibles. En tant qu'univers narratif transmédiatique, Star Wars nous semblait être un exemple typique de corpus massif, d'autant plus que ce corpus est déjà documenté et discuté. Bien qu'un corpus soit sélectionné, celui-ci doit être délimité en fonction de certains critères afin d'obtenir une base de données structurée et pertinente. En effet, quelles formes narratives désirions-nous documenter, et comment permettre des relier entre

elles à partir de *tags*? La première étape a été de faire une présélection des formes médiatiques à l'étude, incluant les productions officielles : livres, bandes dessinées, films, séries TV, jeux vidéo, jeux de rôles et jeux de plateau. Puisqu'elles nous paraissaient incontournables pour décrire la relation entre les produits officiels et ceux des fans dans une analyse de corpus aussi massive — d'autant plus que les données sur les fanfictions et les *fangames* sont facilement accessibles sur Internet. Nous désirions considérer ces œuvres puisque, au-delà du nombre impressionnant de productions faniques associées à Star Wars, il est maintenant évident que les conglomerats médiatiques mobilisent, voire exploitent, les activités des fans pour le développement, le marketing, la publicité, les activités promotionnelles et la distribution de leurs produits, tout en continuant à envoyer des ordonnances de cessation et d'abstention aux créateur-riche-s amateur-riche-s. Sans pour autant élaborer trop longuement sur les enjeux du travail des fans comme de leurs productions (voir Stanfill et Condis 2014), il nous paraissait difficile d'en faire abstraction, surtout que leur impact se fait de plus en plus sentir dans la production officielle. Il suffit de penser à la réception du public quant à *Rogue One* (Edwards 2016) et à la multiplication des *fan edits* dû à la différence entre la bande annonce initiale parue en avril 2016 et la version finale du film sortie en décembre de la même année. Alors que Disney a ensuite tenté de corriger le tir par l'entremise de plusieurs longues entrevues avec le personnel de production, la compagnie en est venue à inviter les fans « to remix and misremember the film in ways they find more pleasing » (Canavan 2018, 285). Considérant l'histoire de Star Wars, il ne s'agit pas de la première instance où les fans ont préféré suppléer par leurs productions et que ces dernières ont été reconnues par la communauté.

Pour ce projet, nous utilisons une base de données construite avec OmekaS. Ce système de gestion de contenus est pensé et développé pour la gestion d'archives et de bibliothèques, ce qui correspond plutôt bien aux enjeux du projet. De plus, il s'agit d'une technologie libre et ouverte, permettant d'avoir accès aux techniques mobilisées et de développer nos propres outils à partir de cette base. Les vocabulaires contrôlés sont construits sur OpenTheso (CNRS 2025), un outil de thésaurus (vocabulaires) collaboratif permettant de discuter les termes employés et de les faire évoluer au besoin. La possibilité de discussion ouverte se limite au contributeur-riche du projet, mais permet tout de même d'ouvrir à d'autres personnes, tout en évitant le *trolling*. De plus, cette collaboration s'applique également aux traductions et définitions des divers termes, permettant d'ouvrir un espace d'une grande richesse pour parvenir à un consensus scientifique.

Une première partie des fiches est consacrée aux informations de base de l'œuvre : le titre, le·la ou les auteur-riche-s, l'instance de publication, la date de publication, l'aire géographique d'origine ainsi que la nature de l'œuvre. Ce dernier point permet de qualifier largement l'œuvre : est-ce un roman ? un film ? un jeu ? À quelle grande catégorie cette

chose appartient-elle? Ensuite, une partie se concentre sur les particularités médiatiques de l'œuvre : cette partie est propre à chaque forme — énoncées ci-dessus. Pour les œuvres vidéo, il s'agira de la durée du contenu, de l'appartenance ou non à une série, le cas échéant du nombre d'épisodes, par exemple. Pour celle de la forme d'un jeu : y a-t-il des cartes? Un plateau? Des pions ou *meeples*? Quel genre de *gameplay* est mobilisé? Enfin, une troisième partie s'intéresse aux caractéristiques de l'imaginaire de l'œuvre : à quels genre et sous-genre appartient-elle? Et à ses particularités propres. Certains corpus nécessitent des informations particulières pour être abordés : pour les réécritures mythologiques, cela peut être les divinités ou les héros présents ; pour Star Wars, notre exemple ici, il s'agira des ères historiques de cet univers. Nous aurions pu nous intéresser aux différents lieux de cet univers (p. ex. Coruscant, la Bordure Extérieure [Outer Rim], l'Espace Sauvage [Wild Space]) ou bien à des personnages emblématiques (p. ex. la famille Skywalker, l'Ordre des Jedi, l'Ordre des Sith, etc.), mais nous avons préféré nous appuyer sur les ères historiques puisque ces dernières se répercutent dans l'ensemble des créations, formelles et faniques, même après le rachat de la franchise par Disney.

Une grande part de ce travail passe par les sites de distribution à destination des libraires, qui permettent d'obtenir des listes — des documents au format CSV notamment — où toutes les publications accessibles sur le sol canadien vont être disponibles. Toutefois, cela pose plusieurs limites : les œuvres non disponibles dans cette aire géographique devront être importées séparément alors que les œuvres qui ne sont plus disponibles vont ainsi demander des recherches plus approfondies. Parfois, les bases de données privées, construites par les fans, servent de mémoire vivante pour les spécialistes qui cherchent à mieux comprendre et étudier des corpus.

Une fois toutes ces données importées dans la base de données, elles ont généralement besoin d'être nettoyées pour être pertinentes et alignées avec les standards voulus. Les champs sont souvent différents en fonction de leur origine et demandent d'être remis en ordre. Il faut généralement aussi ajouter des informations qui ne sont pas utiles aux bases de données d'origines : ainsi les ères historiques des narrations de Star Wars sont souvent disponibles pour les fanfictions, afin de permettre aux lecteur·rices spécialistes de naviguer dans l'univers étendu et tentaculaire ... mais rarement pour les romans, notamment pour les plus anciens, destinés à un public moins spécialiste. Il faut alors les remplir à la main et vérifier globalement les données amassées, souvent avec des informations internes à l'œuvre demandant une étude plus rapprochée, qu'il s'agisse d'une recherche interne, une approche textométrique permettant de faire ressortir les lieux ou époques abordés par exemple. La moindre erreur se répercutant dans les étapes suivantes, il est important de minimiser ces aléas. Cependant, une faible quantité d'imprécision fait partie de la méthode : il faut être conscient de cette marge d'erreur pour pouvoir la prendre en compte lors de l'analyse finale des données.

2.3 Particularités du projet Répertoire

Comme évoqué précédemment, pour que ce projet exprime sa complète utilité, il faut lui poser des questions dont les réponses quantitatives permettent d'obtenir des résultats révélateurs sur le corpus étudié. De nombreuses questions traversent une grande partie des corpus de l'imaginaire : les proportions d'autrices et les évolutions de cette représentation ? La proportion d'œuvres appartenant à des séries par rapport aux œuvres fonctionnant par elles-mêmes ? Les évolutions du nombre de publications des différents sous-genres par an ? Ou encore, la représentation de ces sous-genres en fonction des aires géographiques de publication... Ces questions sont simples, mais peuvent nourrir la réflexion et la recherche sur ces corpus, permettant de montrer concrètement des évolutions globales de notre champ d'études. Gardons en tête que ce ne sont que des exemples, nous n'avons clairement pas le temps de développer toutes les pistes possibles ici. D'autres questions sont plus particulières à un corpus particulier. Pour citer l'exemple de Star Wars, les évolutions des ères historiques représentées sont-elles un reflet des ères présentées dans les films ? Les séries ? Le rachat de la franchise par Disney a-t-il eu des impacts sur les aires historiques présentées ? Sur le nombre de publications ? Sur les supports mobilisés ?

D'un autre côté, cette base de données est destinée à devenir un générateur de réseaux d'œuvres — au sens de Catherine Tauveron (Tauveron 2004) — afin de permettre aux personnes enseignantes ou aux chercheur·se·s en éducation de produire et exporter facilement des corpus ciblés. La recherche nous confronte, notamment pour les œuvres numériques, voire celles qui ne sont pas littéraires, à certaines difficultés. Notamment, il est difficile de trouver des œuvres qui correspondent aux conditions des programmes, et il tout autant ardu de les mettre en relation afin de les rendre utilisables sur le terrain. Des outils comme Répertoire rendent les contenus plus accessibles tout en permettant aux professionnel·le·s et aux chercheur·se·s de gagner en temps et efficacité. En effet, récemment, nous avons été confronté·e·s à une demande : une enseignante cherchait un corpus pour étudier les « récits policiers et d'enquête », mais souhaitait articuler cela à Star Wars qui était à ce moment populaire dans sa classe — notamment avec la sortie de la série *The Mandalorian* (2019–présent). La sélection du sous-genre « policier » et des étiquettes « jeunesse » et « adolescent », toutes deux issues des catégories éditoriales commerciales, ainsi que toutes les œuvres en langue « fre » (française) a permis de lui proposer un corpus raisonné, et elle a pu ensuite travailler à partir de cet ensemble pour construire sa séquence d'enseignement-apprentissage. Si cette voie est encore très embryonnaire, les six mois d'existence du projet s'étant focalisé sur la construction de la base de données pour ce premier ensemble, cette première expérience nous servira de preuve de concept, permettant de cerner tant les besoins — notamment, que les facettes du moteur de recherche correspondent aux catégories du programme — mais aussi un bel espoir d'usabilité puisque l'expérience a finalement été concluante.

3. Observations et résultats : vers des représentations queers dans Star Wars ?

Les données recueillies dans le cadre de ce projet ont permis de tester à la fois la faisabilité dudit projet, mais aussi de remarquer certains traits distinctifs qui précisent les pratiques auctoriales et éditoriales autour de l'univers de Star Wars. Nous avons structuré ces premières données autour de trois catégories : historicité des pratiques, représentation des genres dans les productions officielles, puis la confrontation des œuvres officielles avec celles produites par les fans.

3.1 Observation générale du corpus narratif de Star Wars

La génération, sous forme graphique, des données nous a permis de corroborer nos intuitions de recherche sur l'évolution des œuvres dans l'univers étendu de Star Wars. Le graphique ci-dessous (voir **Figure 1**) montre l'évolution dans le temps des publications officielles dans l'univers de Star Wars. À partir de ce graphique qui recense 2 822 œuvres, nous identifions trois périodes qui marquent des changements significatifs dans la publication des œuvres qui approfondissent l'univers étendu de Star Wars : la première période couvre de 1977 à 1998, la seconde de 1999 à 2012, puis la troisième de 2013 à aujourd'hui.

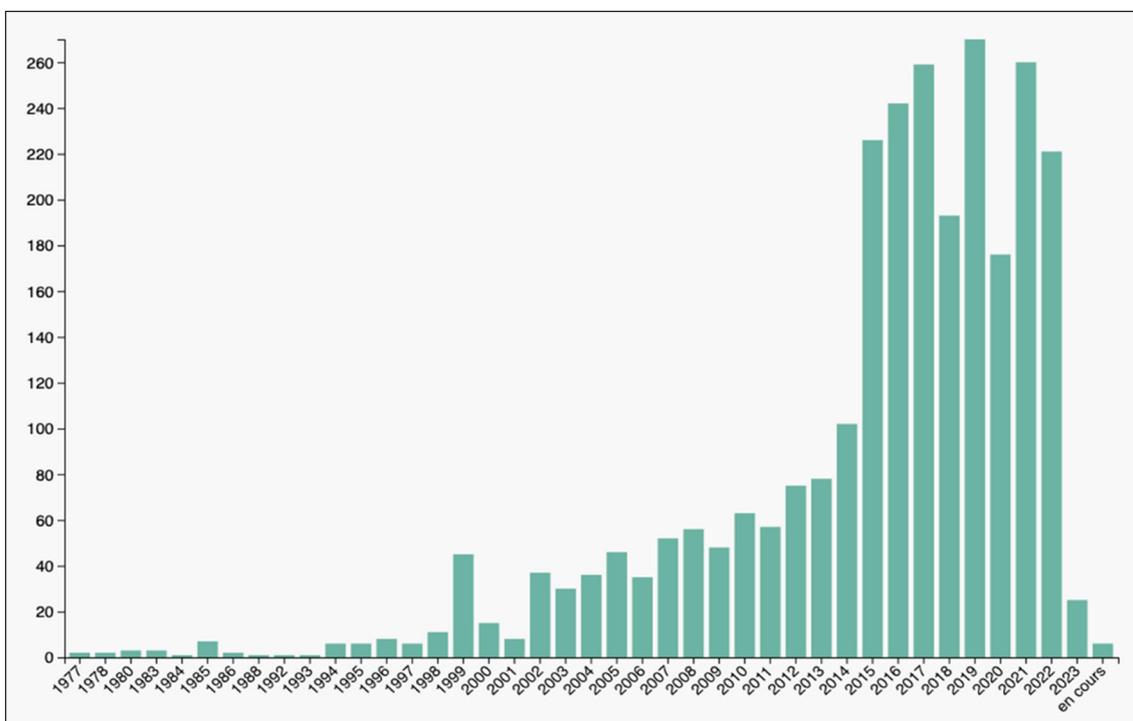


Figure 1 : Nombre de publications recensées par année dans l'univers étendu de Star Wars. Données récoltées au début de l'année 2023.

De 1977 à 1998, nous avons recensé 60 œuvres. Durant cette période, qui débute avec la sortie au cinéma de *Star Wars: Episode IV – A New Hope* (Lucas 1977), nous remarquons une hausse des publications en 1985, deux ans après la sortie de *Star Wars: Episode VI – Return of the Jedi* (Marquand 1983). De même, à partir de 1994, le rythme de publication augmente : plus de la moitié des œuvres (37) durant cette période ont été publiées de 1994 à 1998.

Le renouvellement de l'intérêt pour l'univers de Star Wars se solidifie à partir de 1999. De 1999 à 2012, ce sont 603 œuvres qui ont été publiées, dont 217 ont été publiées entre la sortie de *Star Wars: Episode I – The Phantom Menace* (Lucas 1999) et celle de *Star Wars: Episode III – Revenge of the Sith* (Lucas 2005), année du dernier long métrage avant le rachat de Disney en octobre 2012 (Roark 2023). De 2006 à 2012, 386 œuvres ont été publiées, soulignant l'intérêt, du moins littéraire, pour la sortie après le dernier film de George Lucas.

À partir de 2013, soit quelques mois après la transaction qui officialise Disney comme étant propriétaire de l'univers Star Wars, le nombre d'œuvres publiées explose. De 2013 à aujourd'hui, ce sont au moins 2159 œuvres qui ont été publiées, soit 77 % du corpus recensé. Durant cette période, la moyenne d'œuvres publiées annuellement est de 212,8, à l'exception de 2013, année de transition, et 2020, année marquée par la pandémie de la COVID-19. Évidemment, notre collecte de données ne peut rendre compte du nombre d'œuvres publiées en 2023, puisque la collecte de données s'est déroulée de janvier à juillet 2023.

La mise en graphique des métadonnées sur la parution des ouvrages met en relief l'influence des films dans la production d'œuvres littéraires, mais aussi de quantifier l'effet « Disney » sur les publications. De même, le rassemblement des métadonnées permet d'identifier de manière formelle un enjeu de représentation de genre du côté des créateur-rices qui publient des œuvres dans l'univers étendu de Star Wars (voir **Figure 2**).

Pour les 2 822 œuvres du corpus, nous avons recensé 4 259 auteur-rices, ce qui souligne la place que joue la co-écriture dans l'univers. De ces 4 259 auteur-rices, 3 387 ont été identifiées comme « masculin-es », et 872 comme « féminin-es » : 80 % du corpus a été écrit par des personnes identifiées comme masculines, marquant de manière indélébile l'enjeu de la représentativité au sein des pratiques littéraires dans l'univers de Star Wars. De plus, nous devons aussi souligner que les données ne permettent pas d'observer des identités de genre à l'extérieur de la binarité masculin/féminin, notamment en raison d'un système de classification sur la plupart des sites Internet qui repose essentiellement sur cette binarité.

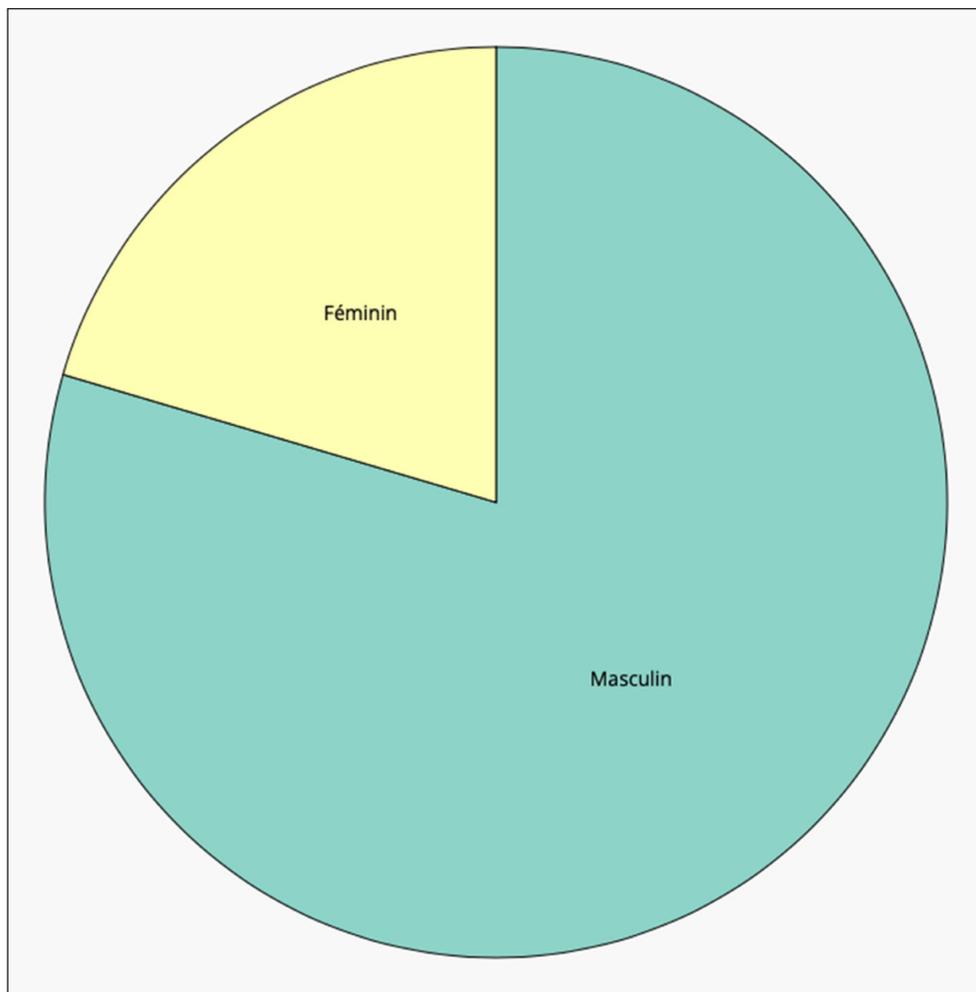


Figure 2 : Représentation des genres des créateur·rice·s. Données récoltées au début de l'année 2023.

3.2 Rapprochement avec les productions de fan

En étudiant les données tirées d'un échantillon des publications quant à Star Wars sur AO3, nous pouvons observer la fluctuation de la publication des fans au fil des ans (voir **Figure 3**) et la mettre en relation avec celle des œuvres officielles de l'univers de Star Wars. Peu de fanfictions de cet univers sont répertoriées avant les années 2010. Toutefois il faut prendre en compte que le site AO3 a été créé en 2008 par l'Organization for Transformative Works, une organisation sans but lucratif et de militantisme fanique, mise sur pied en réponse à la création de sites web qui monétisaient les fanfictions. Notons que les productions répertoriées qui prédatent AO3 proviennent d'autres sites de fans maintenant agrégés à leur répertoire.

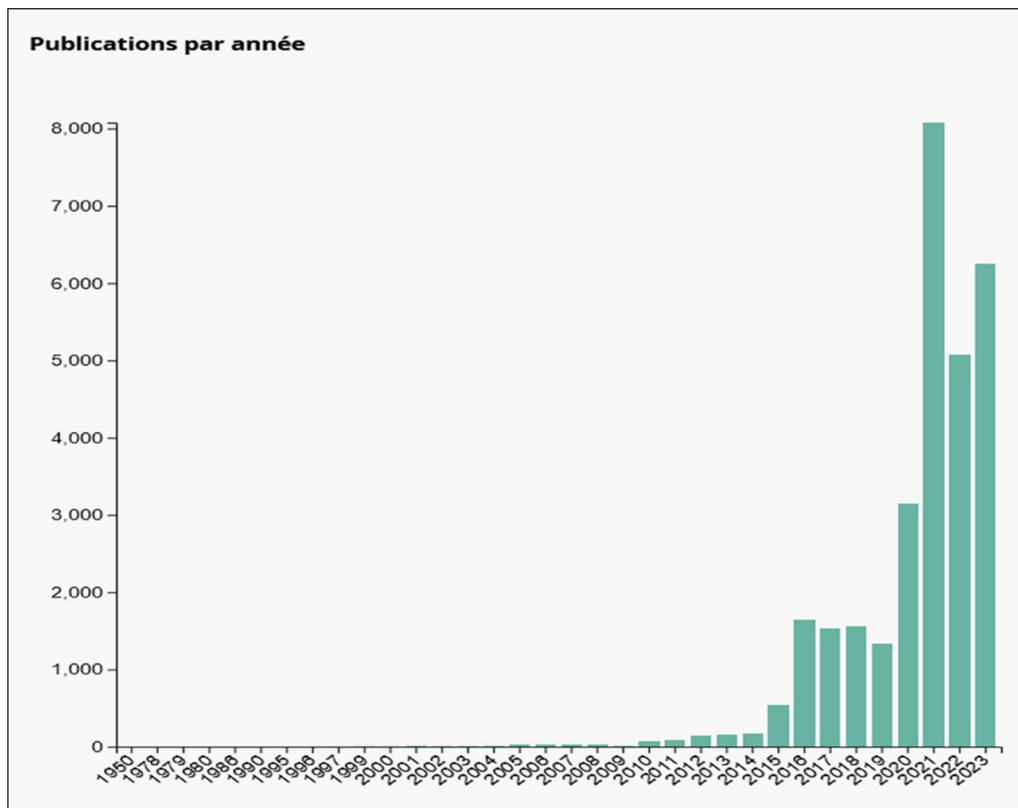


Figure 3 : Nombre de publications de fans recensées par année sur le site AO3, sur un échantillon de 29 950 textes. Données récoltées au début de l'année 2023.

Cependant, nous pouvons remarquer que la production des fans a pris de l'ampleur après la sortie du septième épisode de l'ennéade. En effet, entre 2015 et 2016, le nombre de fanfictions a triplé (passant de 544 à 1645) et est resté constant jusqu'à la sortie de l'épisode 9. Bien que *Star Wars: Episode IX – The Rise of Skywalker* (Abrams 2019) ait reçu des critiques mitigées, le consensus était tout de même clair : « [it] concludes this beloved saga with fan-focused devotion » (Rotten Tomatoes 2019). Ainsi, il n'a pas été surprenant de voir apparaître une panoplie de fanfictions qui prolongent l'histoire de Rey et Kylo Ren, mais également celle de Poe Dameron et Finn (nous y reviendrons plus loin). D'ailleurs, lorsque nous nous intéressons de plus près aux étiquettes (*tags*) de la catégorie « Fandom » (en excluant le tag « All Media Types » qui s'apparente davantage à un *tag parent*), c'est sous « Star Wars Sequel Trilogy » que l'on retrouve le plus grand nombre d'œuvres (15 699 sur les 29 950 recensées, voir **Figure 4**). Alors que la saga des Skywalker a officiellement pris fin, nous pouvons supposer que les fans désirent continuer l'aventure de ces personnages, mais également qu'une nouvelle génération de fans s'est emparée de cette bannière de l'épique *space opera*.

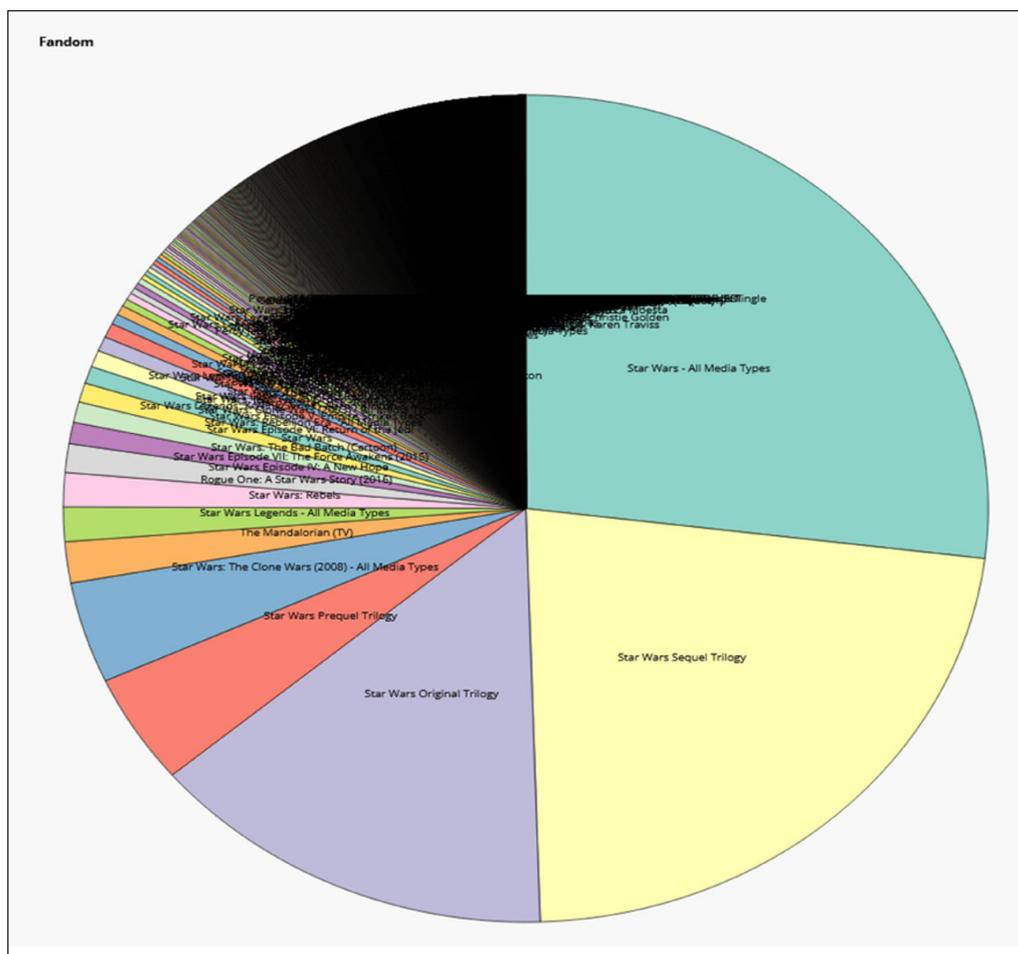


Figure 4 : Distribution des publications de fans recensées selon le tag « Fandom » sur un échantillon de 29 950 textes de AO3. Données récoltées au début de l'année 2023. (Une telle quantité de données dépasse sans doute les capacités des outils de visualisation à produire une représentation lisible, ce qui explique que la qualité du graphique ne soit pas optimale.)

Il est aussi possible de remarquer que l'augmentation des publications officielles depuis 2019 a insufflé un désir de prolonger l'expérience des fans créateur-ices. Si nous nous focalisons seulement sur la production télévisuelle, avec la première saison de la série *The Mandalorian* (Favreau 2019–présent), dont la réception a été des plus élogieuses, et le retour de *Star Wars Rebels* (Kinberg, Filoni, et Beck 2014) pour une dernière saison sur Disney+, il n'est pas surprenant que la quantité de productions de fans soit montée en flèche en 2021 (8 078 textes répertoriés, soit près de 5 000 de plus que l'année précédente), surtout en prenant en compte l'impact de la pandémie de la COVID-19. Les multiples confinements ont été des moments clés de la consommation de masse à domicile (Fize, Le Calvé, et Poux 2022), et il nous paraît logique que ces

temps de latence entre les reprises de la vie quotidienne, jumelés à l'effet de réclusion, aient décuplé les opportunités créatives des fans. Toutefois, ne connaissant pas l'emplacement géographique des créateur-riche-s sur AO3, il serait difficile de faire des corrélations pour prouver de manière significative l'effet des confinements sur la production de fans.

En revanche, il nous est possible de constater qu'il existe une correspondance manifeste entre la création fanique et la production officielle, tout particulièrement lorsqu'il s'agit d'œuvres anticipées par le public. Par exemple, si nous observons les données de publication de 2023 (considérant qu'il s'agit de celles du début de l'année et quelles sont loin d'être complètes), nous pouvons déjà apercevoir l'effet de la troisième saison de *The Mandalorian* (maintenant 1116 textes au sein du *tag* parent « Star Wars »), de la seconde pour *Star Wars: The Bad Batch* (522 textes à présent ; Filoni 2021) et de l'arrivée de la série *Obi-Wan Kenobi* (Crow 2022), trois téléséries particulièrement attendues par les fans (Prange 2021 ; Prange 2022). Cela se perçoit, entre autres, par l'apparition de nouveaux *tags* (p. ex. « Star Wars: Obi-Wan Kenobi (TV) » pour la catégorie « Fandom »), l'augmentation du nombre de *tags* présent pour certains personnages (Obi-Wan Kenobi étant dorénavant le personnage le plus souvent étiqueté, puisque la nouvelle série capte l'attention), mais également la hausse du nombre de *tags* pour diverses relations : le *pairing* « Obi-Wan Kenobi/Anakin Skywalker » a clairement gagné en popularité depuis leur retour à l'écran de Darth Vader et de son ancien maître Jedi.

3.3 Rapprochements queers faits par les fans

Un des éléments particulièrement distincts entre les productions de fan et celles officielles est la représentation de personnages et de relations queers dans les textes. Alors qu'il n'est plus à démontrer que les fanfictions présentent des espaces dans lesquels les fans peuvent développer leur identité, il est devenu apparent que c'est par l'entremise de leurs productions (et de leur consommation) que plusieurs explorent des perspectives non-cis/hétéronormatives quant au genre et à la sexualité (Floegel 2020 ; Hayes et Ball 2009). Cela se présente, entre autres, par le *slash* queer, c'est-à-dire les fanfictions qui mettent en scène une relation amoureuse ou sexuelle entre deux personnages de même sexe qui n'existe pas dans l'œuvre originale. Le terme *slash* vient de l'utilisation de la barre oblique (*slash* en anglais) pour décrire le couple mis en scène (souvent appelé *pairing*), les personnages étant notés X/Y (p. ex. Kylo Ren/Rey, Poe Dameron/Finn). Plutôt que le nom des personnages, il n'est pas rare de voir la lettre M pour « masculin » ou F pour « féminin ». Ainsi, la mention M/M ou F/F prévient le-la lecteur-riche que la fanfiction est du *slash* homosexuel sans pour autant dévoiler le *pairing* (le couple retenu). La barre oblique tend à être parfois remplacée par la lettre

x minuscule ; selon le même principe, les personnages Y et Z seront présentés comme YxZ (p. ex. KyloxHux pour Kylo Ren/Armitage Hux).

Ainsi, dans les options de tri utilisées par la communauté, il est possible d'indiquer précisément ces aspects des fictions et de constater la prégnance des relations entre les personnages des univers. Dans le cadre de notre projet, nous avons relevé les indications de catégories (voir **Figure 5**), puisque nous n'avons pas encore traité la teneur des *tags* dans « Relationship » en détail, dû à leur quantité impressionnante et au manque de conformité au sein même du répertoire d'AO3. En effet, malgré les directives, tant officielles et qu'officieuses, il faut prendre en considération que le site ne cesse d'évoluer, et qu'une certaine liberté existe dans la création des *tags*. La liberté offerte aux fans dans la création et l'utilisation des *tags* leur permet de circonscrire adéquatement le contenu des œuvres, mais cette même liberté devient difficilement adéquate à la visualisation lors d'une collecte de données, tant les *tags* sont nombreux et variés.

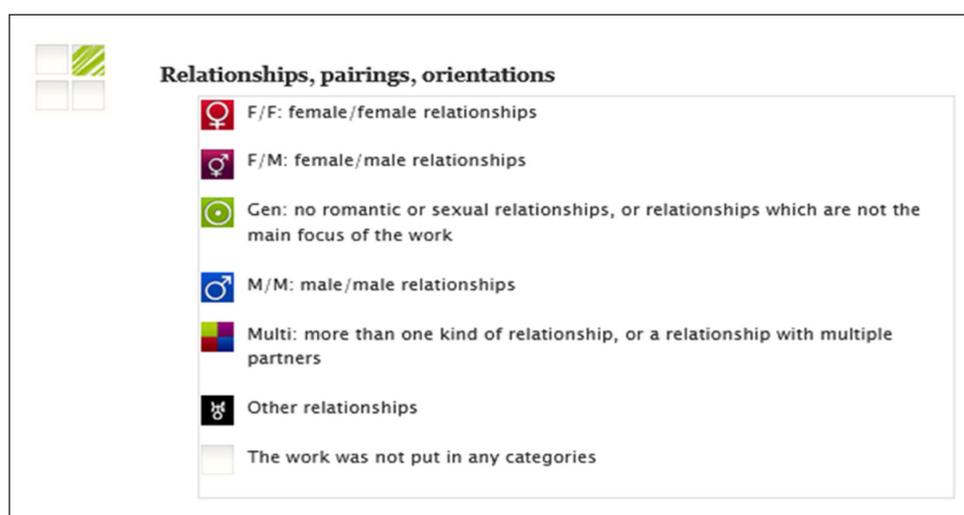


Figure 5 : Légende de la catégorie « Relationships, pairings, orientations » pour permettre de filtrer les fanfictions.

Pour l'instant, nous pouvons constater (voir **Figure 6**) le pourcentage de fanfictions qui possèdent des relations répertoriées dans la catégorie « Relationships, pairings, orientations » pour l'échantillon de 29 950 textes. Pour les 33 274 étiquettes de la catégorie « Relationships, pairings, orientations » de notre échantillon de fanfictions, près de la moitié (43,42 %) présentent des relations hétérosexuelles. (Il est possible d'étiqueter une fanfiction avec plusieurs tags d'une même catégorie, c'est pourquoi le nombre de tags est plus grand que le nombre d'œuvres recensées.) Avec l'engouement autour du couple Kylo Ren/Rey depuis l'épisode 9 et la prépondérance des relations

hétérosexuelles au sein des œuvres officielles (par ex. Leia Organa et Han Solo, Anakin Skywalker et Padmé Amidala), il n'est pas surprenant de voir une plus grande représentation de couples hétéronormés dans les fanfictions. En revanche, il est important de relever la présence des fictions représentant l'homosexualité masculine qui recoupe le quart des fanfictions (25,34 %). Nous pouvons y apercevoir ici le désir queer des fans qui veulent que leurs écrits s'attaquent à des personnages qui reproduisent la masculinité traditionnelle, les hiérarchies de classe et de race, les scripts relationnels traditionnels, et les reconfigurent en récits de sociétés communautaires, d'égalité raciale et de transgression sexuelle (Kustritz 2003). Nous pouvons affirmer que les fanfictions *slash* sont un lieu de prédilection pour la résistance queer puisque, comme Joanna Russ (Russ 1985) l'a souligné dans son essai *Pornography by Women for Women, with Love*, elles combinent autant des traits de sensibilité émotionnelle masculins que féminins. Une telle relation égalitaire, affirme-t-elle, conteste le déséquilibre des pouvoirs généralement observé dans les fanfictions hétéronormatives.

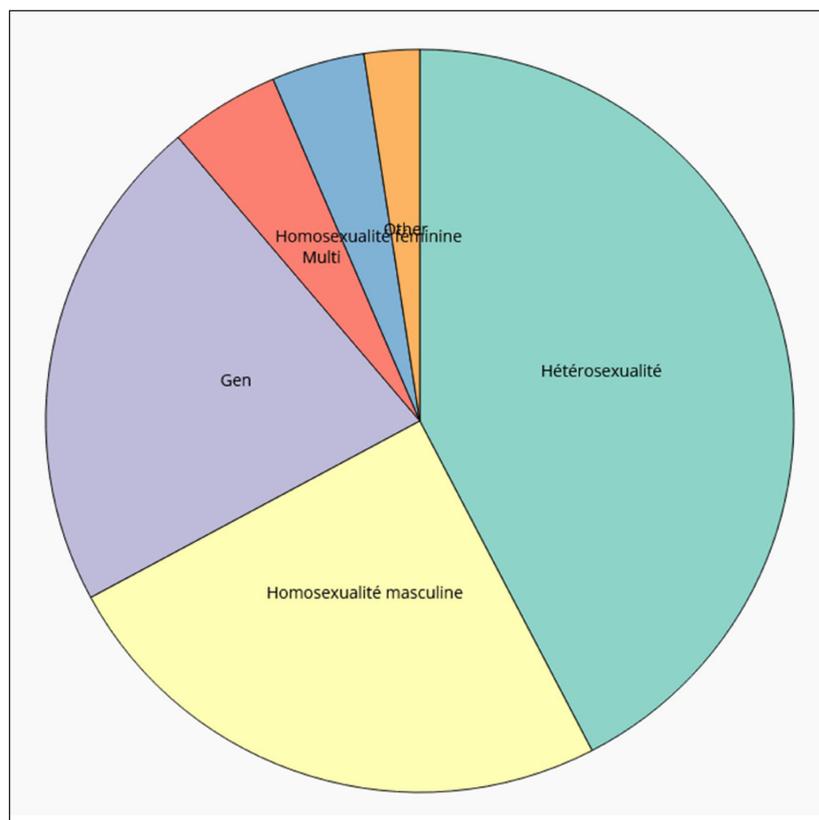


Figure 6 : Publications de fans recensées selon la catégorie « Relationships, pairings, orientations » sur le site AO3, sur un échantillon de 29 950 textes. Données récoltées au début de l'année 2023.

Conclusion

Cette étude de corpus nous a permis de montrer quelques-uns des traits distinctifs des œuvres publiées dans les univers étendus, autant dans celles produites par les canaux officiels que dans celles par les fans. La mise en relation de ces deux corpus encourage la réflexion sur le rôle que joue autant les professionnel·les que les fans dans la création et le développement des univers fictifs, notamment en littératures de l'imaginaire. Cette relation, nous croyons, doit être abordée par le biais de la complémentarité, plutôt que par la reconnaissance d'une disparité, voire d'une opposition, entre les contenus officiels et officieux. En ce sens, le projet Répertoire promeut le dialogue entre les différentes formes que peut revêtir un univers fictif.

Notre collecte de données sur l'univers étendu de Star Wars nous a permis de souligner certaines caractéristiques de la production de cet univers, notamment l'influence des fans et l'impact de l'achat de la franchise par Disney. Ces caractéristiques, lesquelles (re)façonnent constamment l'univers Star Wars, permettent de faire l'histoire de l'évolution de la production de l'intérêt envers les œuvres de son univers. Dans notre perspective de recherche, elles permettent aussi d'approfondir le thème de la représentativité. Si nous soulignons d'emblée les limites d'une analyse du côté des auteur·rice·s, puisque l'information demeure parcellaire, l'étude des représentations queers dans un univers fictif est, quant à elle, bien possible. Elle ouvre sur les formes de représentativité, interroge les formes d'altérité rendues possibles par les littératures et cultures de l'imaginaire, tout en ouvrant le dialogue entre les enjeux contemporains et le contenu des œuvres elles-mêmes.

Avec cette analyse, nous cherchons à encourager davantage d'études sur de larges corpus thématiques qui interrogent autant la corrélation entre les œuvres officielles et celles officieuses que les thématiques et enjeux de représentation au sein de celles-ci. Cette première analyse nous a permis de tester notre méthodologie et d'observer la faisabilité d'une étude statistique sur un large corpus, en dehors des données strictement linguistiques. Une seconde étape dans le projet Répertoire, outre poursuivre l'effort de collecte et de nettoyage, est de travailler au balisage des œuvres et à l'extraction des données afin d'approfondir l'effort collaboratif entre les fans et les auteur·rice·s dans le développement de lieux ou univers imaginaires.

Déclaration d'intérêt

Les auteurs déclarent n'avoir aucun conflit d'intérêts relativement à la rédaction et au contenu de cet article.

Contributions

Contributions des auteur·rice·s

Les contributions des auteurs, décrites à l'aide de la taxonomie CrediT de la Information Standards Organization (NISO), sont les suivantes:

Nom et initiales de l'auteur :

Emmanuelle Lescouet (EL)
Pierre Gabriel Dumoulin (PGD)
Amélie Vallières (AV)

Les auteurs sont classés par ordre décroissant selon l'importance de leur contribution. L'auteure correspondante est EL.

Conceptualisation : EL
Traitement des données : EL
Analyse formelle : EL, AV, PGD
Enquête : EL, AV, PGD
Méthodologie : EL
Ressources : EL
Validation : AV, PGD
Rédaction – Préparation du brouillon original : EL, AV, PGD
Rédaction – Révision et édition : AV, PGD

Contributions éditoriales

Éditeur.e.s-en-chef invité.e.s de l'édition spéciale

Jason Boyd, Université métropolitaine de Toronto, Canada
Bárbara Romero-Ferrón, Université de Malaga, Espagne

Éditeur de section et de traduction

Davide Pafumi, The Journal Incubator, Université de Lethbridge, Canada

Éditrice de la production

Christa Avram, The Journal Incubator, Université de Lethbridge, Canada

Éditeur de fonds d'écran

A K M Iftekhar Khalid, The Journal Incubator, Université de Lethbridge, Canada

Bibliographie

- Abrams, Jeffrey Jacob, réal. 2019. *Star Wars: Episode IX – The Rise of Skywalker*. Lucasfilm Ltd. et Bad Robot Productions.
- Barnabé, Fanny. 2014. « La ludicisation des pratiques d'écriture sur Internet : une étude des fanfictions comme dispositifs jouables ». *Sciences du jeu* 2. Consulté le 26 février 2025. <https://doi.org/10.4000/sdj.310>.
- Besson, Anne. 2015. *Constellations: des mondes fictionnels dans l'imaginaire contemporain*. Paris: CNRS éditions.
- Bréan, Simon. 2012. *La science-fiction en France: théorie et histoire d'une littérature*. Paris: Presses de l'Université Paris Sorbonne.
- . 2022. « Pour un usage externe des théories de la science-fiction ». *ReS Futurae* 20. Consulté le 26 février 2025. <https://doi.org/10.4000/resf.11250>.
- Canavan, Gerry. 2018. « Fandom Edits: “Rogue One” and the New Star Wars ». Dans *Star Wars and the History of Transmedia Storytelling*, dirigé par Sean Guynes et Dan Hassler-Forest, 277–288. Amsterdam: Amsterdam University Press.
- CNRS (Centre National de la Recherche Scientifique). 2025. « OpenTheso ». Consulté le 26 Février 2025. <https://opentheso.huma-num.fr/opentheso/>.
- Coppa, Francesca. 2017. *The Fanfiction Reader: Folk Tales for the Digital Age*. Ann Arbor, MI: University of Michigan Press.
- Crow, Deborah, réal. 2022. *Obi-Wan Kenobi*. Lucasfilm Ltd.
- Dumoulin, Pierre Gabriel. 2024. « Show Me Your True Form! Traduire la salle de classe dans les jeux vidéo ». *Revue de recherches en littérature médiatique multimodale* 20. Consulté le 26 février 2025. <https://revuemultimodalites.com/articles/dumoulin>.
- Edwards, Gareth, réal. 2016. *Rogue One: A Star Wars Story*. Lucasfilm Ltd.
- Ensslin, Astrid. 2014. *Literary Game*. Cambridge, MA: The MIT Press.
- Favreau, John, réal. 2019–présent. *The Mandalorian*. Fairview Entertainment, Golem Creations, Lucasfilm, et Walt Disney Studios.
- Filoni, David, réal. 2021. *Star Wars: The Bad Batch*. Disney+, Lucasfilm Animation, et Lucasfilm.
- Fize, Etienne, Titouan Le Calvé, et César Poux. 2022. « La crise a-t-elle laissé la culture en jachère? Analyses à partir de données bancaires. » *Focus du CAE* 080-2022. Consulté le 26 février 2025. https://www.cae-eco.fr/staticfiles/pdf/cae_focus080.pdf.
- Floegel, Diana. 2020. « “Write the Story You Want to Read”: World-Queering through Slash Fanfiction Creation ». *Journal of documentation* 76 (4) : 785–805. Consulté le 26 février 2025. <https://doi.org/10.1108/JD-11-2019-0217>.
- Guo, Shuojia. 2020. « Return of Fandom in the Digital Age with the Rise of Social Media ». Dans *The Handbook of Research on the Impact of Fandom in Society and Consumerism*, dirigé par Cheng Lu Wang, 193–210. Hershey, PA: IGI Global.

- Hayes, Sharon, et Matthew Ball. 2009. « Queering Cyberspace: Fanfiction Communities as Spaces for Expressing and Exploring Sexuality ». Dans *Queering Paradigms*, dirigé par Bee Scherer, 219–239. Lausanne: Peter Lang Publishing.
- Hellekson, Karen. 2009. « A Fannish Field of Value: Online Fan Gift Culture ». *Cinema Journal* 48 (4) : 113–118. Consulté le 26 février 2025. <https://www.jstor.org/stable/25619733>.
- Hoover, David. L., Jonathan Culpeper, et Kieran O'Halloran, éd. 2014. *Digital Literary Studies. Corpus Approaches to Poetry, Prose, and Drama*. New York : Routledge.
- Houston, Natalie M. 2023. « Distant Reading ». Dans *Technology and Literature*, dirigé par Adam Hammond, 361–376. Cambridge : Cambridge University Press.
- Jenkins, Henry. 2006. *Fans, Bloggers, and Gamers: Exploring Participatory Culture*. New York : New York University Press.
- Jockers, Matthew L. 2013. *Macroanalysis: Digital Methods and Literary History*. Urbana, IL : University of Illinois Press.
- Kinberg, Simon, Dave Filoni, et Carrie Beck, réal. 2014. *Star Wars Rebels*. CGCG Inc., Disney XD, Entertainment Clearances, Lucasfilm Animation, Lucasfilm, The L.A. Studios, et Virtuoso-Sparx.
- Kustritz, Anne. 2003. « Slashing the Romance Narrative ». *The Journal of American Culture*, 26 (3) : 371–384. Consulté le 26 février 2025. <https://doi.org/10.1111/1542-734X.00098>.
- Lescouet, Emmanuelle. 2022. « Gestes de lecture numérique et lecture immersive de science-fiction ». *ResFuturae* 20. Consulté le 26 février 2025. <https://doi.org/10.4000/resf.11364>.
- Lichtenberg, Jacqueline, Sondra Marshak, et Joan Winston. 1975. *Star Trek Lives!* New York : Bantam Books.
- Liu, Alan. 2020. « Toward a Diversity Stack: Digital Humanities and Diversity as Technical Problem ». *PMLA* 135 (1) : 130–151. Consulté le 26 février 2025. <https://doi.org/10.1632/pmla.2020.135.1.130>.
- Lucas, George, réal. 1977. *Star Wars: Episode IV – A New Hope*. Lucasfilm Ltd.
- , réal. 1999. *Star Wars: Episode I – The Phantom Menace*. Lucasfilm Ltd.
- , réal. 2005. *Star Wars: Episode III – Revenge of the Sith*. Lucasfilm Ltd.
- Maci, Stefania M., et Michele Sala. 2022. « Corpus Linguistics and Translation Tools for Digital Humanities : An Introduction ». Dans *Corpus Linguistics and Translation Tools for Digital Humanities: Research Methods and Applications*, dirigé par Stefania M Maci et Michele Sala, 1–16. London : Bloomsbury.
- Marquand, Richard, réal. 1983. *Star Wars: Episode IV – Return of the Jedi*. Lucasfilm Ltd.
- Masure, Anthony. 2017. *Design et humanités numériques*. Paris: Éditions B42.
- Mohany, Simon. 2018. « Cultural Diversity and the Digital Humanities. » *Fudan Journal of the Humanities and Social Sciences*, 11 : 371–388. Consulté le 26 février 2025. <https://doi.org/10.1007/s40647-018-0216-0>.
- Pandemic Studios. 2004. *Star Wars: Battlefront*. LucasArts. Windows, PlayStation 2, Xbox.

Prange, Stephanie. 2021. « “Star Wars: The Bad Batch” Top Anticipated New Show in May – Media Play News ». *Media Play News*, 29 avril. Consulté le 26 février 2025. <https://www.mediaplaynews.com/star-wars-the-bad-batch-top-anticipated-new-show-in-may/>.

———. 2022. « Disney+’s “Obi-Wan Kenobi” Top New Show, Netflix’s “Stranger Things” Top Returning Show Anticipated in May ». *Media Play News*, 28 avril. Consulté le 26 février 2025. <https://www.mediaplaynews.com/disneys-obi-wan-kenobi-top-new-show-netflixs-stranger-things-top-returning-show-anticipated-in-may/>.

Repertorium. 2023. « Carnet de projet ». *Hypotheses.org*. Consulté le 26 février. <https://repertorium.hypotheses.org/carnet-de-projet>.

Roark, Nathaniel. 2023. « When Did Disney Buy Star Wars? ». *Screen Rant*. Consulté le 26 février 2025. <https://screenrant.com/when-did-disney-buy-star-wars/#:~:text=October%2030%2C%202012%20%2D%20Disney%20Acquires,Disney%20in%20a%20press%20release>.

Rotten Tomatoes. 2019. « Star Wars: The Rise of Skywalker Critics ». *Rotten Tomatoes*. Consulté le 26 février 2025. https://www.rottentomatoes.com/m/star_wars_the_rise_of_skywalker.

Russ, Joanna. 1985. *Magic Mommas, Trembling Sisters, Puritans and Perverts: Feminist Essays*. Trumansburg, NY : The Crossing Press.

Saint-Gelais, Richard. 1999. *L'empire du pseudo : modernités de la science-fiction*. Montréal : Nota Bene.

Spivak, Gayatri Chakravorty. 2005. *Death of a Discipline*. New York : Columbia University Press.

Stanfill, Mel, et Megan Condis. 2014. « Fandom and/as Labor ». *Transformative Works and Cultures* 15. Consulté le 26 février 2025. <https://doi.org/10.3983/twc.2014.0593>.

Tauveron, Catherine. 2004. « Fonctions et natures des lectures en réseaux ». *Éduscol*. Consulté le 26 février 2025. eduscol.education.fr/cid46319/fonctions-et-nature-des-lectures-en-reseaux.html.

Underwood, Tom. 2016. « Distant Reading and Recent Intellectual History ». *Debates in Digital Humanities* 44 : 530–533. Consulté le 26 février 2025. <https://dhdebates.gc.cuny.edu/read/untitled/section/3b96956c-aab2-4037-9894-dc4ff9aa1ec5>.

Wuyckens, Géraldine. 2021. « Design Fiction as a Critical Inquiry Method for Developing Media Literacy ». Consulté le 26 février 2025. <https://doi.org/10.13140/RG.2.2.14870.88640>.

