



HAL
open science

Penser (pour) l'écran

Christophe Schuwey

► **To cite this version:**

Christophe Schuwey. Penser (pour) l'écran. Le Verger, 2022, 23, http://cornucopia16.com/wp-content/uploads/2022/04/2_Schuwey_VFF.pdf. hal-04345526

HAL Id: hal-04345526

<https://hal.univ-brest.fr/hal-04345526>

Submitted on 14 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



PENSER (POUR) L'ÉCRAN¹

Christophe SCHUWEY (U. Yale)

À quoi ressembleraient aujourd'hui les humanités numériques si, lors de la conception d'un projet, l'on réfléchissait à ce que l'on voulait créer avant de penser aux données et aux contraintes techniques ? si l'on commençait par prendre une feuille blanche pour dessiner ce que l'on souhaiterait afficher, plutôt que de commencer par réfléchir à la solution technique, aux métadonnées ou aux jeux de balises TEI nécessaires ? si l'on cessait de singer le livre et que l'on exploitait au contraire le potentiel du numérique – le visuel, mais aussi le son – pour créer un autre rapport à nos objets ?

Scénarios d'utilisation

La plupart des services de musique en ligne organisent leur catalogue en catégories simpl(ist)es : titre de la chanson - artiste - album - genre - année². Le modèle correspond à un scénario unique : celui de la musique contemporaine, soit un album original fait de chansons, enregistrées par un artiste unique. Le système révèle rapidement ses limites : dans le cas d'une reprise, par exemple, comment indiquer l'auteur de la version originale ? ou l'auteur de l'adaptation, qui n'est pas forcément l'interprète de la reprise ? faut-il spécifier cela dans le champ « titre » ? ou utiliser le champ « compositeur » ? Mais les artistes ne sont pas nécessairement compositeurs de leurs chansons. Le problème est plus évident encore dans le cas du jazz ou de la musique dite classique. Les spécificités de celle-ci – multitude d'interprètes, prédominance du compositeur, œuvre répartie sur plusieurs pistes, etc. – se plient fort mal au modèle des services de musique en ligne³. Faute d'un schéma de métadonnées adapté, compositeur et interprètes se retrouvent empilés dans le champ « artiste » et des éléments tels que le nom de l'œuvre se retrouvent entassés dans le champ « titre » :

¹ Cette communication reprend des éléments de mon ouvrage *Interface : l'apport des humanités numériques à la littérature*, Neuchâtel, Alphil, 2019. Elle est soutenue par le Whitney et Betty MacMillan Center for International and Area Studies at Yale. Le titre s'inspire d'un cours de Sophie Jausi à l'Université de Fribourg : "Écrire (sur) l'écran".

² Bien sûr, d'autres informations sont disponibles, mais l'interface et la logique de ces services se concentrent autour de ces données primaires.

³ Voir notamment Martin Pichl *et al.*, « Understanding Playlist Creation on Music Streaming Platforms », actes de l'*International Symposium on Multimedia* de 2016 à San Jose, p. 475-480 ; Gregory Camp, « Spotify », *Journal of the Society for American Music*, n° 9, 2015, p. 375-378. Voir la réflexion technique de Martin Gasser *et al.*, « Classical music on the web – users interfaces and data representation », présentation à l'*International Society for Music Information Retrieval Conference*, Málaga, 2015, consulté le 24 mai 2020, URL : <https://pdfs.semanticscholar.org/5bbc/17f103098d0b0154b6aeb0a7a06c71bfd93a.pdf>. Voir également l'article de Ben Sisario, « In Streaming Age, Classical Music Gets Lost in the Metadata », *New York Times*, 23 juin 2019, consulté le 24 mai 2020, URL : <https://www.nytimes.com/2019/06/23/business/media/stream-classical-music-spotify.html>, ainsi que le publi-reportage de Corey Semour qui résume bien les différents enjeux, « Idagio—the Spotify for Classical Music—Has Changed My Life », *Vogue*, 15 mai 2019, consulté le 24 mai 2020, URL : <https://www.vogue.com/article/idagio-spotify-for-classical-music>.



Figure 1. L'exemple d'Apple Music. Les éléments sont rassemblés dans les champs titre (cadre rouge) et artiste (cadre vert). © Apple

L'opération est violente : les informations de l'archive – en l'occurrence, le disque – telles que compositeur, artiste (et rôle de l'artiste), numéro d'œuvre, œuvre (Symphonie concertante), mouvement, sont tordues pour correspondre à une logique exogène, ce qui implique une perte importante et un manque de lisibilité pour les utilisateurs.

Ainsi décrite, la structure de données paraît difficilement justifiable. Comment expliquer son adoption par les principaux services de musique en ligne ? Pour le comprendre, il faut réfléchir à l'utilisation de ces données. Un service de musique en ligne n'est pas une base de données documentaire de l'histoire du disque. Il obéit à deux enjeux. D'une part, les services de musique en ligne ont notamment misé sur une impression d'abondance pour s'imposer face à leurs concurrents : il faut de la masse, du nombre. De l'autre, ils privilégient l'efficacité de la recherche : le but est que l'utilisateur atteigne rapidement le morceau souhaité ou qu'il découvre de nouveaux contenus pertinents⁴. Le service en ligne n'a pas pour but de cataloguer soigneusement tous les enregistrements existants de *Lucia di Lammermoor*, mais bien de proposer des résultats (juste) assez bons pour que l'utilisateur trouve ce qu'il cherche, se contente de la version qu'on lui propose, ou découvre ce qu'il ne croyait pas chercher. En d'autres termes, le scénario d'utilisation principale est celui d'une recherche par artiste, titre d'album, éventuellement de compositeur dans un champ de recherche unique. Dans ce cas de figure, les métadonnées en vrac et redondantes ne sont alors plus vraiment un problème, puisque c'est l'algorithme de recherche qui se charge de les parcourir et de lister les résultats les plus pertinents. Un étiquetage plus détaillé et plus précis ou un schéma de métadonnées plus complexe n'amélioreraient pas nécessairement cet usage et supposerait en revanche un gigantesque travail de catalogage sans bénéfice apparent.

⁴ Voir Anja Nylund Hagen, « Music Streaming the Everyday Life » dans Nowak R., Whelan A. (éds.), *Networked Music Cultures. Pop Music, Culture and Identity*, London, Palgrave Macmillan, 2016, p. 227-245. Ying-Han Chang, Muh-Chyun Tang, « Serendipity with Music Streaming Services: The Mediating Role of User and Task Characteristics » dans G. Chowdhury et al. (éds.), *Transforming Digital Worlds*, New York, Springer, 2018, p. 435-441.



Ce modèle remplit-il toutefois les fonctions attendues ? Lorsque l'on recherche une sonate de Beethoven dans les services en ligne, les intégrales de Fazil Say ou d'Alfred Brendel sont noyées dans les *50 Greatest Pieces of Classical Music* et les *Music to sleep by*. En outre, dès que la recherche devient complexe (par exemple, lorsque l'on précise un numéro de sonate), l'algorithme se perd dans la structure et la qualité des résultats est souvent aléatoire. L'utilisateur se retrouve à devoir parcourir manuellement toutes les pistes d'une intégrale des sonates pour trouver la sonate qu'il souhaite précisément écouter. Enfin, les suggestions sont souvent décevantes : « Classique » n'est pas un genre : entendre un concerto pour piano de Mozart après du John Cage n'est pas nécessairement pertinent. En d'autres termes, même en tenant compte du scénario d'utilisation tel que décrit plus haut, le modèle de données et l'algorithme sont insuffisants pour répondre de manière satisfaisante aux requêtes⁵. À partir de ce constat, certains services (*Idagio*, *Primephonic*) se sont lancés sur le marché en se spécialisant dans la musique classique⁶. Or, la différence avec leurs concurrents réside bien moins dans le catalogue que dans la présentation et la navigation. La musique offerte est, sinon entièrement, du moins largement identique aux autres plateformes, mais les outils de recherche et surtout les suggestions d'écoute sont pensées en fonction des spécificités de la musique classique. Les mêmes données présentées et interrogées différemment donnent ainsi lieu à deux services entièrement différents, pour des publics également différents.

L'exemple de la musique en ligne illustre un constat fondamental : les données informatiques n'existent pas en dehors de leur interprétation et leurs manipulations : leur préparation, leur sélection, leur encodage sont étroitement liés à leur utilisation, aux analyses ou aux affichages que l'on prévoit d'en faire ; bref, les données informatiques se pensent en fonction de l'interface⁷. Difficile en effet de préparer des données ou d'évaluer leur qualité de manière essentielle, en dehors de tout scénario d'utilisation. Leur sens et leur valeur dépendent de la façon dont elles sont présentées et des outils disponibles pour les interroger. Le constat n'est pas un plaidoyer pour une approche utilitariste des données ; il s'agit au contraire de comprendre que l'encodage des données ne peut se penser en dehors d'un scénario d'utilisation et des implications sur la création de projets en humanités numériques. « Numériser X » ne précède pas « afin de Y » : l'un et l'autre sont fondamentalement liés.

L'application des humanités numériques à la littérature illustre toute l'importance d'une telle observation. Qu'il s'agisse d'éditions numériques ou de musique numérique, de réalisations universitaires ou de plateformes commerciales, toute donnée pose en effet un problème d'interface. Même lorsque l'on croit travailler sans penser à la finalité, on obéit en réalité à une structuration préalable inconsciente. Lorsque, dans une édition critique, l'on décide de baliser les chapitres et les paragraphes d'un roman en TEI, sans penser à la finalité de la chose, l'on suit en réalité nos réflexes de littéraires et le scénario d'utilisation par défaut de la TEI, c'est-à-dire une description du livre source au travers de balises prédéfinies, privilégiant certains éléments – les paragraphes – aux dépens des autres, comme la typographie. C'est donc bien la représentation que l'on se fait des données qui détermine leur encodage et leur l'organisation. La qualité des données dépend également de l'utilisation prévue. Songeons ainsi à *Google Books*, dont on connaît la reconnaissance optique erratique.

⁵ Sur les scénarii d'utilisation en musique numérique, voir note 4.

⁶ Corey Semour, art. cit. À l'heure où nous écrivons ces lignes, Apple Music vient d'ailleurs de racheter Primephonic, précisément pour compenser leur interface inadaptée à la musique classique et au jazz.

⁷ Par interface, j'entends ici toute transformation des données en une forme compréhensible par l'humain, qu'il s'agisse d'un texte, de son, de graphique, d'une édition en ligne, des comptages de mots, d'une base de données, etc. À noter qu'il ne s'agit pas du tout ici de diminuer l'importance des modèles informatiques de type MVC ou de contester l'importance de l'interopérabilité et la réutilisabilité des données. Le débat est à un autre niveau. Il s'agit de souligner que la structuration des données obéit nécessairement à un usage imaginé, conscient ou inconscient.

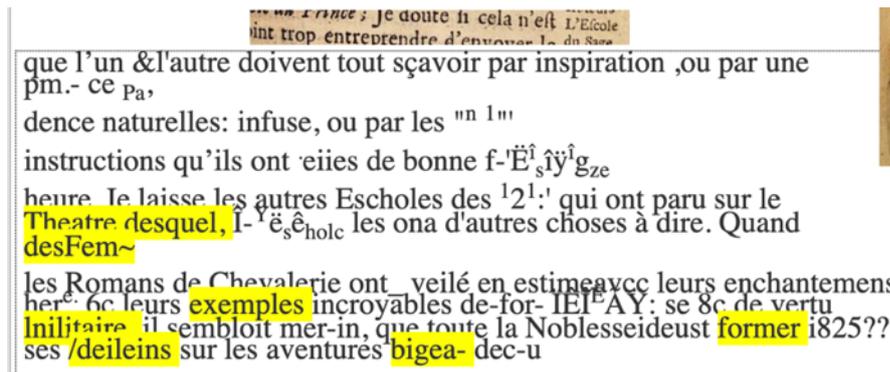


Figure 2. Mode texte de Google Books sur *La Bibliothèque française* de Charles Sorel, Paris, 1664. © Google

Un tel résultat serait inutilisable pour une édition critique : il serait bien moins fastidieux de tout recopier manuellement que de corriger ces données. Celles-ci, très imparfaites, sont-elles alors dépourvues de valeur ? Bien au contraire, mais tout dépend une fois encore des scénarios d'utilisation. Un article sur les algorithmes d'attribution des textes observe en effet que, avec une reconnaissance optique exacte à seulement 20% (un chiffre extrêmement faible), les attributions correspondent déjà aux valeurs de contrôle. Les auteurs et autrices mettent alors en garde contre le fantasme de la donnée parfaite, la qualité de celles-ci étant : « une majorité de chercheurs travaille en présupposant que la qualité des résultats est directement proportionnelle au temps investi dans la préparation des données – en d'autres termes, qu'une analyse de haute qualité ne tolère pas des données de basse qualité⁸ ».

En outre, dans le cas de *Google Books*, la qualité médiocre de la reconnaissance est nécessaire à l'existence même de la ressource. Dans l'état actuel de la technologie, et malgré les progrès importants réalisés dernièrement dans ce domaine, une reconnaissance à 100% nécessiterait en effet une retranscription manuelle et une relecture par plusieurs chercheurs et chercheuses. L'opération est impossible à envisager sur les millions d'ouvrages numérisés par *Google Books*. En acceptant des données médiocres, *Google Books* a en revanche permis la mise à disposition d'un immense corpus dont les bénéfices pour la recherche en littérature comptent parmi les plus incontestables. L'imperfection des données se voit donc compensée par la qualité des algorithmes de recherche, par le nombre d'ouvrages numérisés et la possibilité unique, pour l'utilisateur, d'interroger cette collection en mode plein texte. La reconnaissance optique ci-dessus n'est pas faite pour être lue, puisque c'est le rôle des photographies de l'ouvrage⁹. Une telle recherche ne liste pas non plus tous les livres pertinents, ni tous les livres contenant le terme recherché, et de loin – *Google Books* ne prétend pas être un catalogue exhaustif de bibliothèque – mais elle en répertorie plusieurs et, surtout, fait ressortir des livres oubliés, des sources ou des réemplois textuels significatifs¹⁰. Ajoutons

⁸ Greta Franzini *et al.*, « Attributing Authorship in the Noisy Digitized Correspondence of Jacob and Wilhelm Grimm », *Frontiers in Digital Humanities*, vol. 5, 2018, <https://www.frontiersin.org/article/10.3389/fdigh.2018.00004>, consulté le 24 mai 2020 : « [...] scholars operate under the assumption that the time invested in pre-processing is directly proportional to the quality of the results – in other words, that any high-quality analysis is intolerant to low quality data. [...] a cleanliness above \approx 20% is already sufficient to achieve a higher-than-chance probability of correct binary attribution. »

⁹ À noter en revanche que la mauvaise qualité du texte reconnu empêche la synthèse vocale. En cela, la faible qualité des données de Google constitue un véritable problème puisqu'elle manque une occasion formidable de rendre ces textes accessibles aux personnes souffrant de dyslexie lourde ou de déficiences visuelles.

¹⁰ C'est ce que reconnaissent volontiers les éditeurs et éditrices de l'édition 2010 des *Œuvres complètes* de Molière dans la Bibliothèque de la Pléiade : 70% des nouvelles découvertes proviennent de bibliothèques numériques en ligne et de services comme *Frantext*.



encore que Google semble expérimenter régulièrement de nouvelles interfaces de recherche, preuve de l'importance de cette question.

Enfin, en s'appuyant sur ce corpus tout imparfait, l'outil *Google Ngram* permet de visualiser la fréquence d'utilisation d'un mot par rapport à l'entièreté du corpus numérisé. Ce type de recherche, à peu près impossible à réaliser pour un humain, demande quelques dixièmes de seconde aux serveurs de Google. On peut ainsi observer l'âge de la sensibilité dans la littérature française :

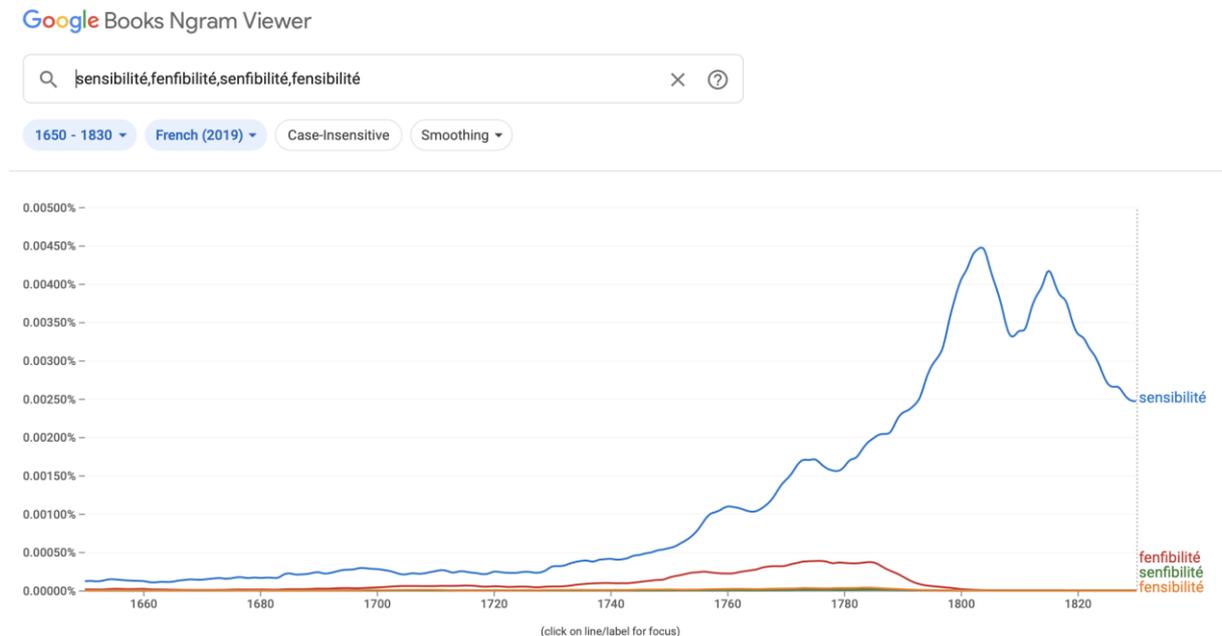


Figure 3. Google Ngram © Google. Requête obtenue à partir de la recherche des termes « sensibilité, fenfibilité, senfibilité, sensibilité » sur le corpus français (état 2019), de 1650 à 1830.

Certes, le résultat ne tient pas compte de l'ensemble de la littérature, mais la courbe de tendance est un indice numérique d'un phénomène observé par ailleurs. Elle permet également d'affiner les tendances, et invite à l'enquête sur certains segments marquants : 1740-1760, ou 1800-1820. On remarque que les mauvaises reconnaissances pèsent très peu sur l'ensemble du corpus : sur des millions d'ouvrages, les reconnaissances erronées (par exemple, les s longs reconnus comme f : « fenfibilité », « sensibilité ») sont trop rares pour fausser la courbe. Seule « fenfibilité » présente un taux d'erreur significatif, sans que cela ne modifie toutefois la tendance générale. Plus l'on remonte dans le temps, et plus le taux d'erreur est important, ce qui remet dans certains cas en cause l'utilité de l'outil, notamment sur la période 1450-1550. Toutes imparfaites qu'elles sont, les données numérisées par Google permettent toutefois de répondre aux questions prévues. En théorie, l'on souhaite toujours des données parfaites ; en pratique, ce souhait aurait empêché l'existence de *Google Ngram*. On serait alors privés d'un outil offrant une vue d'ensemble unique sur l'évolution d'un ou plusieurs termes à travers le temps, et des hypothèses qu'une telle visualisation permet de formuler dans le processus de recherche.

Qu'il s'agisse d'une édition en ligne ou d'une base de données, le risque de ne pas penser l'interface est d'investir des ressources colossales à encoder et éditer des données, à les classer, les répertorier, sans que ces informations ne puissent être utilisées à terme. Pire : il suffit que la question de recherche change, que l'orientation du projet soit modifiée, pour qu'un encodage donné, même supposément exhaustif, apparaisse soudain inadapté. Comment



déterminer ce qui est nécessaire ? Faut-il tout décrire ? Ou faut-il, à l'inverse, plier le projet aux contraintes techniques ? Ce sont les questions discutées dans les pages qui suivent, au travers d'une réflexion sur quelques applications des humanités numériques et des expériences pédagogiques et scientifiques concrètes. Dans un premier temps, je rappellerai la marginalisation de l'interface dans le débat sur les humanités numériques en déconstruisant les présupposés sur lesquels repose cette mise à l'écart et ses conséquences sur le champ. Je discuterai ensuite d'une expérience pédagogique d'édition numérique qui illustre les conséquences que l'on dégage à travailler à partir de l'interface ou, au contraire, à partir du texte, ainsi que l'emprise de la TEI sur le résultat produit. Enfin, je plaiderai pour une pensée fondamentale de l'écran et de ses possibilités dans la pratique des humanités numériques.

INTERFACE = DONNEES

Depuis plusieurs années maintenant, les humanités numériques connaissent un « tournant *design* », selon l'expression de Stéphane Vial¹¹. Le champ prend enfin au sérieux les questions d'affichage, les laboratoires spécialisés dans le design sont plus visibles que jamais et les colloques sur le sujet se multiplient. Toutefois, à observer les appels à communications, les discussions et les politiques gouvernementales, le mouvement est bien loin d'être accompli : pour nombre de praticiens, l'affichage demeure une couche supplémentaire et facultative, une dimension cosmétique, ornementale, par opposition aux vraies humanités numériques. Les seuls affichages unanimement compris et reconnus comme humanités numériques, ce sont les graphiques et autres statistiques textométriques, comme s'il s'agissait des seules visualisations capables de révolutionner le rapport du lecteur à ses objets, comme s'ils n'étaient pas simplement, à l'instar d'une édition numérique ou d'une base de données, une autre visualisation de données. C'est qu'avec les graphiques et les statistiques, le changement de paradigme se voit et se passe d'explications : la littérature entre dans le domaine des sciences dures, justifiant ainsi les financements pour la recherche. Pour beaucoup, détracteurs comme zéloteurs, l'expression « Humanités numériques » signifie ainsi comptages et statistiques. En France, comme aux États-Unis, le débat se concentre majoritairement autour du traitement automatique des textes, de l'intelligence artificielle, du *big data*, en accordant une importance démesurée à une question, somme toute marginale et assez mal posée, que l'on pourrait résumer ainsi : les ordinateurs sont-ils capables ou non de faire de l'analyse littéraire à notre place¹² ?

À côté de ce débat stérile¹³, il est des ressources numériques qui transforment très concrètement le quotidien du chercheur depuis plus de vingt ans : l'accès aux ressources en ligne, la visualisation d'ouvrages dans leur mise en page d'origine, l'usage de moteurs de recherche avancés pour interroger des fonds d'archives ou encore la disponibilité de bases de données spécialisées. L'effort théorique consacré à ces questions est inversement proportionnel à leur impact réel. On réfléchit assez peu à l'effet des moteurs de recherche sur le canon littéraire ; pourtant, lorsqu'une page de résultat mêle indifféremment des ouvrages « mineurs » et « majeurs », c'est tout un rapport au corpus qui s'en trouve modifié. De même, on interroge

¹¹ Stéphane Vial, « Le tournant design des humanités numériques », *Revue française des sciences de l'information et de la communication*, n° 8, 2016, consulté le 24 mars 2020, URL : <https://doi.org/10.4000/rfsic.1828>.

¹² Voir notamment dans les *Chronicles of Higher Education*, la série d'articles regroupée (« les articles regroupés sous » ?) sous l'expression « Digital Humanities War ».

¹³ L'idée de l'ordinateur qui pense est principalement héritée du célèbre article d'Alan Turing, « Computing Machinery and Intelligence », *Mind*, vol. LIX, n° 236, 1950, p. 433-460, qui se dit souvent surpris par l'ordinateur. La surprise est toutefois du côté de l'humain j'ai du mal à comprendre le sens du « toutefois » ici, étonné du résultat des concepts et instructions programmés. Jusqu'à nouvel ordre, aucun ordinateur ne réfléchit, y compris pour ce qui concerne le *deep learning*, dont certains résultats ne sont inexplicables que parce que le nombre de variables en jeu est trop important.



peu la capacité d'une édition en ligne à transformer la perception de l'ouvrage édité. Les projets tendent à choisir un objet à numériser (photographie, scan, information dans des bases de données), puis à discuter du format, du logiciel, du stockage, de l'interopérabilité, de la pérennité, des catégories, des mots-clés et du modèle de données. Et c'est seulement à la fin du projet que l'on réfléchit alors à une manière d'afficher ces données, ou à un usage pertinent de cette numérisation, en espérant soudain que des résultats apparaissent. S'il reste un peu de financement, on engage même un *designer* :

Les chercheurs travaillant sur un projet d'analyse de données peuvent passer 80% du temps du projet à acquérir et nettoyer les données et 20% seulement du temps à travailler sur leurs questions de recherche¹⁴.

Or, toute réflexion sur les questions de recherche recoupe nécessairement la réflexion sur l'interface, puisque tout est affichage et interprétation. L'affichage n'est jamais neutre. La manière de présenter les données ainsi que les outils pour les manipuler ont des conséquences profondes sur leur appréhension et leur utilisation. Comme l'écrit Johanna Drucker :

Les visualisations sont toujours des interprétations – les données n'ont pas de formes visuelles inhérentes qui prédétermineraient leur représentation graphique¹⁵.

Que l'on travaille en mode texte dans un terminal ou que l'on affiche une édition d'un texte à l'interface léchée, que l'on fasse des statistiques dans un logiciel de textométrie austère ou que l'on déploie un moteur de recherche avancé dans une base de données, aucune solution n'est plus neutre que l'autre : il s'agit à chaque fois de la transformation de 0 et de 1 en signaux visuels ou auditifs compréhensibles par l'être humain. Certaines de ces transformations réalisent plus d'opérations que d'autres, certaines interfaces en appellent d'autres (la TEI est déjà un format compréhensible par l'être humain, mais il peut donner lieu ensuite à différents formats et affichages), certaines véhiculent plus de signaux que d'autres, certes. Reste que dans tous les cas, il ne s'agit plus de l'objet original, mais d'une transformation de ceux-ci. Toute transformation induit ainsi une série de perceptions, de présupposés, d'idées qui ne sont absolument pas neutres. Certaines fois, l'interface singe le livre, d'autres fois, elle le déconstruit entièrement, comme le fait par exemple un graphe à partir de centaines d'ouvrages. Dans tous les cas, le texte numérique est un texte autre, et son affichage et sa manipulation informatique en transforment la perception.

Cette transformation nous offre alors l'occasion de penser l'apport des humanités numériques à la littérature. Elle souligne que le rapport que nous entretenons à nos objets est déterminé par 1) la manière dont on y accède, 2) ce que l'on voit et 3) la manière dont on peut manipuler/explore/travailler ce que l'on voit. Malgré la pléthore d'interfaces imaginées, beaucoup (probablement les plus pertinentes) sont encore à inventer. Les graphiques en sont une : voir un livre ou une bibliothèque en courbes, en barres ou en segments constitue moins un résultat en soi qu'une nouvelle perception du livre ou du corpus, qui invite à formuler de nouvelles questions, à repenser nos approches, à confirmer ou infirmer une intuition. L'édition numérique en est une autre. Même incomplète par rapport à ses objectifs initiaux, la mise à disposition d'une large sélection du *Mercurus galant* sur l'OBVIL transforme radicalement la recherche en littérature française du XVII^e siècle¹⁶. L'un des ouvrages les plus importants du

¹⁴ « Studies show that researchers working on data analysis projects can spend up to 80% of project time pre-processing data and only 20% addressing their research question(s). » (Franzini, art. cit.).

¹⁵ « Visualizations are always interpretations – data does not have an inherent visual form that merely gives rise to a graphic expression ». Johanna Drucker, *Graphesis. Visual Forms of Knowledge Production*, Cambridge, Harvard University Press, 2014, p. 7.

¹⁶ Anne Piéjus et Nathalie Berton-Blivet (dir.), *Mercurus galant*, <https://obvil.sorbonne-universite.fr/corpus/mercurus>



siècle de Louis XIV devient soudain aisément accessible aux chercheurs et chercheuses, qui plus est en recherche plein texte, ce qui transforme fondamentalement son accessibilité et sa consultation. Mais il s'agit également d'un choix radical, fondé sur un présupposé textualiste : on perd ici la matérialité de l'ouvrage et la sélection d'extraits opérée omet certaines liaisons entre les articles¹⁷. Le *Mercure galant* devient ainsi un texte (plutôt qu'un périodique illustré à la typographie élégante) ; ou une source d'information (plutôt qu'un vecteur de propagande mêlant nouvelles et poésie, liés sous forme de lettre) ; en tout cas, il devient quelque chose d'autre.

En se substituant au papier blanc moderne et standardisé des éditions critiques, les versions photographiques des ouvrages anciens transforment radicalement la perception du texte et de l'objet. Il faut rappeler en effet ce que représente, pour les spécialistes des siècles anciens, le fait de pouvoir travailler en tout temps avec des images de l'ouvrage et d'accéder à des livres autrefois si rares et difficiles d'accès. Certes, la reproduction en ligne crée de nouvelles illusions d'optique : en standardisant le format des pages, *Google Books*, *Gallica* et *OpenLibrary* invisibilisent par exemple le format des ouvrages. Reste que la possibilité de consulter en un clic un livre rare, ou plusieurs dizaines d'entre eux, de les lire dans leur typographie originale, tout cela est une évolution radicale. Dans le cas d'éditions critiques numériques, l'interface suscite parfois de nouvelles pratiques de lecture, plus proches de celles d'origine, notamment tabulaires¹⁸. En outre, le texte de l'ouvrage retrouve une place centrale, escamote l'intimidation herméneutique des préfaces au profit d'appareils critiques éclatés et resitue l'ouvrage dans la production de l'époque en donnant accès aux ouvrages numérisés.

Dès lors, si la numérisation est de toute façon une transformation, et si l'affichage ne peut être neutre, il devient difficile de penser les données sans penser leur interface. Cette dernière approche soulève en effet plusieurs problèmes. Le premier, c'est de sous-estimer les conséquences de l'affichage sur l'interprétation. Les approches textométriques, en graphiques, ou les nuages de mots encouragent, par leurs présupposés même, une approche textualiste de la littérature¹⁹, aux dépens d'autres questions. Ces options ont contaminé la réflexion sur les politiques étatiques en matière de conservation. Pour *Huma-Num* en France, comme pour *DaSCH* en Suisse, tout l'effort de sauvegarde et de pérennisation porte sur les données, rarement sur leur visualisation, si ce n'est au travers d'une solution standardisée, inadaptée à de nombreux cas de figure. On considère en somme que les données ont du sens indépendamment de leur interface. L'édition en ligne des *Pierres d'Ogham* suffit à défaire ce mythe. Outre la transcription du texte, la photographie de la pierre et sa modélisation 3D sont essentielles pour la compréhension du contenu.

galant/. Sur ce projet, voir, dans ce volume, l'article d'Anne Piéjus, ###. Pour une présentation du Labex OBVIL, voir <https://obvil.sorbonne-universite.fr/obvil/presentation>.

¹⁷ Sur la matérialité, voir Christophe Schuwey, *Un entrepreneur des lettres au XVII^e siècle. Donneau de Visé, de Molière au « Mercure galant »*, Paris, Classiques Garnier, 2020, p. 354-371, ainsi que Barabara Selmeçi, « Donneau de Visé, amateur d'estampes et visionnaire. Le *Mercure galant* en 1686 », *Les Dossiers du GRIHL*, n° 2, 2017, consulté le 24 mars 2020, URL : <http://journals.openedition.org/dossiersgrihl/6781>.

¹⁸ Voir Christophe Schuwey, *Interfaces...*, *op. cit.*, p. 37-44.

¹⁹ Voir notamment l'introduction de Michel Bernard et Baptiste Bohet dans l'excellent *Littérométrie*, Paris, Presses de la Sorbonne nouvelle, 2017.

'the Priests' Hole'), a burial ground and the remains of a sub-circular enclosure. There are several cross-inscribed boulders and small slabs standing in the burial ground and in front of the oratory is a large ogham-inscribed pillar' (Okasha and Forsyth 2001, 138). The ogham stone is said to mark the grave of St Manachán (Bennett and Uí Shíthigh 1995, 13)

Monument

1.68m x 0.46m x 0.28m (Converted from Macalister 1945, 163).

'A cross with equal arms, expanding regularly, is cut on the face, on the dexter side of the inscription, while another cross, with expanding terminals ... is cut on the surface of the opposite face... A grafitto is scratched in a line parallel with the ogham in crudely made half-unicals. It reads FECT QUENILOC followed by a word too defaced to read' (Macalister 1945, 163-4). Note that the 3d scan of this stone showed no trace of this grafitto. Here, as elsewhere, Macalister sees the crosses on this stone as a later attempt at Christianisation but see McManus (1991, 54-61) for a valid counter argument to this theory.

Text

Up. 'The top has been broken and battered away... So far as it goes the inscription is perfectly clear, though some of the vowels are damaged' (Macalister 1945, 163). This reading was confirmed on inspection of the 3d data. Regarding the C of MUC[OI], three of the four scores are visible on the 3d model.

Transliteration

QENILOCI MAQI MAQI-AINIA MUC[OI]

Translation

'of Cellach son of Mac-Áine? descendant of?'

Commentary

CIIC 170. Ballymoreagh (AN BAILE RIABHACH), Co. Kerry

Download Epidoc | 3D PDF | 3D OBJ

Inscription Description 3D View Video Map Epidoc

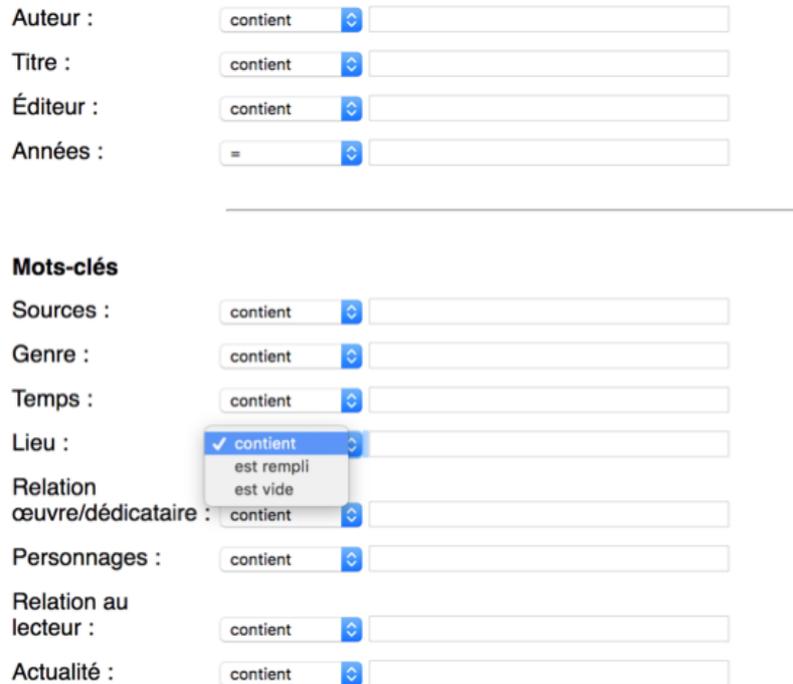
Temple Geal, Ogham Stone by The Discovery Programme

Figure 4. Notice d'une pierre tirée de la collection « Temple Geal » et visualisation 3D © Ogham Project.

En l'état actuel, aucune des plateformes mentionnées n'est équipée pour entretenir un tel projet à long terme.

En deuxième lieu, ne pas accorder suffisamment d'importance à l'affichage amène à développer des projets difficiles à utiliser. Imaginons ainsi un projet, repris d'une histoire vraie, autour de la poésie du roman français qui rassemblerait des préfaces de plus de cinq cents nouvelles et romans européens. L'objectif serait à la fois d'éditer chacune des préfaces et de les classer selon les sujets qu'elles abordent. L'équipe de recherche passe ainsi plusieurs années à élaborer une liste de catégories critiques caractérisant ces différentes thématiques : sources, unités, personnages, lieux, etc. Ainsi, si la préface du roman X traite de la question des unités (de temps, de lieu...) et qu'elle s'y oppose, on spécifie que cette préface 1) parle de la question des unités et 2) qu'elle s'y oppose. Si cette même préface indique les sources du roman et que ces sources sont, par exemple, italiennes, on enregistre dans la base de données que 1) la préface traite des sources et que 2) celles-ci sont italiennes. Le projet rassemble ainsi des données d'une qualité remarquable qui permettent de retrouver immédiatement toutes les préfaces qui traitent de sources italiennes. Ce qui aurait demandé autrefois plusieurs années de

recherche devient ainsi l'affaire de quelques clics. Imaginons maintenant qu'une interface mal conçue rende ces interactions incompréhensibles :



The image shows a search interface with several sections. The first section has four rows: 'Auteur :', 'Titre :', 'Éditeur :', and 'Années :'. Each row has a dropdown menu with 'contient' selected and an empty text input field. The second section is titled 'Mots-clés' and has eight rows: 'Sources :', 'Genre :', 'Temps :', 'Lieu :', 'Relation œuvre/dédicataire :', 'Personnages :', 'Relation au lecteur :', and 'Actualité :'. Each row has a dropdown menu and an empty text input field. The 'Lieu :' dropdown menu is open, showing three options: 'contient' (checked), 'est rempli', and 'est vide'.

Figure 5. La première interface de recherche (concept).

Que signifient les trois options « contient », « est rempli » et « est vide » ? Elles font référence à la structure expliquée ci-dessus. Malheureusement, il est impossible de comprendre leur sens sans avoir en tête le fonctionnement de la base de données – ce que l'on ne peut décentement pas demander à l'utilisateur. « Contient », « est rempli », « est vide » font référence, en fait, à la structure de la base :

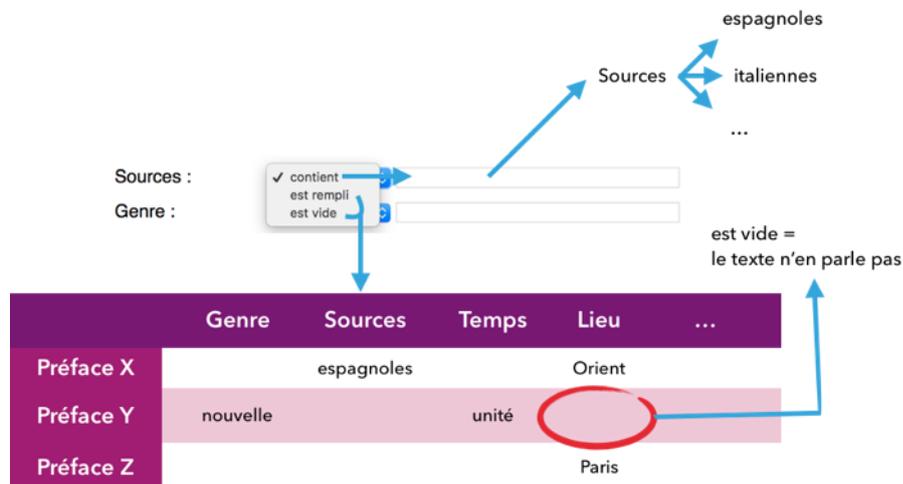
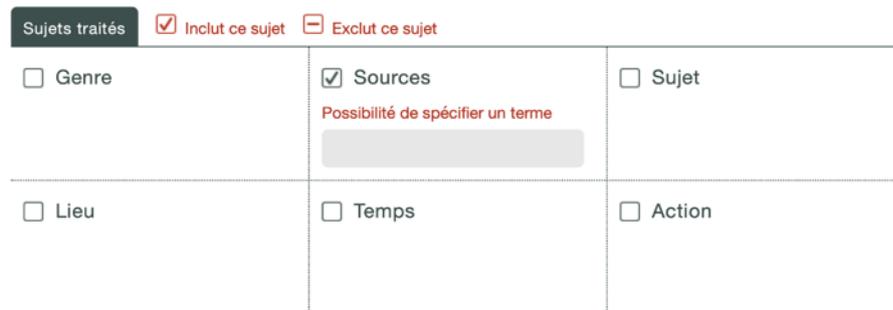


Figure 6. Schéma expliquant le rapport entre l'interface et la base de données.

À chaque préface est associée une série de champs, un par sujet traité (p. ex. « Sources »), et chacun de ces champs peut contenir des précisions sur ce mot clé (p. ex.

« espagnoles » pour « Sources »). On peut donc interroger les préfaces de trois manières : le sujet est-il rempli avec une quelconque information (= la préface traite-t-elle des sources ?), est-il rempli avec une information particulière (= la préface parle-t-elle des sources espagnoles ?), ou au contraire le champ est-il vide (= la préface ne traite pas des sources) ? Ce manque de clarté rend les données du projet peu exploitables. Le problème n'est pas d'ordre esthétique. Il tient au fait que l'interface ne traduit pas de façon compréhensible les possibilités de la base, les intitulés des menus déroulants étant cryptiques. Après des années de travail, le projet a omis une question de fond : comment rendre ce travail utilisable par la communauté scientifique ?

Sans rien modifier aux données existantes, imaginons maintenant une refonte de l'interface qui aurait rendu les données accessibles et manipulables. En qualifiant les différents champs de « mots clés », la précédente interface décrivait la nature de l'information et non son utilité. Puisque ces mots clés déterminaient si une préface traitait ou non d'un sujet, la nouvelle version adopte d'autres éléments d'interaction dont le sens est plus évident – des cases à cocher et à décocher – et une autre terminologie – « sujets traités » plutôt que « mots clés » :



Sujets traités		
<input checked="" type="checkbox"/> Inclut ce sujet <input type="checkbox"/> Exclut ce sujet		
<input type="checkbox"/> Genre	<input checked="" type="checkbox"/> Sources Possibilité de spécifier un terme <input type="text"/>	<input type="checkbox"/> Sujet
<input type="checkbox"/> Lieu	<input type="checkbox"/> Temps	<input type="checkbox"/> Action

Figure 7. La nouvelle interface de recherche (concept).

Cette nouvelle interface indique immédiatement ce qu'il est possible de réaliser à partir du moteur de recherche. La case à cocher indique le sens de l'interaction bien mieux qu'un menu déroulant : l'on souhaite ou non intégrer un paramètre donné, on coche et l'on décoche les sujets pertinents. C'est seulement si un sujet est intégré que l'on peut alors spécifier un terme en particulier : 1) recherche-t-on des préfaces traitant des sources ? 2) un type de sources en particulier ? La zone de texte fait apparaître une liste de suggestions qui affiche les différents mots clés disponibles. Ces quelques améliorations transforment la précédente liste de champs de recherche en une interface plus claire qui, en outre, invite à la découverte grâce aux suggestions. Un peu comme les services de musique en ligne spécialisés dans la musique classique : l'on voit comment, à partir des mêmes données, mais d'une interface à l'autre, la valeur scientifique d'un même ensemble de données se trouve considérablement modifiée.

CE QUE LA TEI FAIT AU TEXTE

Travailler à partir de l'interface, ou travailler à partir des données a des conséquences fondamentales sur le résultat. L'analyse d'une expérience pédagogique récente invitera à réfléchir sur la portée de l'affichage des données dans l'édition critique des textes littéraires au format XML-TEI. Lors d'un séminaire de master entre Yale et l'Université de Washington à Seattle (UW), Geoffrey Turnovsky et moi-même avons proposé à des étudiants de fin de licence et de master d'éditer en ligne des scènes de la *Comédie sans titre* de Boursault, créée en 1683 à la Comédie-Française. Il ne s'agissait pas du tout de refaire l'excellente édition de Marie-

Ange Croft et Françoise Gevrey²⁰, mais plutôt de réfléchir aux potentialités des éditions numériques. La pièce s'inscrivait parfaitement dans les thématiques du cours, et la division en scènes facilitait le travail en petites équipes. Les étudiants de UW et de Yale avaient reçu au préalable un enseignement sur la comédie, sur son contexte et sur ses ressorts comiques. Ils ont alors choisi une scène de la pièce qu'ils souhaitaient éditer et devaient 1) apprendre des rudiments de TEI, 2) transcrire le texte et l'éditer en TEI, 3) imaginer une interface. L'ordre des points 2 et 3 était libre. Les captures d'écran ci-dessous révèlent deux résultats obtenus. Le premier est une remarquable édition critique du texte avec affichage des variantes, l'autre traduit la vitalité du texte de théâtre en une interface inspirée des messageries instantanées. Dans cette dernière, les messages apparaissent au fil du temps, simulant la rédaction de la prochaine réplique par le correspondant :

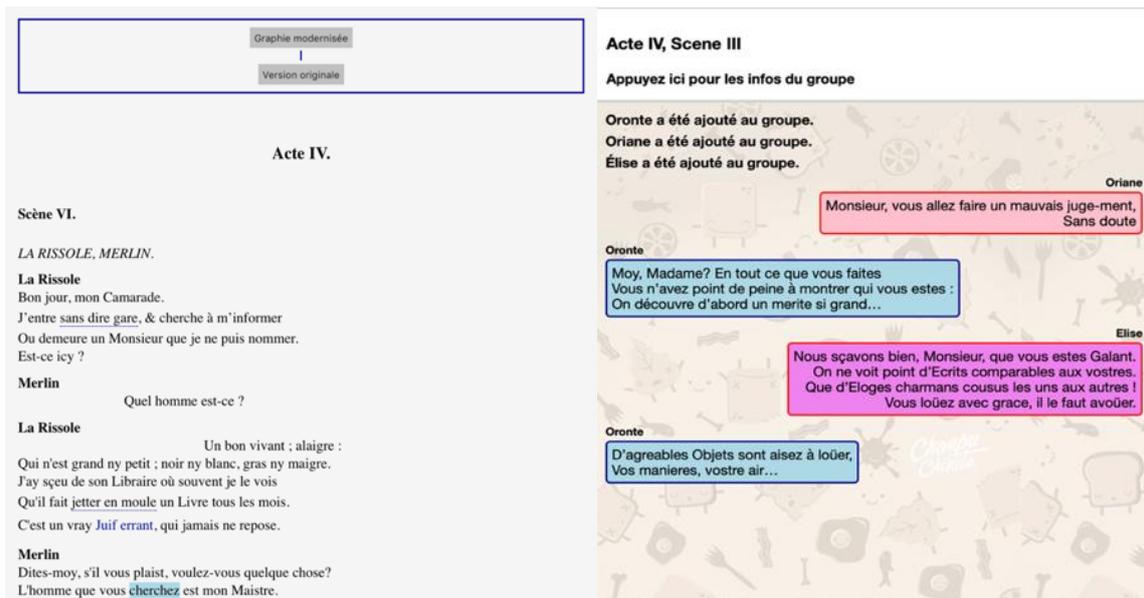


Figure 8. À gauche : interface conçue par Ashley Beeman, Doyle Calhoun et Julee Tanner.
À droite : interface conçue par Abigail Fields, Mara Minion et Nolie Ramsey.

La différence de résultats s'explique par la manière dont les groupes ont abordé le travail : le premier a commencé par éditer le texte en TEI, le second par imaginer une interface. Le semestre suivant, nous avons reconduit le séminaire, avec un nouveau groupe d'étudiants, autour des *Caractères* de La Bruyère. Cette fois-ci, les étudiants avaient pour consigne de commencer par mettre au point une interface. C'est seulement ensuite qu'ils apprenaient les bases de la TEI et qu'ils encodaient leurs textes. Dans les deux projets présentés ci-dessous, la première proposition joue sur la dimension discontinue des *Caractères*, transformant les remarques en cartes à jouer que l'on amasse en collection et que l'on peut rebrasser à volonté. Dans la seconde, on aligne le texte sur l'une de ses traductions anglaises, et l'on met en valeur les différences significatives, dans la lignée d'applications telles que *Juxta* ou *Variance*²¹.

²⁰ *Écrire l'actualité, Edme Boursault spectateur de la cour et de la ville* [édition critique de *La Comédie sans titre*], Reims, Université de Champagne-Ardenne, 2017.

²¹ Logiciels de collation de versions. *Juxta* : <https://www.juxtaoftware.org>, Rudolph Mahrer et Joël Zuffrey (dir.), *Variance*, <http://variance.ch>.

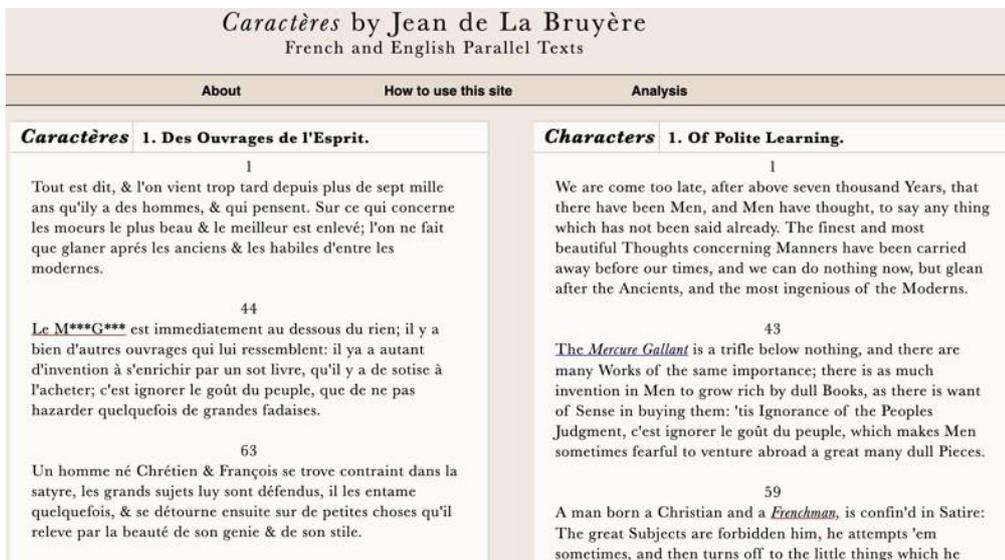
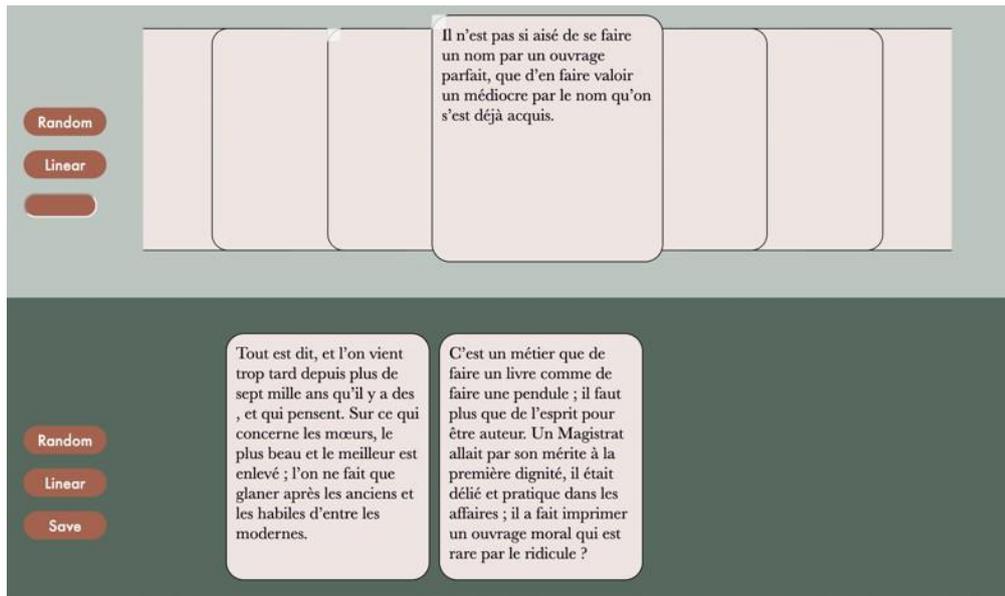


Figure 9. En haut : interface conçue par Nicola Angeli et Kallie Strode.
En bas : interface conçue par Jeanne Sauvage et Julee Tanner.

Partir de l'interface ne signifie donc pas nécessairement que l'on s'éloigne du texte, mais que l'on réfléchit à ce que l'écran peut faire pour celui-ci. Ne doutons pas de l'intérêt d'une interface imitant celle d'une messagerie, ou d'une logique de « cartes à jouer » pour les *Caractères* : de telles propositions stimulent la créativité et créent un nouveau rapport au texte. La première épouse l'essence de la comédie, traduisant bien mieux l'effervescence comique de la scène que ne pourrait le faire une édition papier. La seconde met à mal l'hypothèse d'une construction rigoureuse des *Caractères* : en rebrassant les cartes, on se rend compte rapidement que d'autres ordres fonctionnent tout aussi bien.

Ces expériences révèlent alors les biais de la TEI, qui oriente chercheurs et chercheuses dans une certaine direction²². Le cas des notes illustre parfaitement le problème. En TEI, l'élément <note> hérite directement du fonctionnement de la note de bas de page ou de fin

²² Pour élargir la discussion sur ces questions, voir Anthony Masure, *Design et humanités numériques*, Paris, B42, 2017.



dans un livre. Pour rappel, l'élément <note> se place à l'endroit où l'appel de note devrait apparaître dans le texte, et contient le texte de la note. Dans l'exemple suivant, tiré des *Nouvelles Nouvelles* de Donneau de Visé (1663), la note « Cf. *Les Précieuses ridicules*, sc. I » porte sur « ensemble » :

Il sait, à l'entendre dire, les desseins les plus cachés des tous les princes de l'Europe et de tous leurs ministres. L'autre fait le bel esprit, le galant et le savant tout ensemble<note>Cf. *Les Précieuses ridicules*, sc. I</note> : il s'instruit particulièrement des nouvelles du Parnasse, il sait tous les vers qui se font, tous les livres qui s'impriment, il en dit les beautés et les défauts, même avant qu'ils soient sous la presse et, lorsque l'on le veut écouter, il découvre, si l'on le veut croire, ce que tous les auteurs doivent mettre au jour d'ici à dix ans.

Le système est logique par rapport au modèle du livre. Quid, toutefois, de cette logique pour l'écran ? On pourrait vouloir indiquer, non pas seulement le lieu de l'appel de note, mais préciser sur quel passage porte le commentaire. C'est ce que nous avons fait avec Claude Bourqui dans l'édition critique des *Nouvelles Nouvelles* :

The screenshot shows a digital edition of a text. At the top, there are navigation tabs: 'LE TEXTE', 'PRÉSENTATION', 'DOCUMENTATION', and 'ÉDITION ET PROJET'. On the right, there are logos for 'Yale', 'IN-NE', and 'UNI FR'. The main text area contains a paragraph with a red callout box. The callout box contains the text: 'Cf. *Les Précieuses ridicules*, sc. I : « J'ai un certain valet nommé Mascarille, qui passe au sentiment de beaucoup de gens pour une manière de bel esprit [...] Il se pique ordinairement de galanterie ».' Below the callout box, the text continues: 'tous les vers qui se font, tous les livres qui s'impriment, il en dit les beautés et les défauts, même avant qu'ils soient sous la presse et, lorsque l'on le veut écouter, il découvre, si l'on le veut croire, ce que tous les auteurs doivent mettre au jour d'ici à dix ans. Le troisième est universel, il s'enquiert de tout et raconte de toutes sortes de nouvelles. Il sait les affaires d'État, il sait celles du Parnasse, il sait, de plus, non seulement celles de son quartier, mais encore celles de toute la ville ; en un mot, il ne se passe rien dans le monde'. On the right side, there is a 'TABLE DES MATIÈRES' section with a list of contents: 'PAGE DE TITRE', 'TEXTES LIMINAIRES', 'PREMIÈRE PARTIE', 'DEUXIÈME PARTIE (301)' (with a sub-entry 'Les Nouvellistes [3-301]'), and 'TROISIÈME PARTIE'. The 'DEUXIÈME PARTIE' section is expanded to show sub-entries: 'Élegie [10-13]', 'Conversation des soupçons [20-79]', 'Madrigal pour Julie [86]', and 'Madrigal de la tristesse'.

Figure 10. C. Bourqui, C. Schuwey, *Nouvelles Nouvelles* [1663], 2014, www.nouvellesnouvelles.fr.

Encore fallait-il réfléchir au résultat souhaité avant de réaliser l'encodage. Car la TEI permet bien sûr d'obtenir ce résultat, à condition toutefois de complexifier un peu le modèle. Il suffit d'ajouter un élément parent de type <seg>, qui contient l'élément <note>, et qui identifie le passage sur lequel porte cette dernière. Ainsi, pour que la note porte sur l'ensemble de « L'autre fait le bel esprit, le galant et le savant tout ensemble », le balisage est le suivant :

Il sait, à l'entendre dire, les desseins les plus cachés des tous les princes de l'Europe et de tous leurs ministres. <seg>L'autre fait le bel esprit, le



galant et le savant tout ensemble<note>Cf. Les Précieuses ridicules, sc. I</note><seg> : il s'instruit particulièrement des nouvelles du Parnasse, il sait tous les vers qui se font, tous les livres qui s'impriment, il en dit les beautés et les défauts, même avant qu'ils soient sous la presse et, lorsque l'on le veut écouter, il découvre, si l'on le veut croire, ce que tous les auteurs doivent mettre au jour d'ici à dix ans.

On voit ici à la fois l'adaptabilité de la TEI, qui se plie aux besoins, mais on remarque aussi le contrôle évident qu'elle exerce sur l'édition numérique. Pour descriptives qu'elles soient, la manière dont les balises s'organisent invite d'emblée à caractériser le texte d'une certaine manière, à insister sur certains éléments plutôt que d'autres et à reproduire des logiques du livre imprimé. Il faut donc veiller à ne pas laisser la description des données décider de l'interface de destination.

Car la TEI fascine. Le nombre d'éléments qu'elle offre donne envie de tout coder, de tout décrire et de transformer l'outil (l'encodage) en résultat. La *Folger Library* de Washington propose ainsi les œuvres de Shakespeare avec un encodage maximal, dans lequel chaque mot et chaque espace possède un identifiant unique²³. Cela permet de multiplier les types d'enquêtes. Cet encodage a cependant été largement réalisé automatiquement, et se concentre sur des aspects objectifs du texte : mots, espacements sur la page, etc. Le risque existe d'une pratique solipsiste de la TEI, très satisfaisante par sa rigueur philologique. Le meilleur antidote c'est, une fois encore, de réfléchir très concrètement à ce que l'on veut en faire. Tant que cette question n'a pas été réglée, l'intérêt des balises mises en place est proprement « virtuel »²⁴. Corollairement, l'encodage sans réflexion claire sur la finalité conduit potentiellement à des impasses. Dans le récent projet de Marine Roussillon sur les fêtes de cour, la question de l'encodage s'est posée de manière aiguë²⁵. Quels éléments utiliser pour décrire Louis XIV, lorsqu'il est en roi, ou lorsqu'il joue le personnage de Roland ? Quelles balises utiliser lorsque le texte ne dit que Roland, mais le lecteur sait qu'il s'agit de Louis XIV ? Comment encoder le fait qu'il y a un nom propre, que ce nom propre est un nom de personnage de fiction, mais que ce personnage de fiction, dans cette occurrence, est joué par Louis XIV ? L'attribut « role » ? Mais comment différencier le rôle de théâtre, et le rôle dans l'organisation de la fête (organisateur, artificier) ? Les descriptions du roi et les adresses à celui-ci, comme dans une épître dédicatoire ? Est-ce même du domaine de la TEI ? Les possibilités sont infinies, et les questions parurent rapidement insolubles. La solution fut de revenir aux scénarios d'utilisation, et de réfléchir à ces balises en fonction de ceux-ci, plutôt que de se perdre dans une discussion essentialisante.

LES POSSIBILITES DE L'ÉCRAN

L'exemple des notes révèle ainsi l'intérêt de penser pour et avec l'écran. Si le « tournant *design* » est en train de transformer les humanités numériques, la situation décrite en 2008 par le « Manifeste des humanités numériques » demeure aujourd'hui passablement la même :

[T]out comme l'imprimé reflétait initialement les pratiques de la culture du manuscrit du haut moyen âge, et tout comme le cinéma reflétait les techniques de théâtre, la première vague numérique a reproduit le monde des communications savantes qui avait été graduellement codifié au cours des cinq siècles passés. [...] Désormais, ce monde doit concevoir un avenir dans lequel les caractéristiques

²³ « Download Shakespeare's Plays, Sonnets, and Poems », en ligne, consulté le 4 novembre 2021, URL : <https://shakespeare.folger.edu/download/> rajouter point final?

²⁴ Stéphane Vial, *op. cit.*, « Introduction ».

²⁵ Marine Roussillon (dir.), *Merveilles de la cour*, <https://merveilles17.hypotheses.org>.



spécifiques des média numériques deviennent centrales et dans lequel l'imprimé se voit absorbé dans de nouveaux modes hybrides de communication²⁶.

L'édition numérique demeure prisonnière des logiques du livre. Combien d'exemples reproduisent le modèle de l'édition scientifique papier, avec un texte d'entrée, une préface et une présentation ? Combien de notes de bas de page dans *Cairn*, sur l'OBVIL ou même sur Wikipedia, alors que l'idée d'une note située au bas de la page n'a plus le moindre sens sur l'écran ? Et pourtant, ces ergonomies peu accueillantes sont un gage de sérieux qui rassure chercheurs et chercheuses sur la scientificité de ce qu'ils consultent. Ils constituent ainsi une source de crédibilité pour les éditions numériques en question. La situation enferme les humanités numériques dans une double contrainte : d'un côté, elles sont condamnées à imiter le livre – donc à faire moins bien – pour rassurer la communauté scientifique ; de l'autre, la communauté scientifique les somme de démontrer leur apport par rapport au livre, en ne manquant pas de s'insurger lorsqu'elles ne produisent pas de résultats révolutionnaires que les investissements réalisés devraient amener.

Revenons ainsi aux questions qui ouvraient cet article : à quoi ressembleraient les humanités numériques si nous réfléchissions d'abord à partir de ce que l'on veut afficher ou faire entendre, avant de réfléchir aux contraintes techniques et à adopter naturellement le modèle d'Omeka, tirant tout projet vers une logique d'exposition virtuelle²⁷ ? À quoi ressembleraient nos projets si l'on commençait par dessiner ce que l'on voudrait voir à l'écran, avant de s'interroger sur les jeux de balises nécessaires pour respecter notre objet ? À quoi ressembleraient les humanités numériques si l'on osait exploiter tout le potentiel de l'écran pour créer de nouvelles manières de lire, pour créer un autre rapport à nos textes, et surtout, si l'on reconnaissait la valeur évidente de ces démarches par rapport à l'imitation servile du livre ? Et clarifions ces questions : il ne s'agit pas de ne pas penser à la pérennité, ni surtout d'adopter un logiciel propriétaire ou de bricoler les choses. Mais il s'agit d'y penser ensuite, et d'adapter l'outil au résultat, et non l'inverse.

Nombre de réalisations relèvent de cette approche. Le site *Virtual Paul's Cross Project* propose entre autres choses un sermon de John Donne, non pas sous forme de texte, mais de vidéo et d'audio. On peut ainsi voir entre le prédicateur depuis le *Cross Yard* ou la *Sermon House*, à différentes distances et avec plus ou moins de fidèles assemblés. Le projet redonne ainsi sa dimension fondamentalement performative au sermon, mais soulève aussi la question de son accessibilité et de sa force de persuasion, en fonction de l'endroit d'où l'on écoute. Une édition en ligne peut également être l'occasion de traduire un texte de la première modernité. Traduire, non pas d'une langue à une autre, mais d'un contexte à l'autre, en modernisant pour rendre les intentions d'origine compréhensibles aux lecteurs et lectrices du XXI^e siècle. C'est ce que vise l'édition numérique de *L'Alcoran*, dont l'interface présente le texte à coups de portraits et de phylactères :

²⁶ Quentin Julien-Saavedra et Yves Citton, « Manifeste pour des humanités numériques 2.0 » *Multitudes* 2015, n° 59 (2), §8. Traduit de Johanna Drucker *et al.*, « The digital humanities manifesto 2.0 », 2008, consulté le 24 mars 2020, URL : www.humanitiesblast.com/manifesto/Manifesto_V2.pdf.

²⁷ Certes, Omeka est scriptable et adaptable à des logiques extrêmement diverses. Mais l'injonction d'Humanum à utiliser Omeka repose sur la maintenabilité des projets. Or, le développement de fonctionnalités propres à chaque projet annule le bénéfice, puisque le maintien de ces scripts et fonctionnalités particulières suppose *in fine* un travail spécifique sur le projet.

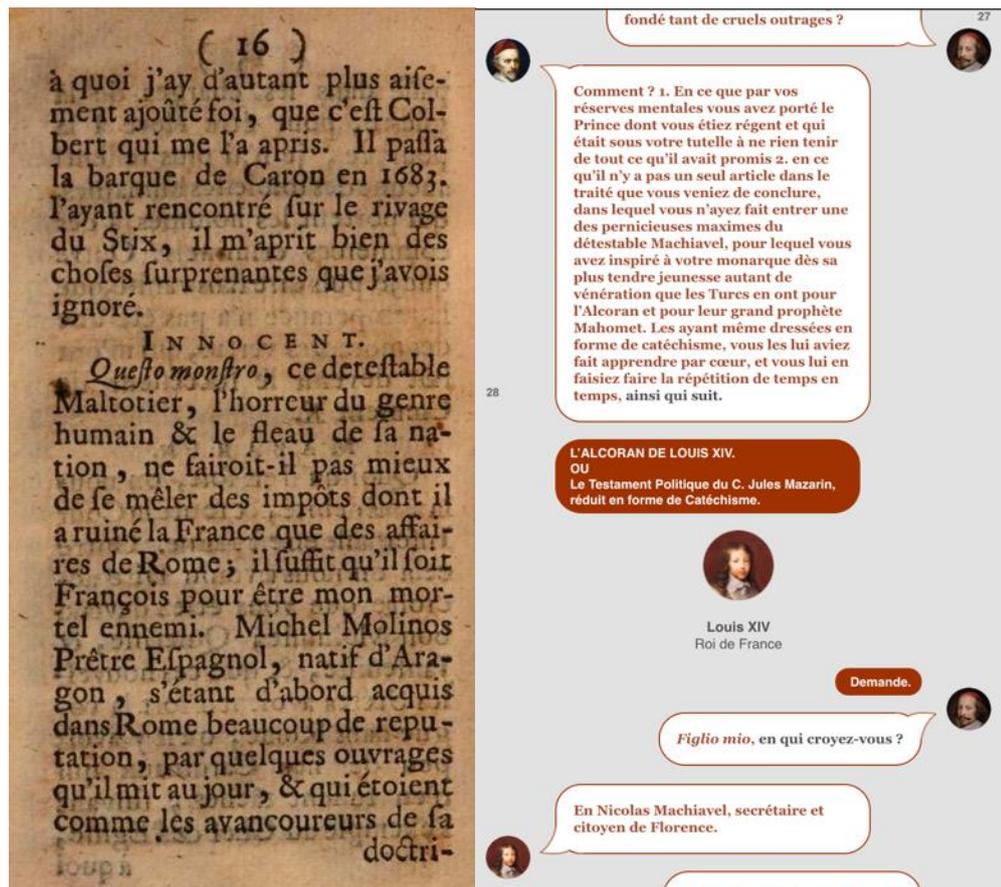


Figure 11. À gauche : *L'Alcoran de Louis XIV*, Roma, Antonio Stampatore [fausse adresse], 1695, consulté en ligne le 24 mai 2020 © Google Books. À droite, édition de K. LaPorta et C. Schuwey, www.dissidence17.fr/alcoran.html, consultée le 24 mai 2020.

L'objectif de cette édition n'est pas de reproduire le livre – pour cela, il y a le *livre*, disponible en bibliothèque et en ligne ! – mais de faire comprendre que ce qui se présente aujourd'hui comme un pamphlet peu engageant a priori était en réalité, pour le public du XVII^e siècle, un texte drôle, rempli de références à l'actualité et qui faisait appel à tout un imaginaire visuel et théâtral. La multiplication de ces portraits et de ces phylactères serait extrêmement coûteuse sur papier, elle ne coûte rien en ligne.

Ces exemples ne disent pas, et de loin, toute la créativité et les possibilités offertes par l'écran. Réfléchir à partir de l'interface plutôt qu'à partir des données, c'est réimaginer nos objets et nos pratiques en fonction du nouveau support, de l'écran et de ses possibilités, inventer de nouvelles manières de communiquer nos résultats et notre vision de la littérature. C'est à cette condition seulement que le numérique fournira une réponse à la question posée dans ce volume par l'équipe *Joyeuses Inventions*, qui dirige ce numéro du *Verger* : comment éditer des corpus qui mettent à mal l'idée même d'édition critique²⁸ ?

Revenons alors, pour conclure, à la musique, en empruntant un parallèle à Marilyn Deegan et Kathryn Sutherland dans leur excellent *Transferred Illusions*²⁹. Dans le monde occidental, la musique numérique est aujourd'hui la norme. Malgré la nostalgie que revêtent

²⁸ Voir dans ce numéro l'article de Magda Campanini, Marine Parra, Carole Primot, Anne Réach-Ngô, Côme Saignol, Miriam Speyer, Sylvie Vervent-Giraud, et Richard Walter : « Le projet *Joyeuses Inventions* : circuler dans le réseau des recueils collectifs de poésies récréatives du XVI^e siècle, une entrée par la porte du *Trésor* ».

²⁹ Marilyn Deegan, Kathryn Sutherland, *Transferred Illusions: Digital Technology and the Forms of Print*, London, Routledge, 2009, p. 100-106.



ces objets, plus personne ne se promène avec un baladeur cassette ou un *discman*. Comment s'explique ce succès ? C'est que la dématérialisation de la musique amène une plus-value évidente : plus besoin de changer de CD et d'emporter sa collection, on peut sauter d'une piste à l'autre aisément sans rencontrer de problèmes mécaniques (qui se souvient du *jogproof* des *discman* ?). La dématérialisation correspond (et détermine) une façon de consommer de la musique : sauter une plage, écouter une chanson sur un disque, une autre sur un autre album. Le livre numérique en revanche, n'a pas connu le même succès, et de loin. Les raisons sont nombreuses, mais l'une d'entre elles tient à l'incapacité d'imaginer des usages pertinents, parce que tous les acteurs se sont acharnés à reproduire le livre, jusqu'à la tour de page sur écran. Si la liseuse répond à certains besoins, elle crée également plus de contraintes qu'elle n'en résout : il faut veiller à sa batterie, on n'est jamais sûr de trouver le livre que l'on veut, elle est fragile (alors que le livre de poche est aisément manipulable), et il est impossible de numériser facilement sa bibliothèque comme on pouvait le faire avec ses disques. Ces questions se posent et doivent se poser pour les humanités numériques, et leur issue dépend avant tout de la place accordée aux scénarios d'utilisation et aux usages ; en somme, à l'interface. *Google Books*, *Gallica*, la mise en ligne d'un ouvrage (même lorsqu'il imite le livre) ou d'une base de données sur les livres imprimés comme le *Universal Short Title Catalogue*³⁰, tous ces projets s'imposent dans le quotidien des chercheurs et chercheuses parce que leur plus-value (notamment en termes d'accessibilité) est évidente. Mais l'on peut aller bien plus loin, faire beaucoup plus. À condition, bien sûr, d'accepter que le numérique transforme nos objets.

Souligner le rôle prépondérant de l'interface est ainsi une façon de repenser l'apport du numérique à la littérature et de prendre la mesure de ses possibilités. L'écran et le son nous permettent de retrouver des pratiques d'histoire de la lecture, du livre, en insistant sur des modes de lecture autre — par exemple, tabulaires, discontinus, etc. — de traduire des objets qui nous sont devenus étrangers, mais aussi, nous offrent l'occasion unique de revoir tous nos présupposés sur la littérature. Une réflexion centrée sur l'interface favorise en outre la créativité, ouvre des possibles et encourage la discussion : l'écran constitue le point de rencontre, entre les données et l'utilisateur, certes, mais aussi entre le littéraire et l'informaticien pour formuler des objectifs communs, s'entendre et se comprendre autour d'un travail en équipe. Les humanités numériques ne sont alors plus une manière de plier notre discipline aux normes des sciences dures, mais au contraire, d'en souligner la vitalité, la créativité, et de rappeler d'une fort belle manière les pouvoirs de la littérature.

³⁰ Andrew Pettegree *et al.* (dir.), *Universal Short Title Catalogue*, en ligne, consulté le 3 novembre 2021, URL : <https://www.ustc.ac.uk>.



TRAVAUX CITES

- BERNARD Michel, BOHET Baptiste, *Littérométrie*, Paris, Presses de la Sorbonne nouvelle, 2017.
- BOURSAULT Edmé, *La Comédie sans titre [1685]*, texte présenté et annoté par Marie-Ange Croft et Françoise Gevrey dans *Ecrire l'actualité, Edme Boursault spectateur de la cour et de la ville*, Reims, Université de Champagne-Ardenne, 2017.
- CAMP, Gregory, « Spotify », *Journal of the Society for American Music*, n° 9, 2015, p. 375-378.
- CHANG Ying-Han, TANG Muh-Chyun, « Serendipity with Music Streaming Services : The Mediating Role of User and Task Characteristics » dans G. Chowdhury, J. McLeod, V. Gillet et P. Willett (éds.), *Transforming Digital Worlds*, New York, Springer, 2018, p. 435-441.
- Chronicles of Higher Education*, série d'articles sous l'étiquette « Digital Humanities War », en ligne, consulté le 25 mars 2020. URL : <https://www.chronicle.com/>.
- DONNEAU DE VISÉ Jean, *Mercure galant*, édition d'Anne Piéjus et Nathalie Berton-Blivet, en ligne, consulté le 25 mai 2020. URL : <http://obvil.sorbonne-universite.site/corpus/mercure-galant/>.
- DRUCKER, Johanna, *Graphesis. Visual Forms of Knowledge Production*, Cambridge, Harvard University Press, 2014.
- FRANZINI Greta, KESTEMONT Mike, ROTARI Gabriela, JANDER Melina, OCHAB Jeremi K., FRANZINI Emily, BYSZUK Joanna, RYBICKI Jan, « Attributing Authorship in the Noisy Digitized Correspondence of Jacob and Wilhelm Grimm », *Frontiers in Digital Humanities*, vol. 5, 2018, en ligne, consulté le 24 mai 2020. URL : <https://www.frontiersin.org/article/10.3389/fdigh.2018.00004>.
- GASSER Martin, ARZT Andreas, GADERMAIER Thassilo, GRACHTEN Maarten, WIDMER Gerhard, « Classical music on the web – users interfaces and data representation », présentation à l'*International Society for Music Information Retrieval Conference*, Málaga, 2015, en ligne, consulté le 25 mars 2020. URL : <https://pdfs.semanticscholar.org/5bbc/17f103098dob0154b6aebo7a06c71bfd93a.pdf>.
- HAGEN Anja Nylund, « Music Streaming the Everyday Life » dans Nowak R., Whelan A. (éds.), *Networked Music Cultures. Pop Music, Culture and Identity*, London, Palgrave Macmillan, 2016, p. 227-245.
- JULIEN-SAAVEDRA Quentin, CITTON, Yves « Manifeste pour des humanités numériques 2.0 » *Multitudes*, n° 59 (2), 2015, §8. Traduit par Schnapp, Jeffrey, Presner Todd, Lunenfeld Peter et Drucker, Johanna, « The Digital Humanities Manifesto 2.0 », 2008, en ligne, consulté le 25 mars 2020. URL : https://www.humanitiesblast.com/manifesto/Manifesto_V2.pdf.
- L'Alcoran de Louis XIV [1695]*, texte présenté et annoté par Kathrina LaPorta et Christophe Schuwey, en ligne, consulté le 25 mars 2020. URL : <http://www.dissidence17.fr/alcoran.html>.
- MASURE Anthony, *Design et humanités numériques*, Paris, B42, 2017.
- MOLIERE, *Œuvres complètes*, texte présenté par Georges Forestier, Claude Bourqui *et al.*, Paris, Gallimard, Bibliothèque de la Pléiade, 2 vol., 2010.



- Ogham in 3D*, en ligne, consulté le 25 mai 2020. URL : <https://ogham.celt.dias.ie/>.
- PETTEGREE Andrew, KEMP Graeme, WEDUWEN Arthur der, Universal Short Title Catalogue, en ligne, consulté le 3 novembre 2021. URL : <https://www.ustc.ac.uk/>.
- PICHL Martin, ZANGERLE Eva, SPECHT Günther, « Understanding Playlist Creation on Music Streaming Platforms », actes de l'*International Symposium on Multimedia* de 2016 à San Jose, p. 475-480.
- SCHUWEY Christophe, *Interfaces : l'apport des humanités numériques à la littérature*, Neuchâtel, Alphil, 2019.
- SCHUWEY Christophe, *Un entrepreneur des lettres au XVII^e siècle. Donneau de Visé, de Molière au « Mercure galant »*, Paris, Classique Garnier, 2020.
- SEMOUR Corey, « Idagio—the Spotify for Classical Music—Has Changed My Life », *Vogue*, 15 mai 2019, consulté le 24 mai 2020. URL : <https://www.vogue.com/article/idagio-spotify-for-classical-music>.
- SISARIO Ben, « In Streaming Age, Classical Music Gets Lost in the Metadata », *New York Times*, 23 juin 2019, <https://www.nytimes.com/2019/06/23/business/media/stream-classical-music-spotify.html>, consulté le 24 mai 2020.
- TURING Alan, « Computing Machinery and Intelligence », *Mind*, vol. LIX, n° 236, 1950, p. 43-460.
- VIAL Stéphane, « Le tournant design des humanités numériques », *Revue française des sciences de l'information et de la communication*, n° 8, 2016, en ligne, consulté le 24 mars 2020. URL : <https://doi.org/10.4000/rfsic.1828>.
- VIAL Stéphane, *L'Être et l'écran. Comment le numérique change la perception*, Paris, PuF, 2013.