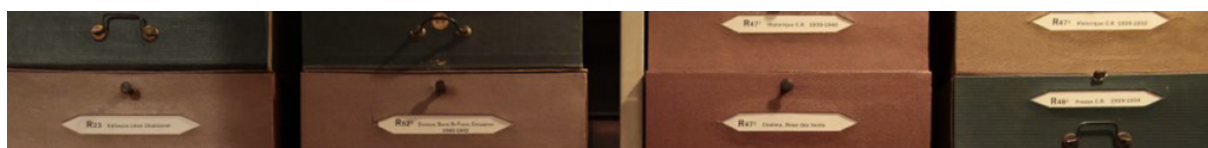


## Introduction

### Bases de données et histoire des spectacles : le temps d'un bilan ?

par Marine Roussillon et Christophe Schuwey



« Humanités numériques ». L'expression, omniprésente dans les appels à projets, les demandes de financement et les profils de poste, demeure problématique. La diversité des pratiques qu'elle recouvre (éditions numériques, bases de données, fouille textuelle...) et les transformations qu'elle induit (consultations d'ouvrages en ligne, recherche plein texte dans des corpus massifs, croisement des critères et des données...), tout cela disparaît bien souvent dans un flou de fantasmes et d'angoisses. Entre théories abstraites et enjeux financiers, désir de *big data* et crainte d'une mise au pas techniciste des humanités, promesses d'outils révolutionnaires et mépris du « gadget », les apports et défis réels de l'informatique appliquée à la littérature sont bien souvent relégués au second plan. Cette situation illustre la tension bien décrite par Pierre Mounier<sup>[1]</sup> entre d'un côté les valeurs des humanités et leur tradition critique, aujourd'hui marginalisées dans la société, et d'un autre, la science et la technologie, promues comme des idéologies au service du contrôle social et d'une industrialisation des pratiques culturelles. L'issue dépend cependant moins d'un débat à tendance essentialisante que de nos pratiques de recherche concrètes. L'informatique est avant tout un outil qui crée de nouvelles interfaces entre les chercheuses, les chercheurs et leurs objets. Les humanités numériques ne sont que ce qu'en font la communauté scientifique, les agences de recherche et les politiques étatiques.

#### Que faisons-nous quand nous faisons des humanités numériques ?

Face aux exigences des agences de recherches, les humanités numériques apparaissent comme un impératif, une façon de « moderniser » nos disciplines anciennes et souvent perçues comme marginales. Il est toutefois possible de se réapproprier nos pratiques en interrogeant les gestes accomplis au cours des deux dernières décennies, leur portée et leur pertinence, afin de remotiver le lien entre les humanités et le numérique. C'est ce à quoi ce dossier souhaite modestement contribuer, dans la continuité de nombreux travaux récents<sup>[2]</sup>. Ni célébration, ni condamnation ; ni théorisation excessive, ni collection de cas disparates, il a été conçu comme une tentative de rassembler des chercheurs et des chercheuses autour de questions et d'expériences concrètes : que faisons-nous quand nous faisons des humanités numériques ? Nos pratiques ont-elles quelque chose en commun ? Dans quelle mesure transforment-elles

l'histoire littéraire et culturelle, de l'accès aux objets à la manipulation des écrits, des questions posées à la disponibilité des ressources ? Débattre du numérique, c'est en réalité questionner le sens, les méthodes et les objectifs de notre recherche.

Nous sommes partis d'un objet parmi d'autres : les bases de données sur l'histoire des spectacles d'Ancien Régime. Ce choix s'inscrit à la fois dans un parcours de recherche – la préparation d'un projet de recherche sur les divertissements de cour, qui supposait un état des lieux des pratiques existantes – et dans une histoire plus large : la base *CÉSAR* (*Calendrier électronique des spectacles d'Ancien Régime*<sup>[31]</sup>) a été l'une des premières bases de données sur ce thème, et l'une des plus largement utilisées, à la fois dans le champ des études théâtrales et de l'histoire littéraire. Près de vingt ans après la création de *CÉSAR*, et dans une période de réflexion sur son avenir, ce dossier est aussi une manière de revenir indirectement sur cette expérience pionnière et sur son héritage.

Les responsables de projets de bases de données que nous avons sollicités ont accepté de partager leur expérience selon un questionnaire commun. Leurs contributions révèlent les apports fondamentaux du numérique à la recherche en histoire du théâtre, mais aussi la variété des défis rencontrés, la pluralité des solutions utilisées, et la grande diversité dans les manières d'écrire l'histoire des spectacles avec des bases de données. Elles mettent au jour les défis majeurs que les technologies numériques posent à nos recherches, à nos disciplines, et plus largement, à nos sociétés.

## **Il y a quelque chose de pourri au royaume des bases de données**

Pour appréhender ces défis, l'histoire de *CÉSAR* est éclairante. En 1988, alors que les premières ressources numériques en histoire des spectacles commençaient à voir le jour, David Trott entreprit la construction d'une base de données sur le théâtre du XVIII<sup>e</sup> siècle. Parallèlement, Barry Russel se lança dans la construction de bases de données, et inaugura en 1995 un site dédié au théâtre de la foire<sup>[4]</sup>. À la fin des années 1990, les deux hommes s'associèrent à Jeffrey Ravel pour mettre en commun leurs travaux. Le « Calendrier électronique des spectacles de l'Ancien Régime » apparut en ligne en 2002 et rencontra un succès immédiat. *CÉSAR* démontrait les bénéfices du numérique pour les humanités : bien conçu, utile, il centralisait des informations autrefois disparates, facilitant leur consultation et leur prise en compte par la recherche.

Cette adoption immédiate de la communauté indiquait cependant autre chose : le numérique était moins une rupture qu'une évolution pour l'histoire des spectacles. Comme l'histoire littéraire et l'histoire du livre<sup>[5]</sup>, la discipline recourait depuis longtemps à des répertoires papier proches du modèle de la base de données. Lors d'une conférence donnée à Dartmouth en 2018, Jeffrey Ravel a rappelé les nombreuses similarités entre le travail des premiers historiens du théâtre français au XVIII<sup>e</sup> siècle – les frères Parfaits – et les pratiques numériques. La nécessité de croiser des sources nombreuses et diverses pour documenter chaque représentation, ainsi que l'influence grandissante des sciences sociales sur la discipline, conduisait les chercheuses et les chercheurs à constituer leurs propres bases de données, sous forme de « fiches » ou de tableaux.

L'outil numérique ouvrait de nouvelles perspectives à ces pratiques bien établies. Il offrait tout d'abord un nouveau mode de publication et de nouveaux moyens d'exploration, plus adaptés à l'objet de recherche que ne l'était le livre imprimé. L'informatique permet en effet de multiplier les liens entre les documents, de les parcourir *via* plusieurs entrées et de les mettre à jour

régulièrement. La comparaison entre *CÉSAR* et le *Répertoire analytique* de Pierre Mélése sur le théâtre à Paris au XVII<sup>e</sup> siècle<sup>[6]</sup> est éclairante : tout essentiel que soit le *Répertoire*, il révèle à chaque page les limites du livre papier. Les extraits sont tronqués, les index, pour nombreux qu'ils sont, demeurent insuffisants, et le contenu est fixé dans le temps puisque, en l'absence de rééditions augmentées, les nouveaux documents découverts depuis la publication en 1934 n'ont pas pu être inclus.

La publication numérique encourageait également le travail collectif. Autour de *CÉSAR*, un groupe de chercheuses et de chercheurs s'est ainsi rapidement constitué, plaçant le site au centre d'un échange d'informations international d'une densité et d'une extension rares pour les études littéraires. En 2003, le site a même ouvert une interface publique permettant aux usagers de suggérer à leur tour des informations et d'enrichir la base.

Partage, collaboration, enrichissement des connaissances et diffusion des savoirs... Si *CÉSAR* concrétise les promesses du numérique, il en illustre aussi les risques. À l'heure où nous écrivons ces lignes, il suffit de se rendre sur la page d'accueil du site pour comprendre qu'il y a quelque chose de pourri au royaume des bases de données. Un message, daté de 2017, indique que « la base de données n'est plus maintenue mais reste parfaitement utilisable [...] ». Quelques fonctionnalités, notamment dans la partie « Livres », peuvent être indisponibles. L'accès aux comptes personnels et à l'espace collaboratif est également désactivé ». *CÉSAR* n'est pas mort, et la rénovation du site, supervisée par Marc Douguet, est en cours. Mais cette interruption longue est révélatrice de la précarité qui caractérise les outils numériques de la recherche. En effet, l'histoire de *CÉSAR* n'est pas simplement l'histoire idyllique d'une communauté de chercheuses et de chercheurs échangeant des connaissances à travers le monde. C'est aussi l'histoire d'une recherche ininterrompue de financements, de la première subvention du AHRB (273 000 £ accordés en 2001 pour trois ans) jusqu'au financement actuel de sa rénovation par le CNRS, l'Institut Universitaire de France et l>IDEX Université Grenoble Alpes. La survie du site et sa maintenance étant toujours menacées par la fin des subventions, les colloques *CÉSAR* ne cessent de faire état de la recherche de collaborations et de financements nouveaux auprès de diverses institutions. Car non seulement la création d'un tel outil coûte cher – bien plus cher que la publication d'un livre imprimé ou que l'organisation d'un colloque – mais en outre, contrairement à un livre imprimé, un site internet n'est jamais publié une fois pour toutes. Pour rester accessible, la base de données doit être régulièrement maintenue. Malgré son utilité évidente pour les études théâtrales, *CÉSAR* est tombée progressivement en décrépitude, au point de devenir difficilement utilisable pendant une longue période, qui doit heureusement s'achever prochainement avec la rénovation en cours.

Lors d'une session intitulée « Print and digital interface »<sup>[7]</sup>, la première question du public a porté sur la pérennité des réalisations. Dans un monde qui s'interroge enfin sur la viabilité de ses modèles économiques et écologiques, l'enjeu en est central. Il remet au premier plan le bilan carbone désastreux du numérique ainsi que ses coûts humains et politiques catastrophiques. Sans solution durable pour la production d'énergie, et tant que les conditions d'extraction des métaux précieux demeureront indécentes, les humanités numériques feront partie du problème. L'enjeu questionne également la contradiction entre le modèle de financement de la recherche par projets, à durée limitée, et les logiques numériques. La pérennité des outils et des publications numériques exige des équipes stables, alors même que la précarisation et l'instabilité du monde académique s'accroissent. En amont, cette précarisation décourage chercheuses et les chercheurs d'investir un temps conséquent dans la recherche et la numérisation des données. En aval, l'impossibilité de garantir la longévité des réalisations permet aux universités de continuer à survaloriser les réalisations papier aux dépens des

réalisation numériques, quand bien même celles-ci contiennent souvent infiniment plus d'informations, d'une qualité équivalente voire supérieure, et qu'elles peuvent être mises à jour en permanence. Comment garantir la pérennité des réalisations numériques quand les financements et les postes sont précaires ? Le constat invite moins à pointer du doigt des responsables introuvables qu'à susciter une remise en question parmi les comités de pilotage et les organismes de financements : les solutions envisagées jusqu'ici, y compris les grandes opérations gouvernementales, sont insuffisantes et inadaptées.

### **Les héritiers de CÉSAR : un état des lieux**

Malgré son histoire mouvementée, *CÉSAR* a fait de nombreuses émules. Il faut bien le reconnaître : depuis des années, la présence d'un volet « humanités numériques » dans les projets de recherche facilite l'obtention de financements. En conséquence, plusieurs chercheuses et chercheurs disposant d'un bagage technique souvent limité ont courageusement repensé leur démarche afin d'y intégrer un volet numérique. La situation a conduit à une démultiplication sauvage des réalisations. Elle offre ainsi de nouvelles ressources précieuses à la communauté scientifique, mais elle véhicule également son lot de réalisations inachevées, rapidement disparues ou difficiles à utiliser, et conduit à un manque d'interopérabilité, de coordination et de mutualisations entre des réalisations parfois très proches. Si les dernières années ont permis de mettre en place un certain nombre de garde-fous, notamment concernant la gestion des données et leur sauvegarde, ceux-ci sont insuffisants (voir ci-dessous) et la nécessité d'une mise en cohérence des réalisations et des pratiques demeure<sup>[8]</sup>.

L'idée de ce dossier doit beaucoup à ce constat : au moment de concevoir un nouveau projet de recherches sur les divertissements à la cour de Louis XIV, il est apparu nécessaire de faire le point, non seulement sur les ressources numérisées existantes<sup>[9]</sup>, mais aussi sur la manière dont d'autres avaient déjà extrait et organisé des données pour étudier les spectacles du XVII<sup>e</sup> siècle. Les échanges informels avec quelques collègues impliqués dans ce domaine nous ont convaincu de l'intérêt de rassembler des expériences significatives qui, nous l'espérons, pourront servir à d'autres. Nous avons donc sollicité les responsables de six projets portant sur une même période – les XVII<sup>e</sup> et XVIII<sup>e</sup> siècle – mais dont la diversité des objectifs et des sources, la disparité des modes de financement et les différences dans l'organisation de la recherche nous ont parues significatives.

Toutes les bases de données présentées ici s'appuient sur des corpus numérisés, à l'exception notable de la base *PerformArt*<sup>[10]</sup>, qui recense et décrit les archives de familles de la haute aristocratie romaine, dont il s'agit d'étudier les pratiques artistiques et le mécénat. Les documents pris en compte par cette base sont extrêmement variés, et ne sont pas tous des textes – on y trouve même des instruments de musique anciens. Son fonctionnement repose sur l'extraction d'informations à partir d'une description fine et seuls certains documents font l'objet d'une transcription intégrale, dans un souci de contextualisation. Le projet *Récital*<sup>[11]</sup> se fonde sur un corpus d'archives, les registres de la Comédie Italienne. Il en propose la transcription et l'édition collaborative en ligne avant d'en extraire les données. Les projets *Mercurie galant*<sup>[12]</sup> et *Théaville*<sup>[13]</sup> reposent quant à eux sur l'édition numérique d'un corpus imprimé (les livraisons d'un périodique du XVII<sup>e</sup> siècle, *Le Mercurie galant*, ou les parodies d'opéra du XVIII<sup>e</sup> siècle) pour en permettre l'exploration. Enfin, les deux derniers projets présentés constituent des anthologies numériques : une collection d'extraits de textes rendant compte de la réception des pièces de théâtre au XVII<sup>e</sup> siècle dans le cas de la base *Naissance de la critique dramatique*<sup>[14]</sup> ; un recueil de textes contemporains de l'œuvre de Molière et entretenant une relation d'intertextualité avec elle pour le site *Molière21*<sup>[15]</sup>. Le geste

caractéristique de ces derniers projets n'est pas tant la divulgation d'un corpus que le rapprochement de textes au sein d'un même ensemble. Ces bases de données articulent donc de manières variables trois gestes principaux : elles rapprochent des objets variés dans un même ensemble (c'est la construction du corpus) ; en extraient des informations (les données) ; et les rendent largement accessibles et interrogeables (c'est la publication<sup>[16]</sup>).

Pour éviter l'éparpillement des récits et permettre la confrontation et le partage des expériences, nous avons proposé aux contributrices et contributeurs un questionnaire, en nous inspirant de la démarche adoptée par la revue *Bien Symboliques/ Symbolic Goods* pour étudier les bases de données littéraires<sup>[17]</sup>. Ce questionnaire (consultable ici) était organisé autour de trois interrogations centrales. La première portait sur la construction des corpus et la production des données. De nombreux travaux ont déjà relevé que cette production n'est ni neutre, ni transparente, et que le numérique a tendance à essentialiser les données, en occultant à la fois leur matérialité et leur processus de production<sup>[18]</sup>. À ce titre, on rappellera toutefois que le numérique constitue moins une rupture qu'une mise en évidence de biais inhérents à la recherche. Le *Répertoire analytique* de Pierre Mélése, pour reprendre cet exemple, essentialisait déjà ses données. En rassemblant, sur des pages blanches à la typographie normalisée, des discours sur les spectacles provenant de supports divers, il faisait oublier les différences contextuelles de ces extraits et les raisons qui présidaient à leur existence. Trois opérations nous ont ainsi semblé devoir faire l'objet d'une vigilance méthodologique particulière : la constitution d'un corpus, sa transformation en données interrogeables automatiquement, et le passage que cette transformation suppose d'une matérialité historique (le plus souvent, celle du papier) à une matérialité nouvelle, le numérique. Il s'agissait de comprendre comment les différents projets avaient pris en charge ces opérations, mais aussi comment ils avaient pu expliciter leur démarche au sein de la base de données, ou avaient au contraire dû se résoudre à l'occulter.

Ce dernier aspect nous a conduit à une deuxième interrogation sur l'élaboration de l'interface et la relation qu'elle construit avec les utilisatrices et les utilisateurs, c'est-à-dire sur les modalités et les enjeux du geste de publication. L'interface conditionne la manière dont les données sont appréhendées et les opérations qui peuvent être réalisées : en mettant en relation les données et les utilisatrices et utilisateurs, elle modèle les pratiques de recherche et le rapport des chercheuses et des chercheurs à leur objet. Nous avons interrogé les contributrices et les contributeurs sur la manière dont leurs interfaces avaient été conçues (ou non) en amont du projet, mais aussi sur les éventuels écarts entre leurs objectifs en concevant leur base de données et la manière dont elle a finalement été utilisée, de manière à mettre en évidence les effets du *design* sur les pratiques de recherche.

Enfin, une troisième série de questions portait sur les financements, les structures et la temporalité de la recherche. Il s'agissait ici de mieux saisir la relation que chercheuses et chercheurs entretenaient avec les technologies mobilisées : qui discute des choix techniques ? Sont-ils subis ou collectivement construits, malgré les difficultés des échanges interdisciplinaires ? Quels effets ont-ils sur les pratiques de recherche, sur la composition des équipes, sur les relations entre leurs membres ? Dans quelle mesure le coût spécifique de ces technologies – sans commune mesure avec le coût usuel de la recherche en humanités – induit-il des modes de financement qui modifient nos pratiques de recherche ? En particulier, comment les différents projets concilient-ils l'importance des coûts et le recours au financement par projet qu'ils impliquent le plus souvent, avec la nécessité d'assurer la pérennité des bases de données ?

À ces questions, les contributions réunies dans ce volume n'apportent pas de réponses toutes faites, directement reproductibles. Les réflexions méthodologiques développées ici et les récits d'expériences sur lesquelles elles s'appuient n'ont pas pour but de constituer un guide de « bonnes pratiques », mais de nourrir la réflexion sur – et la compréhension de – ce que nous faisons, d'explicitier les enjeux d'une pratique qui se généralise et de contribuer à une réflexion collective sur ses enjeux, en lieu et place des bricolages particuliers. À l'issue de ce parcours nécessairement partiel et lacunaire, et sans anticiper sur les conclusions que chacune et chacun pourra en tirer, nous pouvons formuler deux constats et repérer deux défis majeurs.

### **Premier constat : un changement d'échelle**

Le premier constat relève de l'évidence : les technologies numériques permettent d'accroître significativement le nombre d'informations disponibles pour la recherche. Les conséquences sont particulièrement frappantes dans le cas du site *Molière21* : le recours à une base de données intertextuelle permet d'ouvrir l'œuvre de Molière sur l'ensemble de la production écrite contemporaine, transformant radicalement son appréhension. On se rend compte que des ouvrages jugés mineurs par l'histoire littéraire ont largement contribué aux comédies canoniques. Il apparaît clairement que Molière invente moins (au sens moderne du terme) qu'il ne reprend et met en spectacle les dernières publications, modes et sujets d'actualité ; qu'il n'est pas du tout révolutionnaire ou dissident, mais qu'il épouse au contraire les positions consensuelles de la cour, même lorsqu'il paraît la critiquer. Le projet *PerformArt* témoigne d'un même élargissement du terrain de l'enquête. Là où la recherche sur les pratiques artistiques des familles aristocratiques romaines s'était jusqu'à présent intéressée au patronage spécifique de tel pape ou de tel grand seigneur, la base de données permet de quitter l'approche fragmentaire pour produire une vision d'ensemble de l'histoire des spectacles au sein des familles aristocratiques à Rome entre 1644 et 1740. Les différents projets constituent ainsi des variations augmentées sur le thème de la bibliothèque : la collection numérique fonctionne comme une bibliothèque réelle, mais elle est plus accessible et organisée de manière plus complexe, multipliant les liens entre les documents, les regroupements possibles, les chemins de lecture. Il devient alors possible d'obtenir en quelques secondes une information qui aurait nécessité des semaines de recherches quelques années plus tôt.

Ce changement d'échelle défait ainsi une conception ancienne de l'histoire culturelle comme transmission d'un canon. Pour l'histoire des spectacles comme pour l'histoire littéraire<sup>[19]</sup>, la démarche quantitative introduite par les bases de données déconstruit l'idéologie des grands textes et des grands auteurs, les humanités numériques contribuant ainsi à réinsérer les spectacles dans un ensemble de pratiques culturelles plus large et donc représentatif. La base *Théaville* démontre ainsi la vitalité et la circulation d'airs familiers du public, repris, réécrits, chantés avec de nouvelles paroles. Le site *Molière21* met sur le même plan les pièces de Molière et des écrits contemporains longtemps invisibilisés par l'histoire littéraire, *Le Malade imaginaire* dialoguant ainsi avec un roman de Marie-Catherine Desjardins, ou avec un poème de Desmarets de Saint-Sorlin. Ironiquement, cette démarche égalisante contredit le fondement idéologique de la collection Pléiade, alors même que *Molière21* a été conçu à l'origine pour accompagner la nouvelle édition des *Œuvres complètes* en 2010<sup>[20]</sup>. La disponibilité du *Mercure galant* en mode plein texte et sa numérisation sur Gallica et Google Books ne rendent pas seulement sa consultation plus pratique et rapide : elles encouragent le recours à une ressource essentielle du règne de Louis XIV et redonnent une place de premier plan à la matérialité sophistiquée de l'objet<sup>[21]</sup>. La prise en compte de la vie matérielle, des publics, de la musique de scène, des techniques publicitaires ou encore de l'explosion médiatique de la première modernité brouille les frontières artificiellement fixées par les études littéraires. En insérant les

textes et les pièces de théâtre dans des ensembles plus vastes d'écrits, les bases de données contribuent à l'établissement d'une histoire qui envisage les lettres et les arts comme des pratiques sociales parmi d'autres, et s'interroge sur la construction de la littérature, du théâtre ou de l'art au sens large comme institutions<sup>[22]</sup>.

Cet élargissement du terrain de l'enquête a pour deuxième conséquence de favoriser la construction d'une vision d'ensemble du spectacle, englobant les différents arts et techniques qu'il mobilise, les modalités de sa production et de sa réception. Ainsi, la base *Théaville* met en regard les textes des parodies dramatiques d'opéra et les airs qui servent de support à ces parodies, sous la forme de partitions et de musique enregistrée. Seule une telle vision d'ensemble permet de repérer et d'interpréter les parodies. La base *PerformArt* cherche quant à elle à rendre compte de tous les aspects de la performance, depuis la production jusqu'à la réception. L'étendue des disciplines et des savoirs mobilisés est révélatrice : son équipe est constituée de spécialistes d'arts du spectacle, d'histoire, mais aussi d'histoire économique. La pratique des bases de données rencontre ainsi le projet d'une histoire des spectacles comme performances qui fonde les arts du spectacle comme discipline universitaire<sup>[23]</sup>. Elle peut alors apparaître comme un outil privilégié pour cette discipline encore jeune, à la recherche d'outils scientifiques capables de conforter son statut disciplinaire et de lui conférer une forme de légitimité<sup>[24]</sup>.

### **Deuxième constat : la collection comme forme de l'argumentation**

Comme pour toute bibliothèque, l'enjeu méthodologique essentiel des bases de données reste celui de la constitution du corpus. Dans le cas des registres de la Comédie Italienne (*Récital*) ou de la Comédie-Française, la base de données est construite à partir d'un fonds d'archives cohérent et préalablement constitué<sup>[25]</sup>. Mais la plupart du temps, la collection de documents réunie dans la base de données est le résultat d'une opération herméneutique que l'interface tend à occulter, naturalisant ainsi le corpus. La présentation du projet *Mercurie galant* témoigne de cette difficulté et de la prudence méthodologique qu'elle impose. Le projet paraît reposer sur un corpus pré-existant : la collection complète des livraisons du *Mercurie galant* de Donneau de Visé (1672-1710), dans son édition parisienne. Or cet objet n'a pas vraiment d'existence matérielle. Même les collections réputées complètes, comme celles de la BNF, sont en fait lacunaires, certaines partitions et estampes ayant été prélevées sur d'autres exemplaires conservés. Cet état des collections, révélateur de pratiques de lecture, de transmission et d'usages du *Mercurie galant*, est masqué par la reconstruction d'une collection « idéale » complète à partir d'exemplaires épars. De la même manière, le découpage des livraisons du *Mercurie* en articles ou la sélection des articles concernant la vie littéraire et artistique sont des opérations herméneutiques fortes que la présentation du corpus dans la base de données ne permet pas de saisir ni d'interroger. Il en va de même dans la base *Théaville*. En mettant l'accent sur les chansons et airs pour construire son corpus de « parodies d'opéra », elle rappelle l'importance culturelle et scientifique d'un élément souvent ignoré aussi bien de la critique que des metteurs en scène. Ce faisant, elle affirme l'existence d'un genre – la parodie d'opéra – dont elle revendique la valeur : le geste de publication, nullement neutre, est une manière de dire la dignité de l'objet. Comme toute bibliothèque, la base de données ordonne, hiérarchise, distribue de la valeur : en bref, elle argumente.

Cette fonction argumentative n'est donc pas nécessairement un problème méthodologique. Au contraire, elle compte parmi les raisons d'être de certaines bases de données. Le corpus de *Molière 21* et *Naissance de la critique dramatique* sont explicitement constitués pour soutenir une thèse. Dans ce cas, la base de données se présente comme une anthologie de preuves au

service de la démonstration. *Naissance de la critique dramatique* accueille le visiteur avec la liste des genres abritant de la critique théâtrale, tandis que *Molière21* démontre à chacune de ses pages que le théâtre de Molière traite avant tout de l'actualité, que ses sources sont à chercher dans la littérature du XVII<sup>e</sup> siècle, et que le canon ne reflète pas le paysage intellectuel et culturel dans lequel évoluait Molière. Dans les deux cas, le corpus est constitué en fonction d'une finalité démonstrative : les sources de Molière au service de l'interprétation des textes ; les extraits sur *Naissance de la critique dramatique* au service de la thèse de l'apparition d'une nouvelle forme de critique dramatique. Dans les deux cas également, la base de données est adossée à une publication papier (l'édition des *Œuvres complètes* de Molière et le numéro de la revue *Littératures classiques* intitulé *Naissance de la critique dramatique*<sup>[26]</sup>) qui en explicite l'argumentation. C'est aussi la démarche adoptée par la base de données *AGON* recensant les querelles de l'époque moderne en France et en Angleterre<sup>[27]</sup>. Dans cette dernière base, les différentes fiches présentant les querelles sont en outre signées et signalent ainsi clairement leur statut d'interprétation.

Ces derniers cas mettent au jour le fait que la base de données n'est pas seulement un outil pour la recherche. Elle est bien une forme d'écriture de l'histoire, au même titre qu'un ouvrage critique. C'est une proposition qui tire parti des ressources du numériques pour étayer son propos, c'est-à-dire à la fois « jouer cartes sur table », comme l'écrit Claude Bourqui, et revendiquer une forme de scientificité de la démarche, comme le fait Anne-Madeleine Goulet.

### **Premier défi : l'interface**

La contribution de Claude Bourqui révèle cependant combien il est difficile de faire lire une base de données comme une production critique. Tout se passe comme si le support numérique rendait les prises de positions proprement illisibles. Apparaît alors l'enjeu fondamental de l'interface. C'est elle en effet qui définit la manière dont les données sont perçues, et donc interprétées. Et pour cause : non seulement il n'y a aucune perception des données possibles sans qu'une interface traduise la réalité électro-numériques des données pour l'humain<sup>[28]</sup>, mais c'est en outre par le travail sur l'interface que les opérations de constitution du corpus peuvent être rendues visibles ou non. Longtemps négligées, les interfaces sont aujourd'hui une question centrale dans la réflexion sur les humanités numériques<sup>[29]</sup>. Dans l'optique de ce dossier, elles apparaissent comme une réponse forte aux problèmes méthodologiques soulevés par les bases de données, et comme un élément essentiel dans la nouvelle forme d'écriture de l'histoire qui s'y invente.

Une fois encore, il n'y a pas ici de rupture majeure : la discipline littéraire n'a pas attendu l'informatique pour essentialiser les textes et faire oublier leur origine. Les livrets de fêtes de cour ont été traités comme des documents relatant objectivement l'événement alors qu'il s'agit d'opérations de propagande complexes<sup>[30]</sup>. Dans son *Histoire de la folie à l'âge classique*, Michel Foucault a transformé des répliques de comédies en positions philosophiques, sans tenir compte ni du personnage qui les énonçait, ni de leur contexte d'énonciation<sup>[31]</sup>. Une édition critique ou un répertoire papier séparent nécessairement un texte de sa matérialité d'origine, incitant celles et ceux qui les consultent à ignorer le support et la forme du message original. Les humanités numériques héritent ainsi de ces biais plutôt qu'elles ne les créent. Elles disposent en revanche de nouveaux outils pour y répondre. L'interconnexion entre les documents, le renvoi à des photographies de l'ouvrage, les discours de contextualisation, la disposition du texte ou la multiplication des instruments de recherche sont autant de moyens de réinscrire les documents dans leur contexte, de rendre visible la construction des savoirs et le



travail d'interprétation, et de mettre à disposition des chercheuses et chercheurs les outils nécessaires à la critique des extraits présentés.

En affichant une liste de supports sur sa page d'accueil, *Naissance de la critique dramatique* n'illustre pas seulement la diversité de ses sources : elle invite également à prendre en compte le support et les intentions du texte original. Lors d'une recherche plein texte sur la même base, cette dimension passe toutefois au second plan. C'est seulement lorsque l'on consulte un extrait particulier que le support d'origine est rappelé aux lectrices et aux lecteurs. *Théaville*, quant à elle, donne une importance aux chansons au moins équivalente à celle des textes. Elle brise par ailleurs l'unité de l'œuvre en insistant sur la circulation des airs, leurs réemplois et leur popularité. Évoquons encore la pratique de nombreuses éditions critiques numériques, celle de renvoyer à la version photographique de l'ouvrage, compensant ainsi en partie l'essentialisation des données.

Ces rôles fondamentaux de l'interface demeurent toutefois largement sous-estimés dans les pratiques françaises<sup>[32]</sup>. La réflexion à leur sujet est bien souvent absente des projets, condamnant certains d'entre eux à demeurer inexploitable<sup>[33]</sup>. Les raisons de ces échecs sont multiples, allant de l'objectivité illusoire produite par les interfaces minimalistes ou inesthétiques à la méfiance platonicienne face aux représentations et aux images. Le principal problème demeure toutefois celui des politiques gouvernementales et des agences de recherches, qui se concentrent sur les données, aussi bien dans la conception des projets que dans leur pérennisation. Elles exigent ainsi régulièrement un plan de conservation des données sur plusieurs années, mais ne demandent jamais que l'interface soit conçue par un professionnel. Les initiatives étatiques pour la sauvegarde des projets numériques (HumaNum en France, DaSCH en Suisse) se concentrent sur les données, jamais sur leur interface.

## **Deuxième défi : financement et pérennité**

Sans surprise, le deuxième défi qui se dégage de ce dossier est celui de la pérennité des bases de données, c'est-à-dire de leur survie au-delà du temps de leur inauguration. Le problème relève de deux causes convergentes. La première est inhérente au format numérique : les publications numériques sont plus vulnérables que les publications imprimées et leur conservation nécessite une maintenance régulière. L'autre est liée aux choix politiques de financement de la recherche : dans la situation actuelle, les coûts associés au développement d'une base de données sont trop élevés pour être pris en charge par les financements pérennes des laboratoires. La plupart des bases de données sont mises en place dans le cadre de projets financés par des agences pour une durée limitée. Ce temps court entre en contradiction avec le temps long des usages, de la maintenance et de la conservation. Il implique en outre le plus souvent que la partie technique de l'élaboration de la base de données soit confiée à des personnels précaires ou à des prestataires sous contrat : la fin du contrat marque la fin de la collaboration, alors qu'une véritable maîtrise du devenir de la base de données nécessiterait une coopération pérenne. Pour faire face à ces difficultés, les projets *Mercurie galant* et *Récital* ont su s'ancrer dans des institutions de service public (l'équipe IREMUS du CNRS pour le premier ; l'université de Nantes et le laboratoire des sciences du numérique de Nantes pour le second) et ont ainsi pu s'appuyer sur des personnels statutaires. Cependant, le manque de moyens de ces institutions contraint les projets à rechercher très régulièrement des financements complémentaires pour pouvoir poursuivre leur développement.

De la sorte, le problème de la pérennité se pose à différents niveaux : il ne suffit pas d'assurer la conservation des données, mais aussi leur disponibilité – en garantissant la continuité de

l'hébergement – et la maintenance de l'interface, dont on a souligné le rôle essentiel. L'utilisation de formats standards, souvent avancée comme la solution à ces problèmes, ne résout en réalité que celui de la conservation des données. On peut bien élaborer un site dans le respect des standards de la TEI, mais celui-ci nécessitera toujours une institution d'accueil et des mises à jour pour fonctionner. On le constate, la question ne peut pas être résolue à l'échelle de chaque projet par la seule adoption de normes partagées et de pratiques uniformisées. En réalité, les contributions rassemblées ici indiquent toutes que l'obstacle réside avant tout dans la structuration des financements et la répartition des responsabilités institutionnelles.

Sur ce problème épineux, les institutions se renvoient la balle. En France comme en Suisse, les agences de financement de la recherche (ANR, FNS...) ont fait le choix de déléguer la maintenance des réalisations numériques aux universités. Or les moyens informatiques limités de ces dernières ne permettent absolument pas d'assurer cette tâche. Destinés avant tout à assurer la bonne marche des infrastructures, ils n'ont pas été conçus pour entretenir des projets qui se multiplient. Pour éviter de se reposer sur des fonds universitaires, CÉSAR a fait recours à la Très Grande Infrastructure de Recherche HumaNum et sa solution Omeka. C'est aussi le choix que prévoit de faire le projet Perform'Art à l'issue du financement ERC. On peut cependant s'interroger sur les capacités d'HumaNum à absorber les projets qui se multiplient, dans un contexte marqué par le manque criant de personnel.

Pour y parvenir, HumaNum, tout comme le DaSCH, a fait le choix d'une solution technique standardisée devant couvrir tous les besoins et tous les cas de figure. Si l'option est défendable pour la conservation des données, elle minimise – quand elle n'ignore pas tout simplement – leur exploitation, leur visualisation et leur manipulation. Dès lors, soit l'on utilise la solution imposée dans sa forme première, la standardisation annulant bien des possibilités du numérique évoquées ci-dessus ; soit l'on configure l'outil imposé en fonction des nécessités du projet, ce qui suppose le développement de modules spécifiques et pose à nouveau le problème de la mise à jour et de la maintenance. Face à cette alternative, la plupart des projets numériques font le choix de la solution *ad hoc*, développée par un membre du projet, et qui nécessite de recourir régulièrement à un informaticien ou une informaticienne, en comptant sur sa bonne volonté ou bien en faisant la chasse aux financements pour pouvoir rémunérer ses services.

Le modèle de financement par projet est ainsi foncièrement inadapté aux réalités du numérique. Tout en promouvant les nouvelles technologies, les grands organismes de recherches s'obstinent dans un modèle à durée limitée qui ne correspond absolument pas aux contraintes informatiques. Les coûts de départ d'un projet en ligne peuvent être moins élevés que l'impression d'un ouvrage, mais celui-ci requiert un entretien régulier et par conséquent, la possibilité de financer cette maintenance. Si celle-ci est réalisée au bon moment, son coût reste faible. En revanche, le manque prolongé de révisions conduit ultimement à des situations catastrophiques dans lesquelles le site entier doit être repensé. Pour pallier cette situation, certains projets multiplient les demandes de financements et donc le nombre d'institutions partenaires : les projets parviennent ainsi à s'inscrire dans la durée, mais à quel prix ? Le temps dépensé à monter des projets et à faire des bilans est un temps perdu pour la recherche, et le coût humain des situations de précarité que ce mode de financement engendre reste à mesurer.

Il y aurait pourtant une alternative qu'aucune des grandes initiatives européennes n'a voulu prendre en compte. Historiquement, le rôle de la conservation et de l'accès aux ressources est celui des bibliothèques. Certaines d'entre elles remplissent d'ailleurs partiellement cette mission. La BnF assure ainsi un archivage des projets numériques et du web, de même que le projet OpenLibrary. Cela se pense toutefois en termes d'archivage et d'histoire du web, et non

comme le maintien de ressources actives pour la recherche. Les projets numériques ont besoin de nouveaux bibliothécaires. Il y a un nouveau métier de documentaliste à inventer : des personnes disposant de compétences informatiques poussées, et qui s'occuperaient du maintien et de la mise à jour des réalisations numériques hébergées par la bibliothèque. Avec un tel modèle, chacun des projets présents dans ce numéro pourrait aisément être maintenu pendant bien des années au prix de quelques heures annuelles. Un jour peut-être, il sera aussi naturel d'entretenir des réalisations numériques que de prendre soin des livres...

---

## Notes

<sup>[1]</sup> Pierre Mounier, *Les Humanités numériques : une histoire critique*, Paris, Éditions de la Maison des Sciences de l'Homme, 2018.

<sup>[2]</sup> Voir notamment la *Revue d'Historiographie du Théâtre*, n°4, *Études théâtrales et humanités numériques*, Ioana Galleron (dir.), 2018, en ligne sur <https://sht.asso.fr/revue/etudes-theatrales-et-humanites-numeriques/>, et *Bien symboliques / Symbolic Goods*, n°2, *Arpenter la vie littéraire / Surveying literary life*, Claire Ducournau et Anthony Glinoeur (dir.), 2018, en ligne sur <https://www.biens-symboliques.net/207>.

<sup>[3]</sup> *CÉSAR*, en ligne sur <https://cesar.huma-num.fr/cesar2/>.

<sup>[4]</sup> *Le Théâtre de la foire à Paris*, en ligne sur <http://www.foires.univ-nantes.fr>. Le site est hébergé depuis 2005 par l'université de Nantes et maintenu par l'équipe du Centre d'études des théâtres de la foire.

<sup>[5]</sup> Voir Claire Ducournau et Anthony Glinoeur, *op. cit.*, 2018. Pour l'histoire littéraire, ils font notamment référence à Alain Viala, *Naissance de l'écrivain*, Paris, Minuit, 1985, et à Christophe Charle, *La Crise littéraire à l'époque du naturalisme, roman, théâtre, politique*, Paris, Presses de l'ENS, 1979. Pour l'histoire du livre, ils renvoient aux travaux de Roger Chartier et Robert Darnton (p. ex. Roger Chartier, et Henri-Jean Martin, *Histoire de l'édition française*, Paris, Promodis, 1983-1986, et Robert Darnton, *The Case for Books : Past, Present, and Future*, New York, PublicAffairs, 2009).

<sup>[6]</sup> Pierre Mélése, *Répertoire analytique des documents contemporains d'information et de critique concernant le théâtre à Paris sous Louis XIV, 1659-1715*, Paris, Droz, 1934.

<sup>[7]</sup> « Print and Digital Interfaces in Early Modern Literature », panel dirigé par Christophe Schuwey et Geoffrey Turnovsky, MLA, Seattle, 2020.

<sup>[8]</sup> Comme le souligne Clarisse Bardiot, « Arts de la scène et culture analytics », dans Ioana Galleron (dir.), *op. cit.*, en ligne sur <https://sht.asso.fr/arts-de-la-scene-et-culture-analytics/>.

<sup>[9]</sup> Pour un premier recensement, incomplet, voir Marine Roussillon, « Ressources en ligne sur les fêtes de cour (1) : bases de données » et « Ressources en ligne sur les fêtes de cour (2) : catalogues », 2016, en ligne sur *Politiques du Grand-siècle*, <https://pogs.hypotheses.org/292> et <https://pogs.hypotheses.org/312>.

<sup>[10]</sup> *PerformArt*, en ligne sur <https://performart-roma.eu/en/>.

<sup>[11]</sup> *Récital*, en ligne sur <http://recital.univ-nantes.fr/>.

<sup>[12]</sup> *Mercuré galant*, en ligne sur <http://obvil.sorbonne-universite.site/corpus/mercure-galant/>.

<sup>[13]</sup> *Théaville*, en ligne sur <http://www.theaville.org/>.

<sup>[14]</sup> *Naissance de la critique dramatique*, en ligne sur <https://www2.unil.ch/ncd17/>.

<sup>[15]</sup> *Molière21*, en ligne sur <http://moliere.huma-num.fr/>.

<sup>[16]</sup> Sur les nécessaires partis pris liés à l'acte de publication, voir Christian Jouhaud et Alain Viala (dir.), *De la publication : entre Renaissance et Lumières*, Paris, Fayard, 2002.

<sup>[17]</sup> Claire Ducournau et Anthony Glinoyer (dir.), *op. cit.*

<sup>[18]</sup> Voir Matteo Treleani, *Qu'est-ce que le patrimoine numérique ? Une sémiologie de la circulation des archives*, Lormont, Le Bord de l'eau, 2017, en particulier p. 84-88.

<sup>[19]</sup> Claire Ducournau et Anthony Glinoyer (dir.), *op. cit.*, §15.

<sup>[20]</sup> Cela explique peut-être en partie le refus formulé par Gallimard de renvoyer au site depuis l'édition papier. Voir la contribution de Claude Bourqui à ce dossier : RENVOI.

<sup>[21]</sup> Voir Christophe Schuwey, *Un entrepreneur des lettres au XVII<sup>e</sup> siècle : Donneau de Visé, de Molière au « Mercure galant »*, Paris, Classiques Garnier, 2020.

<sup>[22]</sup> Voir Alain Viala, *op. cit.* ; Déborah Blocker, *Instituer un « art » : politiques du théâtre dans la France du premier XVII<sup>e</sup> siècle*, Paris, Champion, 2009 ; Judith Lyon-Caen et Dinah Ribard, *L'historien et la littérature*, Paris, La Découverte, 2010.

<sup>[23]</sup> Ce projet est notamment énoncé par Christian Biet dans « Pour une extension du domaine de la performance (XVII<sup>e</sup>-XXI<sup>e</sup> siècle). Événement théâtral, séance, comparution des instances », *Communications*, n°92, 2013, p. 21-35.

<sup>[24]</sup> C'est ce que propose la contribution d'Anne-Madeleine Goulet. Renvoi.

<sup>[25]</sup> *Le Projet des registres de la Comédie-Française*, en ligne sur <https://www.cfregisters.org/fr/>.

<sup>[26]</sup> *Littératures classiques*, n°89, *Naissance de la critique dramatique*, Lise Michel et Claude Bourqui (dir.), 2016.

<sup>[27]</sup> *Projet Agon. La dispute : cas, querelles, controverses et création à l'époque moderne*, en ligne sur <http://www.agon.paris-sorbonne.fr/fr/>. Voir aussi *Littératures classiques*, n° 81, *Le Temps des querelles*, Jeanne-Marie Houstiou et Alain Viala (dir.), 2013.

<sup>[28]</sup> Stéphane Vial, *L'Être et l'écran : comment le numérique change la perception*, Paris, PUF, 2013, chap. 6.

<sup>[29]</sup> Voir notamment Anthony Masure, *Design et humanités numériques*, Paris, B42, 2017 ; Christophe Schuwey, *Interfaces : L'apport des humanités numériques à la littérature*, Neuchâtel, Livreo-Alphil, 2019 ; Nicolas Thély, « Rôle et enjeux du design graphique », THATCamp Paris, Non-actes des non-conférences des humanités numériques, Paris, Maison des Sciences de l'Homme, 2012 ; Stéphane Vial, *L'Être et l'écran*, Paris, PUF, 2013 et « Le tournant design des humanités numériques », *Revue française des sciences de l'information et de la communication*, n°8, 2016.

<sup>[30]</sup> Marine Roussillon, « Raconter les fêtes de cour : publier, archiver, agir », Bulletin du Centre de recherche du château de Versailles, 2<sup>e</sup> en ligne sur <https://journals.openedition.org/crcv/>, et « Introduction : Récits et imaginaires des fêtes de cour », *Revue d'Histoire du Théâtre*, n°282, 2019, p. 5-28.

<sup>[31]</sup> Michel Foucault, *Histoire de la folie à l'âge classique*, Paris, Gallimard, 1972, notamment p. 104 et p. 260-261.

<sup>[32]</sup> Aux États-Unis, les travaux de Johanna Drucker (notamment *Graphesis*), des initiatives telles que le Design Lab de Stanford ou encore des bourses d'encouragement de projets en humanités numériques destinés au public témoignent d'une culture et de pratiques différentes. Dans le domaine francophone, les chercheurs et chercheuses sur le sujet sont nombreux (voir note 31), mais leur rattachement à des universités étrangères dit bien l'intérêt modéré de la France pour ces questions.

<sup>[33]</sup> Christophe Schuwey, *op. cit.*, chapitre « Faire joli », p. 17-27.