



HAL
open science

Arts de la scène et culture analytics

Clarisse Bardiot

► **To cite this version:**

Clarisse Bardiot. Arts de la scène et culture analytics. Revue d'historiographie du théâtre, 2017, Etudes théâtrales et humanités numériques, 4, pp.11-20. hal-01843789

HAL Id: hal-01843789

<https://hal.science/hal-01843789v1>

Submitted on 18 Jul 2018

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.

Arts de la scène et *culture analytics*

Clarisse Bardiot

Université de Valenciennes

Résumé

Les études théâtrales sont sous-représentées dans les humanités numériques. Elles sont à l'orée de la seconde vague et en marge d'un phénomène en expansion, celui des *culture analytics*. Celles-ci peuvent se définir comme « l'analyse de la culture basée sur les données », dans un contexte où ces dernières sont nombreuses et complexes. Les *culture analytics* reposent sur le traitement algorithmique et la visualisation des données. Que peuvent-elles apporter aux études théâtrales ? Nous essayons de répondre à cette question en présentant des travaux récents, menés à la fois du point de vue de l'histoire du théâtre, de l'étude des textes littéraires et de la mise en scène. À partir de ces travaux, il est possible de développer une nouvelle approche que l'on pourrait nommer « theatre analytics », présentée ici dans ses grandes lignes.

Abstract : Performative arts and culture analytics

After outlining the place of performance and theatrical studies in the digital humanities sphere, this paper focuses on culture analytics. In the context of big data, culture analytics is « the data-driven analysis of culture ». Culture analytics is based on algorithmic approaches and data visualization: what can these methods bring to theatre studies. The paper seeks to answer to this question by presenting recent works conducted in performance arts, theatre history and artworks analysis. Based on these researches, it suggests the development of a new methodological approach, « theatre analytics ».

En 1956, sur les toits de la cité radieuse de Le Corbusier à Marseille, une sculpture cybernétique évolue dans un pas de trois avec deux danseuses. Mobile et autonome, elle répond aux mouvements de ses partenaires, adaptant son comportement en fonction de la lumière et de l'environnement sonore. *CYSP1* (pour les premières lettres de cybernétique et spatio-dynamique) est une œuvre du plasticien Nicolas Schöffer, réalisée en collaboration avec le chorégraphe Maurice Béjart et la société Philips. Avec elle naissent officiellement les *digital performances*, soit des œuvres scéniques faisant intervenir les technologies informatiques¹. Soixante ans plus tard, alors que les arts de la scène ont largement investi le champ des « nouvelles » technologies d'un point de vue esthétique, l'application de ces technologies à leur étude — ce que l'on appelle de manière générale les « humanités numériques » — est un phénomène embryonnaire. Les études théâtrales sont sous-représentées dans les grandes conférences et les revues dédiées aux humanités numériques, où la littérature et l'histoire dominant². Quant à l'enseignement des humanités numériques dans ce champ, en France et ailleurs, il demeure portion congrue. Les humanités numériques ne seraient-elles pas pour les études théâtrales l'étendard prometteur de nouvelles perspectives scientifiques qu'elles sont pour d'autres disciplines ?

Les deux vagues des humanités numériques dans le domaine des études théâtrales

Certes, il existe des projets précurseurs dès les années 1970³, ainsi que des initiatives individuelles et isolées depuis les années 2000, certains chercheurs en art de la scène œuvrant dans le champ des humanités numériques parfois tels Monsieur Jourdain, sans le savoir. Ça et là, des initiatives voient le jour⁴. La plupart d'entre elles concernent l'histoire du théâtre et la littérature, d'autres s'attachent à la reconstitution historique de lieux et de mises en scène dans des environnements numériques. Bien peu s'intéressent à l'étude des représentations et des processus de création contemporains. En 2013-2014, une prise de conscience collective et institutionnelle se structure autour de plusieurs initiatives : organisation d'un THATCamp les 20 et 21 juin 2013 à New York, intitulé « Performing Arts: Presenting and Archiving Performance in the Digital Age » ; puis en novembre d'une *Unconference* intitulée « DH & Theatre Research⁵ » et d'une session consacrée à la question du Big Data, à l'initiative de Doug Reside, dans le cadre du colloque de l'American Society of Theatre Research ; en 2014, création d'un groupe « Digital Humanities in Theatre

¹ Cf. Steve DIXON, *Digital performance: a history of new media in theater, dance, performance art, and installation*, Cambridge, MIT Press, coll. « Leonardo », 2007, 809 p. ; Clarisse BARDIOT, *Les théâtres virtuels*, thèse de doctorat, 2005 ; Clarisse BARDIOT, *Arts de la scène et technologies numériques: les digital performances*, Boulogne, Leonardo/Olats, coll. « Les Basiques », 2013.

² Ces deux disciplines sont également celles qui ont une plus grande antériorité dans le domaine. C'est en partie la raison pour laquelle de nombreux projets relèvent de ces champs. A noter qu'en littérature, c'est le roman qui a été largement favorisé. Comparativement, les pièces de théâtre n'ont pour le moment fait l'objet que de peu de travaux.

³ Voir *The London Stage Project, 1800-1900*, mené par Joseph Donohue.

⁴ Voir par exemple le blog de Robert Davis [<http://www.c19audiences.com/>]

⁵ Informations sur ces manifestations disponibles en ligne : [<http://performingarts2013.thatcamp.org/>] et [<http://dhandtheatrereseach.wordpress.com/>]. Consulté le 14/11/2016.

Research » à la FIRT (Fédération Internationale pour la recherche Théâtrale)⁶ ; la même année, keynote de Nic Leonhardt lors du congrès biennal de la SIBMAS (réseau international du patrimoine culturel des arts du spectacle) intitulée *Digital Humanities and the Performing Arts : Building Communities, Creating Knowledge*. De 2014 à 2016 la thématique devient récurrente et on la retrouve chaque année dans les symposiums cités ci-dessus, avec plus ou moins de publicité. À partir de septembre 2014, Kalle Westerling recense les projets (anglo-saxons) sur son blog⁷. Un premier article de synthèse sur le théâtre classique et les humanités numériques est publié en 2013 en espagnol⁸, puis un second, en anglais et d'ordre plus général dans *Theatre Journal* en 2015⁹. Cette même revue a publié un numéro spécial sur les arts de la scène et les humanités numériques, essentiellement d'un point de vue historiographique, en décembre 2016¹⁰. Toujours en 2016, l'ATHE (Association for Theatre in Higher Education) et l'ASTR créent un prix commun intitulé « Excellence in digital scholarship ».

Côté francophone, les initiatives sont très éparées et apparaissent en filigrane dans d'autres manifestations : séminaire en 2013 « Patrimoines du spectacle : en scène! » (BnF, Labex Arts-H2H Labex Patrima) sur les manuscrits de théâtre et les outils numériques permettant d'en exploiter la diversité et les valoriser ; plusieurs conférences sur le sujet dans le cadre des colloques *Processus de création et archives du spectacle vivant : manque de traces ou risque d'inflation mémorielle ?* (Rennes, octobre 2014), *Remettre en jeu le passé. Métamorphoses du corpus des Registres de la Comédie-Française* (INHA, décembre 2015) ou encore les journées d'études à la MESHS sur le numérique et l'étude des arts de la scène (2016, projet ULISES). À partir de début 2015, en collaboration avec Nicolas Thély et les soutiens de la MESHS et d'Humanistica, j'ai initié le réseau DORRA-DH pour la structuration d'un réseau de chercheurs en arts (au sens large du terme, y compris les études théâtrales) par le développement d'outils en humanités numériques.

La plupart de ces initiatives reposent sur la constitution de bases de données documentaires avec numérisation de fonds d'archives et mise à disposition des documents, parfois publiés avec des interfaces *ad hoc* sur internet ; d'autres s'appuient sur la TEI pour mener à bien un travail savant d'édition de textes littéraires. Ces projets relèvent de la « première vague » des humanités numériques¹¹ consacrée à la numérisation et à la constitution d'infrastructures (bases de données notamment). Dans les autres disciplines des sciences humaines, la première vague a eu lieu fin des années 1990-début des années 2000. Les arts de la scène subissent un décalage temporel, comme le montrent par exemple en France la numérisation des archives des Centres chorégraphiques nationaux à partir de 2008 ou bien la création du Portail des

⁶ L'initiative de ce groupe auquel je participe a été prise lors de la conférence annuelle de la FIRT à Barcelone, en juillet 2013.

⁷ <http://www.westerling.nu/> . Consulté le 14/11/2016.

⁸ Paul SPENCE, « Teatro clásico y humanidades digitales: el cruce entre método, proceso y nuevas tecnologías », *Teatro de Palabras*, août 2013, pp. 9-38.

⁹ Debra CAPLAN, « Notes from the Frontier: Digital Scholarship and the Future of Theatre Studies », *Theatre Journal*, vol. 67 / 2, 2015, pp. 347-359.

¹⁰ *Theatre Journal*, vol. 68 / 4, 2016.

¹¹ Todd PRESNER, « Digital Humanities 2.0: a report on knowledge », *Connexions Project*, 2010.

arts de la marionnette (PAM) en 2009¹². En Irlande, l'ensemble des archives du théâtre national irlandais, l'Abbey Theatre, soit plus d'un million de pages couvrant 110 ans d'histoire, est numérisé à partir de 2012. En Europe, le projet ECLAP – e-library for performing arts – s'emploie de 2010 à 2013 à agréger les contenus numérisés de plus de 35 institutions sur le portail Europeana, tout en les enrichissant selon différentes stratégies (métadonnées, annotations, etc.). Simultanément, dans le sillage des « artistes-archivistes » Merce Cunningham et Robert Wilson, on observe un intérêt croissant pour les archives d'artistes contemporains : numérisation et indexation des archives de Siobhan Davies à partir de 2007, de Pina Bausch à partir de 2010, de la Societas Raffaello Sanzio à partir de 2012. Le retard des arts de la scène sur d'autres disciplines peut s'expliquer par la complexité des fonds d'archives (textes, images mais aussi maquettes, costumes, objets, dont la numérisation est complexe).

Ces campagnes de numérisation ont largement été conçues dans l'optique de conserver des documents fragiles (par exemple les captations filmiques sur supports magnétiques) et de les rendre accessibles au plus grand nombre via des plateformes en ligne, dans un souci de diffusion de la culture. La numérisation des données ne modifie pas les pratiques courantes des chercheurs en arts de la scène, adeptes du *close reading*. Il s'agit pour eux d'une commodité et non pas d'un changement radical de méthodologie. Nous ne sommes aujourd'hui qu'à l'orée de la seconde vague, celle qui va nous permettre d'interagir avec des contenus fraîchement numérisés, voire nativement numériques. Toujours d'après Presner, cette seconde vague, « deeply generative, creating the environments and tools for producing, curating, and interacting with knowledge that is 'born digital' and lives in various digital contexts¹³ » a lieu au tournant des années 2010. Avec cette seconde vague apparaît une autre approche, encore très marginale mais avec un fort potentiel, celle des *culture analytics*. Celle-ci n'est possible qu'à partir du moment où des corpus de données numériques sont constitués et disponibles, ce qui a été l'enjeu de la première vague.

Les culture analytics

Les *culture analytics* sont « l'analyse de la culture basée sur les données »¹⁴. Elles sont nées dans le sillage des recherches de Lev Manovich, lequel a instauré le terme *cultural analytics* (sic) en 2005¹⁵. La culture est considérée ici au sens large du terme : il s'agit autant des interactions sociales que des expressions culturelles. Les données comprennent à la fois les corpus numérisés (textes, images, artefacts) et les données nativement numériques, comme par exemple les tweets, les comptes facebook, les pages internet, etc.

¹² Ces deux initiatives ont bénéficié du plan national de numérisation lancé en 1996 par le Ministère de la Culture et de la Communication français.

¹³ Todd PRESNER, *op. cit.*, p. 6.

¹⁴ « *Culture analytics* is the data-driven analysis of culture ». Culture analytics Long Program, IPAM, UCLA, *White Papers*, juin 2016, inédit.

¹⁵ Pour une présentation par Lev Manovich, voir « The Science of Culture? Social Computing, Digital Humanities and Cultural Analytics » CA: Journal of Cultural Analytics », dans *Journal of Cultural Analytics*, mai 2016, [En ligne : <http://culturalanalytics.org/2016/05/the-science-of-culture-social-computing-digital-humanities-and-cultural-analytics/>]. Consulté le 14/11/2016.

Au printemps 2016, l'institut de mathématiques appliquées (IPAM) à l'université de Californie à Los Angeles a organisé un programme de recherche intitulé *Culture Analytics*¹⁶. Pendant trois mois et demi, une quarantaine de chercheurs venus du monde entier et rassemblés par Tina Eliassi-Rad (Rutgers University, informatique), Mauro Maggioni (Duke University, mathématiques et informatique), Lev Manovich (The Graduate Center, CUNY, informatique), Vwani Roychowdhury (University of California, Los Angeles (UCLA), ingénierie) et Timothy Tangherlini (University of California, Los Angeles (UCLA), Langue et littérature scandinave) ont contribué à la définition de ce nouveau champ de recherche, à la croisée des sciences humaines et sociales, de l'informatique et des mathématiques appliquées. Seule représentante en études théâtrales, j'ai fait partie des chercheurs sélectionnés. À la fin du programme nous nous sommes réunis au Lake Arrowhead pour écrire ensemble le livre blanc des culture analytics et définir ainsi la discipline et ses enjeux. Pour ma part, j'ai contribué au chapitre sur les collaborations multidisciplinaires. Le présent article est largement inspiré de ce contexte et des échanges qui ont eu lieu lors du programme.

Les *culture analytics* sont caractérisées par deux phénomènes : l'utilisation d'algorithmes et la visualisation de données. Les algorithmes permettent de « fouiller » les données, de les analyser, de les agencer selon différents paramètres. La visualisation permet d'avoir un retour graphique de ces analyses dont le résultat se présente sous forme de tableaux chiffrés. Ces deux phénomènes découlent de la masse de données : sans parler de big data, le chercheur en sciences humaines est de plus en plus souvent confronté à des corpus de quelques milliers de documents (textes, images, vidéos, sons) ou à des réseaux (par exemple des comptes facebook ou twitter) qu'il est impossible d'analyser un à un. Les algorithmes liés à la visualisation permettent une lecture de loin (le *distant reading* théorisé par Moretti¹⁷). Cette approche permet d'identifier des motifs récurrents (*patterns*) ou encore des phénomènes particuliers qu'il convient ensuite d'analyser en revenant à la lecture de près. *Close Reading* et *distant reading* ne sont pas exclusives l'une de l'autre. Il ne s'agit pas de substituer le *distant reading* au *close reading* ou encore de sacrifier les méthodologies classiques sur l'autel du progrès informatique. Ces deux méthodes sont complémentaires : l'approche des *culture analytics* permet une boucle rétroactive entre lecture de loin et lecture de près, entre données globales et analyse d'un document spécifique, entre approche quantitative et analyse qualitative.

Que peuvent apporter les *culture analytics* aux études théâtrales ? Si les études théâtrales sont marginales dans le domaine des humanités numériques, elles le sont encore davantage dans le domaine des *culture analytics*. On peut néanmoins distinguer plusieurs approches et terrains de recherche, selon que l'on s'intéresse à l'histoire du théâtre, à l'étude des textes littéraires ou encore à la mise en scène.

¹⁶ Les informations sur le programme sont disponibles sur internet [<http://www.ipam.ucla.edu/programs/long-programs/culture-analytics/>]. Consulté le 14/11/2016.

¹⁷ Franco MORETTI, *Graphes, cartes et arbres : modèles abstraits pour une autre histoire de la littérature*, Paris, Les Prairies ordinaires, 2008.

Les culture analytics appliquées à l'histoire du théâtre

Un premier ensemble d'exemples concerne l'histoire du théâtre, dans ses dimensions politiques, sociales, économiques, avec des analyses qui peuvent couvrir plusieurs décennies ou au contraire détailler une période très courte. Le projet *Visualizing Broadway*¹⁸ conduit par Derek Miller (Harvard) est emblématique de cette démarche. Il s'appuie sur l'Internet Broadway Database, une base de données en ligne qui recense toutes les productions de Broadway depuis ses débuts (soit un siècle de théâtre à New York) et fourmille d'informations : non seulement les acteurs sont recensés mais également tous les corps de métiers impliqués, la durée de l'exploitation d'une pièce, le nombre de reprises, l'évolution du prix des billets, etc. Influencé par les travaux de Franco Moretti, Derek Miller analyse l'histoire de Broadway en prenant en compte l'ensemble des productions (donc aussi et surtout les échecs) et non pas seulement les succès (esthétiques, commerciaux) à partir desquels cette histoire a jusqu'ici été écrite. Cela lui permet par exemple de montrer que les personnes centrales du point de vue de la production théâtrale ne sont pas les grands noms que l'histoire a retenus (metteurs en scène, acteurs, producteurs) mais les scénographes, les costumiers, les régisseurs. Surtout, car cette information était déjà connue, cette analyse permet de montrer les réseaux et les circulations entre les différentes équipes de manière très précise et de répondre à des questions comme par exemple : à combien de productions un scénographe contribue-t-il en moyenne par saison ? Travaille-t-il avec quelques personnes attirées ou au contraire de manière très éclectique ? Est-il lié à un lieu spécifique ou bien au contraire passe-t-il d'un théâtre à un autre ? D'où un appel de la part de Derek Miller à reconsidérer les métiers et les parcours de ces créateurs de l'ombre. De la même manière, en se basant sur les critiques publiées dans les journaux, il fait apparaître les différents thèmes traités dans les productions de Broadway, leurs connexions et leur évolution au cours d'une période donnée. Dans la même veine, Patrick Lonergan commence à mener des analyses sur le théâtre national irlandais, à partir de la numérisation des archives mentionnée plus haut, en commençant par la place des femmes au cours de son histoire.

Debra Caplan s'intéresse à un corpus plus délimité, celui de la troupe Vilna¹⁹, un théâtre expérimental Yiddish qui s'est développé de 1915 à 1936 en Europe et a réuni 290 artistes. À partir des archives de la troupe, elle retrace les relations entre les différentes personnes reliées de près ou de loin à la compagnie. La visualisation de données qui résulte du dépouillement des documents conservés permet d'analyser différents types de connections : professionnelles, familiales (mariages), scolaires (maîtres et élèves), amicales. Ceci afin de répondre à des questions comme : à quelle point la troupe était-elle une affaire de famille ? Comment le réseau a-t-il eu un impact sur la carrière des femmes ? Quelles sont les personnes qui ont le plus d'influence ? Quelles sont les connections de la troupe avec les figures majeures du théâtre, du cinéma, de la littérature ou des arts plastiques de l'époque ? Ce projet fait partie d'un projet plus large, le Digital

¹⁸ <http://visualizingbroadway.com/> Consulté le 14/11/2016.

¹⁹ <https://s3-us-west-1.amazonaws.com/vilnatroupenetwork/index.html> Consulté le 14/11/2016.

Yiddish Theatre Project, dont la vocation est « d'appliquer les outils et les méthodes des humanités numériques à l'étude du théâtre Yiddish²⁰ ».

Dernier exemple : Kate Elswit explore la tournée du ballet américain Caravan en 1941²¹. Son travail se base essentiellement sur une représentation cartographique des données de la compagnie.

Comme on le constate dans ces exemples, toutes les échelles temporelles sont possibles. Tout dépend de la nature des données disponibles et des perspectives de recherche. Ces analyses aboutissent à la représentation de réseaux, de relations complexes (*complex networks*), que la visualisation sous forme de graphes (en utilisant en particulier le logiciel Gephi) et de cartes permet de déchiffrer.

Les culture analytics appliquées à la littérature théâtrale

Le second groupe d'exemples concerne l'analyse des textes. En dépit des promesses initiales, qui annonçaient un rapide développement des études numériques en raison du caractère structuré (et donc informatisable) du texte de théâtre, peu de travaux concernent la littérature théâtrale. Si la TEI est un outil qui permet l'édition numérique savante des textes et manuscrits, d'autres outils d'analyse algorithmique sont également disponibles. L'exemple sans doute le plus célèbre, qui a défrayé récemment la chronique, est l'utilisation de ces méthodes pour réattribuer les pièces de Shakespeare. C'est ainsi que 9 pièces sont considérées dorénavant comme écrites en collaboration avec d'autres auteurs, dont Christophe Marlowe, en plus des 8 qui avaient déjà été identifiées comme telles sur les 44 œuvres attribuées à Shakespeare. Elles sont dorénavant présentées comme issues d'une collaboration dans le *New Oxford Shakespeare* publié fin 2016 - début 2017 par Oxford University Press. Pour parvenir à ce résultat, les chercheurs ont utilisé trois approches de *text mining* différentes afin d'identifier des motifs récurrents, des signatures stylistiques propres à chaque auteur et ainsi confronter Shakespeare à ses collaborateurs.

Le *topic modeling* commence à être utilisé par des chercheurs en études théâtrales. Il s'agit d'une méthode algorithmique qui permet d'identifier dans un ensemble de textes les thèmes récurrents et leur proportion dans chacun des textes. Cette technique est très utilisée pour analyser l'évolution d'un sujet au cours du temps dans un journal. La visualisation de données permet ensuite d'observer globalement le phénomène et de pointer par exemple le moment précis de l'apparition puis de la disparition d'un thème, éventuellement en conjonction avec d'autres phénomènes. Christof Schöch²², de l'université de Würzburg, utilise cette méthode pour caractériser les genres de 750 pièces de théâtre françaises publiées de 1610 à 1810. Il identifie ainsi des thématiques propres à la comédie ou à la tragédie sur l'intégralité du corpus. Plus précisément, cette méthode lui permet d'identifier deux sous-catégories dans les

²⁰ <https://yiddishstage.org> Consulté le 14/11/2016.

²¹ <https://movingbodiesmovingculture.wordpress.com/> Consulté le 14/11/2016.

²² Christof SCHÖCH, « Topic Modeling Genre: An Exploration of French Classical and Enlightenment Drama », *Digital Humanities Quarterly*, à paraître [En ligne : https://zenodo.org/record/48356/files/Schoech_2016_Topic-Modeling-Genre_preprint.pdf]. Consulté le 14/11/2016.

comédies et les tragédies, ou encore de rapprocher certaines comédies avec les tragédies et vice-versa.

Une autre approche est l'utilisation des réseaux pour représenter la structure dramatique des pièces de théâtre, en particulier à partir des interactions entre les personnages. L'un des travaux précurseurs est l'étude de Franco Moretti sur *Hamlet*, où il montre que la visualisation des relations entre les personnages lui permet de passer d'une représentation temporelle à une représentation spatiale, et chemin faisant, d'imaginer que les genres dramatiques puissent être représentés par des formes visuelles spécifiques²³. D'autres chercheurs ont poursuivi cette hypothèse de recherche. Aris Xanthos & Al. proposent une application open source permettant de visualiser les connexions entre les personnages d'une pièce tout en prenant en compte son évolution temporelle. Dans ce travail, il ne s'agit pas de réduire l'ensemble des interactions entre les personnages à une seule image mais de prendre en compte leur évolution dans le temps²⁴.

Les culture analytics appliquées à la mise en scène

Le troisième ensemble concerne l'analyse des représentations scéniques. Sur quelles données s'appuyer ? Autrement dit, comment quantifier ce qui semble a priori irréductible à la mise en chiffres ? La question de la capture, de la saisie des représentations théâtrales et de leur conversion en données est sans doute la plus épineuse qui soit posée dans le domaine des études théâtrales. Dans les deux ensembles précédents – l'histoire et les textes littéraires – le type de données et d'analyses menées sont très classiques dans le domaine des *culture analytics*. Comment transformer en données tangibles et quantifiables des représentations éphémères ? Trois réponses sont proposées : les images, et en particulier les captations vidéo ; la captation de mouvement à même le corps des interprètes ; l'ensemble des traces documentaires issues des œuvres et de leur processus de création.

Dans *Synchronous objects*²⁵, la captation vidéo d'une chorégraphie de William Forsythe, *One Flat thing reproduced*, est le point de départ de toute une série d'analyses. L'image vidéo est annotée graphiquement, analysée avec différents outils algorithmiques empruntés à d'autres disciplines comme la géographie, pour traduire la danse en de multiples visualisations. Ce processus permet de dénoter des aspects singuliers autrement invisibles, de rechercher des motifs, de donner accès à d'autres informations sur la danse : comment comprendre le mouvement autrement que dans la représentation des corps mêmes (ce que fait la captation vidéo) ? Pour Forsythe et ses collaborateurs, la visualisation des

²³ Franco MORETTI, *Network Theory, Plot Analysis, Pamphlets 2*, Literary Lab, Stanford, 2011. [En ligne : Consulté à l'adresse <http://litlab.stanford.edu/LiteraryLabPamphlet2.pdf>]. Consulté le 14/11/2016.

²⁴ Aris XANTHOS, Isaac PANTE, Yannick ROCHAT [et al.], « Visualising the Dynamics of Character Networks », in *Digital Humanities 2016: Conference Abstracts*, Jagiellonian University & Pedagogical University, Kraków, 2016, p. 417-419, [En ligne : <http://dh2016.adho.org/abstracts/407>]. Consulté le 14/11/2016.

²⁵ Norah ZUNIGA SHAW, William FORSYTHE, et Maria PALAZZI, « Synchronous Objects », [En ligne : <http://synchronousobjects.osu.edu>]. Consulté le 14/11/2016 ; Clarisse BARDIOT, « Une autre mémoire : la chorégraphie des données. À propos des objets numériques développés par William Forsythe (Improvisation Technologies, Synchronous objects et Motion Bank) », in Anne Bénichou, (éd.). *Recréer / scripter – Mémoires et transmissions des œuvres performatives et chorégraphiques contemporaines*, éd. Anne Bénichou, Dijon, Les presses du réel, 2015, (« Nouvelles scènes »), p. 235-252.

données n'épuise pas le mouvement mais au contraire suscite d'autres recherches artistiques.

Les recherches sur la captation de mouvement sont riches et fécondes. Il n'est pas question de les retracer ici mais d'indiquer que ces données mènent implicitement à une approche de type *culture analytics*, y compris sur des corpus restreints.

L'utilisation de l'ensemble des traces documentaires issues des œuvres et de leur processus de création est la démarche que je conduis au sein du projet ReKall²⁶. ReKall se base sur les métadonnées présentes dans chacun des documents pour en extraire des informations cruciales (auteur, date de création, lieu de création, mot-clé, etc.). Ces informations sont ensuite utilisées par les outils d'analyse et de visualisation de l'information, afin de révéler des comportements, des usages. En jouant avec les axes et différents filtres qui proposent autant d'angles de vues sur une création donnée, les outils de ReKall permettent d'analyser graphiquement l'œuvre au travers de ses documents et de répondre à des questions telles que : comment une idée a-t-elle circulé d'une personne à une autre au sein d'une équipe ? L'ensemble des collaborateurs est-il présent à toutes les étapes du processus de création ou bien n'interviennent-ils que de manière ponctuelle ? Quels sont les documents les plus importants ?

Theatre analytics

Comme on le voit dans ces exemples empruntés aux trois grands domaines des études théâtrales (histoire, littérature, mise en scène), les *culture analytics* offrent de nouvelles perspectives pour la recherche. Elles permettent non seulement de travailler sur des dimensions de corpus jusque-là impossible à explorer pour le chercheur solitaire mais aussi de proposer des hypothèses de recherche auxquelles seules ces méthodes peuvent répondre. Sans le *text mining*, il aurait été impossible de prouver que certaines des pièces de Shakespeare relevaient en réalité d'une collaboration entre plusieurs auteurs. À partir de ces premières expériences, on comprend l'intérêt de développer aujourd'hui une approche que l'on pourrait nommer « theatre analytics ». Pour paraphraser la définition des *culture analytics*, les *theatre analytics* sont l'analyse des arts de la scène basée sur les données.

Comme pour les *culture analytics*, la question des données est centrale pour les *theatre analytics*. Leur collecte et leur traitement – qui sont loin d'être neutres – sont les prémices à toute entreprise de recherche de ce type. C'est aussi très souvent la partie la plus laborieuse et la plus coûteuse en termes de temps de travail. Un certain nombre de données sont dorénavant et déjà structurées et disponibles, en particulier grâce aux campagnes de numérisation de la première vague des humanités numériques, ou bien parce qu'elles sont nativement numériques et donc plus faciles à assembler. Mais le travail du chercheur ne peut se limiter à la constitution de bases de données, y compris lorsqu'elles sont mises à disposition au format open data ou sous forme de site web. Il ne s'agit là

²⁶ Le logiciel est téléchargeable sur internet : <http://www.rekall.fr>. Voir également BARDIOT, Clarisse, « Video recording and documentation of scenic arts: from the annotation to the visualization of metadata, the example of the ReKall software », *Acoustic Space. Data Drift. Archiving Media and Data Art in the 21st Century*, vol. 14, éd. Rasa Smite, Raitis Smits et Lev Manovich, octobre 2015, p. 159-168.

que d'une première étape qui peut mener vers des recherches classiques (en *close reading*) et des recherches de type *theatre analytics*.

Quelles sont les données des *theatre analytics* ? Impossible de les énumérer toutes ici. Elles vont des comptes twitter et facebook des théâtres aux revues spécialisées en passant par les captations vidéo, les fonds d'archives classiques ou encore les sites internet des compagnies et les données disponibles sur des sites tels que data.bnf. Le numéro de *Theatre Journal* consacré aux humanités numériques montre à quel point les programmes de théâtre sont une ressource précieuse pour mener de telles analyses en histoire du théâtre. Nous évoquerons donc quelques pistes, en plus des fonds spécifiques que nous avons brièvement mentionnés au début de cet article.

Les textes littéraires sont de plus en plus disponibles, qui plus est au format TEI. La seule restriction principale, et elle est de taille, concerne les auteurs contemporains ou les traductions couvertes par le droit d'auteur, ce qui peut restreindre l'accès aux données. Cette question est loin d'être négligeable et est la principale explication au fait que les textes anciens sont aujourd'hui privilégiés par ce type d'approche. Nous touchons ici à un problème de fond pour les données nativement numériques : leur accessibilité pour les chercheurs et le développement des données ouvertes pour la recherche.

Nonobstant, un « gisement » de données passionnant est celui des billetteries des salles de spectacles. Elles recèlent de nombreuses informations socio-économiques pour l'étude des publics, qu'il est possible de croiser avec d'autres informations, comme les politiques culturelles. En Angleterre, à partir d'avril 2016, les institutions financées par le Art Council doivent être en mesure de collecter et de partager des informations statistiques sur les spectateurs. La société Purple Seven leur propose l'infrastructure nécessaire mais également l'interprétation de ces données afin que les théâtres puissent adapter leur publicité, les prix des billets et la programmation de manière personnalisée en fonction des goûts et comportements de chaque spectateur²⁷. On le devine, les données sont ici utilisées pour des questions d'optimisation, au mieux pour le remplissage des salles, au pire pour faire du profit, avec toutes les menaces que cela peut faire peser sur des programmations de type « art et essai ». Les données et leur analyse ne sont jamais neutres. Il est aussi possible d'utiliser ces mêmes informations pour montrer comment peut se structurer la (sur)vie d'une compagnie dans le circuit des institutions publiques et pour comprendre à quel moment et dans quelles circonstances elles ont davantage besoin d'accompagnement.

Dernière piste de travail : les archives visuelles des arts de la scène. Celles-ci concernent les affiches, les programmes, les photographies, les captations vidéo, les croquis de scénographie, de costumes... hier numérisés, aujourd'hui nativement numériques. Lev Manovich a montré comment les *cultural analytics*, devaient s'intéresser à l'analyse de grands corpus d'images et non pas seulement aux textes ou aux données numériques²⁸. Ceci entraîne l'utilisation et la création d'outils informatiques différents de ceux utilisés pour le texte. Les historiens de

²⁷ <http://www.purpleseven.co.uk/data-sharing> Consulté le 14/11/2016.

²⁸ Voir par exemple Lev MANOVICH, « Cultural Analytics: Visualizing Cultural Patterns in the Era of 'More Media' », *Domus*, Spring 2009.

l'art commencent à s'en emparer et de nombreuses connexions et collaborations sont à envisager de ce point de vue avec les études théâtrales²⁹.

La plupart du temps, ces données ne sont pas des « big data » à proprement parler mais des medium data. Tout est une question d'échelle. Et pourtant, nous sommes submergés par ces données qui ne sont pas si « big » que cela : comment à mon échelle individuelle gérer cette (moyenne) masse d'information, dont les variables créent une grande complexité ? Comment y retrouver l'aiguille dans la botte de foin ou en saisir les grandes lignes (et si possible les deux à la fois) ? Et si en plus je pouvais y percevoir des phénomènes dont je n'avais même pas l'intuition ? C'est cela que peuvent offrir les culture analytics aux études théâtrales, avec une diversité d'approches informatiques selon la nature des données et des questions de recherche posées. Le traitement informatique (traitement algorithmique et visualisation) a lieu entre deux temps maîtrisés par les chercheurs en sciences humaines : la constitution des données et l'interprétation des résultats³⁰. Entre les deux ont lieu les différents traitements informatiques. Il est vivement conseillé au chercheur, afin d'éviter autant que faire se peut le phénomène de la « boîte noire », de maîtriser les outils informatiques qu'il utilise ou de collaborer avec des informaticiens, des mathématiciens et des designers spécialisés dans la visualisation de données, comme le montrent les exemples rapidement évoqués dans cet article. Le choix de telle ou telle méthodologie computationnelle puis l'interprétation des résultats requièrent des connaissances préliminaires et un investissement minimal. Dans les faits, ces trois moments forment une boucle : chaque interprétation des résultats est susceptible d'entraîner une modification des données (par exemple de leur structure ou parce que l'on décide de travailler sur une portion du corpus), un ajustement du programme informatique puis de l'interprétation et ainsi de suite. Dans l'interprétation des résultats, l'approche algorithmique et la visualisation de données n'empêchent pas le *close reading* et les méthodes classiques. Bien au contraire : elles le nécessitent. Il faut revenir aux sources, vérifier, expliciter ce que montrent les données. Les *theatre analytics* proposent une autre façon de naviguer dans le corpus, entre micro et macro, entre la loupe et la vue d'avion, entre *close reading* et *distant reading*. La gymnastique consiste à aller de la constitution des données à l'interprétation des résultats, non pas en cercle, dans une course infinie et vaine, mais en spirale, pour faire progresser l'analyse, les hypothèses et les résultats de la recherche. Non pas une spirale infernale où la recherche en études théâtrales serait le dernier rempart aux ordinateurs et au « tout numérique ». Mais une spirale vertueuse pour faire de la recherche en théâtre, car c'est bien de cela qu'il s'agit, non pas du contournement/détournement d'un champ mais au contraire de l'approfondissement de nos connaissances sur une discipline qui s'est trop peu interrogée sur ses méthodologies de recherche.

²⁹ C'est le sens du projet Dorra-DH mentionné plus haut. Voir également la création de la nouvelle revue en 2015 *International Journal for Digital Art History* [<http://www.dah-journal.org/index.html>]. Consulté le 14/11/2016.

³⁰ Pour la représentation d'un workflow typique des culture analytics, cf. David Blei, http://helper.ipam.ucla.edu/publications/caws4/caws4_12767.pdf

Remerciements : cette recherche a en partie été effectuée dans le cadre de la visite de l'auteur en 2016 à l'Institute for Pure and Applied Mathematics (IPAM), soutenu par la National Science Foundation (USA).