

## Entre appréciation et description, les goûts musicaux à l'épreuve de la " data "

Dominique Cotte, Marie Despres-Lonnet, David Vandiedonck, Mathias Heizmann, Bernard Jacquemin

### ► To cite this version:

Dominique Cotte, Marie Despres-Lonnet, David Vandiedonck, Mathias Heizmann, Bernard Jacquemin. Entre appréciation et description, les goûts musicaux à l'épreuve de la " data ". Documents et dispositifs à l'ère post-numérique. Actes du 18e Colloque international sur le Document Électronique (CiDE.18), Université de Montpellier 3 - LERASS, Nov 2015, Montpellier, France. pp.159-170. hal-01281249

HAL Id: hal-01281249

<https://hal.univ-lille.fr/hal-01281249>

Submitted on 1 Mar 2016

**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.

# Entre appréciation et description les goûts musicaux à l'épreuve de la « data »

*Between appreciation and description  
musical tastes challenged by the "data"*

Dominique COTTE, Marie DESPRÉS-LONNET

David VANDIEDONCK, Mathias HEIZMANN

Bernard JACQUEMIN

GERiiCO, Université de Lille (France)

{Dominique.Cotte,Marie.Despres-Lonnet,David.Vandiedonck  
Mathias.Heizmann,Bernard.Jacquemin}@univ-lille3.fr

## Résumé

Cet article décrit le cadre de référence d'un projet de recherche sur l'utilisation des données musicales dans la perspective du web de données et des alignements de référentiels. Il pose la question de la façon dont le mouvement de la « data » interroge la transcription de pratiques, comme l'écoute musicale, en « données » exploitables par des algorithmes de recherche. Il est en effet problématique de tenter d'encadrer, dans des modèles normalisants, des pratiques qui relèvent de paramètres personnels et sociaux complexes, liés au « goût ». C'est pourquoi le projet auquel nous participons actuellement tente de comprendre comment la normalisation informatique, nécessaire pour l'interopérabilité des données, tente de « faire avec » la complexité et l'inventivité de l'écoute musicale comme pratique sociale.

**Mots-clés :** Musique, pratiques musicales, données, catalogues, systèmes de recommandations.

## Abstract

This paper discusses the framework of a research project focusing on the use of music data in the prospect of terminological alignments. It questions the way the "digital data wave" aims to transcribe practices, such as listening to music, into computer usable data. The attempt to frame complex social and personal practices, linked to "tastes", can prove problematic, that is why we are currently involved in a project attempting to understand how the search for a much-needed digital normalization tries to "deal with" the complexity and inventiveness of music listening as a social practice.

**Keywords:** Music, musical practices, data, recommendations, archive, OPAC.

## 1 Introduction

Cette communication vise à présenter les intentions d'un projet de recherche axé sur la représentation de données sur la musique et leur réutilisation via l'application des technologies et méthodes du « web de données » (ontologies, modèles, référentiels, systèmes de recommandations...). Ce projet, intitulé Dorémus<sup>1</sup>, réunit des professionnels de la musique et de sa gestion documentaire, des chercheurs en informatique, des chercheurs en sciences de l'information et de la communication et des entreprises spécialisées dans le domaine de la documentation ou de la mise à disposition de biens culturels. La dimension interdisciplinaire du projet a nécessité une première phase de travail collectif pour poser le cadre épistémologique et scientifique dans lequel s'inscrit notre recherche. Nous proposons, dans cette communication, de revenir sur une partie des questions théoriques que cette première étape nous a permis d'aborder.

Au sein du consortium, notre équipe vise à intégrer, une dimension « usage » dans la mise en œuvre de solutions techniques, en amont et en parallèle de la promesse inscrite dans le « web de données » de faire communiquer entre elles diverses ressources catalographiques sur la production et l'édition musicale et plus largement les « données musicales » disponibles en ligne. Notre objectif est de comprendre comment les usagers abordent aujourd'hui la musique en fonction des multiples situations nouvelles qui leur sont offertes avec les développements récents de la technologie : écoute en mobilité connectée (par comparaison avec la mobilité analogique des premiers « baladeurs » qui ont vu le jour à la fin des années 1970 il y a maintenant presque quarante ans), téléchargements, streaming, partage de playlists, échanges de types réseaux sociaux, constitution de discothèques « virtuelles » et ainsi de suite.

Il s'agit pour nous d'envisager l'écoute musicale comme une pratique. Nous préférons, au concept d'usages cette notion, plus plastique, de pratiques : plutôt qu'à la sociologie des usages, c'est donc dans une perspective pragmatique que s'inscriront nos questionnements et explorations. Nous nous situons ainsi dans une approche critique par rapport à la notion « d'usage », dans le domaine des technologies, qui consiste trop souvent à postuler que des objets ou des systèmes, tels qu'ils ont été conçus, anticiperaient – voire intégreraient comme une affordance (Norman, 1999)<sup>2</sup> – des usages dont il conviendrait ensuite d'observer le bon déroulement. Les pratiques musicales s'ancrent dans des dispositifs sociotechniques, mais elles ne peuvent pas être saisies sous le seul angle des objets techniques qui les équipent. Elles impliquent une dimension sociale et culturelle, se comprennent par rapport à un environnement et une temporalité, une image de soi et des autres, des compétences et des connaissances qui sont mobilisées dans et par l'action.

1. Ce projet financé par l'Agence nationale de la recherche dans le cadre du programme "le numérique au service des industries culturelles et éditoriales (DS0703)" – ANR 14-CE-0020, réunit les partenaires suivants : la BnF, la Philharmonie de Paris, Radio France, les laboratoires Eurecom, Gerico et Lirmm, ainsi que les sociétés Meaning Engines et Ourouk.

2. On dit à Donald Norman ce concept, avec lequel il a pris ensuite une certaine distance, et qui laisse entendre que l'usage d'une interface est spontanément compris par l'utilisateur, qu'il se laisse « saisir » en quelque sorte par ce qu'il montre.

Les données catalographiques (qu'il s'agisse ici du catalogue des éditeurs ou des institutions de référence) rendent peu compte de ces dimensions, et on nous dira qu'elles ne sont pas là pour ça. Certes, mais le mouvement d'exposition de la « data » amène aujourd'hui à mettre en scène ces données catalographiques dans un univers de la « recommandation », du partage (web dit 3.0) qui en change l'environnement et interroge les modèles traditionnels de description de la matière musicale.

Ce sont ces nouveaux environnements que nous souhaitons interroger, car c'est dans l'emboîtement composite de lieux, d'espaces, d'objets techniques, supports et prothèses, tenus ensemble par des relations situées dans des mondes sociaux, que se dessine et se joue la possibilité d'une relation à la musique. C'est l'écoute, comme acte performatif faisant advenir la musique, qui solidarise ces objets, ces lieux et ces acteurs, et c'est ce composite (Le Marec, 2003) que nous voulons observer.

## 2 L'objet d'écoute dans sa dimension documentaire

Notre travail s'inscrit dans un projet sur les « données », reprenant ainsi un terme consacré dans le langage courant de l'industrie informatique d'aujourd'hui, mais chargé en plus d'une forte dimension culturelle, économique et sociale sous la figure de la data, que l'on décline en « big », « smart », « open », « linked »...

Nous nous trouvons soumis à une sorte d'injonction paradoxale : il faut produire et valoriser de la donnée, dans des quantités si grandes que l'on parle de « données massives » (pour traduire Big Data), sans que l'on se demande réellement pour quoi faire et quel peut en être l'intérêt. Évidemment tout phénomène peut générer une quantité astronomique de données. Un capteur corporel peut enregistrer en temps réel notre flux sanguin, notre rythme cardiaque, notre sudation et autres phénomènes biologiques, à partir de quoi on peut imaginer toutes sortes d'usage sanitaires mais aussi sociaux (mesure du stress). On se trouverait ainsi à la tête de milliards de données, dont il est supposé que leur « traitement » permettrait non seulement d'analyser, d'expliquer, mais aussi de prédire et d'anticiper toutes sortes de phénomènes (comportementaux, sociaux...)

Cette approche ne laisse cependant pas d'être problématique, car elle gomme toute la dimension liée au travail de construction qui a, au fil du temps produit les données en question. Bruno Latour (2001) proposait déjà il y a trente ans de parler des « obtenues » plutôt que des « données » (cité dans Noyer, Carmes, 2013), car justement, les collections d'éléments discrets servant à recenser et décrire tel ou tel phénomène ne nous sont jamais « donnés ». Toute activité d'analyse vise à prélever dans le réel des éléments qu'il semble pertinent de retenir à l'aune de la théorie qui la sous-tend, puis à les agencer entre eux dans des modèles de représentation qui permettent de faire avancer la connaissance. Cet arbre que je vois peut être étudié et analysé sous de nombreux regards différents, du biologiste à l'agriculteur en passant par le service des espaces verts de la commune, le garde forestier... Chacun de ces acteurs peut avoir intérêt à prélever, de cet objet des traits significatifs que l'on peut qualifier de « données » dès lors qu'elles sont susceptibles de s'inscrire dans un champ descriptif : taille, espèce, forme des feuilles, âge, lieu d'acclimata-

tion et ainsi de suite. Les motifs qui président au choix de ces critères mais aussi et surtout à leurs différents agencements, sont situés, car ils répondent à des projets différents en fonction des situations et des acteurs qui s'organisent pour la collecte de ces « données ». La date de plantation de l'arbre relève du fait objectif et peut se retrouver telle quelle dans chacun des « jeux de données » possibles, mais elle n'a ni la même signification ni la même organisation pour le botaniste ou pour le marchand de bois. Ce même indicateur va se trouver inséré dans un système de références singulier, induisant par exemple pour le marchand de bois un prix ou un marché particulier.

Dans sa sémantique même, le terme « données » est fondamentalement ambigu, car il apparaît comme un synonyme de « déjà-là ». Évoquer la « collecte » de données c'est renvoyer à une logique d'économie de cueillette où il suffirait de disposer des bons capteurs pour réunir des collections de ressources fiables et exploitables. Or, les « données » sont toujours des construits épistémologiques, prélevés à partir de « déjà-là » qui sont des éléments de la nature ou de la société. La seule délimitation de ces éléments relève déjà d'une construction théorique (Davallon, 2004) et donc d'un regard qui peut varier d'une culture à une autre. En ce qui concerne la musique, il s'agirait alors de remonter cette chaîne des « déjà-là » pour définir un élément de départ, originaire, dont la complexion s'accorde avec une description en éléments simples que l'on pourrait qualifier de « données ». Mais cette demande n'est déjà pas si simple : qu'est-ce qui est premier : une « œuvre » ? Un air ? Toute musique ne prend-elle pas la suite d'une autre musique et donc aussi la part d'inspiration qui peut résider dans la première ? Les ethnomusicologues peinent à distinguer ce qu'est une « œuvre » par rapport à des chants qui n'ont ni auteur, ni titre mais qui se définissent plutôt par rapport à une situation sociale, un événement, un rite. Dans ce cas, le titre attribué par le chercheur (« chant de foulon », « chant pour mariage de groupe »...) est un élément rapporté.

On comprend aisément, à partir de ce seul exemple, les difficultés qui peuvent naître lorsque l'on projette d'englober dans un système de classification, sous le mot « musique », des productions aussi dissemblables que les *Études pour piano de Ligeti* enregistrées par Pierre Laurent Aimard en 2001 et les *Chants Pygmées Aka* captés par Simha Arom dans les années soixante. En réalité, ce n'est pas tant le concept de musique (à condition de ne pas confondre « musique » et « œuvre musicale ») qui pose problème, que celui des pratiques musicales. Si, comme le rappelle Gilbert Rouget, « les Pygmées chasseurs-cueilleurs de la grande forêt équatoriale africaine sont connus pour être des gens qui chantent et dansent » (Rouget, 2004), cela ne signifie pas nécessairement qu'ils fassent de la musique dans l'acception occidentale de ce mot. Le terme même, par l'orientation de la pensée qu'il induit, ne permet pas de rendre compte du point de vue de la culture source. Dans le cas de certains chants pygmées, il s'agit d'assurer le succès d'une chasse en l'accompagnant d'une pratique musicale intégrant chants et danse. Et quand bien même on se trouverait dans le cas d'une pratique ludique rappelant les *Essercizi* de Scarlatti, il est impossible de considérer le résultat de cette activité comme une œuvre musicale : « la musique en tant qu'événement sonore n'est en fait pas le but principal [du groupe considéré]. Ces gens ne sont pas occupés à "faire de la musique" [...] Ce que le groupe a manifeste-

*ment en vue, c'est le plaisir pris collectivement à la produire, bien plus que le produit lui-même*» (Rouget, 2004).

Tout le problème tient dans la difficulté à dire ce qu'est la musique en tant qu'objet construit, cet objet étant grandement lié aux définitions qu'en donnent les cultures. En Occident, nous considérons volontiers la musique sous l'angle de ses productions, ce qui nous permet notamment d'envisager une « ontologie » partant de la production pour aller vers les œuvres : musique → musique classique → sonate → Sonate au clair de lune. Or, une telle chaîne est loin d'aller de soi. Ainsi l'action même peut être décrite mais son objet ne peut être considéré du point de vue occidental sans produire une série de malentendus, malentendus encore accentués lorsque deux types de productions, radicalement dissemblables, se retrouvent juxtaposés. D'où par exemple l'utilisation par Gilbert Rouget dans *La Musique et la transe* (1980) du verbe « musiquer » qui met l'accent, selon son auteur, « sur l'action même, vue indépendamment de son résultat : aux yeux de l'ethnomusicologue la manière de faire de la musique importe tout autant que la musique elle-même » (Rouget, 2004).

### 3 Les données musicales, une affaire d'écritures.

Si l'on se réfère aux musiques plus savantes ou commerciales, qui ont la particularité d'être écrites et éditées, nous avons affaire à des formes d'écriture qui encapsulent un certain nombre de « données ». Le constat selon lequel chacune des institutions partenaires du projet décrit différemment un même objet n'a rien d'étonnant en soi. Une telle analyse, comparant le traitement d'un même document audiovisuel dans les catalogues de l'Ina et de la BnF a déjà été menée (Després-Lonnet 2014). Le catalogage est une activité qui s'exerce en fonction d'objectifs différents, là patrimonial, ici pour une utilisation à des fins commerciales ou par des chercheurs. Chaque institution porte un regard situé sur les ressources qu'elle détient. Au-delà de cet aspect, il est important de remonter à des sources beaucoup plus lointaines, constitutives mêmes du jeu de « données » que l'on juge représentatif d'une œuvre musicale. Si l'on prend la musique classique, il n'est pas toujours acquis que le compositeur donne un titre à sa composition, laquelle est parfois titrée par l'éditeur de la partition.

On retrouve cette situation dans l'ensemble du corpus de la musique savante, et ce depuis toujours. Ainsi, pour des raisons pratiques et commerciales la Sonate pour piano n° 27 n°2 en ut dièse mineur de Beethoven s'est vu accoler le nom *Mondscheinsonate, Sonate au clair de lune*.

La musique savante se retrouve donc avec deux systèmes qui cohabitent ou se superposent : une description formelle (Sonate [pour] piano [en] ut dièse mineur [,] opus 27 n°2) et un titrage (*Mondscheinsonate*).

La musicologie décompose le libellé d'une œuvre en éléments tenant à la forme (sonate, concerto, cantate, messe...), à l'ensemble d'exécution (trio, quatuor...), au langage (tonal, sériel...) à la tonalité (ré mineur). Tous ces éléments font l'objet de traces écrites qui accompagnent l'écriture musicale proprement dite et qui en

constituent en quelque sorte le paratexte.

À cela s'ajoute un paratexte éditorial qui pourra faire figurer, évidemment, le nom du compositeur ou son rôle s'il s'agit d'arrangements ou de variations à partir d'une œuvre existante, une date et un lieu d'édition (différents de la date de composition), auquel il convient d'ajouter les mouvements et les tonalités qui forment souvent le plan de l'œuvre (les mouvements d'une sonate par exemple), ainsi que les éléments textuels portés par le compositeur à l'intention de l'interprète (indication de jeu, texte poétique ; Liszt n'hésite pas, par exemple, à reporter le texte poétique dans un lied transcrit au piano) ou les indications diverses qu'un chef d'orchestre peut noter et qui constitueront une sorte d'appareil critique.

L'édition phonographique « joue » d'un certain nombre de supports pour enrober l'œuvre musicale enregistrée dans un ensemble qui est tout à la fois informatif, juridique, commercial, et marketing : photos et visuels pour les pochettes et les livrets, paroles des chansons, mentions légales et crédits, biographies et photos des artistes, informations musicologiques et techniques (durée...)

Les documentalistes/discothécaires, qui ont pour tâche d'établir les catalogues d'institutions utilisent ces supports comme des adjuvants documentaires à l'objet qu'il s'agit de décrire, à savoir la composition musicale concernée. Cataloguer « disque en mains » c'est *prélever* dans un appareillage documentaire déjà constitué et savamment construit un certain nombre d'éléments qui répondent à un « modèle de données » cible qui a lui-même pour objectif de classer et retrouver la musique en fonction de différents usages.

Nous utilisons ici le terme « modèle de données » de façon volontairement anachronique et décalée, car ce n'est pas ainsi que s'exprimaient les discothécaires qui remplissaient, à la main ou à la machine à écrire, les fiches cartonnées destinées à constituer le catalogue. Pourtant, ces fiches étaient toujours structurées de la même manière pour recenser les éléments minimaux que l'on jugeait utile de consigner pour le classement et la recherche des œuvres musicales : titre de l'œuvre s'il est connu, nom du compositeur, éléments musicologiques, genre, cote, etc. L'informatique documentaire, en s'émancipant du cadre étroit matérialisé par l'espace du carton et en promettant des tris calculés, a suscité une inflation des « champs » de description, accroissant ainsi la capacité descriptive de l'outil et générant un supplément de « données » exploitables.

Avec le numérique, c'est l'objet éditorial lui-même qui s'évanouit dans la mesure où il se décompose en une série d'éléments qui ne sont pas forcément unifiés comme ils peuvent l'être dans la matérialité de l'objet physique édité : le coffret avec les disques porteurs des sons, le livret explicatif, les photos, les paroles et les informations musicologiques. Dans les dispositifs d'écoute en ligne, le fichier musical reste le plus souvent habillé d'éléments résiduels du support édité (visuel de la pochette réduit à la taille d'une vignette, nom du créateur « principal », date de production, d'enregistrement, ou de mise en ligne). Éléments sélectionnés et saisis de façon souvent erratique. En outre, le paratexte (paroles des œuvres chantées, notices, photos des interprètes...) est généralement totalement évacué. Ce que l'objet disque soli-

darisait se trouve dispersé sur la toile<sup>3</sup>. Pour retrouver ces différents éléments, l'internaute doit aujourd'hui naviguer d'un réseau social à un site de compilation de paroles de chansons, d'un site de label à un blog amateur/expert présentant la discographie d'un artiste etc.

Or, aussi bien l'industrie phonographique que le public restent attachés à la forme « album », c'est-à-dire à un objet éditorial complet qui regroupe des éléments sonores, textuels et visuels, informatifs et descriptifs dans un seul ensemble. Même l'industrie de la variété qui fonctionne à l'unité « chanson » ou « morceau » annonce la sortie « dans les bacs » de telle nouvelle création d'un artiste.

Ce n'est pas le lieu ici de retracer l'histoire matérielle de ces supports, mais on rappellera seulement que ce sont souvent des contraintes techniques qui se traduisent dans des formes matérielles données : la galette du microsillon appelle une pochette carrée pour ranger l'objet, la surface de la pochette appelle de l'information et de l'illustration et ainsi de suite. Il est tentant de penser que la musique dite « dématérialisée » transporte avec elle, en un seul flux, tous les éléments utiles à sa description, à sa recherche, et partant, à son écoute. Or, la situation n'est pas si simple pour deux raisons. La première, que nous avons déjà abordée est que ces flux sont loin d'être organisés dans un processus d'écriture contrôlée et donc de produire des contenus acceptables, « propres » au sens documentaire du terme et immédiatement versables dans des applications : les documentalistes/discothécaires ne font pas que décrire ; ils normalisent, ils enrichissent, ils unifient la description des morceaux de musique. Les noms d'artistes et de musiciens ne sont pas orthographiés de manière identique selon la langue de l'éditeur (Chostakovitch/Shostakovitch) ; les tonalités sont exprimées différemment en français et en allemand (do/A, ré/B...), et ainsi de suite.

La deuxième raison est que l'on confond ce travail de traitement pour la mise à disposition et la façon dont l'utilisateur final, réel, se manifeste vis-à-vis des outils de présentation et de recherche. Notre tâche dans le projet consiste effectivement à observer et réfléchir non pas aux seuls « usages » que l'on fait des outils mais aux *pratiques* au sein desquelles s'insère tel ou tel outil de recherche ou de recommandation. Il nous semble en effet important de considérer l'amateur (le consommateur ?) de musique comme un être multiple, immergé dans le monde et dont la posture n'est pas réduite à la station assise devant un écran ou à la manipulation d'un smartphone.

Le contexte social, temporel et spatial de la pratique n'est pas un surplus informationnel mais une dimension constitutive de l'expérience d'écoute musicale. Le « feuilletage » des supports dans un bac chez un disquaire, l'écoute à la radio, l'échange de disques ou de fichiers avec des amis ou dans le cercle familial, la réalisation de playlists pour une fête, la fréquentation de brocantes à la recherche de vinyles, la lecture de critiques ou de presse spécialisée, sont autant de scènes sociales dignes d'attention pour le chercheur, tout comme, bien sûr, la consultation de plates-formes de diffusion en ligne ou de réseaux sociaux, sur mobile, ordinateur

3. L'ambition de sites comme BBC Music (<http://www.bbc.co.uk/>) est de réunir, dans la logique du web de données, différents éléments sur un artiste sur une même application, en allant puiser des ressources sur le net (par exemple la biographie de l'artiste sur Wikipédia).



fixe ou tablette.

## 4 La logique du « déjà-là » dans le web sémantique

La plupart des grands projets connus dans le domaine de la culture (comme par exemple [data.bnf.fr](http://data.bnf.fr)) se fondent sur des jeux de données déjà existants, puisqu'il s'agit d'ouvrir via des formats interopérables des données catalographiques contenues dans des bases de données fermées et propriétaires.

Il ne paraît pas exagéré, face à l'appréciation du caractère quasiment « révolutionnaire » du web de données, de dire que celui-ci n'apporte finalement pas grand-chose d'autre qu'une manière de ré-agencer les données et, partant, de les mettre en valeur en multipliant les vues sur des jeux de données, et en créant de nouvelles interrelations porteuses de sens. Ce qui n'est déjà pas rien. Il convient ici de distinguer ce niveau, cette couche du système, qui repose sur des modèles et formats comme RDF, du niveau de captation ou de fabrication de la matière première, c'est-à-dire des données.

Nous relevons deux origines majeures pour les matériaux auxquels s'appliquent ces technologies.

### Une origine dans la masse des données déjà existantes

La première est celle des données déjà existantes, telles qu'on peut les trouver dans des environnements structurés : listes, tableurs, bases de données, et qui, soumises aux traitements du web de données, trouvent une nouvelle visibilité et dialoguent entre elles de manière plus riche grâce à des réagencements possibles. Pour prendre un exemple, les catalogues des grandes bibliothèques patrimoniales, dont les notices sont organisées selon le format MARC<sup>4</sup> et ses différentes déclinaisons, contiennent d'énormes quantités de données dont la plupart ne sont pas ou mal exploitées par les moteurs de recherche embarqués dans les catalogues. Par exemple, le nom d'une personne peut figurer dans le champ préface, mais il n'existe pas de relation réversible qui permette d'attribuer le rôle de « préfacier » à cet auteur. Dès lors que dans un modèle (dit « ontologie ») on a précisé qu'un auteur de préface est un préfacier, il est facile de ressortir, parmi les facettes d'un écrivain, plusieurs rôles, dont celui de préfacier<sup>5</sup> (ou traducteur, auteur de lettres...)

De ce fait, en ouvrant ainsi les catalogues, en les reliant entre eux, en y associant des éléments documentaires (textes, images, sons) avec les bibliothèques numériques, le web de données réalise une des promesses de l'internet et du web tout court, celle de constituer une *toile* dont le suivi des fils permet de se promener au sein d'un ensemble de connaissances inter-reliées.

4. L'acronyme pour MACHine Readable Catalogue, rend bien l'ambition informatique qui présidait à la première automatisation des catalogues de bibliothèque : normaliser pour permettre aux machines de "lire" les informations.

5. Il n'en demeure pas moins étrange de voir attribuer à François Villon le rôle de "parolier" de Georges Brassens (<http://data.bnf.fr/documents-by-rdt/11928414/510/page1>).

## Une origine dans les données produites automatiquement

Ce deuxième grand volume de matériaux est produit par la numérisation des signaux et la multiplication des capteurs qui permettent de produire de la donnée à partir d'appareils de mesure, fixes et surtout mobiles. Le fait que des outils conçus à l'origine pour des conversations vocales mobiles soient devenus de véritables usines à produire du signal et à le transformer en données est un bouleversement bien plus grand que ce qu'offre le « web de données ». Celui-ci, comme nous l'avons dit, suit un sillon dont la trace est entamée dès la première conception du web<sup>6</sup> ; en revanche l'hyperproduction de signaux est un phénomène en rupture avec tout ce que l'on pouvait connaître jusqu'à présent. Cette masse d'indicateurs produits se divise à son tour en plusieurs grandes catégories selon le mode de captation envisagé.

- Le premier sous-groupe concerne tout ce qui peut être capté automatiquement par des machines. Par exemple la transformation de signaux naturels en traces interprétables. Le sismographe, le baromètre, le thermomètre, transforment en signes décryptables selon des conventions des phénomènes naturels. Ils savent le faire sur une base analogique en utilisant des matériaux sensibles comme le mercure.

L'électronique et la micro-électronique ont évidemment démultiplié de manière colossale cette production automatique qui peut s'appliquer au contrôle d'infrastructures ou d'installations sensibles, à l'exploration astronomique, à la mesure de consommation d'énergie, et, nouveauté des nouveautés, à tout type de comportement humain dès lors que celui-ci est capté par un artefact numérique et s'exerce sur un réseau. La géolocalisation par les téléphones mobiles ou autres objets connectés, le monitoring du corps humain via des bracelets, des textiles dits « intelligents », voire des puces implantées dans l'organisme, permettent d'ajouter le vivant à la masse des domaines captés et surveillés.

- Ceci nous amène au deuxième sous-groupe, qui concerne la masse des informations émises volontairement par les internautes et stockées sur des conteneurs informatiques, que l'on peut là aussi séparer grossièrement en deux grandes catégories. Tout d'abord les données saisies dans des formulaires qui en précisent la nature aux yeux de l'internaute et à ceux des gestionnaires des sites (comme par exemple le nom ou l'adresse explicitement demandés dans le cadre d'une inscription sur un site ou d'un achat en ligne). Ces données forment ce que Fanny Georges appelle l'identité déclarative (Georges, 2009). La seconde catégorie est principalement collectée par l'entremise des réseaux sociaux. Tweets, photos, commentaires, évaluations et avis..., la liste est longue et connue.

C'est là que notre distinction entre le pur automatique et le volontaire mérite d'être affinée. En premier lieu, l'acte « volontaire » (ex. poster une photo sur un site de tourisme collaboratif) est un acte doublement encadré, techniquement et sémiotiquement. Techniquement parlant, la façon dont est capturée la trace, dont elle est stockée, traitée puis devient un objet de commerce échappe complètement à la

6. La continuité est aussi assurée dans la personne de ses promoteurs, puisque dans les deux cas, c'est l'omniprésente figure de Tim Berners-Lee qui constitue la référence.

connaissance moyenne de l'internaute et donc au contributeur de la trace. L'interprétation généralement quantitative de ces données vise à construire une identité « calculée » (Georges, 2009) des individus dont ils retrouvent une trace partielle sous la forme par exemple d'une performance relative dans l'activité associée (nombre d'amis, de followers, d'avis...).

D'un point de vue sémiotique, l'action se fait et se met en scène à partir d'archi-textes, c'est-à-dire d'outils d'écriture de l'écriture qui la guident et la dominent. Or, sous ce qui se présente comme des dispositifs d'écriture se cachent aussi (surtout) des dispositifs sophistiqués de captation de données. Nous avons employé plus haut le terme d'« informations » pour caractériser les matériaux que laissent sciemment les internautes sur les différentes formes de réseaux sociaux et sur le web en général. En effet, dire que l'on « aime » telle nouvelle, signaler un spectacle ou un article de journal, poster une photo d'un monument visité ne fait pas appel à la notion de « donnée » au sens statistique et informatique du terme. Finalement, dans le système de captation sophistiqué monté par l'industrie numérique, ce n'est pas tant le contenu informatif qui prime, mais la trace elle-même et ce qu'elle induit. Or, sur internet, l'écoute musicale laisse des traces, que ce soit sous la forme des *playlists*, des recommandations ou du renvoi aux amateurs de leurs propres recherches dans les moteurs des plates-formes qui permet de leur suggérer d'écouter tel ou tel artiste ou d'élargir dans telle direction le champ de leur goût musical.

La « donnée » n'émerge pas tant ici du contenu que du parcours. Ce qui est important c'est qu'à tel moment, Monsieur X ou Madame Y a consulté tel produit sur une boutique en ligne, a montré qu'il « y était » en postant une photo de voyage, a « liké » un profil sur un réseau social et ainsi de suite. C'est, non pas en marge, mais dans les à-côtés de cette activité communicationnelle que tout se joue. Il n'y a donc alors rien d'étonnant à ce que ce soit le modèle du marketing qui s'impose comme cadre de références exclusif pour toute opération d'échange de données, de partage d'informations et de consommation culturelle. Car la musique, puisque c'est le terrain de notre projet, est à la fois un acquis anthropologique, une pratique sociale, un bien culturel et une marchandise portée par une industrie qui cherche à renouveler ses codes et ses modalités de production. Que, sous ce dernier aspect, (à travers ses supports : le disque, le concert, le fichier...) elle fasse l'objet d'une promotion marketing, notamment sous la forme d'outils visant à la « recommandation » n'a rien d'étonnant. En revanche, notre questionnement, en tant que chercheurs en sciences de l'information et de la communication, consiste à se demander si le goût musical, l'influence, le partage, peuvent être si aisément interprétés par des algorithmes exploitant des jeux de données.

## 5 Conclusion et ouverture sur le projet

Les auditeurs professionnels, comme les critiques, produisent, dans le cadre de leur activité d'écriture, des traces de leur écoute, mais celles-ci ne sont qu'une forme, particulièrement saillante, de ce qui caractérise l'activité d'écoute en tant que telle (et la distingue de la simple perception). Dans ce cadre, nous interrogeons la place

laissée aujourd'hui aux internautes pour « habiter » les lieux de la musique en ligne.

Comment se manifestent les tactiques des « usagers-clients » pour produire, ensemble, ici et maintenant, le sens et les formes de leur rapport à la musique ? Les sites d'écoute en ligne offrent une vitrine et des contenus innombrables, mais ces nouveaux lieux peuvent-ils être considérés comme des espaces de pratiques ? (Després-Lonnet 2014) Sont-ils des lieux pratiqués, appropriés, et ne se réduisant pas à leur seule matérialité et aux stratégies des offreurs ?

Pour Peter Szendy (Szendy, 2001) l'écoute est à la fois adressée et signée. Ce n'est pas seulement une activité discrète, extérieure à l'œuvre musicale et qui ne laisse pas de trace. Peter Szendy traque les traces objectives de l'écoute qui permettent cette étrange faculté « d'écouter écouter ». . . Il s'attache principalement aux arrangements et transcriptions (de Liszt, Busoni, etc.) définies comme des formes objectives d'une écoute signée, remettant du même coup en question l'idée d'une œuvre « déjà là », d'une œuvre en soi.

Cette perspective est intéressante dans la mesure où aujourd'hui, l'espace de la réception est peuplé, envahi même, de traces innombrables : captations testimoniales au Smartphone de concerts, traces discursives sur des expériences d'écoute, playlists, favoris, définition de soi dans les profils des réseaux sociaux par le biais de choix musicaux, sons et vidéos recommandés, publiés sur Facebook ou tweetés, etc. Les sites d'écoute en ligne tentent, pour certains, de proposer des fonctionnalités interactionnelles et sociales. Comment ces fonctionnalités sont-elles investies par les internautes ? Comment les dimensions relationnelles et transactionnelles de l'acte de recommandation, de l'adresse (« tu dois écouter cela ! ») s'accommodent-elles (et accommodent-elles) les propositions et recommandations algorithmiques ?

## 6 Références

- Becker, H. (1988). *Les mondes de l'art*, Flammarion, Paris.
- Certeau, M. (de). (1990). Pratiques d'espace. In *L'invention du quotidien. 1. arts de faire*, Gallimard, Paris, 137-191.
- Cotte, D. (2011). *Émergences et transformations des formes médiatiques*. Hermès-Lavoisier, Paris.
- Davallon, J. (2004). Objet concret, objet scientifique, objet de recherche. *Hermès*, num. 38, 30-37.
- Debruyne, F. (2012). Le disquaire et ses usagers. Du magasin au site web. *Communication & langages*, num. 173, 49-65.
- Després-Lonnet, M. (2014). *Temps et Lieux de la documentation : transformation des contextes interprétatifs à l'ère d'internet*. ANRT, Lille.
- Georges F. (2009). Représentation de soi et identité numérique. Une approche sémiotique et quantitative de l'emprise culturelle du web 2.0. *Réseaux*, num. 154, 165-193.
- Hennion, A. (1993). *La passion musicale, une sociologie de la médiation*. Métailié,

Paris.

Jaujou, N. (2009). *Accords mineurs. De l'usage de catégories musicales*. Presses universitaires de Bordeaux, Pessac.

Latour B., (2001). *L'espoir de Pandore : pour une version réaliste de l'activité scientifique*, Paris, La Découverte.

Le Marec, J. Babou, I. (2003). De l'étude des usages à une théorie des composites : objets, relations et normes en bibliothèque, In *Lire, écrire, récrire : objets, signes et pratiques des médias informatisés*, Souchier, E., Le Marec, J., Jeanneret, Y. (Eds.). Éditions de la BPI/Centre Pompidou, Paris.

Merzeau, L. (2014). Le flâneur impatient. *Médium*, num. 41, 20-29.

Norman, A., (1999). Affordance, conventions and design. *Interactions*, vol. 6, num. 3, 38-43.

Noyer J.-M., Carmes M. (2013). L'irrésistible montée de l'algorithmique : méthodes et concepts en SHS <Sic\_00917623>

Rouget, G. (2004), L'efficacité musicale : musiquer pour survivre, *L'Homme*, num. 171-172, 27-52.

Szendy, P. (2001). *Écoute, une histoire de nos oreilles*. Éditions de Minuit, Paris.

Vandiedonck, D. (1999). *Qu'est-ce qui fait tourner le disque classique ?* PUS, Lille.