

Détection de périodes musicales d'une collection de musique par apprentissage

Rémy Kessler¹ Nicolas Béchet² Audrey Laplante¹ Dominic Forest¹
(1) Université de Montréal
C.P. 6128, succursale Centre-ville, Montréal H3C 3J7, Canada
IRISA, UMR 6074, INSA Rennes
{remy.kessler, audrey.laplante, dominic.forest}@umontreal.ca
nicolas.bechet@irisa.fr

Résumé. Dans ces travaux, nous présentons une approche afin d'étiqueter une large collection de chansons francophones. Nous avons développé une interface utilisant les paroles comme point d'entrée afin d'explorer cette collection de musique avec des filtres en fonction de chaque période musicale. Dans un premier temps, nous avons collecté paroles et métadonnées de différentes sources sur le Web. Nous présentons dans cet article une méthode originale permettant d'attribuer de manière automatique la décennie de sortie des chansons de notre collection. Basée sur un système évalué au cours d'une des campagnes DEFT, l'approche combine fouille de textes et apprentissage supervisé et aborde la problématique comme une tâche de classification multi classes. Nous avons par la suite enrichi le modèle d'un certain nombre de traits supplémentaires tels que les tags sociaux afin d'observer leur influence sur les résultats.

Abstract. In this paper, we present an approach to label a large collection of songs in French by decade. We have developed an information visualization interface that allows users to browse the collection searching for lyrics with musical periods dependent filters. We first harvested lyrics and metadata from various sources on the web. We present in this article an original method to automatically assign the decade of songs for our collection. The original system was developed for a DEFT challenge and combined text mining and machine learning with a multi class approach. We subsequently enriched the model with additional features such as social tags to determine their impact on the results.

Mots-clés : Fouille de textes, apprentissage supervisé, paroles de chansons, tags sociaux.

Keywords: text mining, machine learning, lyrics, social tagging.

1 Introduction

Considérant la grande quantité d'informations textuelles sur la musique pouvant être extraite du Web, ce qui inclut métadonnées diverses et paroles de chanson, de nouvelles avenues pour l'étude de la chanson populaire s'ouvrent aux chercheurs. Au moyen de techniques de fouille de textes, il devient désormais possible de développer des systèmes permettant aux chercheurs d'étendre leur étude des paroles de chansons à de larges corpus. Le projet IMAGE¹ propose d'utiliser les paroles comme point d'entrée afin d'explorer une collection de musique. Dans nos précédents travaux (Kessler *et al.*, 2014a), nous présentions une approche de fouille de textes ainsi qu'une interface permettant la recherche et l'exploration d'une large collection de chansons francophones sur la base des thématiques abordées dans leurs paroles. Les applications de notre démarche sont nombreuses et variées dans le contexte du Web et de la diffusion de contenu multimédia (analyse de la proximité entre les artistes sur la base des thématiques abordées dans leurs chansons, analyse diachronique, etc.) ainsi qu'auprès des sociologues, musicologues qui étudient la musique populaire. Notre corpus incluant des métadonnées riches sur les chansons qui le composent, nous avons développé des fonctionnalités supplémentaires. Nous avons associé à chaque chanson une période musicale (années 80, 90, etc.) afin de visualiser les chansons en fonction de chaque période. Au cours de ce processus, nous avons dû écarter une grande quantité de chansons de notre corpus initial dans la mesure où nous ne disposions pas d'information concernant leur date de sortie. Afin de compléter les informations contenues dans la collection et d'éviter une tâche fastidieuse d'étiquetage manuel, nous avons décidé d'exploiter le reste de la collection ainsi que les métadonnées dans le but d'étiqueter automatiquement la collection. Après une présentation des travaux réalisés par d'autres chercheurs dans ce domaine, nous présenterons succinctement la méthode de collecte du corpus ainsi que des statistiques. Nous aborderons ensuite les traitements appliqués aux paroles et métadonnées avant de développer notre approche pour détecter automatiquement la décennie de chaque chanson.

¹Indexation de Musique À Grande Échelle

2 Travaux connexes

Étant donné l'abondance d'informations musicales disponibles en accès libre sur le Web, il n'est pas surprenant de constater qu'un grand nombre de chercheurs aient développé des outils pour collecter et exploiter ces informations dans le repérage de la musique (voir Li *et al.* (2012) pour une revue exhaustive de la littérature sur ce sujet). Ainsi, McKay *et al.* (2010) utilisent les paroles de chansons en combinaison avec d'autres données telles que descripteurs acoustiques et symboliques, contenu culturel tiré du Web afin d'améliorer la classification automatique par genre. Même si l'approche sous forme de sacs de mots reste la plus répandue pour classer les chansons par genre, Mayer *et al.* (2008) utilisent des caractéristiques additionnelles telles que les rimes ou les catégories grammaticales. Des chercheurs ont également démontré la pertinence d'utiliser les paroles pour réaliser d'autres tâches, notamment pour la détection d'émotions (Hu *et al.*, 2009; Van Zaanen & Kanters, 2010), ou encore dans le but d'établir le degré de similarité entre artistes à l'aide d'une analyse sémantique des paroles (Logan *et al.*, 2004). Müller *et al.* (2007) proposent un moteur de recherche sur les paroles de chanson permettant d'accéder directement à la partie de l'enregistrement audio correspondant à sa requête au moyen d'un alignement automatique entre paroles et bande sonore. L'utilisation des tags sociaux² dans le cadre des données musicales est devenue pratique courante afin d'améliorer la description des ressources en ligne. En effet, ils offrent un grand volume d'informations textuelles qui va bien au-delà des métadonnées standards. Levy & Sandler (2008) montrent leur efficacité dans un système de recherche tout en soulignant les problèmes de polysémie³ ou encore de synonymie⁴. De la même façon, Laurier *et al.* (2008) les utilisent afin de constituer une référence pour évaluer leur système tandis que (Trant, 2009) décrit une méthodologie permettant la construction automatique d'une taxonomie hiérarchique en fonction de leurs fréquences. Même si les tags sociaux ont des caractéristiques particulières, nous pensons qu'ils sont une source de données d'apprentissage pertinente. À notre connaissance, il n'existe pas de travaux visant à détecter la période de création de chansons ou d'autres pièces musicales. Néanmoins, l'analyse des informations temporelles est souvent une composante essentielle dans la compréhension de textes et est utile dans un grand nombre d'applications de recherche et organisation d'informations comme le souligne (Alonso, 2008; Kanhabua, 2009). La tâche de variation diachronique a ainsi fait l'objet de deux campagnes d'évaluation DEFT⁵. Différentes approches statistiques et symboliques, en combinaison avec des techniques d'apprentissage, ont été développées par les participants afin de détecter la période de textes journalistiques. Génereux (2010) a ainsi pu mettre en évidence des variations de fréquences ou des mots saillants pour chaque décennie tandis que Raymond & Claveau (2011) ont pris en compte les variations de la ponctuation au cours du temps. Au niveau linguistique, une belle étude de réformes orthographiques a été effectuée par Albert *et al.* (2010), permettant aux auteurs de concevoir des filtres pour chaque période spécifique. Des ressources externes ont aussi été utilisées, telles que des entités d'événements nommés (Monceaux & Tartier, 2010) ou encore une base de données avec la date de naissance de personnages célèbres (García-Fernandez *et al.*, 2011). Même si la sémantique de la musique reste difficile à détecter comme le montre les travaux précédents, nous pensons que les tags sociaux et les paroles de chansons, au travers des styles ou genres musicaux, fournissent des descripteurs pertinents.

3 Constitution d'un corpus de chansons

Dans le domaine musical, le jeu de données le plus imposant est le *Million Song Dataset* (MSD) (Bertin-Mahieux *et al.*, 2011) qui, comme son nom l'indique, contient plus d'un million de chansons. Il est accompagné des paroles de plus de deux cent mille chansons représentées sous forme de matrice sac de mots, le *MusiXmatch Dataset*, lequel n'inclut malheureusement qu'une faible portion de chansons en français. Nous avons donc choisi de constituer notre jeu de données. À l'aide d'une liste de pays francophones, nous avons utilisé les API de différents entrepôts de données musicales (*LyricWiki*, *EchoNest*, *LastFM*, *Paroles.net* et *musiXmatch*,) afin de collecter un grand nombre de données sur chaque chanson. Comme il s'est avéré difficile de trouver des informations fiables sur la langue des chansons, nous avons comparé les paroles collectées à des antidiCTIONNAIRES dans plusieurs langues (anglais, français, etc.) afin de détecter automatiquement la langue de chaque chanson. Le processus de collecte d'informations est détaillé dans (Kessler *et al.*, 2014b). Cette méthode nous a permis de collecter les paroles de chanson ainsi que de riches métadonnées, incluant les tags sociaux de Last.fm et diverses données bibliographiques (p. ex. : nom du parolier, du compositeur) couvrant une période de 1950 à nos jours. Comme il est courant pour les artistes de sortir plusieurs versions d'un même titre (single, compilations, live), nous avons dû retirer beaucoup de doublons de notre collection. Pour ces expériences, nous avons extrait un sous-ensemble de 5000

²Les tags sociaux sont les mots clés libres que les utilisateurs d'un service attribuent aux ressources (p. ex. : chanson, album ou artiste), notamment dans le but de les repérer par la suite.

³le tag « love » peut correspondre à une chanson favorite aussi bien qu'une chanson parlant d'amour

⁴la désignation de musique électronique peut être retrouvée au travers de différents termes tels que « musique électro », « musique électronique » ou encore « électro » tout court

⁵Défi Fouille de Textes <http://deft.limsi.fr/>

chansons de cette collection. Nous n'avons retenu que des chansons en français pour lesquelles nous avons une date de sortie dans les métadonnées. Nous avons ensuite procédé à un échantillonnage stratifié visant une répartition homogène des chansons par période. Le tableau 1 présente quelques statistiques descriptives de ce sous-ensemble.

	<i>avant prétraitements linguistiques</i>	<i>après prétraitements linguistiques</i>
Nombre total de mots	1 166 445	506 082
Nombre total de tags sociaux différents	4 417	1 363
Moyenne de mots par chanson	225,68	100,22
Moyenne de tags sociaux par chanson	1,92	1,73

TAB. 1 – statistiques du sous-ensemble de la collection.

4 Méthodologie

4.1 Filtrages et prétraitements linguistiques

Certains différents prétraitements linguistiques sont effectués : conversion des majuscules en minuscules, retrait des chiffres et des nombres (numériques et/ou textuels), des accents et des symboles (par exemple « \$ », « # », « * »). Afin d'éviter l'introduction de bruit dans le modèle, nous utilisons un antidictionnaire composé de verbes auxiliaires et d'autres mots fonctionnels couramment utilisés comme marqueurs discursifs ou comme conjonctions de coordination et/ou de subordination. Nous avons par ailleurs enrichi l'antidictionnaire de termes extraits de la langue populaire du Québec et de France que l'on retrouve fréquemment dans les chansons (p. ex. : « té » (tu es), « chu » (je suis), « c'te » (cette)). Nous appliquons par la suite un processus de lemmatisation simple⁶ afin de réduire considérablement les dimensions de l'espace tout en augmentant la fréquence des termes canoniques.

4.2 Une approche par Boosting

Nous avons utilisé l'algorithme *AdaBoost* qui combine les votes pondérés d'une multitude d'apprenants faibles. L'algorithme a été implémenté dans l'outil de classification à large-marge *BoosTexter* (Schapire & Singer, 2000). Il permet de fournir au classifieur des données numériques, mais également des données textuelles sous forme de n grammes de mots. Nous avons utilisé l'outil *IcsiBoost* (Favre *et al.*, 2007), qui est l'implémentation open-source de cet outil et qui présente l'avantage de pouvoir fournir un score de confiance pour chaque exemple à classifier. Le système de base pour ces expérimentations a été évalué lors de la campagne DEFT 2010 sur la tâche de variation diachronique comme un des systèmes initiaux (Rk_icsiboost) de (Oger *et al.*, 2010) et sera présenté comme baseline (appelé par la suite *SB*) dans la section 5. Les paramètres d'entrée du système sont les entrées lexicales de chaque document, les paroles des chansons ainsi que le titre avec une représentation sous forme de n grammes de mots dont nous avons fait varier la taille entre 1 et 5. Nous avons par la suite enrichi le modèle d'un certain nombre de traits supplémentaires afin d'observer leur influence sur les résultats. Nous présentons ci-dessous l'ensemble de ces traits regroupés en fonction de leur type :

- Traits *EchoNest* : Il s'agit d'un ensemble de valeurs numériques pour chaque chanson transmise par le site *EchoNest*. Différentes caractéristiques pour chaque chanson sont ainsi ajoutées telles que « danceability », « Hotness », « loudness » ou encore tempo.
- Collaborateurs : Chaque chanson étant associée à des collaborateurs (parolier, compositeur, arrangeur, musiciens, etc.), ceux-ci ont été regroupés dans un champ textuel unique. Nous avons cependant fixé la taille des n grammes à 3 afin de conserver l'association entre le type de collaboration et le nom du collaborateur (par exemple, « parolier Martial Tricoche »).
- Chansons et artistes similaires : Nous avons regroupé dans ces traits l'ensemble des chansons et artistes similaires associés à chaque chanson selon les informations issues des sites *LastFM* et *EchoNest*.
- Tags sociaux : Les tags sociaux associés à chaque chanson ont été regroupés en un champ textuel unique avec une taille de n grammes de 1. Les prétraitements classiques sont appliqués (antidictionnaire et lemmatisation) en français ainsi qu'en anglais étant donné que plusieurs utilisateurs du site Last.fm attribuent des tags dans cette langue, même pour les chansons francophones.

Nous présenterons en section 5 les différents résultats obtenus pour chacun des types de traits ainsi que pour leurs combinaisons.

⁶La lemmatisation simple consiste à trouver la racine des verbes fléchis et à ramener les mots pluriels et/ou féminins au masculins singulier.

4.2.1 Utilisation de motifs séquentiels

Toujours dans l'idée de détecter des descripteurs additionnels sur la base des informations textuelles à notre disposition, nous avons eu recours à l'utilisation des motifs séquentiels. Nous introduisons dans un premier temps ces derniers avant de présenter la méthode mise en œuvre afin d'exploiter les motifs séquentiels avec des données textuelles.

Définitions. Introduite par (Agrawal & Srikant, 1995), l'extraction de motifs séquentiels est un champ de la fouille de données visant à extraire des régularités dans un jeu de données se présentant sous la forme de séquences. Ce jeu de données est appelé *base de données séquentielles (SDB)*. Les auteurs ont introduit la notion de *séquence* notée $s = \langle I_1 \dots I_m \rangle$, comme étant une liste ordonnée d'*itemsets*. Un itemset, noté $I = (i_1 \dots i_n)$ est un ensemble de littéraux appelés *items*. Ainsi, la séquence $\langle (a) (a b c) (a c) (d) \rangle$ est ici composée de quatre itemsets, dont le deuxième itemset comporte trois items a , b et c . Une séquence $S_1 = \langle I_1 \dots I_n \rangle$ est *incluse* dans une séquence $S_2 = \langle I'_1 \dots I'_m \rangle$ si il existe des entiers $1 \leq j_1 < \dots < j_n \leq m$ tel que $I_1 \subseteq I'_{j_1}, \dots, I_n \subseteq I'_{j_n}$. On dit alors que S_2 *supporte* S_1 et que S_1 est une *sous-séquence* de S_2 . Par exemple, $\langle (a)(a c) \rangle \subseteq \langle (a)(a b c)(a c)(d) \rangle$. Le *support* d'un motif séquentiel P (notée $sup(P)$) est le nombre de séquences de la *SDB* supportant P divisé par le nombre de séquences de la *SDB*. Finalement, un motif séquentiel fréquent est dit *maximal* si il n'existe aucun motif pouvant être inclus dans ce dernier.

Utilisation des motifs pour les textes des chansons. Dans notre cas, chaque chanson est considérée comme une *SDB*. Ainsi, chaque vers représente une séquence. Les itemsets constituant les séquences sont alors construits à partir des mots de chaque vers de la manière suivante. Pour chaque mot, un lemme et une catégorie grammaticale sont extraits afin de former un itemset⁷. Considérons par exemple le vers suivant extrait d'une chanson : "*Sauver une âme*". La séquence résultante sera alors : $\langle (Sauver VERBE) (un PRONOM) (âme NOM) \rangle$.

Afin d'obtenir de nouveaux descripteurs caractérisant les chansons nous procédons alors comme suit : 1) construire les *SDB* pour chaque chanson, 2) extraire les motifs séquentiels maximaux à partir des *SDB*, 3) conserver uniquement les k motifs les plus fréquents afin de représenter chaque chanson. Finalement, une mesure de similarité est en cours de réalisation afin de mesurer la proximité entre deux motifs séquentiels ce qui nous permettra de calculer des médoïdes représentant les différentes classes. Toute nouvelle chanson, une fois les motifs séquentiels la caractérisant extraits, pourra alors être comparée à ces médoïdes afin de lui attribuer automatiquement une classe.

5 Expériences

5.1 Protocole expérimental

Afin d'entraîner les classifieurs, une étiquette a été associée à chaque chanson en fonction de sa période musicale (années 80, 90, etc.). Compte tenu du peu de documents pour ces périodes dans le corpus, les chansons des années 50, 60 et 70 ont été ramenées à une étiquette unique, années 50-70. Nous avons évalué cette approche avec une validation croisée classique sur 5 partitions du corpus. Chaque chanson est ainsi répartie aléatoirement dans une des partitions, avec un nombre équilibré de chansons pour chaque période. Nous utilisons 4 sous-corpus en tant que données d'apprentissage (trois sont utilisés en apprentissage tandis que le dernier permet de calibrer le classifieur) et le cinquième corpus est réservé au test : 5 jeux de tests tournants sont ainsi créés. Chacune des expériences a été évaluée sur les corpus de test en utilisant les mesures classiques de Précision, Rappel et F-score des documents bien classés, moyennées sur toutes les classes (avec $\beta = 1$ afin de ne privilégier ni la précision ni le rappel (Goutte & Gaussier, 2005)). Le score final d'un test est ensuite calculé en faisant la macro moyenne de tous les scores obtenus sur chacun des ensembles de test. À titre de comparaison, nous avons inclus la mesure d'évaluation développée par (Grouin *et al.*, 2011) permettant d'évaluer le résultat obtenu sur la base d'une fenêtre de 15 ans autour de l'année de référence.

5.2 Résultats

En observant les résultats du tableau 2, on constate que le système classique (SB) obtient un score de 0,38 en Précision et un F-score de 0,54, ce qui est globalement faible, mais comparable aux résultats de la campagne DEFT, même si le nombre de classes était plus important. L'utilisation d'un antidiCTIONNAIRE et d'une étape de lemmatisation apporte peu individuellement, mais permet d'améliorer les résultats des autres expériences. Concernant les traits EchoNest, ceux-ci apportent peu de gain, mais nous expliquons cela par le nombre conséquent de documents pour lesquels nous n'avons pas ces données. L'apport des autres traits (COL, ACS) reste aussi faible et peu significatif même si l'on observe que les performances sont conservées lors des combinaisons.

L'étude détaillée des résultats montre qu'en dehors des paroles de chansons, les tags sociaux sont le premier trait pris en compte lors de la prise de décision suivi par les motifs ainsi que le nom des collaborateurs. Néanmoins, une observation

⁷Les lemmes et catégories grammaticales sont extraits en utilisant l'outil Treetagger

	Précision	Rappel	F-score	Mesure DEFT 2011
Sans prétraitements linguistiques				
Système de base (SB)	0,383	0,974	0,550	0,279
SB+traits EchoNest (TE)	0,388	0,954	0,551	0,280
SB+Collaborateurs (COL)	0,387	0,961	0,551	0,277
SB+Tags sociaux (TS)	0,518	0,759	0,613	0,286
SB+Artistes/Chansons similaires (ACS)	0,383	0,974	0,550	0,279
SB+TE+TS	0,533	0,752	0,620	0,280
SB+TE+TS+COL	0,511	0,773	0,615	0,280
SB+TE+TS+COL+ACS	0,511	0,773	0,615	0,281
Avec prétraitements linguistiques				
Système de base (SB)	0,383	0,969	0,549	0,278
SB+traits EchoNest (TE)	0,388	0,971	0,554	0,282
SB+Collaborateurs (COL)	0,388	0,959	0,553	0,281
SB+Tags sociaux (TS)	0,541	0,758	0,629	0,284
SB+Artistes/Chansons similaires (ACS)	0,383	0,969	0,549	0,278
SB+TE+TS	0,533	0,752	0,620	0,280
SB+TE+TS+COL	0,404	0,899	0,556	0,277
SB+TE+TS+COL+ACS	0,545	0,754	0,630	0,290

TAB. 2 – Ensemble des résultats obtenus par chaque exécution du système.

plus approfondie des erreurs commises montre que l'utilisation de certains tags sociaux comme « français » ou encore « live » influence de façon négative le système. De même, la fréquence importante du tag social « star » pour de nombreux artistes sur des périodes différentes semble conduire le système à commettre des erreurs lors de l'attribution de la décennie. Même si certains tags à caractère personnel comme « favoritos » apportent peu d'information, d'autres tels que « party 2010 » ou encore « seen in 2013 » permettent au système de déterminer la bonne période. En outre, l'utilisation de tags tels que « Disco francophone », « 1960s », « french celtic rap », ou encore les noms d'artistes permettent de réduire de façon significative les erreurs commises par le système. Différentes expériences complémentaires sont en cours afin d'expliquer et améliorer ces résultats. Les résultats obtenus à l'aide de la mesure d'évaluation de la campagne 2011 confirment ces tendances même si les tâches ne sont pas identiques (détection de décennie versus année exacte). On notera cependant les scores relativement bas ainsi que la faible variation entre les différentes exécutions du système.

6 Conclusion et perspectives

Dans cet article, nous avons présenté une approche combinant fouille de texte et apprentissage afin de retrouver les périodes musicales manquantes d'une large collection. Plus spécifiquement, l'approche repose d'abord sur l'utilisation d'algorithmes de fouille de textes ainsi que sur des techniques d'apprentissage machine. Notre collection est représentée sous forme de sacs de mots complétés par des traits supplémentaires avant de les classer à l'aide d'un algorithme de boosting. Les résultats obtenus montrent que l'utilisation des paroles et des tags sociaux sont une source intéressante afin d'identifier les liens entre les différentes périodes musicales fournissant une base d'apprentissage non négligeable. Finalement, les premiers tests basés sur les motifs séquentiels présentent des résultats encourageants, mais nécessitent des expériences complémentaires. Nous prévoyons de continuer à augmenter le corpus afin de confirmer ces hypothèses et envisageons d'autres expériences comme la détection de parolier ou encore de styles musicaux propres à chaque artiste.

Remerciements

Les auteurs tiennent à remercier le Conseil de Recherches en Sciences Humaines (CRSH), qui ont partiellement financé ces travaux, le LIA pour la mise à disposition de ressources, Pierre-François Marteau pour sa collaboration ainsi que Mélanie Couderc pour ses relectures attentives.

Références

AGRAWAL R. & SRIKANT R. (1995). Mining sequential patterns. In *Int. Conf. on Data Engineering* : IEEE.

- ALBERT P., BADIN F., DELORME M., DEVOS N., PAPAZOGLU S. & SIMARD J. (2010). Décennie d'un article de journal par analyse statistique et lexicale. In *Actes DEFT 2010*, p. 85–97, Montréal, Canada.
- ALONSO O. R. (2008). *Temporal information retrieval*. PhD thesis, University of California at Davis, Davis, CA, USA. Adviser-Gertz, Michael.
- BERTIN-MAHIEUX T., ELLIS D., WHITMAN B. & LAMERE P. (2011). The million Song Dataset. In *ISMIR 2011, Miami, Florida*, p. 591–596.
- FAVRE B., HAKKANI-TÜR D. & CUENDET S. (2007). Icsiboost.
- GARCÍA-FERNANDEZ A., LIGOZAT A.-L., DINARELLI M. & BERNHARD D. (2011). Méthodes pour l'archéologie linguistique : datation par combinaison d'indices temporels. In *Actes DEFT 2011*.
- GÉNÉREUX M. (2010). Classification de textes en comparant les fréquences lexicales. In *Actes DEFT 2010*, p. 57–68, Montréal, Canada.
- GOUTTE C. & GAUSSIER E. (2005). A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. *ECIR 2005*, p. 345–359.
- GROUIN C., FOREST D., PAROUBEK P. & ZWEIGENBAUM P. (2011). Présentation et résultats du défi fouille de texte DEFT2011 quand un article de presse a-t-il été écrit ? *Actes du septième Défi Fouille de Textes*, p.3.
- HU X., DOWNIE J. S. & EHMANN A. F. (2009). Lyric Text Mining in Music Mood Classification. (5,049), 411–416.
- KANHABUA N. (2009). Exploiting temporal information in retrieval of archived documents. In *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2009, Boston, MA, USA, July 19-23, 2009*, p. 848.
- KESSLER R., LAPLANTE A. & FOREST D. (2014a). Encore des mots, toujours des mots : fouille de textes et visualisation de l'information pour l'exploration et l'analyse d'une collection de chansons en français. In *Journées d'Analyse statistique des Données Textuelles (JADT 2014), Paris, (à paraître)*.
- KESSLER R., LAPLANTE A. & FOREST D. (2014b). Exploration d'une collection de chansons à partir d'une interface de visualisation basée sur une analyse des paroles. volume EGC, RNTI-E-26, p. 112–123.
- LAURIER C., GRIVOLLA J. & HERRERA P. (2008). Multimodal Music Mood Classification Using Audio and Lyrics. In *Machine Learning and Applications, 2008 ICMLA'08*, p. 688–693.
- LEVY M. & SANDLER M. (2008). Learning Latent Semantic Models for Music from Social Tags. *Journal of New Music Research*, (2), 137–150.
- LI T., MITSUNORI O. & TZANETAKIS G. (2012). *Music Data Mining*, volume 21. CRC Press Llc.
- LOGAN B., KOSITSKY A. & MORENO P. (2004). Semantic Analysis of Song Lyrics. In *ICME'04, IEEE International Conference*, p. 827–830.
- MAYER R., NEUMAYER R. & RAUBER A. (2008). Rhyme and Style Features for Musical Genre Classification by Song Lyrics. In *ISMIR 2008*, p. 337–342.
- MCKAY C., BURGOYNE J. A., HOCKMAN J., SMITH J., VIGLIENSONI G. & FUJINAGA I. (2010). Evaluating the Genre Classification Performance of Lyrical Features relative to Audio, Symbolic and Cultural Features. In *ISMIR 2010*, volume 10, p. 213–218.
- MONCEAUX L. & TARTIER A. (2010). Utilisation d'outils linguistiques pour trouver la date ou l'origine d'un fragment textuel. In *Actes DEFT 2010*, p. 21–34, Montréal, Canada.
- MÜLLER M., KURTH F., DAMM D., FREMEREY C. & CLAUSEN M. (2007). Lyrics-Based Audio Retrieval and Multimodal Navigation in Music Collections. In *Research and Advanced Technology for Digital Libraries*, p. 112–123.
- OGER S., ROUVIER M., CAMELIN N., KESSLER R., LEFÈVRE F. & TORRES-MORENO J. M. (2010). Système du LIA pour la campagne DEFT'10 : datation et localisation d'articles de presse francophones. In *Actes DEFT 2010*, p. 69–83, Montréal, Canada.
- RAYMOND C. & CLAVEAU V. (2011). Participation de l'IRISA à DEFT2011 : expériences avec des approches d'apprentissage supervisé et non supervisé. In *Actes DEFT 2011*.
- SCHAPIRE R. E. & SINGER Y. (2000). BoosTexter : A boosting-based system for text categorization. *Machine Learning*, p. 135–168.
- TRANT J. (2009). Studying social tagging and folksonomy : A review and framework.
- VAN ZAAANEN M. & KANTERS P. (2010). Automatic Mood Classification using tf* idf based on Lyrics. In *Proceedings of the 11th Int. Conf. on Music Information Retrieval*, p. 75–80.