

## **Réseaux sociaux et indexation de la musique, vers de nouveaux modes d'évaluation par l'usage**

**Résumé :** La communication décrit un projet de recherche en cours soutenu par l'ISCC. Le projet se propose d'examiner les problèmes liés à la production sociale des métadonnées de la musique eu égard aux nouveaux dispositifs et nouvelles pratiques numériques l'indexation et les métadonnées de la musique. Sont décrits le contexte et la motivation de l'étude, un bref rappel de concepts relevant de l'étude. La communication se penche par ailleurs sur l'évaluation de la production musicale et sa réception dans le cadre des réseaux sociaux. Le questionnaire ainsi que les premiers résultats sont décrits.

**Abstract:** This paper describes an ongoing study supported by ISCC. The project examines issues surrounding social metadata creation for music in light of new devices and digital practices, indexing, and music metadata. The context and rationale for the study are described, as well as a summary of relevant concepts. The evaluation of music production and its reception in social networks are also covered. The questionnaire and the preliminary results will be presented.

### **1. Introduction et cadre de l'étude**

Dans cette communication nous présentons les travaux en cours du projet ICOMUSIC qui s'inscrit dans le cadre de projets exploratoires soutenus par l'ISCC (Institut des Sciences de la Communication du CNRS), 2009-2010.

Le projet se propose d'examiner les problèmes liés à la production sociale des métadonnées de la musique eu égard aux nouveaux dispositifs et nouvelles pratiques numériques. Notre terrain d'étude étant l'indexation des œuvres musicales, nous y identifions les dispositifs (lieux et formes de production) et les pratiques afin de statuer des modes de descriptions et d'indexations. Il convient de procéder à une évaluation de l'existant (forme et modalité de balisage, etc,...) afin d'analyser l'usage, la réception et l'appropriation de cette indexation. Le consortium a une longue expertise dans le domaine de l'évaluation de l'usage et mettra son expérience à profit.

Dans le cadre de ce projet le travail, que nous menons, a consisté dans un premier temps, à faire un aperçu historique et technique de la complexité à discrétiser un objet sonore, d'un point de vue quantitatif et qualitatif. Nous avons ensuite étudié les métadonnées afin de comparer celles relevant d'une indexation centralisée et « contrôlée » par rapport à celles propres à l'indexation sociale. Nous étudions également les pratiques informationnelles émergentes liées à cette nouvelle forme d'indexation. Cela permet d'apprécier dans quelle mesure il est pertinent de numériser le fait musical grâce aux techniques émergentes d'indexation sociale. Dans cette partie nous faisons actuellement une évaluation des modes d'indexation de sites (maison de disques, communautés virtuelles musicales, webradios, etc.....) qui pratiquent ce type d'indexation. Notre travail ne pourra pas ignorer l'existant (indexation de l'information textuelle, information olfactive, ). Nous sommes actuellement en train de rassembler les informations relatives aux formats de stockage de la musique, tel que le MP3, qui enrichissent la numérisation de l'œuvre musicale avec des métadonnées. Dans cette partie il sera nécessaire de faire une évaluation des modes d'indexation de sites (communautés virtuelles musicales,

webradios, blogs, sites amateurs, etc.,...) qui pratiquent ce type d'indexation.

## **2. Métadonnées de la musique et Web social .**

Une nouvelle génération de sites associés au concept Web 2 .0. a permis l'apparition de nouvelles pratiques et usages basés sur la collaboration et la participation. Face à la forte demande les éditeurs et producteurs de contenus offrent de nombreux outils d'aide à la navigation et au tagging collaboratif. Depuis l'intégration du premier système de tagging collaboratif sur le site Del.icio.us en 2003, le tagging a connu une évolution rapide et s'est progressivement diffusé à d'autres sites de partage de contenus tels que Flickr en 2004 et Youtube et Dailymotion en 2005 (Crepel, 2008).

Le mode d'indexation pratiqué dans ces sites est une classification collaborative produite sur la base de l'indexation libre des tags appelé communément « folksonomie ». Celles-ci désignent un processus de classification collaborative par des mots-clés librement choisis. Ce mode d'indexation dite sociale, par opposition à l'indexation normée (métadonnées normées, listes d'autorité, thesaurus,..) trouve ses origines dans des pratiques classiques (organisation et classification des connaissances) mais aussi dans des pratiques métier d'ordre techniciste (conservation et accès aux documents, métadonnées, Dublin Core). Il constitue une des bases de développements actuels sur un web sans cesse en évolution (web 2.0, ontologies sociales, web sémantique et socio-sémantique, ..). Ces pratiques disposent aujourd'hui d'outils et d'interfaces permettant un « balisage » entièrement subjectivé (par opposition à des balises normées, émanant d'une autorité et ayant été validées).

Ce mode d'indexation a fait l'objet d'évaluations et d'appréciations entreprises par de nombreux chercheurs. Les folksonomies et d'autres techniques d'indexation sociale seraient pérennes d'après Galzot (Galzot et al. 2006) pour trois raisons dont nous relevons ici les plus importantes :

- Raison économique : les outils de recherche, et particulièrement les trois grands, développent une offre de portails personnalisés dans laquelle sont intégrées de manière plus ou moins naturelle et intuitive, la gestion de ces communautés d'intérêt et la possibilité de soumettre tout type de document
- Des raisons techniques : les interfaces souvent très intuitives accessibles à tous en ajoutant une dimension de partage, de capitalisation et de diffusion en temps réel.
- Une autre raison pour une éventuelle pérennisation de ce type de classification est le faible coût cognitif. Les auteurs expliquent qu'en raison de la complexité de l'activité de catégorisation ou d'indexation dans un contexte documentaire professionnel nécessitent l'usage de métadonnées normalisés alors qu'une indexation et catégorisation dites sociales ne nécessitent qu'un faible coût cognitif.

Nous nous proposons dans cette communication de faire état de l'évaluation que nous avons menée afin de tester nos hypothèses quant à la usage, l'appropriation des réseaux sociaux sites amateurs et tout autre dispositif utilisé par les usagers qui sont les producteurs (auto-produit) , les amateurs, les blogs, etc. ..

## **3. Évaluation de la production musicale et sa réception dans le cadre des réseaux sociaux**

### **3.1. Le paradigme évaluation**

L'évaluation est définie comme étant une relation qui vise à déterminer un indice de satisfaction exprimé par un usager d'un système d'information. Si l'évaluation est un

exercice d'appréciation des performances d'un système assorti d'un indice de satisfaction, cet indice ne peut être calculé que si l'observation est effectuée par rapport à des besoins et des référentiels de type intrinsèque ou extrinsèque. La diversité des besoins et des référentiels génère plusieurs types et degrés d'évaluation. En ce qui concerne l'évaluation de l'indexation pratiquée par les indexeurs de sites dédiées à l'information musicales (communautés virtuelles musicales, webradios, blogs, sites amateurs, etc....) nous retenons deux types d'évaluation : une évaluation qualitative et une évaluation quantitative

L'évaluation des modes d'indexation de sites dédiés à l'information musicale s'inscrit dans un cadre plus large, celui de l'innovation technologique et de ses rapports à la société.

Notre démarche d'observation des sites s'appuie sur une longue expérience (conduite des deux projets) d'évaluation d'outils de traitement d'information textuelle (outils d'extraction de ressources terminologies dans divers contextes relevant du traitement de l'information, construction de référentiels terminologiques pour l'indexation et la recherche d'information). Nos travaux dans ce domaine nous ont permis de dégager des concepts féconds et de définir un cadre théorique et méthodologique pour l'évaluation d'outils issus de la technologie linguistique (Mustafa El Hadi, 2004 ; Timimi et al. 2008). Nous étendrons cette expérience à l'évaluation des sites d'information musicale (SIM).

L'Évaluation *qualitative* qu'on oppose traditionnellement à l'évaluation *quantitative* (Cavazza, 1993 ; Spark-Jones, 1996), est la forme la plus élémentaire d'évaluation des systèmes de traitement d'information textuelle, c'est-à-dire un jugement de valeur sur leur fonctionnement faisant intervenir des concepts descriptifs propres au paradigme d'évaluation. Le principal inconvénient de ces jugements est leur caractère parfois subjectif. L'évaluation qualitative peut cependant avoir un intérêt dans le domaine de la conception des systèmes, par son caractère comparatif. Dans le cas général des évaluations qualitatives, l'opérateur analysera les textes manuellement pour fournir des données de référence. L'évaluation quantitative repose sur la définition de grandeurs mesurables ou *métriques*, permettant la comparaison entre systèmes (*évaluation horizontale*) mais aussi entre différentes versions d'un même système (*évaluation verticale*) (Ibid).

Pour adapter cette définition au contexte d'indexation de la musique, nous avons choisi deux types d'évaluation, une évaluation quantitative et une évaluation qualitative. Cette dernière impliquera nécessairement une autre forme d'évaluation qualitative, l'*évaluation par adéquation*. Elle se définit comme l'application de l'adéquation d'un système aux besoins de l'utilisateur, aux performances requises et au coût de l'usage prévu. Il convient aussi dans ce type d'évaluation de définir des profils d'utilisateurs.

### **3.2 Vers des évaluations de SIM fondées sur l'usage**

Concernant l'emploi des termes, usage (note), *usager* et *utilisateur*, nous adoptons la position de Y. Le Coadic (1997). L'auteur a établi une distinction entre les termes *usage* et *utilisation* ; D'après lui "*l'usage* est une activité sociale " et "*l'utilisation* est synonyme d'usage pratique", ((1997) p. 19).

D'autres définitions du concept « usage » l'opposent à un autre concept pertinent pour notre étude est celui de « pratique » compte tenu de notre intérêt pour les pratiques d'indexation musicales. Annette Béguin (Béguin, 2004, p. 13), définit « pratique » et l'oppose à usage « Les pratiques sont des conduites finalisées, individuelles ou

collectives, figées ou adaptatives, socialement situées, inscrites dans une temporalité, sous-tendues par des représentations, des savoirs, une logique et un raisonnement, marquées par une appréciation de soi et des autres, et révélatrices d'une culture qu'elles enrichissent éventuellement en retour ».

« Comme *pratique, usage* implique une dimension temporelle, une forme d'habitude (*les us et coutumes*), mais surtout et fondamentalement un objet : *user de abuser de, user, utile*. Je définirais l'usage comme une conduite située face à des objets. Cette définition implique une dimension sociale et culturelle, comme pour les pratiques, mais elle implique aussi une dimension technique de relation à l'objet. Autrement dit, les usages s'inscrivent dans des pratiques mais pour analyser les usages, il faut prendre spécifiquement en compte la rencontre d'un sujet avec les objets », (Béguin, 2006).

De nombreuses études prennent en charge la dimension ou le rôle des usagers dans le développement d'une technologie. Elles décrivent comment les utilisateurs consomment, modifient, domestiquent, re-configurent et même résistent à une technologie.

L'expérience montre toutefois que ces approches ne suffisent pas à prédire complètement les usages effectifs qu'inventeront les utilisateurs et qui ne sont pas toujours anticipés. Le succès de la diffusion des SMS, par exemple, est typique d'une technologie dont le succès dépend des usages et compétences (y compris linguistiques) qu'ont inventés les utilisateurs, en tant qu'individus en interaction les uns avec les autres.

L'évaluation par l'usage de l'indexation de métadonnées de la musique (indexation sociale et indexation normée) est au cœur de notre réflexion. L'approche méthodologique que nous avons adoptée pour étudier l'usage des métadonnées est une approche hybride que nous développerons dans la quatrième partie. Après cette brève étude du paradigme de l'évaluation, sa méthodologie et ses motivations nous menons un travail d'investigation théorique des notions couramment utilisés en Sciences de l'information, notions qui reviennent fréquemment lorsque nous abordons la question de l'évaluation qui sont au cœur de cette proposition : la notion d'*usage* et, *dévaluation par l'usage*, de *dispositif* ; d'*appropriation*, d'*adoption* et donc de *détournement d'usage* pour voir dans quelle mesure il y a un continuum entre ces notions et comment elles s'articulent autour de notre étude.

#### **4. Méthodologie dévaluation adoptée et premiers résultats**

La méthodologie a fait appel à une approche d'évaluation hybride (quantitative, qualitative).

Le questionnaire visait à prendre un aperçu sur les consommateurs de la musique et leurs pratiques informationnelles au sein des réseaux sociaux (site, blogs, web radios,). La population n'est pas ciblée selon des critères socio-professionnels mais selon d'autres critères qui seront détaillés ultérieurement.

Les espaces privilégiés de recherche sur l'Internet sont les sites non-marchands reconnus (Myspace, Last.fm, Blogs, etc.). Ceci basé sur des raisons de confiance envers ces sites et diversité offerte.

Nous avons adopté une approche qualitative et quantitative, mettant en œuvre notamment une approche infométrique des données recueillies. A cette fin nous avons élaboré un questionnaire visant à prendre un aperçu sur les consommateurs de la musique et les usagers de sites amateurs bibliothèques musicales dans leur stratégie de recherche. La population n'est pas ciblée selon des critères socio-professionnels, et le questionnaire a

été soumis à des usagers directement dans des magasins de distribution de la musique. Une plateforme en ligne est également disponible ( <http://icomusicquestionnaire.c.la/> ). Les espaces privilégiés de recherche sur l'Internet sont les sites marchands reconnus (Amazon, iTunes, etc.) et les sites de références qui à la base étaient non-marchands (Myspace, Last.fm, Blogs, etc.). Ceci basé sur des raisons de confiance envers ces sites et diversité offerte.

Les espaces privilégiés de recherche sur l'Internet sont les sites non-marchands (Myspace, Last.fm, Blogs, etc.). Ceci basé sur des raisons de confiance envers ces sites et diversité offerte.

Notre questionnaire portera sur la connaissance qu'a la population des métadonnées :

- 1. par ouïe dire ;
- 2. via l'usage des métadonnées ou en navigant ;
- 3. par moteur de recherche.

Leurs critères sont essentiellement et respectivement : efficacité, rapidité, curiosité, confiance et simplicité.

Leur stratégie de recherche leur paraît globalement très pertinente, et simple à mener. Cependant en terme d'évaluation, la moitié déclare comme étant insuffisamment satisfaite des résultats obtenus. En effet, ils estiment que la plupart des résultats obtenus de prime abord sont inutiles. Ils préfèrent toutefois accorder plus d'importance au temps passé à rechercher et à la souplesse du moyen de chercher (moteur de recherche : « il suffit de taper ce qu'on veut »), plutôt qu'à l'efficacité de recherche.

Les deux tiers trouvent facilement ce qu'ils recherchent, et ceux qui n'y parviennent pas trouvent suffisamment d'aide pour mener à bien leurs recherches. Néanmoins, un désir d'autonomie se manifeste, il leur semble préférable de parvenir seul au terme d'une recherche, même s'ils accordent de l'importance au conseil extérieur (critère de curiosité et de découverte).

Une même proportion d'usagers met en ligne de la musique, et seulement 68% accorde de l'importance aux tags de ces chansons. Leur façon de tagguer est, par ordre de préférence :

- folksonomie ;
- indexation libre ;
- indexation classique.

Les raisons ici sont très simples : d'une part, ceux qui ne tagguent pas clament à l'inutilité et à la perte de temps, alors qu'au contraire les autres reconnaissent l'utilité et le rôle des métadonnées. Leur souci principal étant que la chanson puisse être facilement retrouvée par autrui.

Les dernières questions sont très révélatrices d'un désir de découverte important. Si 66% de la population interrogée se fait conseiller des chansons par son entourage, seul 44% juge pertinentes les chansons proposées ; mais pour autant ne souhaitent pas conseiller, orienter l'entourage à affiner les propositions sous couvert de curiosité et de découverte.

## **5. Conclusions et perspectives**

Ces indexations, libres ou semi-libres, sont des enjeux à la fois déontologiques et industriels. En effet libres ou semi-libres, elles sont basées sur une confiance octroyée à l'utilisateur. Il est attendu que celui-ci « joue le jeu », renseigne avec pertinence les champs,

et permette un réel enrichissement de l'expérience de la navigation dans les réseaux sociaux dédiées à l'information musicale. Des détournements d'usage ont lieux, qu'ils soient attendus (espérés) ou non, ils sont bien présents. Paradoxe entre intentions d'usages, détournements d'usages et usages convenus. Garant d'autonomie et de dispense institutionnelle d'autorité, l'outil d'indexation professionnel (les métadonnées) est détourné de son usage originel et tend à un devenir modifié distinct. L'utilisateur s'approprié les métadonnées comme espace visible sur la page (tag cloud) pouvant posséder une finalité enrichie. Or ce devenir, même s'il relève en apparence d'un détournement d'usage, se meut en convenu et devient ainsi la « règle ». Nous tenterons de vérifier cette hypothèse dans nos travaux en cours.

## 6. Références bibliographiques

Béguin, A. (2006) Séminaire Pratiques et usages, Geriico-Ecole Doctorale SHS Nord., 18 décembre 2006

Béguin, A. (2004), La communication graphique, In : Critique de la raison graphique, *Hermès*, N° 39, Paris, 2004.

Beuscart, J.S, (2007) « Les transformation de l'intermédiation musicale, la construction de l'offre commerciale de musique en ligne en France ». In : *Réseaux* vol. 25----n° 141-142/2007, p.143-176.

Beuscart, J.S., (2008), « Sociabilité en ligne, notoriété virtuelle et caractère artistique : les usages de MySpace par les musiciens contemporains autoproduits », In : Cardon, D. (dir) (2008) *Réseaux* ; vol 26---N° 152/2008, p. 139-168.

Cardon, D. (dir) (2008) *Réseaux sociaux de l'Internet*, *Réseaux* ; vol 26---N° 152/2008, dossier coordonné par Dominique Cardon, Hermès-Lavoisier.

Cavazza M. (1993), Méthodes d'évaluation des logiciels incorporant des technologies d'informatique linguistique, Rapport MRE-DIST, Paris.

Chartron, G., Broudoux E., (dir) (2006). *Document numérique et société*, Actes de la conférence organisée dans le cadre de la Semaine du document numérique à Fribourg (Suisse) les 20 et 21 septembre 2006. ADBS Éditions, 2006. 344 p. Collection Sciences et techniques de l'information, Fribourg.

Crepel, M. (2008), « Les folksonomies comme support émergent de navigation sociale et de structuration de l'information sur le Web ». In : In : Cardon, D. (dir) (2008), vol 26---N° 152/2008, p. 169-2004.

Debaecker, J. Mustafa El Hadi, W.(2010) « L'indexation de la musique, vers de nouvelles approches à l'ère des réseaux sociaux », In Papy, F. (dir), *Evolutions socio-techniques des bibliothèques numériques* à paraître.

Ertzscheid, O. Gallezot, G. (2006), « Etude exploratoire des pratiques d'indexation sociale comme une renégociation des espaces documentaires. Vers un nouveau big bang documentaire ? », In : Document numérique et société sous la dir. de Ghislaine Chartron et Evelyne Broudoux, p. 291-306.

Frith, S. Le Guern, Ph. (dir), 2007, Sociologie des musiques populaires, Réseaux vol. 25--  
--n° 141-142/2007.

Le Coadic, Y. (1997). Le besoin d'information, Nathan, Paris 1997

Mustafa El Hadi, W. (2004) « Acquisition de ressources terminologiques ». In :  
*Evaluation des Systèmes de traitement de l'information*, sous la direction de Stéphane  
Chaudiron, Hermès, pp. 149-169

Timimi, I., Mustafa El Hadi, W. (2008), "CESART - Campagne d'Evaluation de  
Systèmes d'Acquisition de Ressources Terminologiques» in Chaudiron, S., Choukri,  
K. (dir.), *L'évaluation en technologies de la langue : les campagnes Evalda*, Paris :  
Hermès, p. 71-91.

Sparck-Jones K., Gallier, J.R. (1996). Evaluating Natural Language Processing Systems  
An Analysis and Review, Springer, Berlin.