

# CLASSER LES SUPPORTS POUR PRESERVER LES CONNAISSANCES : L'EXEMPLE DES 'PATCHS' DANS LE DOMAINE DE LA CREATION MUSICALE CONTEMPORAINE

Bouchra Lamrini, Jérôme Barthélemy, Alain Bonardi, Rousseaux Francis

▶ **To cite this version:**

Bouchra Lamrini, Jérôme Barthélemy, Alain Bonardi, Rousseaux Francis. CLASSER LES SUPPORTS POUR PRESERVER LES CONNAISSANCES : L'EXEMPLE DES 'PATCHS' DANS LE DOMAINE DE LA CREATION MUSICALE CONTEMPORAINE. Journées d'Informatique Musicale, May 2011, Saint-Etienne, France. hal-03104732

**HAL Id: hal-03104732**
**https://hal.science/hal-03104732**

Submitted on 9 Jan 2021

**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.

# CLASSER LES SUPPORTS POUR PRESERVER LES CONNAISSANCES : L'EXEMPLE DES 'PATCHS' DANS LE DOMAINE DE LA CREATION MUSICALE CONTEMPORAINE

*Bouchra Lamrini* *Jérôme Barthélemy* *Alain Bonardi* *Francis Rousseaux*
IRCAM, 1, place Igor Stravinsky, 75004 Paris
bouchra.lamrini@ircam.fr

**RÉSUMÉ**

Le document numérique est présent partout dans les activités de nos sociétés contemporaines où les échanges médiatisés s'accroissent. Les tâches de constitution et de classification des documents numériques deviennent complexes, notamment dans un flux numérique de gestion électronique de documents où les outils informatiques disponibles ne sont pas toujours adéquats. Nous proposons dans cet article, une approche d'analyse et de classification des patchs de la musique interactive. Soumis aux difficultés de la préservation, ces processus temps réel de traitement sonore sont considérés souvent comme des véritables documents numériques, ils sont à la fois supports de création et supports de constitution de connaissances dans la création artistique contemporaine. En s'appuyant sur l'analyse des expressions algébriques contenues dans des documents transcrits à partir des processus existants, conçus dans l'environnement Max/MSP, un lien systématique entre classification conceptuelle et classification par apprentissage automatique est donc établi afin d'établir les prémices d'une organologie des traitements musicaux et audio numériques.

## 1. INTRODUCTION

Une partie importante des œuvres de musique contemporaine fait appel aux technologies électroacoustiques et numériques aussi bien pour leur composition que pour leur exécution. Depuis une trentaine d'années, se sont multipliés les dispositifs et les outils logiciels permettant de manipuler les sons en temps réel. Parmi ces dispositifs figure le patch comme processus numérique temps réel utilisé en musique contemporaine (performances, concerts en live, etc.) et dans une approche de modélisation du dialogue musical homme-machine [3, 4, 5]. C'est un ensemble d'informations qui représente l'une des seules traces de l'activité créatrice traduisant d'une certaine manière une pensée créative.

Dans la perspective d'une pérennisation à long terme de la création musicale, le problème se pose de savoir préserver ce processus temps réel d'une manière à le rendre indépendant de l'environnement technique sous-jacent (hardware et software). Pour répondre à cette question, nous proposons une modalité de préservation permettant de transcrire les processus existants, les analyser et les classer afin de constituer des bases de connaissances du domaine en jeu. Le passage progressif des processus transcrits à l'ontologie prônée par cette méthodologie est organisé en prenant en compte la structure des documents, où les traitements et objets sont transcrits au moyen du langage FAUST[1] (Functional Audio STreams), et des éléments porteurs d'information. Le découpage structurel de chaque document correspondant à une caractérisation sémantique de son contenu. L'idée est donc de tirer profit de la structure et de la hiérarchisation du document pour souligner la complémentarité entre identification de concepts (objets recherchés) et extraction de relations. Les différentes réflexions exposées pour l'analyse et la classification de ces documents, font appel à des concepts et des techniques provenant du TAL (Traitement Automatique des Langues) et de l'Intelligence Artificielle.

Ce travail de recherche se place au cœur des échanges entre classification, terminologie, ontologie et représentation de connaissances. Il amène une réflexion sur les divers paliers à envisager dans une telle démarche de modélisation de connaissances textuelles pour des objectifs de construction d'ontologie de domaine et préservation d'une certaine authenticité de l'exécution d'une œuvre. Dans ce qui suit, nous exposons la problématique et les défis menés pour la documentation de ces processus numériques afin de préserver une certaine authenticité vis-à-vis de l'exécution d'une œuvre. Nous présentons ensuite notre démarche et nos perspectives envisagées en termes d'analyse et de classification.

## 2. LE PATCH, PROBLEMATIQUE ET DEFIS DE SA DOCUMENTATION

Il convient tout d'abord de noter que les actions de préservation concernant les objets numériques mis en œuvre dans une création artistique doivent examiner la nature de l'œuvre originale, et les relations de l'objet numérique avec l'œuvre, avant d'évaluer les actions à effectuer pour le préserver. Il s'agit bien de préserver l'ensemble des connaissances permettant dans le futur d'apprécier le contexte dans lequel l'objet numérique est

[1] http://faust.grame.fr/

mis en œuvre : les informations portant sur l'environnement technique numérique, mais aussi l'environnement immédiat (par exemple, les caractéristiques de la salle de concert), les périphériques utilisés (microphone, hauts-parleurs...). On devra aussi tenir compte d'une caractéristique de la démarche artistique, qui est une tendance à repousser les limites d'une technologie donnée, afin de la transcender pour obtenir de nouveaux effets. Il s'agit ici de préserver les intentions qui sous tendent la démarche créatrice. Une approche ou un modèle global et structuré de préservation devra impérativement tenir compte de ces considérations.

Le patch lui-même, document numérique porteur d'informations sur la conception musicale et technique d'une œuvre musicale, est confronté à diverses difficultés (fragilité des supports de stockage, obsolescence des standards techniques, l'instabilité des patchs vis-à-vis de leur complexité, etc.) qui compliquent la tâche de préservation. Le processus de documentation comporte généralement trois activités : la recherche (repérer et identifier les données pertinentes), la préservation (pérenniser les données) et la diffusion (rendre ces données disponibles et les transférer en connaissances utilisables). Une pratique documentaire rigoureuse est d'autant plus essentielle que les sources d'information sur la création contemporaine sont non structurées, et notamment pour les œuvres elles-mêmes. Compte tenu des défis communs à la préservation des documents et de l'art numérique, la recherche dans ces documents est bien entendu pertinente. En matière de documents numériques, Margaret Hedstrom [13] fait une recommandation intéressante : « Préserver le contenu, le contexte et la structure tout en préservant la capacité d'afficher, de lier et de manipuler les objets numériques ». La plus grande difficulté réside probablement dans le second aspect, car cela suppose de garder l'accessibilité à une multitude de logiciels et de systèmes d'exploitation. Hedstrom ajoute également : « La préservation de matériel numérique nécessite souvent la transformation complexe et coûteuse d'objets numériques, pour qu'ils demeurent des représentations authentiques de la version originale ainsi que des sources utiles aux fins d'analyse et de recherche ». Le paradoxe de la préservation de documents numériques réside entre les expressions « transformations » et « représentations authentiques de l'original ». Les spécialistes proposent de plus en plus le concept d'émulation [19] et d'enveloppe contextuelle parmi les stratégies d'archivage numérique. Ce concept est une solution potentielle au problème des documents numériques dépendant des logiciels et, ce qui est plus important, du matériel requis pour y avoir accès. Ce concept permet également de conserver l'entière capacité de traitement des données.

Cette approche suppose de placer le document, conservé dans sa forme d'origine, dans une enveloppe virtuelle contenant toutes les instructions nécessaires pour sa récupération, son affichage et son traitement. Les instructions expliquent comment lier le document à un ensemble d'émulateurs qui servent de passerelle entre le document, qui peut demeurer stable, et le contexte technologique en constante évolution. Ainsi, plutôt que de tenter de modifier une multitude de documents, les gestionnaires d'archives ou de collections numériques n'ont qu'à mettre à jour les émulateurs. Une fiche technique renfermant d'importantes spécifications peut être ajoutée au document sous forme de métadonnées. Cela facilite le repérage de documents qui nécessiteront une certaine intervention en vue de leur conservation future. Grâce aux métadonnées, la description du document peut être incluse dans le document même, faisant de ce dernier sa propre fiche de catalogue. Il existe cependant un inconvénient : les métadonnées peuvent être aussi imposantes sinon plus que le document décrit.

Les patchs, comme documents de représentation de connaissances, sont source d'intérêt pour la communauté des chercheurs en musicologie et en informatique musicale. En matière d'œuvre, nous citons quelques travaux, cités dans [3, 4] menés pour la maintenance des patchs selon les quatre actions, préservation, émulation migration et virtualisation. Depuis quelques années, la préservation des arts de la performance fait l'objet de travaux scientifiques [6, 21, 22]. Nous citons également l'exemple de projet de collaboration 'Mustica' entre l'IRCAM[2] et l'INA[3] [1] sur la production d'une base de données ouverte aux organismes souhaitant remonter des œuvres contemporaines. L'article de Bullock et Coccioli [8] présente l'exemple de l'œuvre *Madonna of winter and spring* et la tentative d'émulation d'un modèle Yamaha TX816 (pour la synthèse sonore), qui n'existe plus, par des patchs créés sous PureData. A ces exemples s'ajoutent les travaux de [18] dont les auteurs recommandent de sauvegarder les patchs réalisés par MaxMSP au format texte plutôt qu'en codage binaire. Enfin, nous citons le travail d'Andrew Gerzso (cité dans [5]) qui a été réalisé au sien de l'IRCAM sur la recherche des représentations indépendantes d'une implémentation technique, dans le cadre de l'œuvre *Anthèmes II* de Pierre Boulez.

## 3. METHODOLOGIE DE CONSTITUTION ET DE PRESERVATION DE CONNAISSANCES MUSICALES

### 3.1. Cadre de travail

Le travail de recherche que nous présentons considère le document numérique comme porteur de connaissances et doté d'une intentionnalité qui le construit ou le reconstruit pour une préservation à long terme. Dans cette vision du document comme ensemble de signes, nous portons plus particulièrement notre intérêt sur son analyse et sa classification, c'est-à-dire sur les objets qui le constituent. Pour illustrer ce besoin de

[2] http://www.ircam.fr/
[3] http://www.ina.fr/

préservation du document, le projet ASTREE dans lequel s'inscrit notre travail, a permis une réflexion nouvelle sur la documentation et la préservation des processus temps réel vis-à-vis de l'histoire d'une création et l'exécution d'une œuvre de l'art performance. Cette réflexion consiste à transcrire les processus existants, conçus dans les environnements tels que Max/MSP ou PureData, dans le langage fonctionnel Faust (Functional AUdio STreams) [17]. Cette réflexion se résume par les trois points suivants :

⇒ Génération automatique de documentation sous une forme indépendante de toute technologie (hardware et software). Cette documentation permettra une réimplémentation entièrement manuelle du processus originel.

⇒ Génération d'un code optimisé et indépendant des librairies et des systèmes de traitement de signal.

⇒ Application des techniques de data mining sur les expressions algébriques issues de FAUST et de la documentation afin de constituer des connaissances sur les processus en question.

FAUST est un langage formel dédié à la synthèse et au traitement du signal. Via un concept familier de bloc-diagramme, il permet de décrire facilement des processeurs de signaux qui traitent et transforment des signaux en entrée pour produire des signaux en sortie. L'utilisateur dispose de modules de traitements élémentaires qu'il combine de différentes manières pour obtenir le traitement souhaité. Le compilateur FAUST utilise ensuite cette description pour écrire automatiquement un programme C++ équivalent. Des techniques d'optimisation particulières permettent d'engendrer un code C++ de qualité dont l'efficacité est généralement comparable à celle d'un programme écrit à la main. FAUST peut donc être utilisé seul, pour écrire des applications audio autonomes, mais également en complément de systèmes existants comme Alsa (*Advanced Linus Sound System*), LADSPA (*Linux Audio Developer's Simple Plugin*), PureData, etc.). Un bloc-diagramme peut être représenté comme un graphique G = (N, C) où N est un jeu de nœuds, c'est-à-dire les blocs du diagramme et *C* l'ensemble de connexions entre ces blocs.

La figure 1 présente la fonction 'Envelop' qui permet de générer un bruit blanc contrôlé par une enveloppe. Plusieurs paramètres sont utilisés pour ajuster la forme de l'enveloppe : la durée d'attaque, la durée de décroissance, le pourcentage de volume maintenu, et la durée du relâchement.

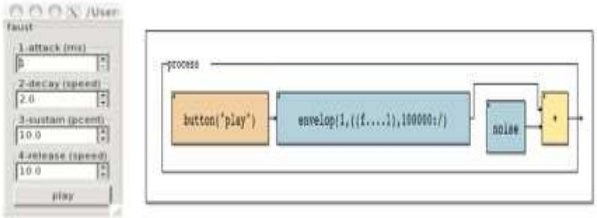



**Figure 1.** Schéma de représentation du contrôle par une enveloppe d'un bruit blanc au moyen de FAUST.

### 3.2. Présentation du corpus

Les objets pris en compte dans nos travaux sont des objets issus des expressions algébriques contenues dans la documentation transcrite par le langage fonctionnel FAUST. Pour illustrer la nature du problème à résoudre, nous avons choisi un exemple de documentation sur un filtre passe band 'BPF' comme un processus temps réel conçu dans Max/MSP et transcrit dans le langage FAUST. La figure 2 présente les types de structures contenues dans le corpus en question. Les concepts que l'on cherche à repérer dans la documentation se trouvent dans le code FAUST, la documentation générée en latex, les trois librairies importées (maxmsp, music, et math) et les commentaires inclus dans les codes et les librairies.

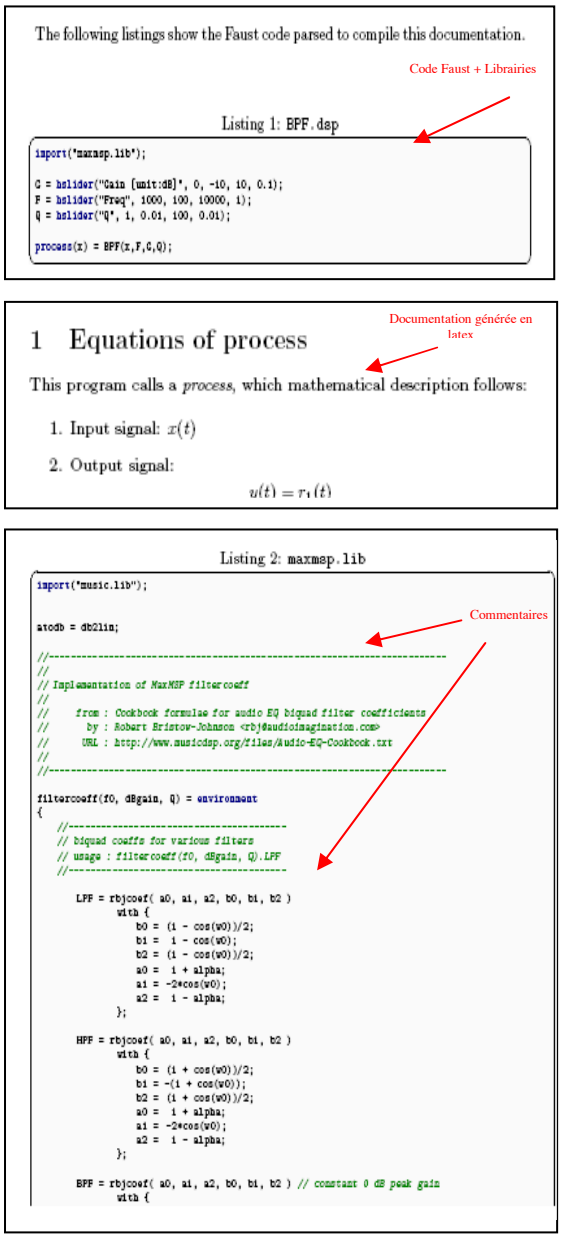



**Figure 2.** Structures contenues dans le corpus à analyser.

Si nous nous intéressons par exemple seulement aux

objets provenant de Max/MSP, les trois fragments du document permettent de fournir une information (Figure 3) sur les objets clés :
- le filtre passe-bande 'BPF' présent dans la librairie 'maxmsp.lib'. Ceci est donné par l'expression décrivant le processus général 'process (x) = BPF(x, F, G, Q)'.
- présence d'une entrée x(t), ceci est indiqué dans le texte descriptif et par l'expression algébrique du processus.
- les coefficients du filtre, ici le gain 'G', la fréquence 'F' et le facteur de qualité 'Q'.
- les expressions fonctionnelles qui lient le processus aux coefficients du filtre comme la présence de la fonction de transfert 'biquad'. La présence de cette fonction entraine également la présence d'autres paramètres de conversion et de fonctions de calcul comme les fonctions trigonométriques.
- et également un domaine de variation de chaque coefficient approprié au filtre en question (Exp : *''Freq'', 1000, 100, 10000, 1*).

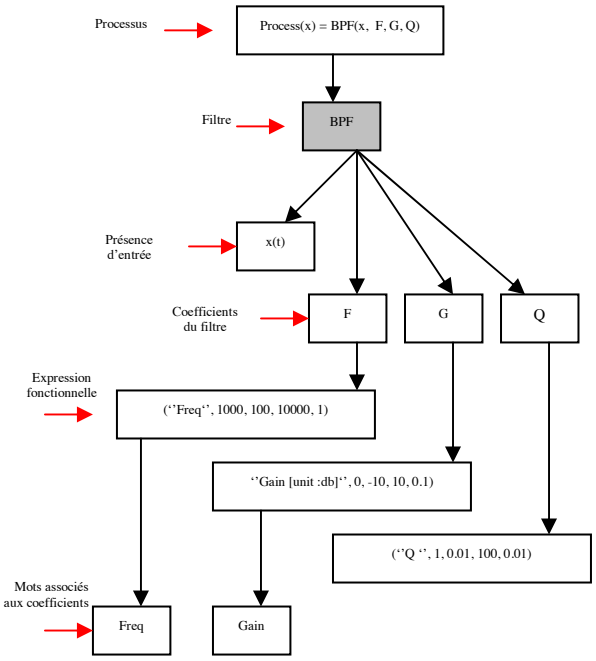



**Figure 3.** Diagramme illustrant les différents termes clés qui peuvent fournir une connaissance sur le filtre passe-bande 'BPF : Band Pass Filter '.

Les processus dans leur expressions algébriques peuvent être vus comme des objets clés qui génèrent une sorte de connaissances appropriées aux termes recherchés comme 'Filter, BPF, APF, Oscillator, Osc...', et les termes 'Q, G, F,...' comme des attributs de ces termes recherchés. En s'appuyant sur les différents documents transcrits et normalisés à présent, nous avons commencé à constituer une base de connaissances sur les objets clés, les termes associés ainsi que d'autres données permettant de fournir une information sur la proximité, la corrélation des objets, la proximité des classes objets, la distribution des objets autour d'un objet de la même classe (chaque objet a une distance proportionnelle à son degré d'association). En classification conceptuelle ascendante par exemple (car c'est la plus recommandée pour des raisons de complexité algorithmique et 'd'explicabilité' pour la validation par l'utilisateur en cours d'apprentissage), l'algorithme général consiste à réunir successivement par paires, les objets proches, ou les objets et les classes objets proches, afin de former des hiérarchies ou des graphes de classes d'objets.

### 3.3. Méthodologie d'analyse et de classification

Sur notre problème d'analyse et de classification des patchs, un système d'extraction d'information classique seul n'est pas capable d'analyser les différents segments du corpus en question, notamment sa structure qui s'appuie en général sur des expressions algébriques, combinées d'une certaine manière à décrire les entrées et les sorties des processus réalisés d'une part, et les fonctions de transformation effectuée entre ces entrées et ces sorties d'autre part. En effet, pour analyser une phrase, un système d'extraction d'information effectue successivement une analyse lexicale (segmentation de la phrase en chaines de caractères qui représentent des mots), une analyse morpho-syntaxique (étiquetage des mots par leur catégorie syntaxique et association de chaque mot à sa forme canonique..), une analyse syntaxique (analyse de la structure de la phrase), et en fin une analyse sémantique (compréhension du sens des mots et des relations entre les mots). Or, notre document numérique est le résultat d'une transcription à base d'un langage fonctionnel. Sous sa forme observable, ce document peut être considéré comme une suite de mots ordonnés selon des règles qui ne reflètent pas forcément l'ordre dans lequel les mots s'appliquent les un aux autres pour former l'interprétation sémantique fonctionnelle.

Comme nous l'avons mentionné auparavant, notre démarche propose une nouvelle réflexion de mise au point d'une méthode de classification permettant de former de manière automatique une ontologie conceptuelle du domaine musical en termes des traitements musicaux et audio numériques. De manière générale, la définition d'une méthode de classification conceptuelle repose sur deux éléments : la définition d'une distance permettant de comparer les objets à classer, et d'autre part la définition d'un algorithme de classification qui construit la structure arborescente proprement dite. Dans le domaine de la formation des ontologies conceptuelles, la notion de distance entre mots a été étudiée en TAL pour construire les classes sémantiques qui forment les nœuds de ces catégories suivant une approche ascendante ou descendante, tandis que les algorithmes de classification ont été plus largement étudiés en analyse de données et en apprentissage automatique. Dans ce contexte, nombreux outils destinés à l'acquisition automatique/ou semi-automatique de classes sémantiques visant à regrouper de termes proches, sont élaborés. De point de vue sémantique, cette notion de proximité est généralement

basée sur des mesures de distance entre termes en fonction du degré de ressemblance des contextes dans lesquels ils apparaissent. Les descriptions et les régularités recherchées des contextes des termes dépendent de l'approche menée. Ainsi, les contextes peuvent être purement graphiques, i.e. sous formes de co-occurrences dans une fenêtre de mots [20, 7, 9]. Le choix de la distance appropriée pour un corpus est donc un problème posé et peu étudié [12]. La majorité des travaux ne s'intéressent qu'à la seule évaluation de la tâche finale pour laquelle l'apprentissage est effectué. Les critères de cette évaluation restent quantitatifs et donnent rarement lieu à des études comparatives [10]. En revanche, la caractérisation des résultats des méthodes fournit les outils d'aide aux choix de la plus pertinente ou la mise au point de nouvelles méthodologies. Ceci est pareil pour les algorithmes de classification. Elaborer une méthodologie ou un outil permettant de mettre au point un algorithme de classification conceptuelle pour former des ontologies en fonction de la tâche est une tâche ardue. Par ailleurs, les travaux effectués en classification conceptuelle [11, 2] n'ont pas trouvé de propos en apprentissage à partir de corpus. Certes, il faut souligner que la construction d'une ontologie dans le domaine du TAL exige une phase de choix et d'adaptation des algorithmes existants aux problèmes posés par le corpus autour de la distance, voir développer de nouveaux outils et moyens méthodologiques.

Dans le cadre de ce travail, notre méthodologie envisagée d'analyse et de classification s'appuie sur deux pistes complémentaires, qui s'emboîtent pour donner une classification finale des objets en question, et notamment jugée par l'expert du domaine (Figure 4).

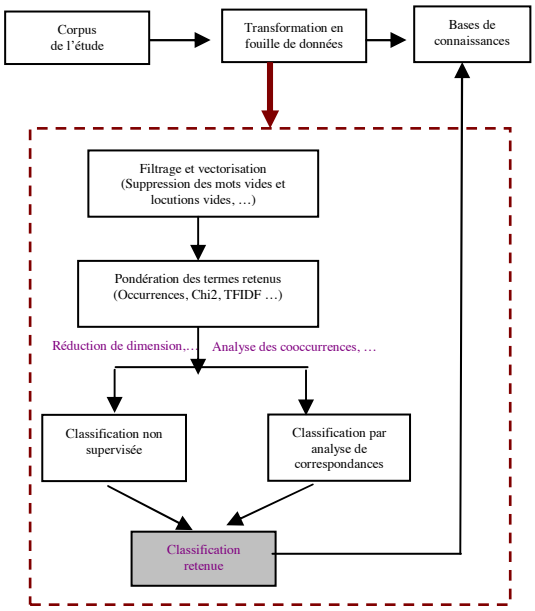



**Figure 4.** La méthodologie proposée pour l'analyse et la classification des patchs.

Notre idée dans un premier temps est de réaliser une série de prétraitements classiques au moyen des outils du TAL afin d'éliminer les objets non-clés pour notre étude et fixer le type de pondération que l'on doit affecter aux objets retenus pour la classification. Cette phase de prétraitement est guidée notamment par l'expert du domaine. L'intervention de l'expert permettra de valider chaque étape avant la validation finale de la classification attendue par notre étude. Ensuite, on a deux voies à parcourir pour avoir une classification satisfaisante selon nos objectifs fixés au début de l'étude. La première voie est de poursuivre les analyses par les mêmes outils du TAL tels que : les analyses des cooccurrences d'objets-clés, par exemple faire des comparaisons entre les paires d'objets-clés de la famille Filter : 'APF-BPF' et 'APF-HPF', et également l'analyse d'autres objets associés à ces objets-clés ; les analyses thématiques des unités de contexte, autrement dit opérer et modeler les objets qui émergent des différents unités du corpus et qui sont décrits à travers leur vocabulaire caractéristique, c'est-à-dire à travers des ensembles de mots-clés (lemmes ou catégories) co-occurrents. Ces objets émergents peuvent être utilisés pour obtenir de nouvelles variables qui peuvent être utilisées dans des analyses ultérieures. Et enfin les analyses comparatives des sous-ensembles du corpus qui donneront lieu à une classification par une analyse des correspondances. Comme toutes les techniques factorielles, les analyses comparatives permettent l'extraction de facteurs qui ont la propriété de récapituler d'une façon organisée l'information significative contenue dans les innombrables cellules des tableaux de données. En outre, cette technique d'analyse permet la représentation graphique, dans un ou plusieurs espaces, des points qui détectent les objets en lignes et colonnes, et qui dans notre cas sont les entités linguistiques (mots, lemmes, segments de texte, textes, etc.). Les résultats d'analyses permettront d'évaluer des rapports de proximité/distance - ou de similitude/différence - entre les objets considérés.

Pour la deuxième voie, nous envisageons de réduire la dimension de corpus par l'application de l'une des techniques d'analyse factorielle citées auparavant, la plus adaptée pour notre cas est la version neuronale de l'ACP (Analyse en composantes Principales), appelée AGH (Algorithme Hebbien Généralisé ). Cela permettra de projeter les documents dans un espace beaucoup plus compact avec des concepts plus significatifs. En effet ces nouvelles dimensions représentent essentiellement des corrélations entre termes dans un corpus donné, révélant de la sorte les thèmes principaux du corpus presque aussi bien qu'une classification des documents. Ensuite, nous utilisons les vecteurs dans le nouvel espace (donnés par les valeurs de sortie du réseau AHG) pour effectuer une classification non supervisée des documents. Les algorithmes les plus connus sont soit l'algorithme des centres mobiles (k-means) qui emploie une distance euclidienne pour ne pas avoir à normaliser les vecteurs, soit les réseaux de neurones artificiels par les cartes auto-organisatrices de Kohonen [14]. Nous citons ici les travaux de Nguyen et Zreik [15, 16, 23] sur le développement du système Hyperling pour reconnaitre les langues dominantes dans un site Web

multilingue et dont les deux algorithmes K-means et SOM ont été implémentés et testés avec succès.

## 4. CONCLUSION

A travers ce travail, nous avons présenté le rôle du patch comme document numérique dans la création artistique contemporaine, ainsi que les différentes approches mises au point pour leur maintenance dont la classification est l'une des solutions retenue en termes de préservation des œuvres de l'art de la performance. La méthodologie de classification et de préservation de connaissances développée est une nouvelle réflexion destinée à analyser la nature de ces processus temps réel et d'en relever les propriétés cachées, et ceci dans le but final de constituer des connaissances permettant de générer une organologie des traitements sonores temps réel qui contribuera ainsi à des pratiques de préservation de l'œuvre temps réel. Certes, les sources documentaires et les applications issues du TAL pour construire de façon plus ou moins automatique des ontologies de type hiérarchies conceptuelles à partir de corpus analysés, s'accroissent et deviennent applicables à une large variété de problèmes. A l'opposé, il manque les outils et les méthodologies qui permettent d'évaluer et comparer les points faibles et forts des différentes approches pour un corpus spécifique à un domaine et une tâche donnée. Notre méthodologie propose donc un modèle pour formaliser et modéliser des connaissances musicales en intégrant des techniques provenant du TAL pour la classification conceptuelle et celles provenant de l'intelligence artificielle pour analyser et évaluer les classes qui vont former l'ontologie attendue en fin de ce travail.

Les travaux en cours et futurs se concentrent sur les moyens qui permettent d'obtenir et d'évaluer les classes attendues par nos objectifs, à savoir les critères de constitution du corpus d'entrée, les distances, et les critères d'évaluation des résultats, nivaux d'abstraction (lié aux objets qui doivent être isolés de la classification, etc.).

## 5. REMERCIEMENTS

Les auteurs tiennent à remercier les partenaires du projet ASTREE, et notamment ceux du Centre National du Création Musicale Grame, à l'origine du langage FAUST, et de l'outil de génération automatique de documentation sur lequel notre travail se base

## 6. REFERENCES

[1] Bachimont, B., Blanchette, J. F., Gerszo, A., Swetland, Lescurieux, O., Mahoudeaux, P., N. Donin, N., Teasley, J. "Preserving Interactive Digital Music: A report on the Mustica Research Initiative", In Proceedings of the Third International Conference on WEB Delivering of Music (WEB'03), Leeds, England, 2003.

[2] Bisson, G. "Conceptual Clustering in a First Order Logic Representation", In Proceedings of 10th European Conference on Artificial Intelligence (ECAI'92), 458-462, Vienna, 1992.

[3] Bonardi, A., Barthélemy, J. "Le patch comme document numérique : support de création et de constitution de connaissances pour les arts de la performance", 10ième Colloque International sur le Document Electronique, 2-4 juillet, Nancy, France, 2007.

[4] Bonardi, A., and J. Barthélemy, J. "The preservation, emulation, migration, and virtualization of live electronics for performing arts: An overview of musical and technical issues", Journal on Computing and Cultural Héritage (JOCCH), 1(1), Article 6, 2008.

[5] Bonardi, A., Serra, M. H, Fingerhut, M. "Documentation musicale et outils hypermédias", Deuxième Colloque International sur le Document Electronique (CIDE'99), 5-7 juillet, 295-309, Damas, Syrie, 1999.

[6] Bosma, H. "Documentation and Publication of Electroacoustic Compositions at NEAR", In Proceedings of the Electroacoustic Music Studies Network International Conference (EMS 05), Montreal, Canada, 2005.

[7] Brown, P. F., Della Pietra, V.J., DeSouza, P.V, Lai, J.C., Mercer, R.L. "Class-based n-gram models of natural language", Computational Linguistic, 18(4), 283-298, 1992.

[8] Bullock, J., Coccioli, L. "Modernising Live Electronics Technology in the Works of Jonathan Harvey", In Proceedings of the International Computer Music Conference, Barcelona, Spain, 2005.

[9] Church, K. W., Hanks, P. "Word Association Norms, Mutual Information, and Lexicography", In proceedings of the 27th Annual Meeting of the Association for Computational Linguistics, 76-83, 1989.

[10] Dagan, I., Lee, L., Pereira, F. "Similarity-Based Methods For Word-Sense Disambiguation". In proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL'96), 56-63, 1996.

[11] Fisher, D. H. "Knowledge Acquisition via Incremental Conceptual Clustering", In Machine Learning Journal, 2, 139-172, 1989.

[12] Habert, B., Nazarenko, A. Salem, A. "Les linguistiques de corpus", Ed Armand Collin, 1997.

[13] Hedstrom, M. "Digital Preservation: A Time Bomb for Digital Libraries", Computers and the Humanities, 31, 189–202, 1998.

[14] Kohonen, T. ”Self-organization of very large document collections: state of the art”, In Proceeding of ICANN'98, London, 1998.

[15] Nguyen, D., Zreik, K. “HYPERLING : Système de reconnaissance et de classification des hyperdocuments multilingues”, In International Conference in Computer Science-Research, Innovation and Vision of the Future (RIVF05), 21-24 February 2005,University of CANTHO, Vietnam, 2005.

[16] Nguyen, D. “Nouvelle méthode syntagmatique de vectorisation appliquée au Self-organizing map des textes vietnamiens’’, JEP-TALN-RECITAL (Rencontre des Etudiants Chercheurs en Informatique pour le Traitement Automatique des Langues), 19-22 avril 2004, Fès, Maroc, 2004.

[17] Orlarey, Y., Fober. D., Letz. S. “FAUST : an efficient Functional Approach to DSP Programming”, New Computational Paradigms For Computer Music, Delatour, France, 2009.

[18] Puckette, M. “New Public-Domain Realizations of Standard Pieces for Instruments and Live Electronic*s”,* In Proceedings of the International Computer Music Conference, Miami, 2004.

[19] Rothenberg, J. “Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation”, Published by ECPA (European Commission for Preservation and Access), Edo H. Dooijes. Amsterdam, 1998. http://www.clir.org/pubs/reports/rothenberg/contents.html

[20] Sparck Jones, K., Barber, E. B., “What makes an automatic keywords classification effective?”, Journal of the ASIS, 18, 166-175, 1971.

[21] Teruggi, D. “Preserving and Diffusing”, Journal of New Music Research, 30 (4), 403-405, 2001.

[22] Tiffon, V. “Les musiques mixtes entre pérennité et obsolescence”, Revue Musurgia, XII/3, Paris, 2005.

[23] Zreik, K., Nguyen, D. “Catégorisation de documents multilingue : le système Hyperling”, In CIDE.8, 25-28 Mai 2005, Mission Culturelle Française, Beyrouth, Liban, 2005.