

Management des connaissances dans le domaine du patrimoine culturel

Stefan du Château, Danielle Boulanger, Eunika Mercier-Laurent

MODEME Centre de Recherche IAE Université de Lyon. 6, av Albert Thomas F-69008 Lyon

1 Introduction

Nous présentons une application innovante de management des connaissances dans le domaine du patrimoine historique.

Ce système vise à offrir aux experts du patrimoine culturel des outils permettant : le recensement d'objets historiques sur le terrain à l'aide d'enregistrements audio, la « traduction » de ces enregistrements en texte, l'extraction des informations selon un modèle de connaissances prédéfini et la recherche pertinente des œuvres et de leurs contextes spatio-temporels.

L'innovation majeure de cette application est la connexion dans un système hybride de technologies du traitement du signal, du traitement du langage naturel et de la modélisation et découverte des connaissances. Elle ouvre la possibilité aux experts d'enregistrer oralement des informations qu'ils détiennent sur l'œuvre.

Cette application doit répondre à deux exigences, d'une part permettre de collecter des informations correspondantes à un cahier des charges précis défini par le SDI¹ (Verdier, 1999), d'autre part répondre aux exigences d'un système de management des connaissances.

2 La démarche

La méthode appliquée est KM global (Amidon et al., 2006). Notre système doit couvrir le cycle d'acquisition et de la modélisation des connaissances sur l'œuvre du patrimoine culturel. Il doit également permettre l'exploitation des connaissances et notamment la recherche pertinente.

La première partie est composée de quatre modules:

1. Acquisition vocale de la description d'une œuvre à l'aide d'un dictaphone, un Pocket PC ou un téléphone portable.
2. Transcription automatique de l'enregistrement audio en texte ASCII. à l'aide de logiciel DRAGON², configuré avec un vocabulaire spécifique.
3. Recherche et extraction d'informations à partir de texte à l'aide du module développé à partir du logiciel XIP³
4. Validation par un expert, des informations trouvées dans l'étape précédente.

Les descripteurs issus de l'étape 3 et 4 alimenteront, une base de données, et guidés par des ontologies partielles, seront utilisés pour la modélisation des connaissances en patrimoine culturel.

3 Travaux effectués

Le module d'acquisition et retranscription a été testé sur le terrain. Pour tester sa robustesse et l'exactitude nous avons effectué les expérimentations avec des chercheurs en patrimoine culturel dans un bruit environnant et avec des accents différents. Les chercheurs connaissaient l'œuvre qu'ils devaient décrire. Ces premières expérimentations sont encourageantes : l'exactitude du système varie entre 90% à 98% de reconnaissance correcte. La qualité des informations obtenues dépend des résultats de la retranscription et de la complexité de la description de l'œuvre. Si le texte est correct et si l'information est contenue dans le texte, celle-ci sera retrouvée par le système.

Le module d'extraction d'information est en grande partie terminé et permet d'obtenir automatiquement les informations correspondantes au (SDI). Il reste à terminer le module de gestion des connaissances ainsi que le module de recherche pertinente.

¹ système descriptif de l'inventaire du patrimoine

² <http://www.nuance.fr/naturallyspeaking/>

³ Xerox Incremental Parser, développé par AïtMokhtar, Chanod et Roux .

4 Indexation semi-automatique à partir du texte transcrit

La connaissance recueillie sur l'œuvre est partielle, car elle est valable dans un laps de temps et ne peut pas être limitée à une grille descriptive figée par un cahier des charges réalisé pour un type d'application donnée.

Notre système doit assurer l'extraction d'informations en rapport avec le SDI, mais en plus la gestion d'un Système à base de connaissances. Les connaissances qui intéressent les chercheurs en patrimoine sont : la manière dont un objet a été fabriqué, par qui, quand, dans quel but, des transformations et des déplacements qu'il a subis, son état de conservation et des matériaux utilisés. Voici quelques concepts que nous-pouvons dégager : Temps, Lieu, Acteur (Personne), Etat de conservation. Certains de ces concepts peuvent être liés les uns aux autres, comme état de conservation et temps, transformations et temps, déplacements et lieu, transformation et personne. La référence dans ce domaine est le modèle conceptuel pour la modélisation des connaissances sur le patrimoine culturel CIDOC-CRM (Doerr, 2006), que nous utiliserons dans notre projet. Le cœur du modèle est constitué de l'entité temporelle exprimant la dépendance entre le temps et les différents événements dans la vie de l'œuvre.

Le SDI permet d'exprimer facilement des informations comme : auteur, appellation, matériaux (...), En revanche l'information qui concerne les différents déplacements que l'œuvre a subis, son impossible à exprimer sauf en texte libre mélangé avec d'autre type d'informations dans le champ historique. La même information peut être sans problème exprimée grâce à l'ontologie CIDOC-CRM que présente la fig 1.

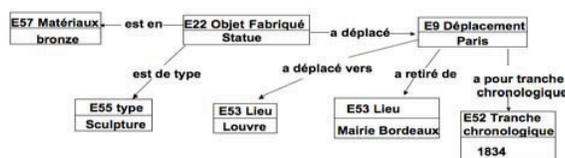


Fig. 1- Exemple de modélisation des connaissances d'une sculpture dans CIDOC-CRM

Le passage de modèle défini par le (SDI) vers l'ontologie CIDOC-CRM se fera grâce à la recherche des correspondances entre les champs du (SDI), dont le contenu peut être considéré comme l'instance d'une des classes de l'ontologie CRM. Pour les cas où cette correspondance ne pourrait pas être assurée, parce que l'information n'existe pas dans le (SDI), il faudra l'extraire du texte retranscrit, à condition que le locuteur ait pris le soin de la dicter, sinon il faudra la saisir au moment de la validation de l'information extraite automatiquement par le système.

5 Conclusion

L'originalité de notre système se situe au niveau du lien qu'il établit entre plusieurs domaines de recherche, comme le traitement du signal, acquisition et modélisation de connaissances, traitement automatique de la langue et management des connaissances. Le module d'acquisition audio permet d'éviter le passage de notes papier vers la saisie clavier de ces notes et donc éviter la perte du temps. Il fournit à l'expert un outil, qui lui permet d'acquérir des connaissances directement observables sur le terrain, qu'il est capable d'interpréter en se basant sur ses connaissances antérieures. La modélisation de la connaissance sous forme d'ontologie et les coopérations ontologiques permettront d'apporter de la souplesse et l'extensibilité ainsi que l'amélioration de l'interrogation requise dans notre domaine de recherche.

References

- Amidon, D. M., P. Formica, and E. Mercier-Laurent (2006). Knowledge Economics Emerging Principles, Practices and Policies, Tartu University Press, volume 2, ch VII.
- Doerr M, Crofts N, Gill T, Stead S, Stiff M (editors) (2006), Definition of the CIDOC Conceptual Reference Model, October 2006.
- Verdier, H. (1999). Système descriptif des objets mobiliers. Paris : Editions du Patrimoine.

Summary

This document presents our work on a definition and experimentation of a voice interface for cultural heritage inventory. This hybrid system includes signal processing, natural language