



**HAL**  
open science

## Capitalisation des connaissances sur l'objet image du patrimoine : acception de partage et de communication par les acteurs

Nabil Khemiri, Sahbi Sidhom, Malek Ghenima

### ► To cite this version:

Nabil Khemiri, Sahbi Sidhom, Malek Ghenima. Capitalisation des connaissances sur l'objet image du patrimoine : acception de partage et de communication par les acteurs. 2nd International Conference on Information Systems and Economic Intelligence - SIIE 2009, Malek GHENIMA (ESCE Université la Manouba - Tunisie) and Sahbi SIDHOM (Nancy Université - France), Feb 2009, Hammamet, Tunisie. pp. 913-931. inria-00426374

**HAL Id: inria-00426374**

**<https://inria.hal.science/inria-00426374v1>**

Submitted on 25 Oct 2009

**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.

# Capitalisation des connaissances sur l'objet image du patrimoine : acceptation de partage et de communication par les acteurs

*Nabil KHEMIRI*

*Doctorant*

Université la Manouba  
Tunisie

[Nabil.Khemiri@yahoo.fr](mailto:Nabil.Khemiri@yahoo.fr)

*Sahbi SIDHOM*

*MCF & Chercheur*

LORIA & Nancy Université  
France

[Sahbi.Sidhom@loria.fr](mailto:Sahbi.Sidhom@loria.fr)

*Malek GHENIMA*

*MA & Chercheur*

ESCE & Université la Manouba  
Tunisie

[Malek.Ghenima@escem.rnu.tn](mailto:Malek.Ghenima@escem.rnu.tn)

## Résumé

Nos propos portent sur la structuration du contenu sémantique pour une banque de données sur les images, dans le but de partager du savoir, du savoir-faire et des connaissances entre les acteurs. Dans ce contexte, le Laboratoire "Régions et ressources patrimoniales de Tunisie" (LRRPT) veut lancer cette acceptation de partage et de communication entre les acteurs du domaine sur l'objet image du patrimoine tunisien.

Dans ce projet, la mise en pratique d'un processus d'intelligence économique était essentielle, dans le but de répondre aux besoins décisionnels du directeur du laboratoire. Nous avons proposé un cadre méthodologique aux acteurs du laboratoire pour capitaliser des ressources informationnelles autour d'un objet complexe : l'image du patrimoine. Sur le plan pratique, nous avons développé un système d'information, apparenté à une photothèque, pour faciliter la disposition des ressources, la recherche d'information sur les contenus et leur enrichissement. Le système est fondé sur une architecture fonctionnelle pour faire communiquer les processus d'indexation, d'annotation et de recherche d'information.

### Mots clés

Système d'information, intelligence économique, capitalisation, indexation, annotation, recherche d'information, image du patrimoine, management des connaissances.

## Abstract

Our subject concerns the structuring of semantic content for image database in order to share knowledge, expertise and know-how between actors. In this context, the Laboratory "Regions and patrimony resources in Tunisia" (LRRPT) wants to release this acceptance of sharing and communication between the actors of the domain and the image object of tunisian patrimony.

In this project, we use the economic intelligence process in order to answer the decision-making needs of the Lab director. The proposal is a methodological framework way for the Lab actors to capitalize information resources around a complex object: the patrimony image. In a practical way, as solution, we proposed an information system, with an interface similar to a digital-library, to facilitate the provision, the information retrieval and the enrichment of resources.

The system is based on a functional architecture to provide the indexing, annotation and information retrieval process.

### Keywords

Information system, business intelligence, capitalization, indexing, annotation, information retrieval, patrimony image, knowledge management.

## 1. Introduction et position du problème

Depuis 2002, le Laboratoire "Régions et ressources patrimoniales de Tunisie" (LRRPT), à la Faculté des Lettres, des Arts et des Humanités, de l'université la Manouba a initié un projet sur les ressources numériques du patrimoine tunisien. Ce projet s'inscrit dans la stratégie de mise en œuvre nationale pour doter l'université tunisienne d'une banque de données du patrimoine couvrant de multiples monuments, objets, composants ou symboles traditionnels et historiques en Tunisie [Larguèche, 2003 et 2008]. A cette initiative, il était question, de réaliser des prises de vue photographiques en numérique sur le patrimoine dans plusieurs sites cibles et régions du pays. La détermination des lieux patrimoniaux couvre les musées, les « souks », les médinas, les lieux de culte, etc. et de s'étendre aux évènements de types folklorique, culturel, traditionnel et religieux.

Etant convaincus de l'importance du projet, le fait de doter les chercheurs et certains départements ministériels, des meilleurs outils de travail, rendra un service considérable aux acteurs associés. En exemple, la proposition de ressources électroniques consultables du patrimoine tunisien peut apporter une réponse au problème. Egalement, une photothèque ouverte aux acteurs du laboratoire et aux usagers enrichira les horizons d'exploitation et des études sur l'objet du patrimoine. C'est par le fait que la photo du patrimoine étant une information historique, culturelle et sociale, permettra aux acteurs concernés de mener des recherches, des investigations ou des études comparatives conjointement avec d'autres ressources autour de l'objet du patrimoine et de sa valeur.

Le cadre de ce projet nous permet donc de formuler le contexte de la recherche qui est lié à la capitalisation des informations sur l'objet patrimoine. Par la médiation informationnelle sur les contenus (ie. mémorisation [Elmasri et Navathe, 1989], indexation et recherche d'informations), le management des connaissances [Gardarin, 2003] sur l'image du patrimoine constituera une mémoire numérique [Serre-Floersheim, 1993] pour les futures générations du pays.

Ainsi, notre problématique on la présente ainsi : « comment faciliter l'exploitation des images du patrimoine tunisien ? ». Cette problématique sous-entend des enjeux, à savoir :

- (i) Comment structurer des informations autour de l'image du patrimoine ?
- (ii) Par quel moyen faciliter la traçabilité des contenus informationnels sur l'image ? et comment se matérialise l'apport des usagers ?
- (iii) Comment rendre possible la recherche d'information dans un contexte d'aide à la décision ?

Le présent article traite de l'élaboration d'une telle problématique : c'est le contexte de la recherche en §2. Des éléments de réponse conséquents pour ses enjeux et le cadre d'application en §3. L'élaboration d'un outil de capitalisation et de managements des connaissances sur l'objet image du patrimoine comme principaux résultats en §4.

## 2. Contexte de la recherche

Dans nos orientations de recherche sur le management des connaissances, nous nous intéressons dans ce projet à l'étude des moyens de diffusion et de valorisation des objets du patrimoine, tout particulièrement sur l'image et son contenu. En effet, la valorisation de l'image du patrimoine peut se faire à travers des informations recueillies dans des documents associés, comme la notice bibliographique et les annotations. Dans une notice bibliographique, on peut collecter des informations construites par l'indexeur humain de type signalétique (titre, production, lieu, date, etc.) et de type analytique ou analyse de contenu (chapeau, résumé, thèmes, descripteurs, etc.). Quant à l'annotation, elle peut collecter des informations objectives ou subjectives émanant de acteurs ou usagers de l'image. Ces annotateurs peuvent être des historiens, chercheurs du patrimoine, experts, architectes, etc. pour enrichir le contenu par leur interprétation, évaluation ou vote.

Par la médiation de l'utilisateur, un premier objectif comporte l'insertion de l'information sur le contenu de l'image. Cette activité est essentielle car, d'une part, l'image n'a pas de valeur interprétative en soi, et d'autre part pour résoudre le problème d'analyse du contenu [Baird, 2003], le recours aux informations de traçabilité [Amadeilh et al., 2005] des différentes lectures s'avère utile. Par la médiation de l'outil, le second objectif se compose de l'identification puis de l'extraction [Allan et al., 2002] de l'information pertinente dans les contenus associés à l'image. Ce processus est approprié pour répondre à des besoins informationnels et pour contribuer à résoudre un problème décisionnel dans un contexte d'intelligence économique (IE).

En se posant les questions du « pourquoi », pourquoi on collecte de l'information sur les contenus images du patrimoine ? , de « qui », qui insère les informations sur l'image ? et du « comment », comment répondre à un besoin informationnel sur l'image ? , nous cherchons finalement à mettre en place une banque de données fonctionnelles pour les acteurs de l'image du patrimoine. Cette ressource de données nécessite des enrichissements [Bruckmann, 2006], directement, par des analyses de contenu de l'indexeur humain (ie. le spécialiste de l'image du patrimoine) et, indirectement, par les annotations des usagers (ie. des acteurs compétents ou non sur la valeur de l'image du patrimoine). Ainsi, il est question d'améliorer la recherche d'information (RI) dans le système sur les contenus avec l'apport des informations à valeur ajoutée. La valeur ajoutée n'est autre qu'une information tangible créée par l'acteur dans son contexte. Le contexte en question ne peut viser que l'environnement décisionnel cherché par l'acteur ou l'expression de son besoin informationnel.

Afin de mieux expliquer notre méthodologie de travail, nous présentons dans le contexte d'IE un processus structuré qui va nous permettre de réussir l'intégration d'un tel processus pour répondre au mieux à notre problématique. Dans ce contexte, nous retenons parmi les définitions celle de [David et Thiery, 2001], qui considèrent que le recours à l'IE s'impose quand il s'agit d'étudier des processus impliqués dans la production des indicateurs interprétables pour la prise de décision, tout en se fondant sur des informations internes et externes à l'organisation. Dans le processus d'IE, une clarification sur la définition du problème décisionnel est faite pour extraire les besoins informationnels des acteurs impliqués. Les besoins informationnels vont invoquer une démarche de veille à l'intérieur du processus d'IE.

Ainsi, le processus d'IE adopté est celui issu des travaux de l'équipe de recherche SITE du LORIA [SITE, 2008]. Il se présente comme suit :

- 1- Identification d'un problème décisionnel ;
- 2- Traduction du problème décisionnel en un problème de recherche d'informations ;
- 3- Identification des sources pertinentes d'information ;
- 4- Collecte des informations pertinentes ;
- 5- Analyse des informations collectées pour extraire des indicateurs pour la décision ;
- 6- Interprétation des indicateurs ;
- 7- Prise de décision ;

Ce processus comporte plusieurs problématiques dont celles qui intéressent notre travail :

- (i) la problématique liée au processus de collecte d'information qui couvre les phases (en partie 2- puis 3- et 4-), impliquant principalement les veilleurs et les concepteurs du système d'information ;
- (ii) la problématique liée à la pertinence de l'information qui couvre les phases (3-, 4-, et en partie 5-), -, impliquant principalement les veilleurs ;
- (iii) la problématique liée au processus décisionnel qui couvre les phases (1-, 2-, et 6-, 7-), dont l'acteur principal est le décideur ;

Dans ce cadre méthodologique en fondant l'approche sur le processus d'IE, notre contribution est d'apporter des réponses au problème décisionnel du directeur du Laboratoire LRRPT. Elle est sujette de la mise en place d'un système d'information (SI) pour rendre possible l'accès, le partage et l'enrichissement des ressources du patrimoine. Le SI va prendre en compte les attentes des acteurs internes et externes du laboratoire qui sont dérivées de leur besoin informationnel. C'est ce que nous comptons développer en terme de cahier des charges orienté IE pour le décideur dans la section suivante.

### **3. Cadre d'application de l'IE**

#### **3.1. Caractérisation de l'enjeu décisionnel**

Une méthodologie d'IE s'inscrit dans un projet qui comprend un ensemble d'actions que l'organisation réalise, *dans les meilleures conditions* [Jakobiak, 2001] *de délais et de coûts*, pour être en mesure de résoudre des problèmes décisionnels. Certes, tous les problèmes ne sont pas identiques et donc ne nécessitent pas des interventions de même nature pour l'organisation.

Or, de la perception du problème par le directeur du laboratoire (ie. décideur) à la mise en chantier d'un projet par les acteurs et associés (ie. veilleurs, analystes et concepteurs de système d'information [Volle, 2001], [Juillet, 2007]), il existe une phase très importante en amont du problème [Bouaka, 2004] : dans cette phase, il est question, d'une part, que le décideur définisse le problème en terme d'enjeu et, d'autre part, que les acteurs et associés s'approprient du problème, qu'ils le traduisissent en indicateurs, puis en problème de

recherche d'information et enfin en objectifs de recherche. Cette traduction permettra aux acteurs et associés du projet du décideur d'obtenir :

1. une contextualisation de la demande du décideur,
2. une évaluation entre l'objectif du décideur et une formulation explicite de la demande,
3. une objectivation des buts à atteindre.

Au terme de cette acception, on peut expliciter l'enjeu décisionnel par l'expression d'une règle déclarative à trois prédicats : objet, signal et hypothèse, selon [KISLIN, 2005] :

- **Objet** : ce qui caractérise l'objet de l'environnement d'une organisation sur lequel on peut agir ;
- **Signal** : ce qui provoque le problème décisionnel ;
- **Hypothèse** : ce qui correspond au risque conséquent dans la mesure de l'inertie face au problème.

A l'issu de cette caractérisation, l'enjeu est formulé par la règle suivante :

**Si { non Action\_sur(objet) et Etat\_de(signal, actif) } Alors Risque\_de(hypothèse) ;**

Dans la continuité de cette définition de l'enjeu, nous avons également clarifié d'autres enjeux concernant le veilleur et celui du concepteur du système. Le recouvrement des trois enjeux avec objet, signal et hypothèse propres à chaque type d'acteurs qu'on arrive à cerner la complexité multidimensionnelle du projet et de l'objet d'étude en IE.

Afin d'illustrer nos propos, nous allons présenter le cadre d'application de cette approche IE sur notre projet pour construire un cahier des charges spécifique.

### **3.2. Cahier des charges orienté IE pour le décideur**

En terme de cahier des charges « décisionnel », l'identification des attentes du directeur du laboratoire non explicitées de sa part en tout début nous a demandé de cumuler une double activité de veilleur et de concepteur d'un système d'information. Tout d'abord, nous avons explicité la demande du directeur du laboratoire pour répondre au mieux à son besoin décisionnel. Ensuite, c'est la phase de recueil des données pour formaliser et ainsi envisager des solutions adéquates au problème posé en terme d'enjeux, à savoir :

#### **❖ Contexte général :**

Le Laboratoire **LRRPT** a initié le projet, depuis 2002, sur la capitalisation des ressources du patrimoine tunisien. Ce projet s'inscrit dans la stratégie de mise en œuvre d'un dispositif pour les étudiants, les enseignants et les chercheurs sur l'objet du patrimoine. Les ressources numériques seront rendues accessibles par une banque de données pour la valorisation et la diffusion des richesses matérielles et immatérielles du patrimoine tunisien.

#### **❖ Environnement :**

Les publics concernés sont universitaires (enseignants, chercheurs, étudiants) et non universitaires (professionnels du patrimoine, de l'architecture, de l'art ou culture, etc.) de la banque des ressources numériques du patrimoine.

❖ **Formulation initiale :**

Par besoin d'interopérabilité, le directeur du Laboratoire **LRRPT** désire savoir comment rendre accessible prioritairement aux enseignants, aux chercheurs et aux étudiants, les ressources du patrimoine afin de participer à la diffusion et à la valorisation de l'objet patrimoine tunisien.


❖ **Enjeu décisionnel :**

- **Objet :** procédé (méthode et technique) pour capitaliser des ressources dans une banque de données numériques sur l'image du patrimoine tunisien.
- **Signal :** niveau d'existence insuffisant, sans une véritable méthodologie de capitalisation des connaissances sur l'image du patrimoine et sans outils pour fédérer les actions des différents acteurs du laboratoire.
- **Hypothèse :** si le directeur ne trouve pas un procédé pour rendre accessible au public (universitaire et autres) une banque d'images du patrimoine, alors il ne pourra pas diffuser, partager et enrichir des connaissances sur l'objet du patrimoine pour les futures générations.

❖ **Formulation finale :**

Pour doter l'université d'une banque de données couvrant de multiples informations sur le patrimoine, le directeur du laboratoire cherche prioritairement une méthode et une application interopérable pour rendre accessible au public universitaire les ressources du patrimoine culturel tunisien et leur valorisation par les usagers.

❖ **Informations à traiter sur l'objet du patrimoine :**

<ul style="list-style-type: none"> <li>- <b>Nom :</b> Ardoise</li> <li>- <b>Propriété intellectuelle :</b> LRRPT</li> <li>- <b>Type image :</b> JPEG</li> <li>- <b>Taille image :</b> 1024x1544 pxls.</li> <li>- <b>Source image :</b> photo numérique</li> <li>- <b>Résolution image :</b> 72 dpi</li> <li>- <b>Date d'insertion-base :</b> 13-05-2007</li> </ul> <p style="text-align: center;"><b>Exemple 1 :</b> Propriétés de l'image du patrimoine : "Ardoise".</p>	 <p style="text-align: center;"><b>Figure1 : (LRRPT ©)</b> Ardoise pour l'apprentissage du Coran.</p>
<ul style="list-style-type: none"> <li>- <b>Titre :</b> Ardoise</li> <li>- <b>Date de la prise de vue :</b> 23/04/2005</li> <li>- <b>Auteur :</b> N. SAADI</li> <li>- <b>Résumé :</b> Ardoise utilisée pour l'apprentissage du Coran</li> </ul>	<ul style="list-style-type: none"> <li>- <b>Annotateur :</b> N. KHEMIRI</li> <li>- <b>Objectif :</b> remarque</li> <li>- <b>Section annotée :</b> image</li> <li>- <b>Type annotation :</b> texte</li> <li>- <b>Annotation :</b> cette ardoise en bois est</li> </ul>

<ul style="list-style-type: none"> <li>- <b>Photographe</b> : N. SAADI</li> <li>- <b>Résumé du photographe</b> : Ardoise utilisée pour l'apprentissage du Coran</li> <li>- <b>Référence</b> : Cd 2PKEF</li> <li>- <b>Mots clés</b> : enseignement, Coran, kuttab,</li> <li>- <b>Localisation géographique</b> : Kef</li> <li>- <b>Sous localisation</b> : Kef</li> <li>- <b>Catégorie du patrimoine</b> : Muséal</li> <li>- <b>Sous catégorie du patrimoine</b> : Musée des Arts et des traditions populaires</li> <li>- <b>Style/école</b> :</li> <li>- <b>Epoque</b> : Médiévale</li> <li>- <b>Fabriquant</b> : Ecole coranique de la grande mosquée de Kef</li> <li>- <b>Matériaux</b> : planche en bois (50x30cm)</li> </ul> <p style="text-align: center;"><b>Exemple 2 :</b> Notice bibliographique associée à l'image "Ardoise".</p>	<p>utilisée pour l'apprentissage du Coran selon la méthode traditionnelle : l'écriture des paroles, la répétition et le par cœur pour mémoriser.</p> <ul style="list-style-type: none"> <li>- <b>Liens</b> : ardoise2.jpg, eleve1.jpg, ecriture1.jpg</li> </ul> <p style="text-align: center;"><b>Exemple 3 :</b> Annotation sur l'image "Ardoise".</p>
---	---

#### ❖ **Objet du patrimoine :**

- **Les monuments** : œuvre architecturale, sculpture ou peinture de monument, élément ou structure à caractère archéologique, inscription, grotte et groupe d'éléments qui ont une valeur universelle exceptionnelle du point de vue de l'histoire, de l'art ou de la science ;
- **Les ensembles** : groupe de constructions isolées ou réunies, qui en raison de leur architecture, de leur unité, ou de leur intégration dans le paysage, ont une valeur universelle exceptionnelle du point de vue de l'histoire, de l'art ou de la science ;
- **Les sites** : œuvres de l'homme ou œuvres conjuguées de l'homme et de la nature, ainsi que les zones y compris les sites archéologiques qui ont une valeur universelle exceptionnelle du point de vue historique, esthétique, ethnologique ou anthropologique.

#### ❖ **Solution envisagée :**

Le système d'information qu'on se propose de construire, prendra en compte des informations formelles sur les propriétés physiques de l'image (cf. Figure1 et Exemple1), des informations de contenu dans la notice bibliographique (cf. Exemple2). Quant aux informations informelles, elles sont issues des annotations (cf. Exemple3) des usagers du système (ie. les remarques, points de vues, votes, etc. des annotateurs). Ces annotations peuvent être étendues aux différents types de document : l'image, la notice et les annotations.

La conception d'un SI ouvert aux ressources informationnelles sur l'image du patrimoine a été motivée par le caractère stratégique qu'impose l'actuelle organisation du laboratoire. Dans ce domaine, la tendance actuelle de l'ensemble des tâches exige, de manière symétrique, la conception d'un système d'information intégré. Cela implique la maîtrise d'une information très détaillée, dense et hautement intégrée, et une conception adaptée avec des ressources de gestion : « dans le processus de développement du système d'information, la conception est la tâche la plus créative mais aussi la plus difficile » [Rolland et al, 1988], [AFOLABI, 2007].



## ❖ **Exploitation : projet de la photothèque du patrimoine**

Dans notre projet, l'utilisateur du système consulte les images de patrimoine, ce type de document source, en lui-même, ne renvoie pas à des informations symboliques comme dans un texte écrit. Il y a donc un obstacle pour l'accès immédiat à l'information de contenu. C'est pourquoi l'apport de l'indexeur humain pour ce type de document est incontournable dans le but de restituer une part d'une lecture « rationnelle » et « objective » sur le contenu de l'image du patrimoine.

Certes, l'information formelle n'enlève pas les ambiguïtés sur le contenu lors d'une consultation et reste une ressource limitée pour un processus de recherche d'information. Ce type de ressource ne renvoie pas toutes les informations pour répondre aux divers besoins des usagers. C'est en se tournant vers l'enrichissement des ressources par les annotations qu'on peut y remédier et permettre de lever les ambiguïtés tout en valorisant le point de vue de l'utilisateur et ses interprétations sur le contenu.

Dans la section suivante, nous détaillons l'intégration des objets d'analyse [Blaha et Premerlani, 1998] de l'acteur lors du processus d'indexation avec les objets interprétatifs du processus d'annotation afin de capitaliser des connaissances sur les ressources images du patrimoine et ainsi enrichir le processus de recherche d'informations dans ses résultats.

## **4. Principaux résultats du projet**

### **4.1. Processus d'indexation :**

La consensus à considérer dans un processus d'indexation automatique est, d'une part, d'éviter les erreurs humaines liées aux aspects subjectifs d'une indexation intellectuelle, le décalage dans cette pratique entre les indexeurs et, d'autre part, d'accélérer le processus sur de gros volumes de données tout en fondant l'indexation par similarité à une pratique quasi-humaine et régulière dans ses résultats pour relever les concepts clés d'un contenu.

Le choix d'un processus d'indexation est dépendant de plusieurs critères tels que: le choix de l'approche selon le type du document à indexer, la précision attendue d'un système d'indexation et le contexte interprétatif d'un contenu qui est nécessairement lié au processus de recherche d'information.

Le corpus d'images a été construit par des étudiants en master, des historiens et des experts dans le domaine du patrimoine tunisien au laboratoire LRRPT. Les notices bibliographiques servent à représenter l'indexation de l'image par des informations signalétiques et par des propriétés sur l'analyse de contenu, comme: le résumé, les mots-clés, la localisation géographique, la catégorie du patrimoine, l'époque, les thèmes, etc.

Dans ce projet, le processus d'indexation implémenté consiste à relever des informations sémantiques sur l'image [Brini et al., 2005], [Cottin, 2004], spécifiquement dans des champs textuels propres à l'analyse du contenu. Dans cette approche, nous avons procédé par filtrage de concepts sémantiques (ie. termes et multi-termes) à l'aide d'une grammaire régulière en éliminant les mots asémantiques à partir d'un dictionnaire de mots vides, ie. les mots outils de la langue, (cf. Figure 2). Certes, l'approche est limitée, mais dans l'avenir nous comptons

étendre la grammaire à celle des syntagmes nominaux [Sidhom, 2002] plus riche d'un point de vue sémantique en utilisant un dictionnaire complet de la langue. Dans cette première version, Le filtrage permet d'alimenter le fichier inverse, ie. liste de couples (terme, documents associés), comme source de connaissances sur les documents indexés : images, notices bibliographiques, annotations.

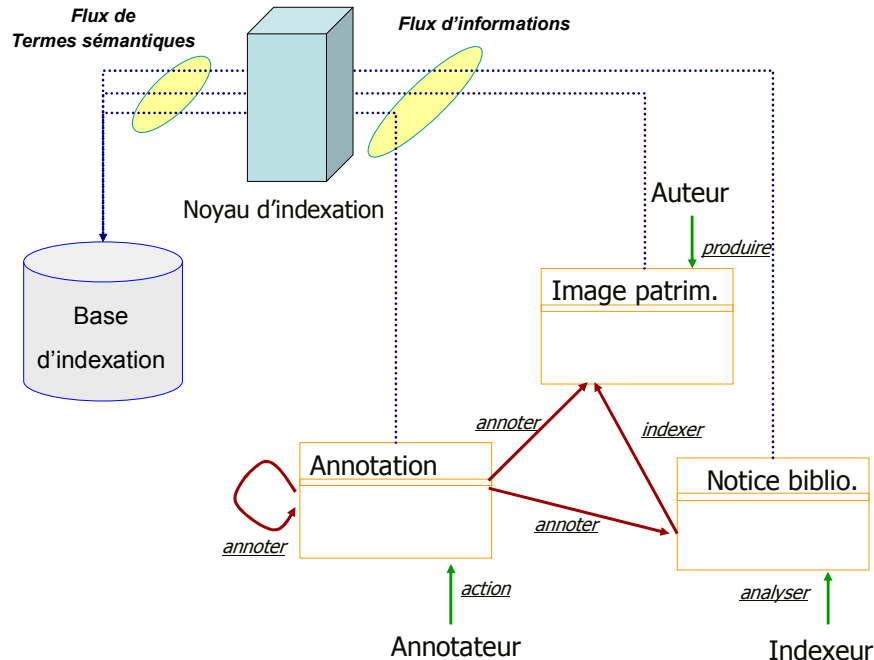


Figure 2 : Processus d'indexation [Sidhom, 2008].

En relation avec l'indexation dans les notices, l'information annotée nous a servi à la lecture des sources (images du patrimoine) et à l'identification des contenus pertinents : c'est une valorisation faite par l'implication des usagers.

#### 4.2. Processus d'annotation :

Dans un contexte pratique, l'annotation d'un contenu, quelque soit la typologie du document associé, est une trace de l'activité intellectuelle sur ce dernier. Cette activité peut témoigner le passage d'un usager (ie. l'annotateur) pour rajouter des informations explicites ou implicites, comme l'expression de son opinion, de son interprétation ou de son évaluation d'un contenu. La matérialisation de l'annotation [Ciravegna et al., 2002] peut être étendu du message écrit, à la parole (audio), audiovisuel (parole et image), etc.

Pour notre projet, nous retenons seulement l'annotation comme message écrit afin de faciliter la connexion avec le processus d'indexation et finaliser son exploitation comme ressource d'indexation ou valeur ajoutée sur un contenu.

Cette activité d'annotation est riche d'un point de vue sémantique et qui peut être dotée d'un outil adapté à des fins de partage d'informations, de mémorisation des connaissances ou d'aide à la décision. L'objet annotation porte donc, d'une part, sur les actions de l'annotateur par rapport au contenu et, d'autre part, sur le contenu d'annotation en terme de valeur informationnelle. Il nous semble, de ce point de vue, que le développement du processus

d'annotation sur l'image du patrimoine est incontournable à l'enrichissement des activités collaboratives entre les acteurs du laboratoire.

De ce fait, la valeur ajoutée sur un contenu informationnel est nécessairement induite du rapport entre l'action de l'annotateur et son objet annotation [Sidhom, 2008]. Sur l'image du patrimoine, c'est un apport riche par l'implication de l'utilisateur en terme de : (i) prospective exploratoire de l'information, (ii) pérennisation dynamique sur l'objet du patrimoine et (iii) inachèvement prolongé de l'information par l'option de récursivité dans l'annotation.

L'annotation peut être d'origine interne ou externe : les acteurs internes ou associés du laboratoire, les usagers internes ou externes de l'université et ceux qui se connectent à notre application comme système d'information sur le patrimoine tunisien (cf. Figure 3).

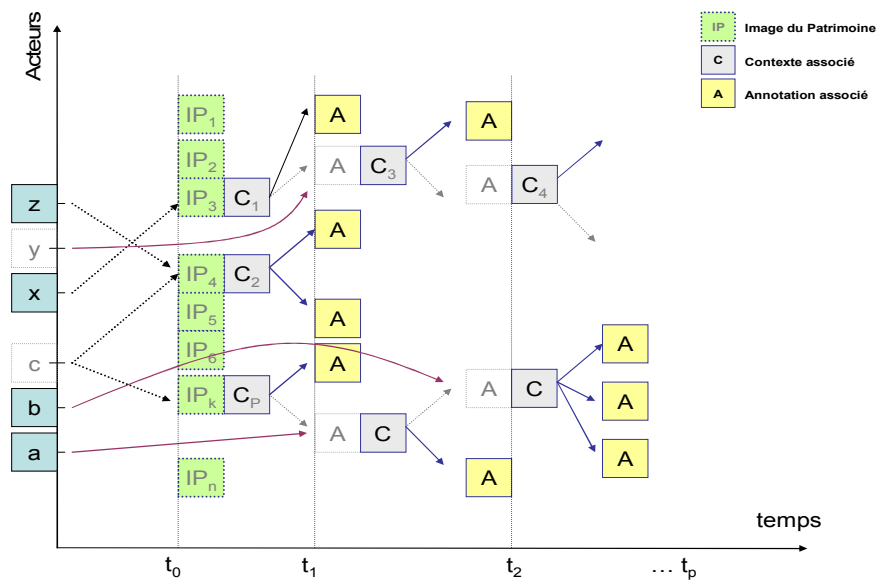


Figure 3 : Processus d'annotation et la relation (Image, Contexte, Annotation).

Dans une approche d'automatisation de l'annotation, l'explicitation des fonctions à attribuer est essentielle pour relier (a.) le contexte à l'exploitation de l'information annotée et pour relier (b.) l'objet annotation au profil de son auteur. Les propriétés, ainsi définies par [Sidhom et David, 2006], [Robert, 2007], permettent de clarifier le concept annotation, à savoir :

- le contexte d'annotation : pour créer une nouvelle annotation ou le suivi d'une ancienne, tout en spécifiant s'il s'agit d'une requête, d'une recherche d'information, d'une interprétation, etc. ;
- la source à annoter : de spécifier de quel document à annoter (l'image du patrimoine, la notice bibliographique associée ou l'annotation d'un acteur du laboratoire) ;
- le profil de l'annotateur : il est question de créer ou de rappeler le profil de l'annotateur en activité et de spécifier ses actions antérieures (annoter, rechercher des informations, formaliser une requête, etc.) ;
- l'annotation : elle comporte l'objet annotation et ses propriétés ;

Dans cette considération, l'exploitation des contenus d'annotation passe par le même principe d'indexation pour alimenter la base d'indexation du système.

Un tel avantage, dans les propriétés du concept annotation, va servir le processus de recherche d'information à mieux répondre aux besoins informationnels des acteurs du laboratoire.

### 4.3. Processus de Recherche d'Informations :

Aussi bien l'indexation que l'annotation, les deux processus antérieurs contribuent à l'alimentation des connaissances pour la recherche d'information (RI). Trouver l'information pertinente est une opération de qualité [Boucher et al., 2005], [Schen et al., 2005], qui reste délicate au cœur des activités du management des connaissances et en intelligence économique.

Dans le processus RI, on utilise une fonction d'appariement [Dufournaud et al., 2000], [Veltkamp et Tanase, 2000] qui extrait les mots sémantiques de la requête de l'utilisateur, sur les mêmes principes de base du processus d'indexation, puis de les comparer aux termes stockés dans la base d'indexation. Le résultat de l'appariement permet de proposer les ressources concernées : les images du patrimoine, les notices et les annotations associées.

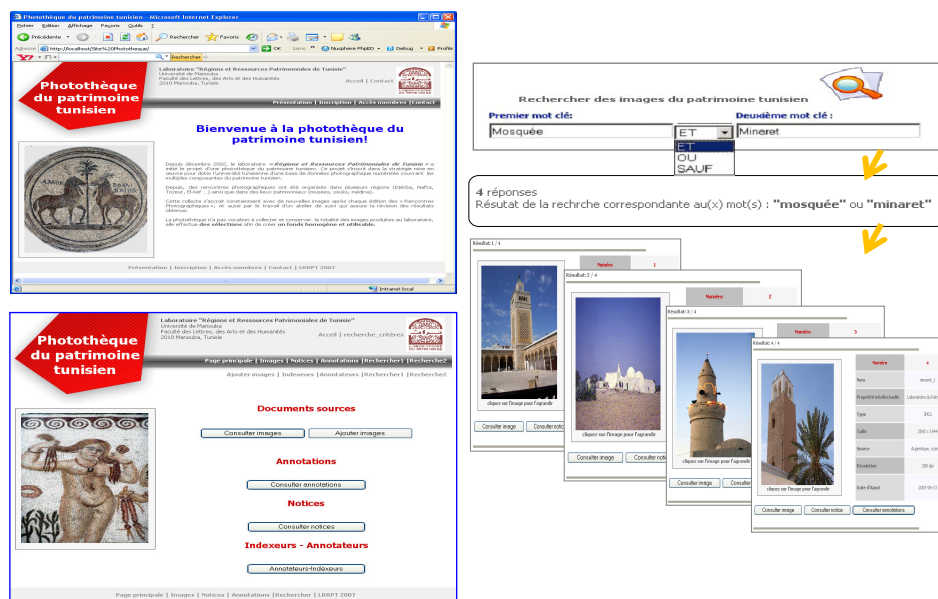


Figure 4 : Fonctionnalités de l'application - photothèque du patrimoine [KHEMIRI, 2007].

Pour nous, le noyau de RI est construit dans le but de répondre aux mieux à un besoin informationnel de l'utilisateur. En effet, son rôle principal est de sélectionner les documents adéquats à la formulation de la requête. On passe de l'exploration des champs spécifiques des documents aux champs d'indexation dans les notices puis aux éléments à valeurs ajoutées dans les annotations (cf. Figure 4).

Dans la logique opératoire, la requête de l'utilisateur porte sur un contenu sémantique à retrouver dans la banque de données [Batini et al., 1992]. Cela nécessite la détermination des concepts clés [Rastier, 1999] dans la requête qui peuvent être partagés par les concepts des documents ou sémantiquement proches. Egalement, il est question de l'implémentation des fonctions d'appariement pour trouver et mesurer la distance sémantique entre les concepts

selon des traits morphologiques ou syntaxiques et les difficultés posées par la synonymie, la polysémie et l'homonymie.

L'avantage dans une telle acception, l'intégration des champs sources, notices et annotations dans l'indexation, pourra étendre les fonctionnalités d'un processus RI dans l'avenir, en passant de l'image aux documents multimédia du patrimoine.

## **5. Conclusion et perspectives**

La pratique d'une méthodologie d'IE était nécessaire dans ce travail pour répondre au problème décisionnel du directeur du laboratoire LRRPT. La solution qui consiste à capitaliser des connaissances sur l'image du patrimoine tunisien n'était pas si évidente en tout début de l'expression du besoin. Egalement, l'idée de se projeter vers la conception d'un SI qui facilite l'accès, le partage et le management des connaissances sur les contenus de l'image du patrimoine avec l'implication des acteurs [Tamine and Bahsoun, 2006] qui n'était du tout dans le pré-requis du projet.

Dans ce projet, nous avons proposé une solution pour une problématique qui touche à la fois la recherche d'information et l'aide à la décision sur des objets du patrimoine. Les solutions apportées permettent au laboratoire de disposer d'une méthodologie d'IE et d'outils opératoires afin de valoriser les activités autour de l'objet patrimoine : les clarifications apportées dans le cahier des charges orienté IE. Ce cahier des charges constitue l'interface de communication entre le directeur du laboratoire, les acteurs associés dans le projet et le concepteur du SI pour caractériser les enjeux du problème décisionnel.

Dans ce laboratoire, le directeur et ses collaborateurs enseignants-chercheurs, professionnels-experts, doctorants et post-doctorants partagent les mêmes ressources informationnelles et contribuent à la valorisation du patrimoine. Les ressources du patrimoine sont diverses et le recours au support numérique s'avère incontournable pour constituer une banque de données. L'outil proposé a été construit dans l'optique pour trouver les ressources et capitaliser les connaissances autour de l'objet image du patrimoine.

Ainsi, l'orientation vers une architecture SI [Cauvet et Rosenthal, 2001] ouverte sur le web était dans l'objectif de disposer des ressources tout en assurant l'intégrité des informations par rapport à la contribution des acteurs. L'outil d'indexation fondé sur l'approche des mots sémantiques de la langue en connexion avec les processus d'annotation et de recherche d'information a contribué au management des contenus : l'interface du SI s'apparente à une photothèque du patrimoine (cf. Figure 4).

Au stade du développement actuel du processus d'annotation, l'amélioration de l'outil sera nécessaire pour supporter en l'occurrence de nouveaux types d'annotation : de l'écrit à la parole ou à l'image (ie. l'annotation multimédia). Egalement, la banque de données du patrimoine peut s'ouvrir dans l'avenir sur des nouvelles ressources : de l'image à la séquence multimédia du patrimoine.

Certes, dans une première version du SI et de ses outils testés, les contributions et les améliorations ne sont pas arrêtées et ne le seront pas. Le laboratoire dans son environnement continue à développer de nouvelles ressources et à observer de nouveaux besoins.

## Remerciements :

Cet article fait référence à un travail réalisé collaborativement avec le laboratoire LRRPT sous la direction du Pr. Abdelhamid Larguèche, que nous remercions ainsi que les membres contributeurs pour la récupération des données sur l'image du patrimoine.

## 6. Bibliographie

- [Afolabi, 2007] AFOLABI B. (2007). La conception et l'adaptation de la structure d'un système d'intelligence économique par l'observation des comportements de l'utilisateur, Thèse, Sciences de l'Information et de la Communication, Université Nancy 2, France.
- [Allan et al., 2002] J. Allan et al. (2002). Challenges in information retrieval and language modeling. In Workshop held at the center for intelligent information retrieval, September 2002.
- [Amadeilh et al., 2005] F. Amadeilh, P. Laublet, and J.-L. Minel. (2005). Document annotation and ontology population from linguistic extractions. In K-CAP '05: Proceeding of the 3rd international conference on knowledge capture, page 161-168, New York, NY, USA, 2005. ACM Press
- [Baird, 2003] BAIRD H. (2003). Digital Libraries and Document Image Analysis, International Conference on Document Analysis and Recognition (ICDAR), vol. 1, 2003, pp. 2-14.
- [Batini et al., 1992] BATINI, C., Ceri, S., Navathe, B. (1992) « Conceptual Database Design – An Entity- Relationship Approach », Benjamin/Cummings, 1992.
- [Blaaha et Premerlani, 1998] BLAHA, M., Premerlani, J. (1998). « Object-Oriented Modeling and Design for Database Applications », Prentice-Hall, 1998.
- [Bouaka, 2004] Bouaka N. (2004). Développement d'un modèle pour l'explicitation d'un problème décisionnel: un outil d'aide à la décision dans un contexte d'intelligence économique, Thèse, Sciences de l'Information et de la Communication, Université Nancy 2, France
- [Boucher et al., 2005] Boucher A., Ha. T.D, Lan L.T (2005), «Classification vs recherche d'information : vers une caractérisation des bases d'images », 12èmes Rencontres de la Société Francophone de Classification (SFC), mai/juin 2005, Montréal (Canada).
- [Brini et al., 2005] A. Brini, M. Boughanem et D. Dubois. (2005). A model information retrieval based on possibilistic network. In Proceeding of the symposium on String Processing and Information Retrieval (SPIRE), pages 271-282, 2005.
- [Bruckmann, 2006] Bruckmann, Denis (2006) « L'image fixe », BBF, 1993, n° 5, p. 56 [en ligne] <<http://bbf.enssib.fr>> Consulté le 24 novembre 2006.
- [Cauvet et Rosenthal, 2001] Cauvet C., Rosenthal-Sabroux C. (2001). Ingénierie des systèmes d'information. Hermes Sciences Publications.
- [Ciravegna et al., 2002] F. Ciravegna, A. Dingli, Y. Wilks, and D.Petrelli. (2002). Amilcare: adaptive information extraction for document annotation. In SIGIR '02: Proceedings of the 25th annual international ACM SIGIR conference on Research redundancy, page 367- 368, New York, NY, USA, 2002. ACM Press.
- [Cottin, 2004] Stéphane Cottin. (2004). « Méthode d'indexation de fonds d'images fixes - catalogage des photographies », 2004, [en ligne] <<http://www.servicedoc.info/Methodes-d-indexation-de-fonds-d.html> > Consulté le 24 novembre 2006
- [David et Thiery, 2001] David, Amos et Thiery, Odile. (2001). Prise en compte du profil de l'utilisateur dans un Système d'Information Stratégique. In Veille Stratégique Scientifique et Technologique - VSST'2001.
- [Dufournaud et al., 2000] Y. Dufournaud, C. Schmid, and R. Horaud. (2000). Appariement d'images à des échelles différentes. In 12ème Congrès AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, Paris (France), Fév. 2000.

- [Elmasri et Navathe, 1989] ELMASRI, R., Navathe, S.B. (1989). « Fundamentals of Database Systems ». The Benjamin/Cummings, Redwood City. 1989.
- [Gardarin, 2003] Gardarin, G. (2003). Bases de données : Objet et relationnel. Eds. Eyrolles. Avril 2003.
- [Jakobiak, 2001] JAKOBIAK, F. (2001). « L'intelligence économique en pratique », éditions d'organisation, 2001.
- [Juillet, 2007] Alain Juillet. (2007). Intelligence économique: « Nous manquons cruellement d'outils informatiques d'origine française ou européenne ». in 01 DSI (n° 10) [INTERVIEW par : Nicolas Arpagian, le 01/03/2005].
- [Larguèche, 2003] Abdelhamid Larguèche. (2003). « Patrimoine ethnologique entre muséification et folklorisation ». in Colloque Dynamiques culturelles, patrimoine et mondialisation. 29 septembre - 05 octobre 2003, Avignon.  
(source d'information: <http://calenda.revues.org/nouvelle3203.html>, Février 2008)
- [Larguèche, 2008] Abdelhamid Larguèche. (2008). « Histoire, Patrimoine et Enjeux identitaires en Tunisie ». in l'Année du Maghreb, Aix, 2008
- [Rastier, 1999] Rastier, F. (1999). De la signification au sens – pour une sémiotique sans ontologie. In revue texto, 1999.
- [Robert, 2007] Robert, C. (2007). L'annotation pour la recherche d'information dans le contexte d'intelligence économique, Thèse, Sciences de l'Information et de la Communication, Université Nancy 2, France
- [Schen et al., 2005] X. Shen, Tan B. et Zhai J. C. (2005). Context-sensitive information retrieval using implicit feedback. In Proceeding of the 29th annual international ACM SIGIR Conference on Research and development in Information retrieval, pages 43-50, August 2005.
- [Serre-Floersheim, 1993] SERRE-FLOERSHEIM, Dominique. (1993). Quand les images vous prennent au mot ou comment décrypter les images. Paris: Éd. d'Organisation Université.
- [Sidhom et David, 2006] SIDHOM Sahbi, DAVID Amos. (2006). Automatic indexing of multimedia document as starting point to annotation process. In : 9th International ISKO Conference on Knowledge Organization for a Global Learning Society. Ed. Ergon Verlag, 4-7 July 2006 Vienna, Austria.
- [Sidhom, 2002] SIDHOM, Sahbi. (2002). " Plate-forme d'analyse morpho-syntaxique pour l'indexation automatique et la recherche d'information: de l'écrit vers la gestion des connaissances.", Thèse de Doctorat à l'Université Claude Bernard Lyon1, France, Mars 2002.
- [Sidhom, 2008] SIDHOM, Sahbi (2008). Approche conceptuelle par un processus d'annotation pour la représentation et la valorisation de contenus informationnels en intelligence économique (IE). In actes de la conférence internationale (SIIE'2008) sur les Systèmes d'information et l'Intelligence économique. 13-15 février 2008 à Hammamet (Tunisie).
- [Tamine and Bahsoun, 2006] L. Tamine et W. Bahsoun. (2006). Définition d'un profil multi-dimensionnel de l'utilisateur : Vers une technique basée sur l'interaction entre dimensions. In Actes de la 3ème Conférence en Recherche d'Information et Application (CORIA), Page A paraître, Lyon Mars 15-17 2006.
- [Veltkamp et Tanase, 2000] VELTKAMP R., TANASE M. (2000). « Content-Based Image Retrieval Systems: A Survey », rapport no UU-CS-2000-34, 2000, Dep. of Computing Science, Utrecht University.
- [Volle, 2001] Volle M. (2001). Economie des nouvelles technologies, Chapitre 11, Quelques indications sur les systèmes d'information. in Economica 2001.