

Ergonomie de la Recherche d'Information (RI) sur le Web : quel modèle de l'activité ?

Manuel Zacklad, Jean-Pierre Cahier, Nadia Gauducheau, Hedi Zaher, et al.

ISTIT (FRE CNRS 2732) /Equipe Tech-CICO

Ecole d'été du GDR TICS, Batz sur mer : 2005

Université de Technologie de Troyes

Présentation

- Équipe pluridisciplinaire Technologie de la Coopération pour l'Innovation et le Changement Organisationnel (UTT)
- Trois axes :
 - Modèles et méthodes pour l'ingénierie et la gestion des connaissances collectives
 - Environnements pour la coopération (Computer Supported Cooperative Work)
 - Analyse linguistique, sociologique et gestionnaire des usages collectifs des TIC

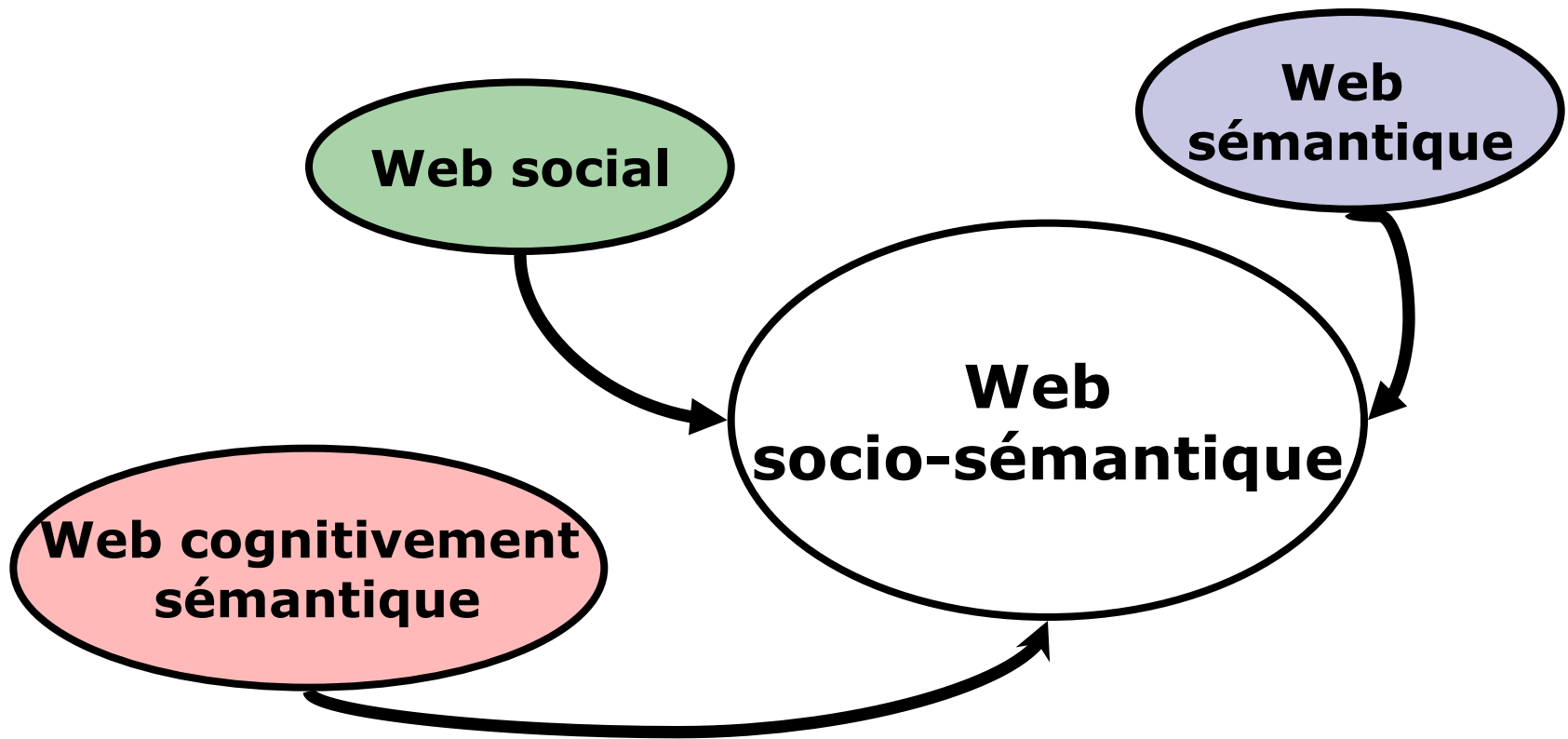
Actions de recherche personnelles

- Théorie des communautés d'action & TIC
- Théorie transactionnelle de l'action
- Gestion & Économie des services (approche régulationniste) & communication
- Stratégies de gestion des connaissances (basées sur le pragmatisme)
- Documents pour l'action et documentarisation (coopération et productions sémiotiques)
- Ingénierie, usages et ergonomie du Web Socio Sémantique
- Domaines d'applications divers :
 - Innovation de service et gestion des connaissances dans les grandes entreprises de réseaux (GDF, EDF, FT...) et l'ingénierie collaborative
 - Travail en réseau au sein de communautés « bottom up » : réseau de santé, usage des TIC au sein des PVD, lien avec le DD...

Grands usages du Web

- Communication synchrone, adressée, directe : chat, téléphonie...
- Communication asynchrone adressée ou non : mail, forum...
- **Communication asynchrone non adressée**: weblog, articles divers, journaux (production sémiotiques), catalogues (description d'entités physique)...

Positionnement du Web Socio Sémantique



Web Socio Sémantique

- Un Web qui s'adresse à des communautés d'utilisateurs poursuivant des objectifs similaires
- Un Web social qui participe de la construction d'une représentation structurée du domaine et du collectif
- Une structuration progressive des réseaux sémantiques gérés par le collectif (présentant des enjeux pour le réseau social lui-même)
- Approche qui peut être pertinente sur des marchés électroniques notamment dans le cadre de « produits complexes »
- Importance des activités de **recherche d'information**

Deux approches dominantes pour le soutien aux activités de « recherche d'information »

- Les moteurs de recherche se généralisent en offrant un accès direct au contenu
 - Versus :
- Des systèmes « d'annuaires » ou de « catalogues » plus ou moins formalisés (ontologies) offrant une représentation cartographique de l'information (« cartes de thèmes »)
 - Selon cette approche nous développons à Tech-CICO :
- Un système de gestion collaborative « schémas de classification » ou « d'ontologies sémiotiques » (*Agorae*), basé sur une proposition de norme de représentation de l'information (*HyperTopic*) et offrant une représentation cartographique des « thèmes » (topics)

Objectif de notre programme de recherche

- Méthodologie de conception d'annuaires – catalogues - listes basés sur des ontologies sémiotiques et étude de la pertinence de ce mode d'accès à l'information sur le Web
- Une technologie simple du point de vue algorithmique mais complexe et importante du point de vue des enjeux cognitifs et sociaux associés: construction de listes, de hiérarchies, des liens transverses, de représentations mi-textuelles mi-graphique de l'information...

Exemple de l'interface d'Agorae dans le contexte de l'expérience réalisée

UTT - Catalogue des UVs

Consulter

[Accéder à Structurer et Contributer](#)

[Accueil](#) > [Point de vue](#)

Consigne 0 : Dans le cadre d'un stage d'été, vous souhaitez renforcer vos compétences en choisissant une UV parmi celles proposées à l'UTT. Pendant ce stage vous devrez assurer la Communication pour des événements organisés par Conservatoire de Troyes, et en particulier faire preuve de bonnes connaissances en musicologie. Quelle UV vous serait nécessaire ?

1- SAVOIRS FONDAMENTAUX (6)

- 1- Gestion, organisation et économie (20)
- 2- Informatique, réseaux et télécommunications (9)
- 3- Autres Sciences de l'ingénieur (5)
- 4- Mathématiques et Logique (5)
- 5- Autres Sciences sociales et Humanités (16)
- 6- Sciences de la Nature et de la Vie (3)

3-METIER et FONCTION (14)

- 01- Management général de l'entreprise (1)
- 02- Métiers commerciaux (3)
- 03- Conception et développement du produit (10)
- 04- Logistique, Gestion de Production et Productique (4)
- 05- Métiers de l'Informatique et de l'Organisation (10)
- 06- Ressources Humaines (6)
- 07- Services financiers
- 08- Communications d'entreprise interne et externe (3)

2- METHODES, TECHNIQUES et OUTILS (6)

- 1- Méthodes de l'ingénieur (12)
- 2- Techniques managériales et commerciales (6)
- 3- Techniques financières (3)
- 4- Outils et techniques industrielles (7)
- 5- Outils et techniques informatiques (4)
- 6- Méthodes adressant les aspects humains et organisationnels (21)

4-PROJET (3)

- 1- Projet lié à l'entreprise (20)
- 2-Projet personnel - Savoir-être (11)
- 3- Projet social (4)

La Recherche d'Information (RI) du point de vue des sciences de l'information et de l'ergonomie cognitive

Contexte des activités de RI

- Nature de l'information recherchée, deux cas de figure :
 - Les entités recherchées sont elles-mêmes de nature sémiotique et sont de plus en plus directement accessibles (plus ou moins partiellement) à travers « le système de recherche » : l'utilisateur peut les expérimenter (recherche de document, musique... l'entité est « l'information » – **RI directe**)
 - Les entités recherchées sont de nature artefactuelle et ne sont que pointées par le système de recherche : l'utilisateur ne peut directement les expérimenter – achat de produit technique, de voyage... – **RI indirecte** : c'est l'entité « matérielle » qui est recherchée à travers sa description (une information sur l'entité)

Connaissances préalables de l'utilisateur

- Deux cas de figures :
 - RI **limitée** : l'utilisateur connaît exactement l'entité recherchée et souhaite essentiellement la « localiser »
 - RI **ouverte** : l'utilisateur ne sait pas quelle entité peut répondre à son besoin -> paradoxe de la RI
 - La RI permet **à la fois** d'affiner le besoin (voire de remettre en cause celui-ci) et de trouver des entités y répondant
 - Conclusion d'un cycle de RI :
 - l'entité est jugée satisfaisante vis-à-vis du besoin (succès a priori)
 - pas d'entité trouvée répondant au besoin (mais persistance de la croyance en l'existence de l'entité)
 - remise en cause du besoin

<p>Accès direct ou non à l'entité -></p> <p>Ouverture de la recherche</p>	<p>RI directe (on recherche de l'info pour elle-même)</p>	<p>RI indirecte (on recherche de l'info sur un objet non exclusivement informationnel)</p>
<p>RI fermée (on connaît l'entité et on cherche sa localisation)</p>	<p>Où trouver ce « produit informationnel » X, comment le télécharger ?</p>	<p>Où acheter ce « produit » X, comment le commander ?</p>
<p>RI ouverte (on ne sais pas quelle entité est susceptible de répondre au besoin – voire si elle existe)</p>	<p>Quel est le « produit informationnel » qui peut répondre à mon besoin ? (voire le type de produit)</p>	<p>Quel est le « produit » qui peut répondre à mon besoin ? (voire le type de produit)</p>

Approche de la psychologie cognitive : la RI vue comme une activité de RP (modèle de Rouet et Tricot)

- Le processus **d'évaluation** a pour fonction d'identifier les informations qui manquent pour effectuer la tâche (information(s) cible(s)) -> représentation du but, comparaison de l'information disponible au but et élaboration des critères qui vont guider la RI. Certains critères sont déclaratifs (domaine de la RI) et d'autres procéduraux (actions à réaliser sur le système).
- Le processus de **sélection** de l'information pertinente correspond à une prise de décision en référence aux critères élaborés dans la phase évaluative -> l'utilisateur calcule une valeur d'intérêt pour chaque catégorie d'information qui aboutit ou pas à l'examen de cette catégorie d'information.
- Le processus de **traitement** de l'information recherchée met en jeu principalement une activité de lecture -> compréhension de texte à laquelle viennent s'ajouter d'autres tâches (décision quand à la pertinence du contenu, etc.).
- Les cycles se succèdent jusqu'à l'atteinte du but. Des mécanismes de gestion cognitive de l'activité accompagnent ces processus (planification, contrôle, régulation).

Modélisation de l'activité de RI

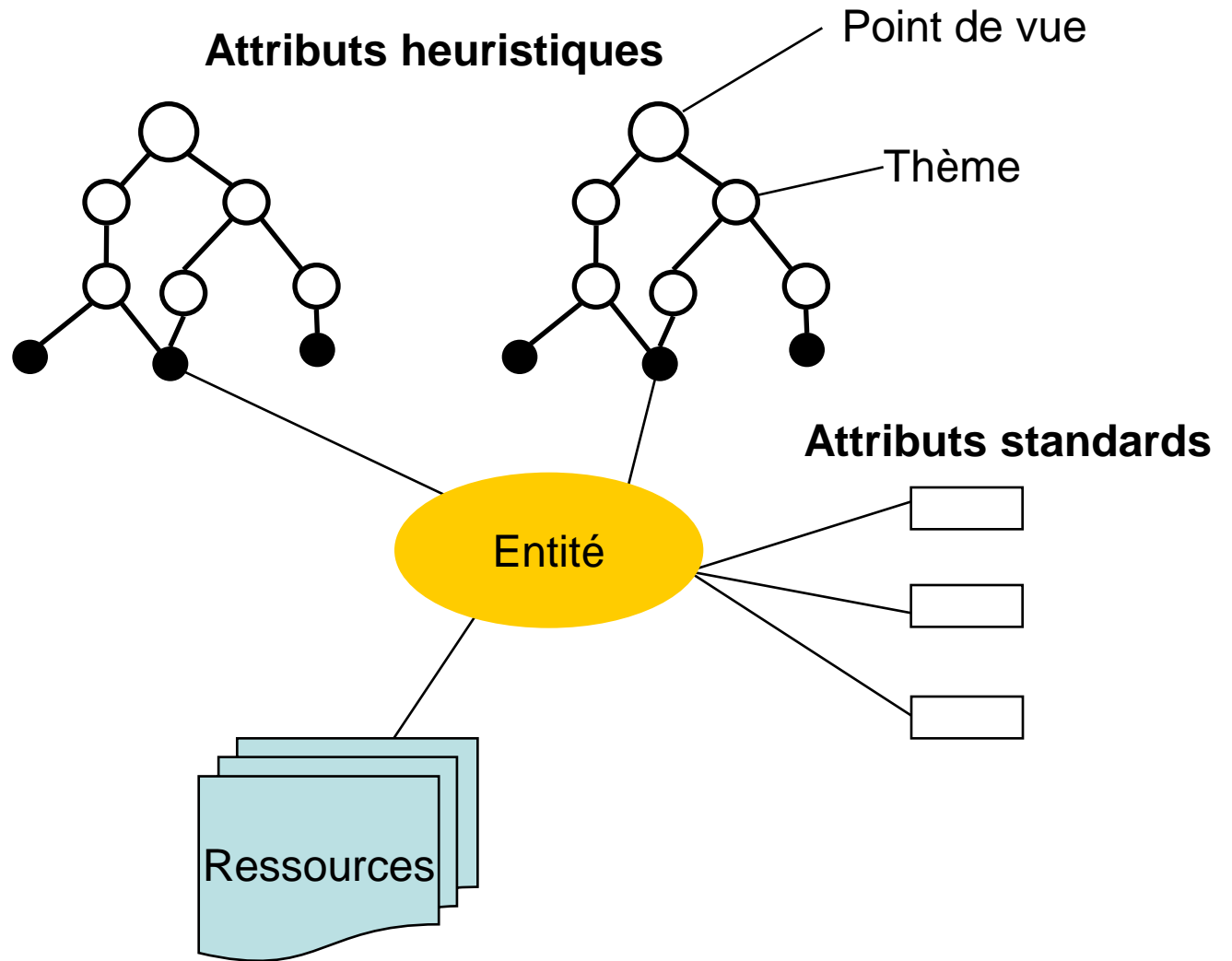
- Une modélisation à deux niveaux « des représentations » associées à l'activité (approche typique en ingénierie des connaissances)
- Modélisation de niveau « épistémique » de l'information: entités, attributs, ressources informationnelle...
- Sur ce premier niveau description une modélisation orientée domaine d'activité: situation d'usage, caractéristiques fonctionnelles, technique, extraits du bien informationnel...

Intégration de la RI dans un cycle de recherche/consommation global

- Identification d'un besoin en référence à des situations sociales (type d'entité)
- Recherche de l'entité par stratégie de résolution de problème spécifique impliquant de la manipulation d'information (RI spécifique)
- Usage effectif de l'entité (lecture d'un livre, utilisation de l'appareil photo...)
- Évaluation différée du succès de la RI (ex-post)

Zoom sur la phase de RI proprement dite

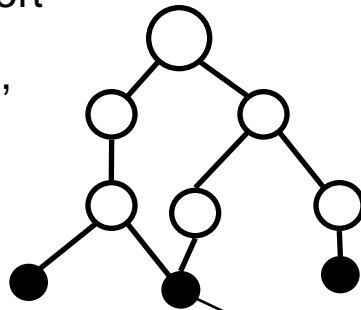
- **Situations sociales** -> **activités** «médiatisées»
- **Activités** médiatisée -> recours à un média (entité) = objet/ instrument/ artefact, dans diverses situations **d'usage** (« besoins » liés à l'instrument)
- **Situations d'usages** (besoin) -> **caractéristiques fonctionnelles** (confort d'utilisation, fiabilité technique, innovation, design, etc...) [accès à l'entité]
- **Sous ensemble de caractéristiques fonctionnelles** -> **caractéristiques techniques** intrinsèques et extrinsèques (services) et administratives [accès à l'entité]
- **Accès direct** à des **extraits du contenu** d'une entité (expérience partielle) ou **accès indirect** à des images des descriptions détaillées, des comptes rendus d'utilisation
- **Sélection** de l'entité (évaluation ex-ante)
- **Évaluation ex-post** par l'usage



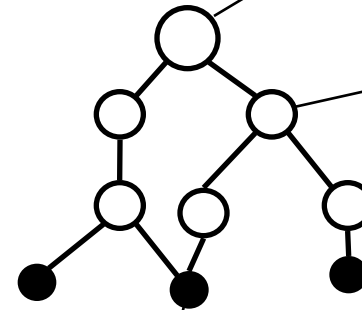
Langage HyperTopic

Caractéristiques fonctionnelles (confort d'utilisation, fiabilité technique, innovation, design, etc...)

Attributs heuristiques

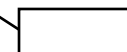
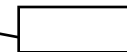
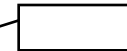
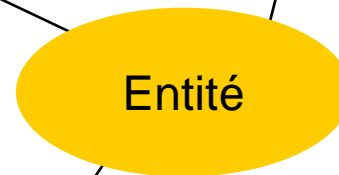


Point de vue

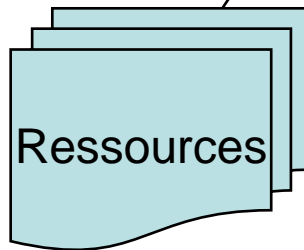


Thème

Attributs standards



Accès direct à des **extraits du contenu** d'une entité (expérience partielle) ou **accès indirect** à des images, des descriptions détaillées, des comptes rendus d'utilisation

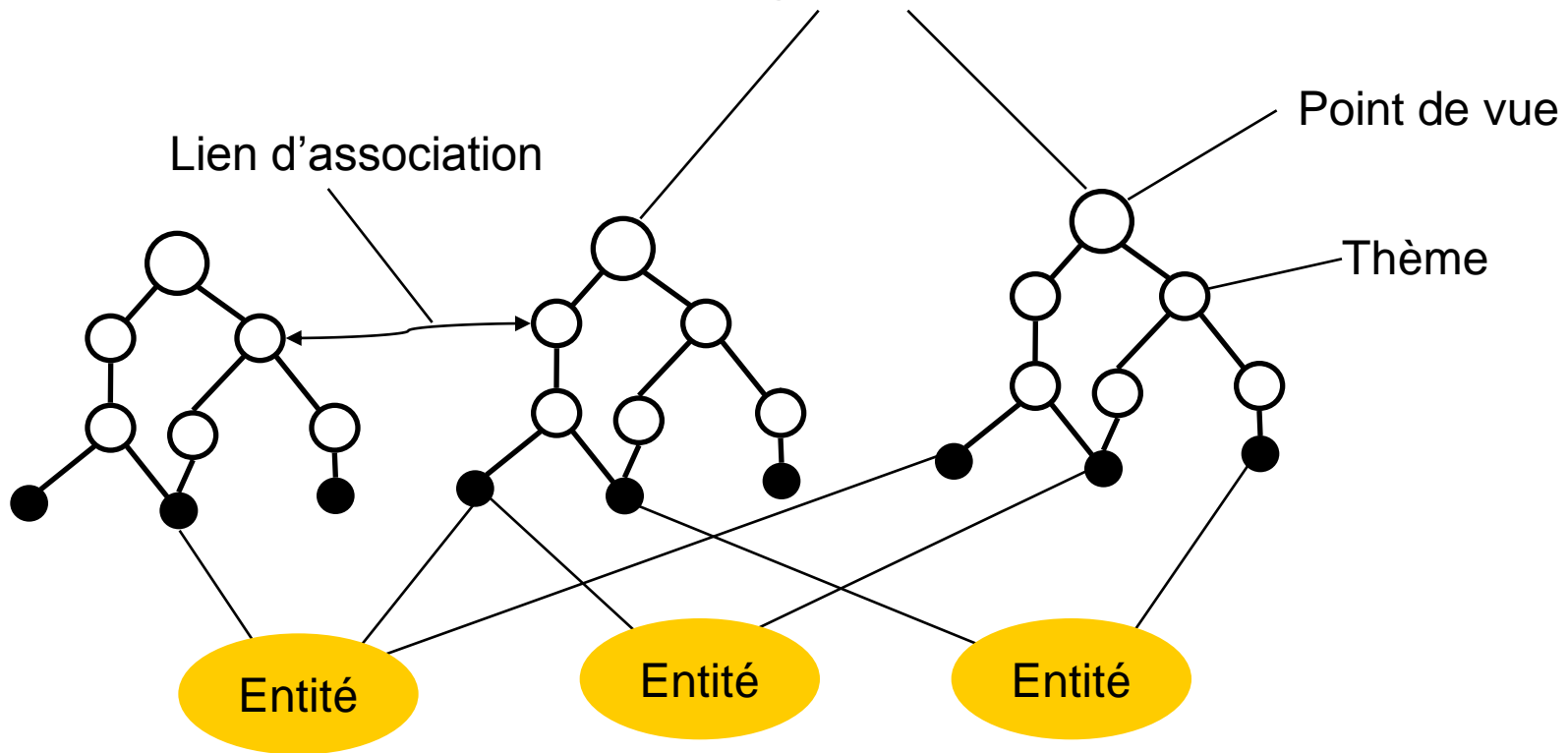


Caractéristiques techniques intrinsèques et extrinsèques (services) et administratives

Modélisation adaptée au domaine d'activité

Attributs heuristiques (ontologies sémiotiques)

Chaque point de vue représente une classe de préoccupation mais aussi un groupe d'intérêt différent

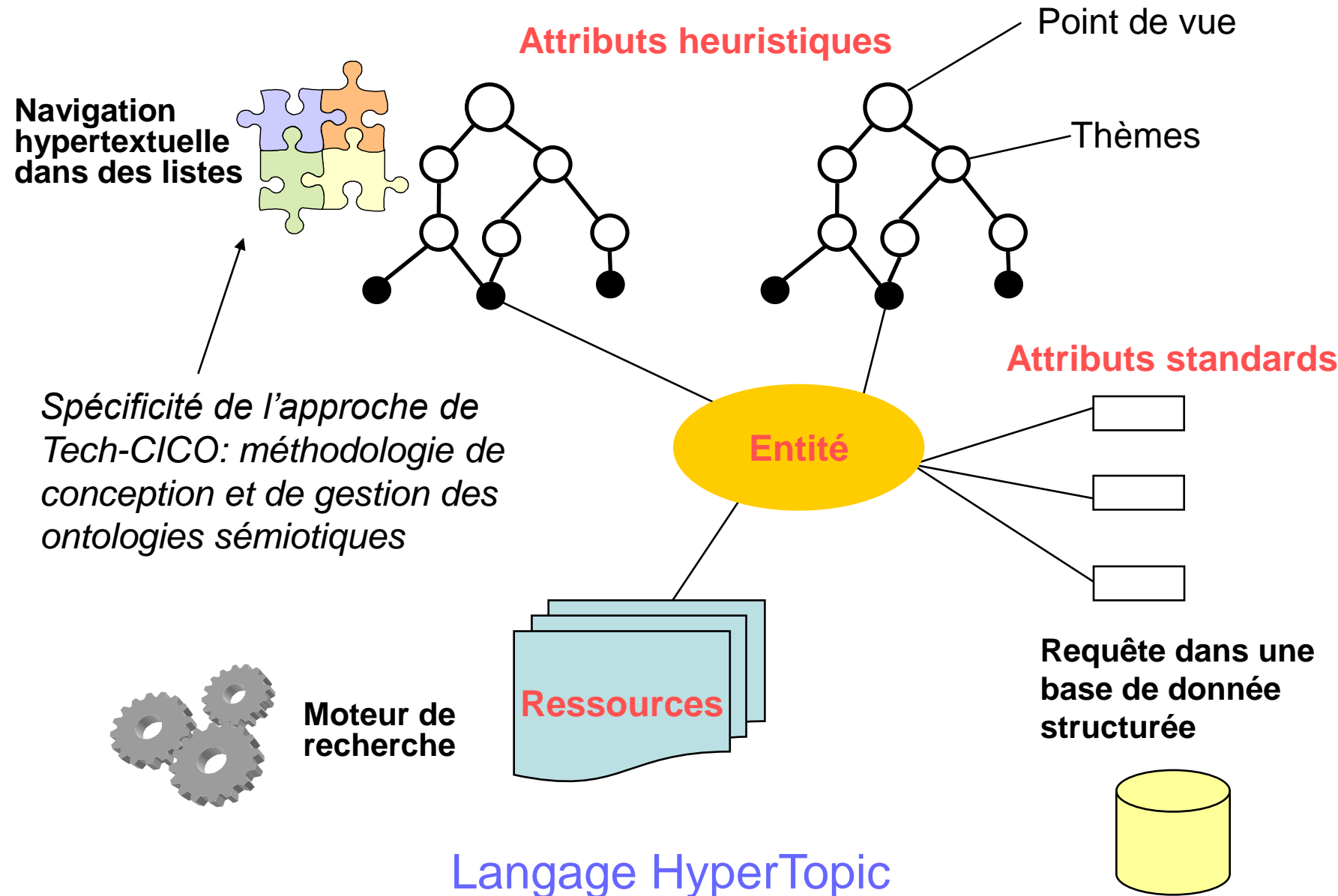


Description des entités

- **Attributs standards:** critères reconnus de manière consensuelle représentés dans des schémas entités associations ou des ontologies formelles (pour l'interopérabilité entre bases de données)
- **Attributs heuristiques:** gérés dans des listes plus ou moins hiérarchisées -> thésaurus, ontologies sémiotiques...
- **Ressources** (description directe ou indirecte): URI, documents, images...

Trois familles de technologies associées aux modes de description de l'entité

- **Attributs standards**: requête dans des bases de données ou inférence dans des ontologies formelles
- **Ressources** (description directe ou indirecte): moteur de recherche, technique d'analyse linguistiques
- **Attributs heuristiques** : ou navigation « manuelle » souvent « hypertextuelle » dans des listes ou des cartes (organisés sous la forme d'ontologies sémiotiques ou schémas de classification)

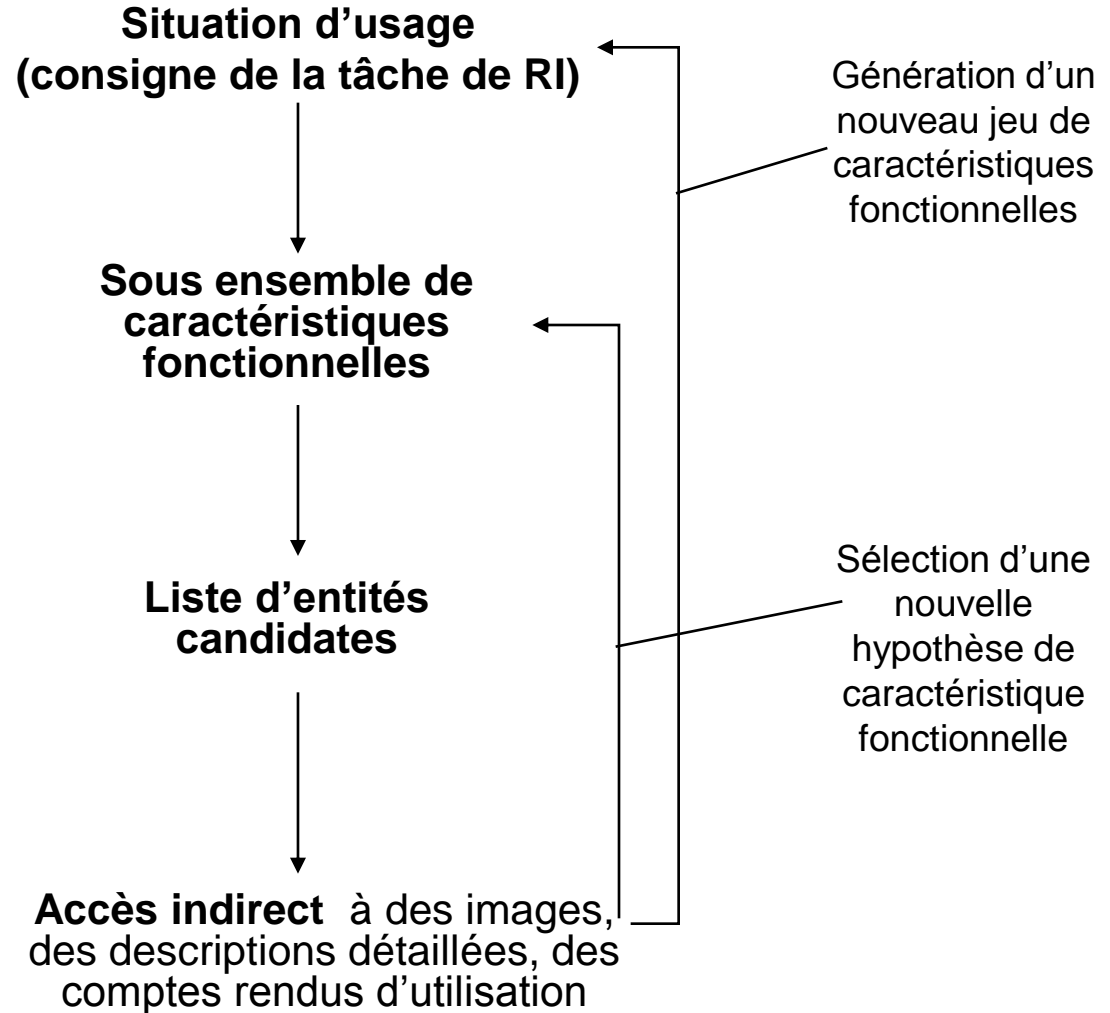
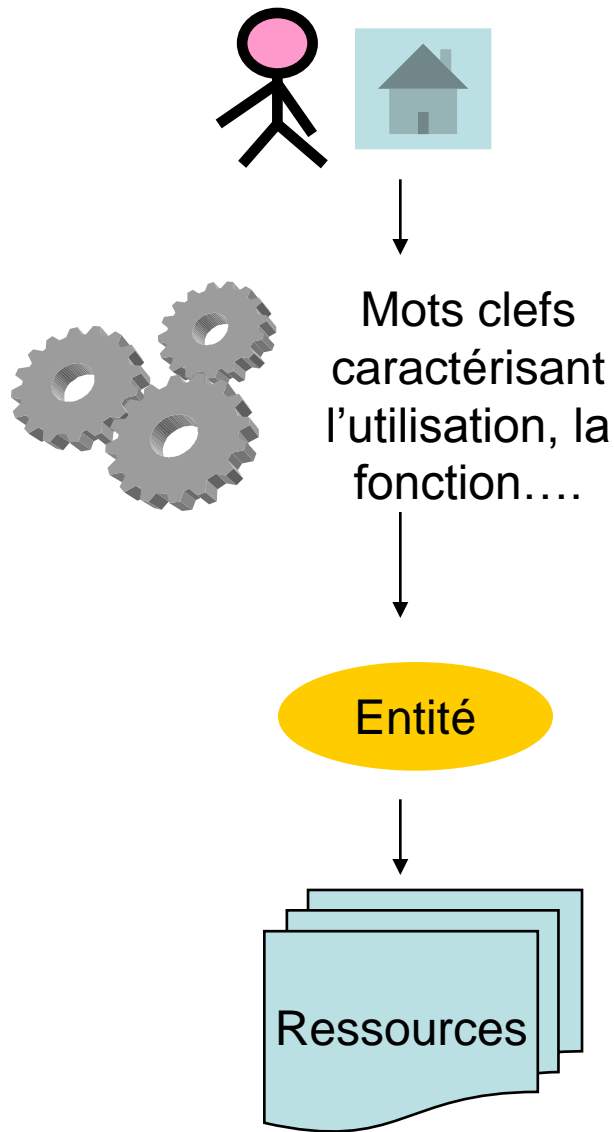


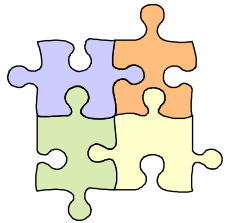
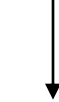
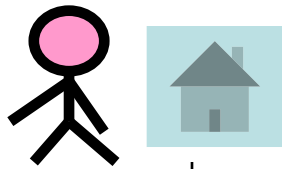
Questionnement : comparaison moteur vs ontologies sémiotiques

- L'utilisation d'ontologies sémiotiques demande un « investissement définitionnel » important. Y a-t-il un retour sur cet investissement ? (par rapport à un moteur)
- Quels types de représentations graphiques sont les plus adaptées pour la navigation?
- Quels types de processus cognitifs situés & distribués sont mobilisés pour la recherche de l'entité?
- Quels type de connaissances sont requises pour l'usage du système? Quels apprentissages se mettent en place?
- Comment faire gérer le système de classification par le collectif? Comment le faire évoluer par la délibération.

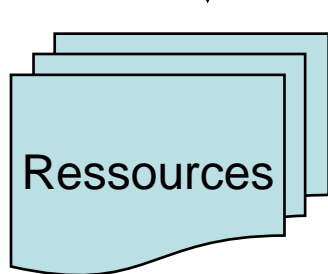
Principe de l'étude

- Comparaison de l'efficacité de la recherche effectuée :
 - à l'aide d'un moteur de recherche
 - versus
 - des schémas de classification multi-points de vue
- sur la base d'un corpus contrôlé





Schémas de classification accessibles via une carte de thème Agorae



Situation d'usage (consigne de la tâche de RI)



Sous ensemble de caractéristiques fonctionnelles

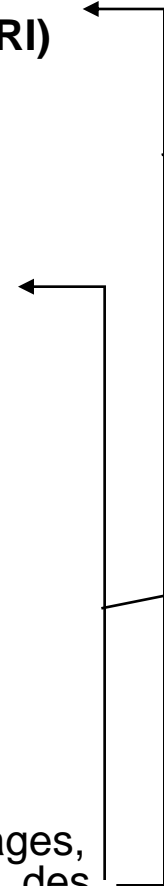


Nom de l'entité



Accès indirect à des images, des descriptions détaillées, des comptes rendus d'utilisation

Génération d'un nouveau jeu de caractéristiques fonctionnelles



Sélection d'une nouvelle hypothèse de caractéristique fonctionnelle

UTILISATION D'AGORAE (ontologies sémiotiques)

Hypothèses sur les avantages des schémas de classification

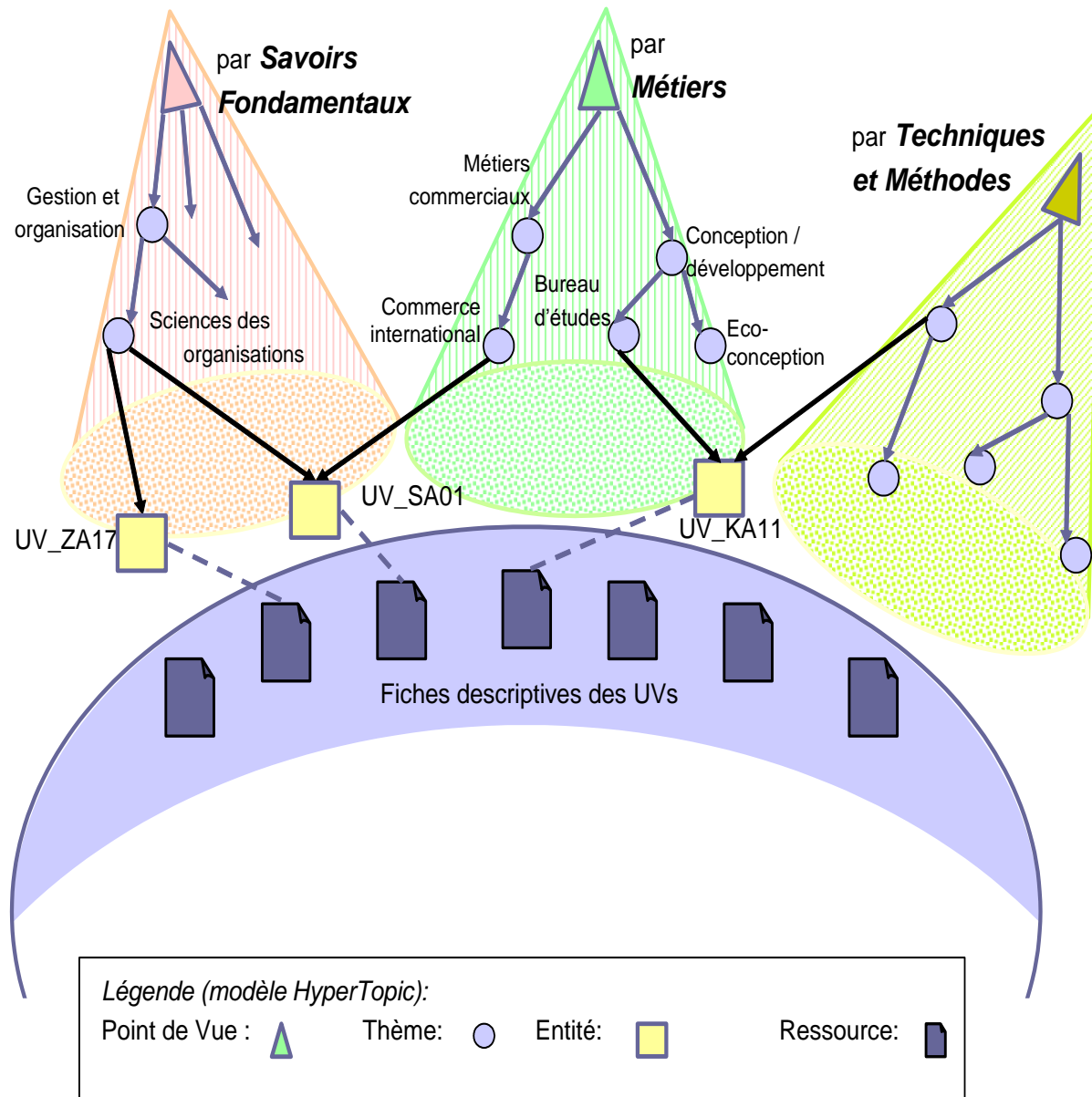
- Les schémas de classification permettent de bénéficier des l'investissement définitionnel effectué par les communautés d'intérêt qui les ont réalisé
- Ils favorisent l'apprentissage « méta-informationnel » (points de vue sémantiques défendus par différents groupes – une composante de l'expertise spécifique en RI)
- Ils favorisent l'exploration et la découverte du réseau sémantique à l'intérieur de chaque point de vue (p.e caractéristiques fonctionnelles mises en avant selon chaque point de vue)

Hypothèse sur les avantages de la représentation graphique des cartes de thèmes

- Externalisation de la représentation du réseau sémantique : mise en évidence des thèmes et de leurs associations
- Explicitation des relations de subsomption hiérarchique et d'association...
- ...correspondant à la mise en évidence d'un plan de recherche (arbre de décision) pour la sélection des entités pertinentes

Protocole de l'expérience

- Vingt-deux étudiants de l'université de technologie de Troyes
- Tâche : choix d'une UV dans un catalogue informatisé
- Utilisation de l'outil Agorae vs utilisation d'un moteur
- Une seule UV permet à chaque fois de répondre à la demande
- Temps limité de 30 minutes pour effectuer la tâche



UTT - Catalogue des UVs

Consulter

[Accéder à Structurer et Contributer](#)

[Accueil](#) > [Point de vue](#)

Consigne 0 : Dans le cadre d'un stage d'été, vous souhaitez renforcer vos compétences en choisissant une UV parmi celles proposées à l'UTT. Pendant ce stage vous devrez assurer la Communication pour des événements organisés par Conservatoire de Troyes, et en particulier faire preuve de bonnes connaissances en musicologie. Quelle UV vous serait nécessaire ?

1- SAVOIRS FONDAMENTAUX (6)

- 1- Gestion, organisation et économie (20)
- 2- Informatique, réseaux et télécommunications (9)
- 3- Autres Sciences de l'ingénieur (5)
- 4- Mathématiques et Logique (5)
- 5- Autres Sciences sociales et Humanités (16)
- 6- Sciences de la Nature et de la Vie (3)

3-METIER et FONCTION (14)

- 01- Management général de l'entreprise (1)
- 02- Métiers commerciaux (3)
- 03- Conception et développement du produit (10)
- 04- Logistique, Gestion de Production et Productique (4)
- 05- Métiers de l'Informatique et de l'Organisation (10)
- 06- Ressources Humaines (6)
- 07- Services financiers
- 08- Communications d'entreprise interne et externe (3)

2- METHODES, TECHNIQUES et OUTILS (6)

- 1- Méthodes de l'ingénieur (12)
- 2- Techniques managériales et commerciales (6)
- 3- Techniques financières (3)
- 4- Outils et techniques industrielles (7)
- 5- Outils et techniques informatiques (4)
- 6- Méthodes adressant les aspects humains et organisationnels (21)

4-PROJET (3)

- 1- Projet lié à l'entreprise (20)
- 2-Projet personnel - Savoir-être (11)
- 3- Projet social (4)



Search	
<input type="text"/>	<input type="button" value="Go"/> display <input type="text" value="10"/> <input type="button" value="▼"/> results
<input checked="" type="radio"/> and operator	<input type="radio"/> exact phrase <input type="radio"/> or operator

Please fill the search form field.



Consigne 0 : Dans le cadre d'un stage d'été, vous souhaitez renforcer vos compétences en choisissant une UV parmi celles proposées à l'UTT. Pendant ce stage vous devrez assurer la Communication pour des événements organisés par Conservatoire de Troyes, et en particulier faire preuve de bonnes connaissances en musicologie. Quelle UV vous serait nécessaire ?

[Voir le panier](#)

Résultats

	Nombre de réponses fournies		Nombre de réponses correctes		Taux de réussite (réponse correctes / réponses fournies)	
	Moyenne	Ecart type	Moyenne	Ecart type	Moyenne	Ecart type
Agoræ	8,1	2,2	4,3	2	0,53	0,2
Moteur	11,3	1	4	1,3	0,35	0,1

↑
Différence significative

Résultats

- En temps limité :
 - Les utilisateurs du moteur fournissent plus de réponse mais plus souvent inexactes
 - Les réponses fournies grâce aux schémas de classification sont plus précises mais le temps de recherche est plus long

Conclusion : Coopération médiatisée par le Web Socio Sémantique

- Échanger et permettre la constitution d'ontologies sémiotiques multipoints de vue
- Définition d'une sémantique partagée
- Contribuer à la structuration du collectif autour d'un réseau sémantique partagé
- Comprendre les mécanismes cognitifs et sociaux associés à cette organisation