

IFT-66529 Ingénierie des connaissances

Nicole Tourigny, Ph.D.

Département d'informatique et de génie logiciel

Université Laval

Hiver 2003

© Tourigny (2003)

IFT-66529 Ingénierie des connaissances

- Jean CHARLET, Gilles KASSEL, Manuel ZACKLAD et Didier BOURIGAULT, « Ingénierie des connaissances: recherches et perspectives », *Ingénierie des connaissances. Évolutions récentes et nouveaux défis*, Eyrolles, Paris, 2000, pp. 1-22.

© Tourigny (2003) IFt-66529 IC

2

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- Motivation
 - Point sur les recherches françaises
 - <http://www.irit.fr/GRACO/>
 - Bulletin de l'AFIA
- Objectifs
 - Définir IC
 - Proposer une synthèse
 - Points de repère sur le domaine et ses évolutions
 - Principaux acquis
 - Réflexions
 - Nouveaux défis
 - Index thématique

© Tourigny (2003) IFt-66529 IC

3

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- Définition

L'IC est « l'étude des concepts, méthodes et techniques permettant de modéliser et/ou d'acquérir les connaissances pour des systèmes réalisant ou aidant des humains à réaliser des tâches se formalisant a priori peu ou pas. »

© Tourigny (2003) IFt-66529 IC

4

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- $IC \subseteq IA$
- IA : but ultime
 - Construire des artefacts informatiques visant à aider des humains dans des tâches difficiles qui requièrent la mise en œuvre de connaissances tant pour l'humain que pour la machine.

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- Méthodes, techniques et outils d'IC
 - Modélisation cognitive, modélisation conceptuelle, langages de modélisation et de représentation, méthodes de travail sur corpus, etc.
- Domaines de l'IC
 - Acquisition à partir de textes
 - Recherche d'informations sur le Web
 - Gestion et capitalisation des connaissances en entreprise
 - Définition d'indicateurs de gestion et leur mise en œuvre des systèmes d'information
 - etc.

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- Acquisition des connaissances : tendances actuelles
 - Des SE aux SC pour l'élaboration d'un système utilisant des connaissances
 - ~~X~~ Modélisation psychologique des connaissances d'un expert
 - ~~X~~ Transmission de connaissances
 - \Rightarrow Construction coopérative d'un modèle de connaissances, avec 1 ou plusieurs experts

© Tourigny (2003) IFt-66529 IC

7

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- Acquisition des connaissances : tendances actuelles
 - Au cœur de la conception d'un SC : modélisation des connaissances
 - Sources : textes, ...
 - Transcriptions d'entrevues d'experts
 - Corpus de textes liés aux objets visés par l'activité experte (descriptions de centrales nucléaires et leur maintenance pour assistance à la maintenance, ...)
 - Modélisation
 - Transformation des descriptions textuelles en descriptions formelles pour mise en œuvre des connaissances

© Tourigny (2003) IFt-66529 IC

8

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- **Réflexions sur l'IC : au carrefour de plusieurs réflexions**
 - Linguistique : pour étudier la formulation linguistique des connaissances
 - Terminologie et ontologie : pour dégager les concepts
 - Psychologie : pour élaborer des méthodes d'élicitation
 - Logique : pour élaborer des modèles formels
 - Informatique : pour opérationnaliser les modèles formels
 - Sémiotique: pour interpréter et s'approprier le comportement du système

⇒ **Fécondation réciproque**
⇒ **Évolution du paradigme**

© Tourigny (2003) IFt-66529 IC

9

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- **Réflexions sur l'IC : les SC « redéfinis »**
 - Manipulation de représentations symboliques selon règles formalisées lors de la modélisation des connaissances
 - Représentations exprimées à l'aide de primitives
 - Primitives qui renvoient à des notions du domaine
 - Manipulation des primitives en respectant la grammaire du système formel dans lequel elles s'inscrivent
 - Supports des expressions symboliques : support papier évoluant vers support numérique
 - Manière dont les utilisateurs s'approprient le système et lui attribuent du sens

⇒ **SC : systèmes sémiotiques de manipulation d'inscriptions symboliques, dont le fonctionnement informatique doit permettre à un utilisateur d'interpréter et de comprendre le système dans le cadre de son activité et de ses usages, dans les termes du domaine**

© Tourigny (2003) IFt-66529 IC

10

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

• Réflexions sur l'IC : les SC « redéfinis »

- SC : système technique plongé dans un système d'usages
- Élaboration de SC : ingénierie fondée sur la manipulation de l'inscription symbolique de connaissances
- IC : définition opérationnelle
 - Étude de modèles symboliques formels plongés dans des systèmes d'usage; ingénierie informatique et logique de modèles en fonction des usages qu'ils rendent possibles et des appropriations qu'ils permettent
 - Élaboration d'outils formels ou techniques
 - logiques de descriptions, langages de description documentaire (SGML, XML), raisonnement à partir de cas
 - Mise en perspective des concepts d'usage, issus ou non des sciences cognitives
 - raisonnement classificatoire, navigation documentaire, exploration d'hypothèses dans un domaine donné

© Tourigny (2003) IFt-66529 IC

11

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

• Évolutions récentes et nouveaux défis

- Nombreux champs d'applications de l'IC
- En continuelle évolution : « les nouveaux outils que sont les SC constituent de nouveaux usages qui à leur tour influencent leur conception »
- Élargissement de la notion de SC : « tout système de manipulation d'inscriptions si cette manipulation est appréhendée et explicitement modélisée en fonction des usages et des fonctions qu'elle autorise »
- Ouverture de l'IC à de nombreux domaines parfois ignorés de l'AC
- Systèmes d'inscriptions mobilisant des traitements systématiques reposant sur une modélisation : Les SC de l'IC incluent ceux de l'AC.

© Tourigny (2003) IFt-66529 IC

12

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

• **Évolutions récentes et nouveaux défis : thèmes**

- Thèmes fondateurs et toujours actifs de l'IC
 - Modélisation des connaissances
 - Langages de modélisation et de représentation
- Thèmes depuis 1992 : textes, terminologie et ontologies
 - Constitutions d'ontologies comme fondements de la modélisation
 - Étude de corpus de textes comme sources privilégiées de connaissances et de modélisation
- Raisonnement à partir de cas
- Thèmes plus récents
 - Hypertexte, hypermédia et Web (moyens-outils)
 - Gestion des connaissances (savoir-faire métier)
 - Systèmes d'Information pour le travail coopératif et méthodes cognitives (gestion)

© Tourigny (2003) IFt-66529 IC

13

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

• **Évolutions récentes et nouveaux défis : modélisation des connaissances (MC)**

- **Référence incontournable en IC : KADS/CommonKADS**
 - Structure en couches, MRP, tâches, « réutilisabilité », etc.
 - Critiques constructives nombreuses
 - Difficultés rencontrées lors de la mise en œuvre de méthodes de modélisation conceptuelle
 - Pertinence cognitive d'une méthode de modélisation et de son utilisation dans un SC
- Travaux de MC
 - Problème : difficulté à réutiliser des méthodes de RP génériques
 - Combinaison de la méthode KOD (méthode ascendante de modélisation) appliquée à un corpus) et de la méthode KADS (modélisation utilisant des MRP génériques) (Duribreux-Cocquebert et Houriez)

© Tourigny (2003) IFt-66529 IC

14

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- **Évolutions récentes et nouveaux défis : modélisation des connaissances (MC)**
 - Problème : difficulté à réutiliser des méthodes de RP génériques
 - Construction d'une solution verticale dans laquelle on propose une démarche méthodique, réaliste et économiquement viable de la « réutilisabilité » pour un ensemble d'applications suffisamment voisines dans une classe de problèmes donnée. Modèle à la KADS, bibliothèque de composants dédiés à une classe de thèmes considérée. Techniques d'implantation issues de l'IA, fonctions de génération de code, interfaces et bibliothèques graphiques, etc. (Talbi et Laublet)
 - Construction de modèles de RP à partir d'ontologies du domaine spécifiées formellement. Modèles construits à partir de composants réutilisables (Reynaud et Tort)

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- **Évolutions récentes et nouveaux défis : modélisation des connaissances (MC)**
 - Problème : expertise complexe et mal formalisée
 - Nécessité de créativité de la part de l'IC. Stratégies d'alimentation de bovins. Nécessité de stimuler le processus d'AC. Représentations graphiques des situations. (Girard)
 - Problème : Autres aspects de MC
 - Système interactif pour la construction d'applications de traitement d'images. Applications modélisées dans un modèle de tâches ; les experts les sélectionnent ou les assemblent pour construire l'application de traitement d'images désirée. Réutilisabilité. Raisonnement à partir de cas. (Ficet-Cauchard, Revenu, Porquet et Clouard)
 - Modélisation des connaissances de savoir-faire dans le domaine de la conception mécanique assistée par les connaissances. Validation des connaissances. Description de la validation à un niveau générique dans des métamodèles, d'où la mise en place de procédures de validation pour des classes d'objets. (Sellini et Yvars)

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

• **Évolutions récentes et nouveaux défis : langages de modélisation et de représentation**

- Définition de langages pour la conception de SC simulant une activité de RP : champ très actif
- Règles de production (1970-1985).
- Transfert des connaissances / construction coopérative de modèles de connaissances : modéliser le monde réel et les processus de RDP. Modèles de connaissances composés de modèles du domaine et du modèle de la tâche. Modèles du domaine : descriptions de systèmes du monde réel selon différents points de vues (fonctionnel, structurel, causal, empiriques, etc.). Modèles de la tâche: buts et MRP.
- Limites des langages des règles de production : une structure de données et plusieurs types de connaissances. Connaissances du domaine et connaissances de RP codées implicitement, d'où limites des capacités explicatives. Nécessité de langages plus déclaratifs permettant de spécifier séparément les modèles du domaine et le modèle de la tâche.

© Tourigny (2003) IFt-66529 IC

17

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

• **Évolutions récentes et nouveaux défis : langages de modélisation et de représentation**

- Langages avec multiformalisme : chaque formalisme est adapté aux particularités des connaissances qu'il doit représenter.
- Langages de modélisation (notation informelle), langages de représentation (syntaxe définie par des règles formelles). Caractéristiques des langages de représentation : héritage de propriétés, classification automatique de descriptions de concepts, etc.
- Langage ZOLA (Tchounikine, Istenes et Trichet) Approche flexible de l'opérationnalisation du MC. L'IC peut adapter les primitives offertes par défaut par le langage aux particularités de l'expertise à modéliser. Validation du MC. Prototypage au niveau des connaissances.
- Langage DEF* (Kassel, Barry et Abel) Langage d'implantation. Conciliation de la programmation déclarative et des temps de développement et d'exécution performants pour l'application.
- Langage LISA dans un contexte industriel. Langage utilisé à EDF pour développer des fonctions d'aide à l'opérateur dans le domaine de l'exploitation de réseaux électriques. (Jacob, Krivine et Montclar)

© Tourigny (2003) IFt-66529 IC

18

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

• **Évolutions récentes et nouveaux défis : langages de modélisation et de représentation**

- Autres projets développés : MACAO (Aussenac-Gilles et Matta), COKACE (Dieng et Corby), TASK (Pierret et Talon)
- Représentation objet des connaissances et génie logiciel objet
- Comparaison entre GC et logiques de description (Nobécourt).
- Représentation des connaissances et exploitation des connaissances. Systèmes de fouille de données appliquée à l'épidémiologie des cancers de l'enfant. (Simon et Napoli)
- Aspects récents de l'évolution du GL et ponts avec l'IC. Objet pour la modélisation d'un composant logiciel vers l'objet pour la modélisation du monde et l'objet pour la modélisation de procédures. (Bézivin)

© Tourigny (2003) IFt-66529 IC

19

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

• **Évolutions récentes et nouveaux défis : textes, terminologie et ontologies**

- Développement des outils de bureautique et des réseaux de communication. Augmentation et diversification des applications basées sur des ressources lexicales ou conceptuels du domaine de spécialité : système d'aide à la décision, système de gestion de données techniques, gestion des connaissances, système d'accès à la documentation technique, moteur de recherche thématique pour la veille technologique sur le Web, etc. Masse de données textuelles. Nécessité d'outils pour le traitement automatique de ces connaissances. Nouveaux défis pour l'IC : acquisition des connaissances à partir de textes
- Nécessité d'un rapprochement entre linguistique et IC. SC privilégiée: textes (documentation technique, ouvrages, textes d'entrevues, etc.). Nécessité d'analyser des documents textuels lors de l'AC. D'où travaux de recherche par des linguistes et des terminologues avec cognitivien (Bourigault et Condamines)

© Tourigny (2003) IFt-66529 IC

20

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- **Évolutions récentes et nouveaux défis : textes, terminologie et ontologies**
 - L'AC à partir de textes doit être assistée par des outils d'analyse de corpus. Aide au cogniticien grâce aux outils d'analyse de textes dans la phase d'AC et de MC à partir de textes. Aide à l'AC. Faciliter la tâche d'analyse du corpus textuel. Combinaison de techniques : analyse syntaxique, repérage de marqueurs, analyse statistique.
 - L'utilisation d'outils textuels doit être spécifiée dans un cadre méthodologique global. Nécessité d'une réflexion méthodologique poussée. Réalisation de guides méthodologiques et de plates-formes logicielles facilitant l'intégration afin de permettre une mise en œuvre efficace des outils proposés.
 - Une réflexion théorique sur la notion d'ontologie s'impose. Bases de connaissances terminologiques et ontologie régionale.

© Tourigny (2003) IFt-66529 IC

21

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- **Évolutions récentes et nouveaux défis : raisonnement à partir de cas**
 - RPC : remémoration, adaptation, apprentissage.
 - Description des cas. Indexation des cas. Synthèse de molécules organiques. Cycle de RPC (Lieber et Napoli)
 - Utilisation du RPC pour l'aide au diagnostic médical (Christel, Le Bozec, Jaulent et Zapletal)
 - Aide à la recherche de documents sur le Web grâce au RPC (Corvaisier, Mille et Pinon)

© Tourigny (2003) IFt-66529 IC

22

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- **Évolutions récentes et nouveaux défis :
hypertexte, hypermédia et Web**

- 1990 : IA et hypertextes
- Normalisation et opérationnalisation de l'hypertexte par le Web.
Diffusion de langages de l'ingénierie documentaire: SGML, XML.

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

- **Évolutions récentes et nouveaux défis : gestion des
connaissances**

- Domaine récent de recherche (Roscoff 1997). Besoins exprimés par le monde industriel. Gestion des connaissances : ensemble des savoirs et savoir-faire en actions mobilisées par les acteurs de l'entreprise pour lui permettre d'atteindre ses objectifs (produire des biens ou des services). Les connaissances détenues par les employés constituent un capital immatériel qu'il convient de gérer au même titre que le capital financier, les machines et les biens immobiliers. D'où le terme de capitalisation des connaissances.
- Modélisation des connaissances. Prise en compte de différentes formes de savoirs (tacites, dans la tête des spécialistes, textuels, iconographiques)

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

• **Évolutions récentes et nouveaux défis : systèmes d'information pour le travail coopératif et méthodes cognitives**

- SI : rôle dans la gestion des entreprises et plus particulièrement dans le pilotage des activités à partir d'indicateurs mis à jour régulièrement. Difficulté de définir des indicateurs de gestion pour des activités plus intellectuelles et informelles. IC utilisée pour formaliser l'activité. Mutations profondes de l'organisation des entreprises. Fonctions de pilotage de l'activité de plus en plus intégrées aux postes opérationnels. Opérateurs engagés dans des dialogues complexes avec des applications informatiques. L'IC mobilisée pour la conception de ces dialogues.
- Travail coopératif. Logiciels de « *groupware* ». Gestion par projet. Méthodes pour permettre d'explicitier les savoir-faire de spécialistes.

© Tourigny (2003) IFt-66529 IC

25

Ingénierie des connaissances: recherches et perspectives

(Charlet *et al.* 2000)

• **Construction d'un index thématique**

- Notions référencées dans l'index. Collection d'articles: plusieurs auteurs. 150 000 mots. Logiciel d'extraction de terminologie LEXTER. Réseau terminologique d'environ 15 000 candidats termes. Interface HTL. Sélection des entrées parmi les candidats, structuration des entrées retenues. Choix du niveau des renvois.
- Index très riche (570 entrées). Termes nombreux et précis. Renvois vers des articles ou des parties d'articles (section et sous-section plutôt que no de page).

© Tourigny (2003) IFt-66529 IC

26