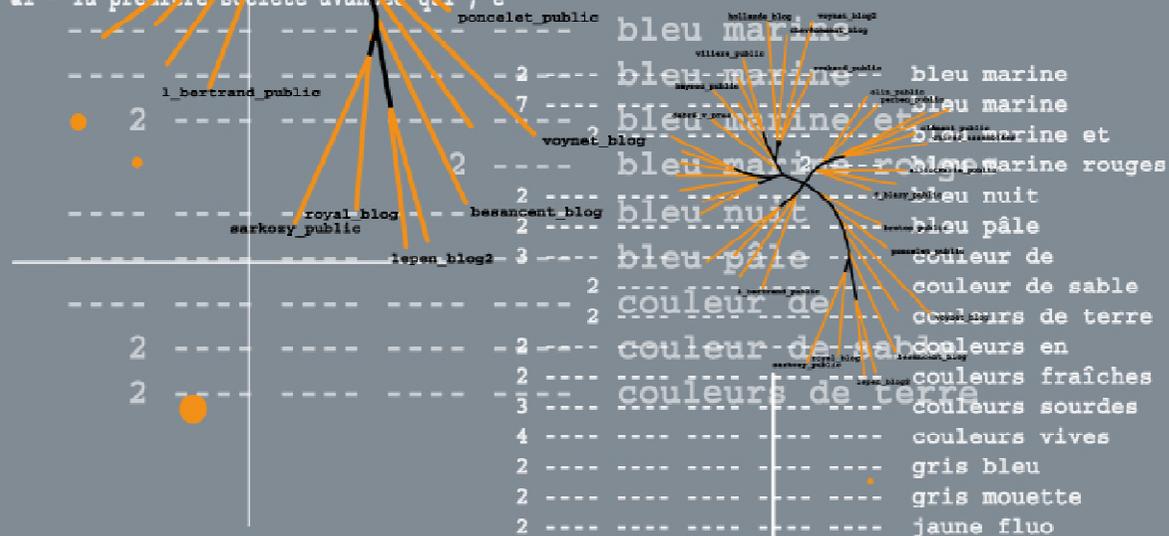


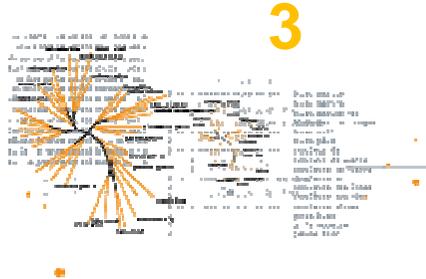
24 mai 2013

la liberté , la société , et donné à ce
, elle a miné la société sous tous ses a
mbreux cas où la société civile ne peut
état maîtres public société civile . cela n
- on dire que la société civile s' est
au motif que la société spectaculain public
l'écrasé public la société se délite , ell
terrorisme , la société civile italienne
niché dans notre société civile ? peut -
isation de notre société , nous ne sauri
adical de notre société qui , en se tra
eille , et notre société illuminée s' é
al - la première société avancée qui , e

Amani DAKNOU

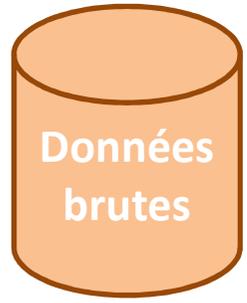
Post-Doc Céditec



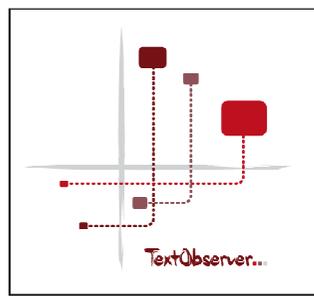


De l'analyse à la visualisation des données Web

- Volumineuses
- Massives
- Hétérogènes
- Structurées ou non



- Informations:
- Interprétables
 - Intelligibles

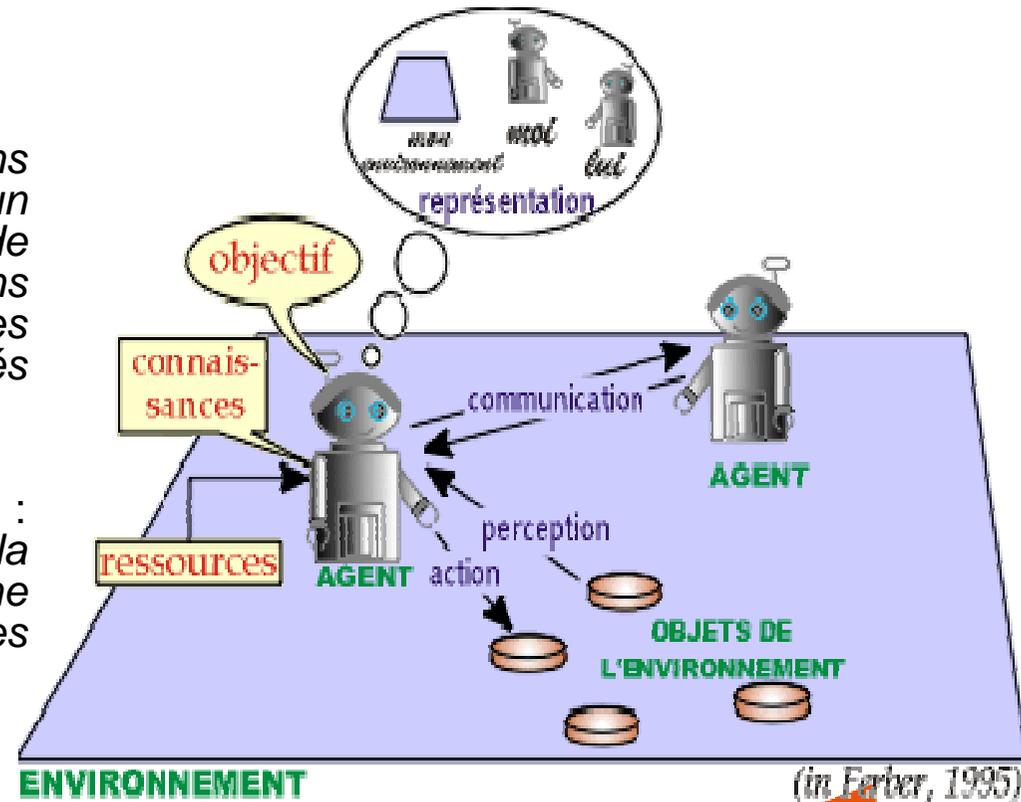


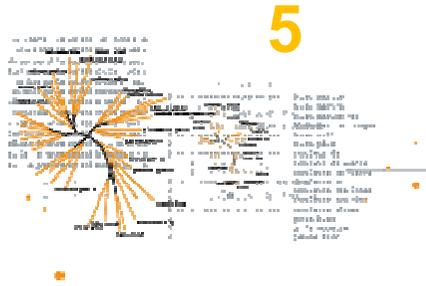
Visualisations



Définitions

- **Agents Logiciels:** «Applications informatiques, situées dans un environnement, capables d'agir de manière autonome et flexible dans le but de résoudre les problèmes pour lesquelles elles ont été conçues » [JENNINGS98].
- **Système Multi-Agents (SMA) :** «Les SMA consistent à étudier la résolution de problèmes par une collection d'agents autonomes » [MOULIN96].





Types d'utilisation des SMA

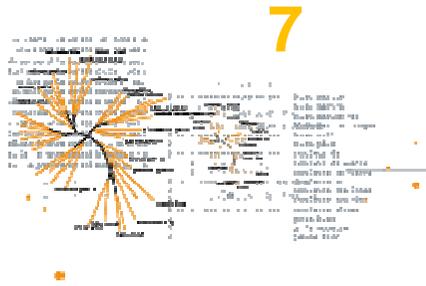
- La simulations ou la modélisation de phénomènes complexes.
- La résolution de problèmes et prise de décision.
- La conception de programmes.

6



Intérêt de l'approche SMA

- La distribution des données et du contrôle
- Le dynamisme de l'environnement
- Processus de communication structuré
- Plateformes de communication normalisées
- Mécanismes de raisonnement, d'apprentissage, de négociation.
- Méthodes d'analyse

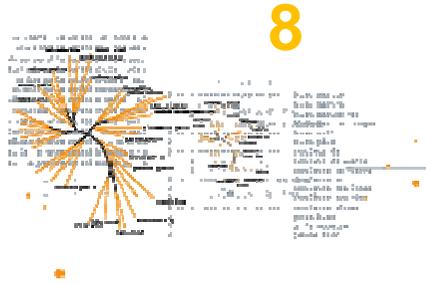


Pourquoi adopter un SMA?

Les sources des informations sont:

- Distribuées
- Hétérogènes
- Dynamiques

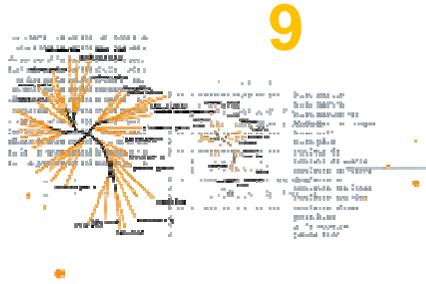
→ une bonne option pour concevoir un système de recueil et de construction d'un corpus thématique tiré du web.



Domaines d'application des SMA

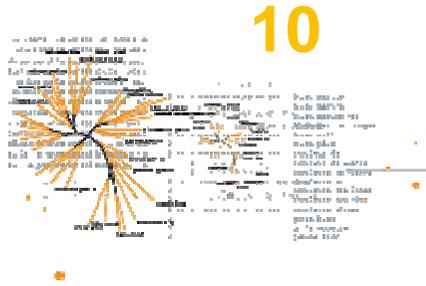
- *La Santé*

- La supervision des malades
- Amélioration des soins données aux patients palliatifs
- Coordination des transplantation d'organes
- L'optimisation de la prise en charge des patients au SU. [thèse 2011 , Daknou]



Domaines d'application des SMA

- L'Énergie
 - L'achat de la puissance dans une manière intelligente
 - La gestion des réseaux
- L'Industrie
 - L'automatisation des processus et de la production
 - La logistique
 - Les robots coopératifs



Domaines d'application des SMA

- La Communication (y compris télécommunication)
 - La gestion de réseaux
 - Le commerce électronique
 - La maison intelligente
 - Les services du réseau personnel
 - Le calcul mobile



Domaines d'application des SMA

- *Le transport*
 - La logistique
 - Le support de la mobilité
 - L'information du voyage

- *Les Composants*
 - L'automatisation de la production
 - Les smart cartes intelligentes



Domaines d'application des SMA

- *L'Information*

- L'assistance personnelle
- La recherche d'Information
- L'accélération des flux de données au travers Internet
- La gestion du workflow
- La maison intelligente



Les fonctionnalités d'un agent logiciel

- être capable de rechercher une information demandée.
- être capable de choisir l'information par rapport aux préférences utilisateur.
- être capable de réduire le volume d'informations destinées à l'utilisateur.

14



Création automatique de corpus thématique

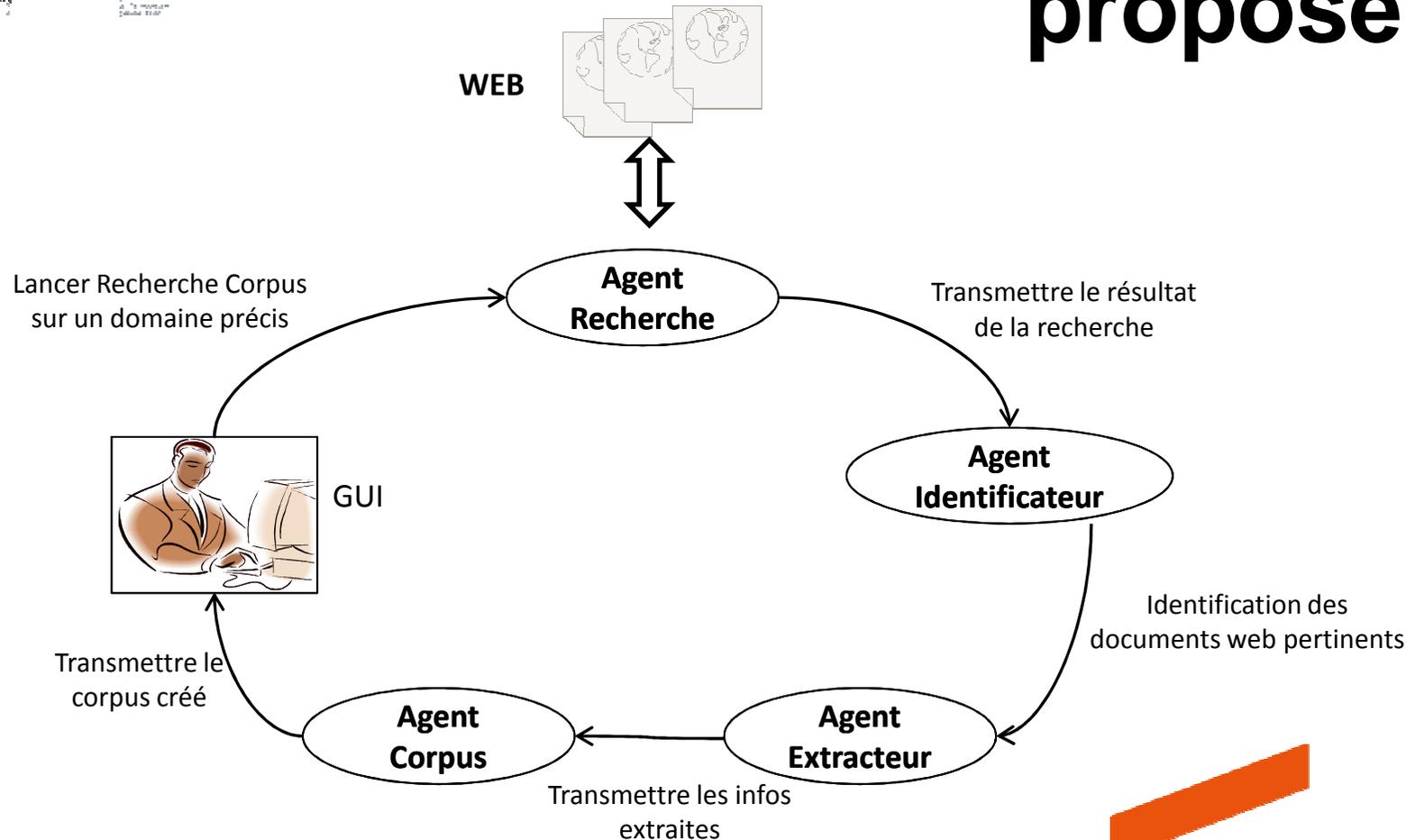
Récolter toutes les pages correspondant à la thématique qui nous intéresse

Analyser leur structure

Extraire l'information de chacune de ces pages pour former un corpus.



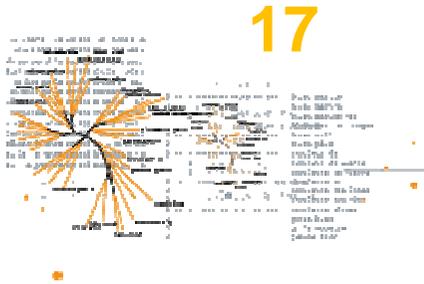
Architecture du SMA proposé





Agent Recherche

- Interroger les moteurs externes sur le Web par exemple Google
- Sélectionner les pages Web
- Transmettre le résultat à l'agent identificateur



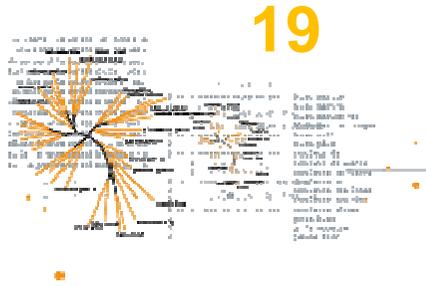
Agent Identificateur

- Validation des pages Web:
 - vérifier si elles sont au format HTML
 - Vérifier si elles sont accessibles
- Identification du contenu, du titre, des liens et des emails disponibles dans la page
- Transfert des pages prêtes à l'extraction des informations pertinentes à l'agent extracteur



Agent Extracteur

- Extraire des pages web validées l'ensemble des informations adéquates au thème demandé.
- Choisir les textes à inclure dans un corpus.
- Transmettre le résultat à l'agent corpus.



Agent Corpus

- Construire le corpus:
 - Appliquer les paramètres choisis par l'utilisateur: taille, type ...
 - Choisir des regroupements assez homogènes
- Transmettre le corpus:
 - Lancer Textobsever / Webobserver pour le traitement automatique
 - Sauvegarder le corpus



Intérêt d'un tel système

- Disposer un corpus à la demande
- Construction automatique d'un corpus thématique tiré du web
- Exploration en temps réel du corpus



Conclusions

- JM.Leblanc, A.Daknou « ***TextObserver/WebObserver : Propositions ergonomiques pour l'exploration et l'exploitation des données textuelles multidimensionnelles*** », 12e édition de la Conférence Internationale « Hypertextes et Hypermédias. Produits, outils et méthodes », ***H2PTM 2013***.
- JM.Leblanc, A.Daknou « ***TextObserver: outil de visualisation dynamique des données et résultats textométriques*** », 16e Colloque International sur le Document Électronique, ***CiDE 2013***.

22



Merci pour votre attention

Amani DAKNOU
adaknou@gmail.com

L'utilisation d'agents communicants pour l'extraction et la visualisation de corpus textuels tirés du web