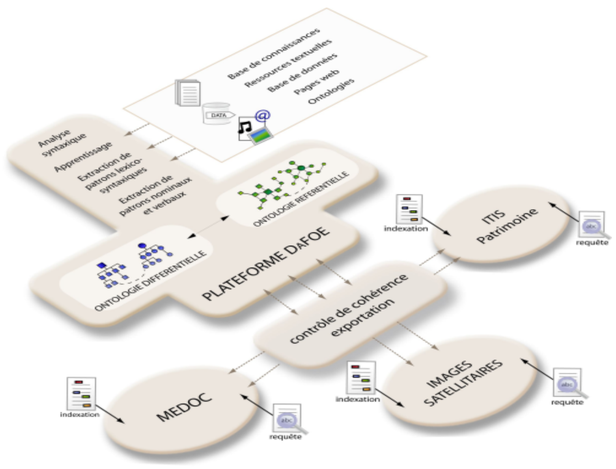
Apport de l'analyse ergonomique des besoins à la spécification d'un éditeur d'ontologie

Nathalie Aussenac-Gilles, Sandrine Cazabat - Institut de Recherche en Informatique de Toulouse, CNRS et Université de Toulouse
Audrey Baneyx, Jean Charlet - INSERM, UMR_S 872, eq. 20, Paris
Nadia Nadah - Université de Technologie de Compiègne - Heudiasyc (UMR CNRS 6599)
Sylvie Szulman - Laboratoire d'Informatique de l'université Paris-Nord, CNRS et Université Paris XIII
Bernard Vatant - MONDECA, Paris

Dafoe
Financé par Agence Nationale de la Recherche ANR

Présentation du projet DaFOE



Objectifs

- Méthode et plateforme technique pour la conception d'ontologies
- Modélisation à partir du domaine
- Ontologies catégoriales
- Evolution et maintenance des ontologies
- Utiliser les textes
  - pour faciliter la tâche de l'ontologue
  - pour faciliter l'utilisation et la réutilisation

Partenaires

- Inserm, U872, éq. 20 (porteur)
- UTC (Heudiasyc)
- IRIT (CSC)
- LIPN
- LISI
- ENST (GET)
- SUPELEC
- Société MONDECA

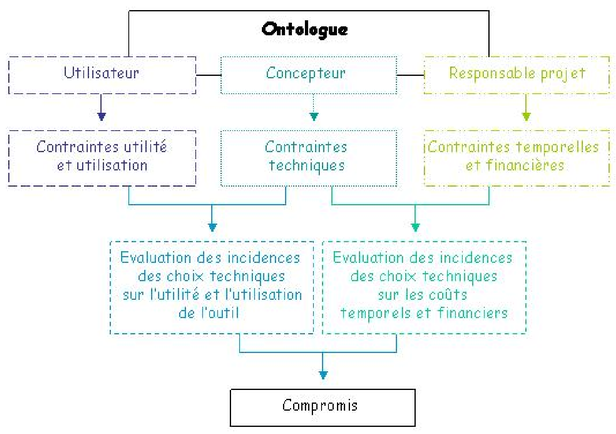
Pourquoi une intervention ergonomique dans le projet DaFOE ?

Atouts de l'intervention ergonomique

- Accroitre l'implication des utilisateurs dans la conception,
- Définir un référentiel opératif commun.
- Mieux partager connaissances, décisions et solutions.
- Déboucher sur une conception participative
- Impliquer les partenaires comme co-concepteurs.

Constats initiaux

- Pluralité des rôles des partenaires : chercheurs, utilisateurs, chefs de projet ;
- Multiplicité de plateformes existantes ;
- Révision des besoins formulés lors du montage du projet.



Étude des environnements de construction d'ontologies

Étude des environnements existants

- Utilisation expérimentale de plusieurs logiciels : Protégé, Swoop, DOE, Terminae, ...
- Étude bibliographique
- Grille d'analyse renseignée à partir du retour d'expérience des partenaires

Grille d'analyse des outils existants

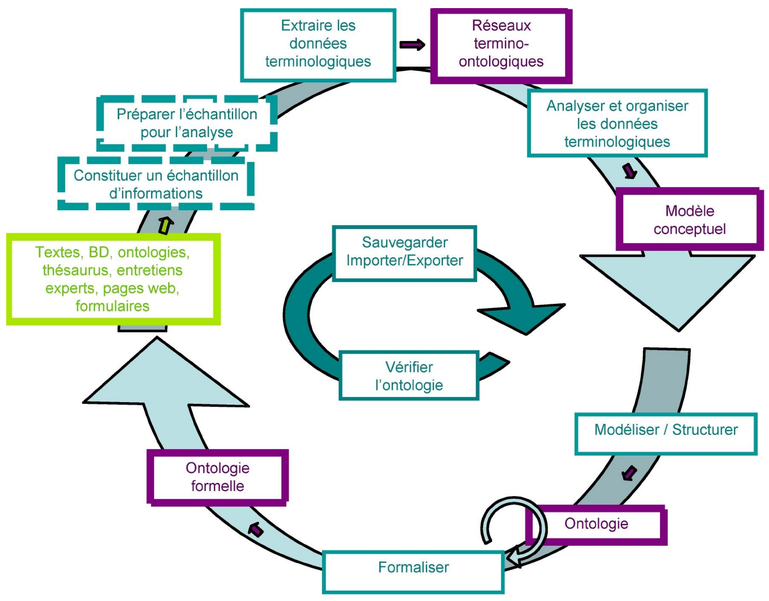
Table 1 Grille d'étude des outils existants (extrait)

Critères généraux
Source, Contact, Site Web, version, Disponibilité, Langage de base, plateforme…
Architecture de l'outil
Extensibilité (plugins…), Stockage de l'ontologie (format fichier, BD…), SW Architecture
Interopérabilité entre outils / entre ontologies
Langage(s) d'import, Langage(s) d'export, Interopérabilité avec d'autres outils, Support Web, Disponibilité sous forme d'API
Méthodologie
Support méthodologique déclaré (papier)
Utilisation du logiciel guidée par la méthode (vs fonctionnalités indépendantes)
Fonctionnalités : Critères relatifs aux inférences
Langage axiomatique, moteur d'inférence intégré ou accessible, Vérification de la consistance, classification automatique
Fonctionnalités – visualisation du modèle
Visualisation graphique de la taxonomie, de l'ontologie, Zoom, filtrage, éditeur graphique
Fonctionnalités – sources possibles et outils pour les intégrer
Bibliothèque d'ontologies, réutilisation d'ontologies externes, support à la fusion, à l'alignement, Outils d'analyse de textes, réutilisation de thesaurus, de terminologies, de BD
Caractérisation des modèles possibles / produits
Degré de formalisation (expressivité…), Présence de relations autres que hiérarchiques, composante terminologique, Gestion du multilinguisme, Passage à l'échelle
Fonctionnalités : Interopérabilité interne
Travail collaboratif (alertes, multi-utilisateurs, …), Gestion des versions
Fonctionnalités : Module Analyse de textes
extraction terminologique et de relations (automatique / non automatique, indep/intégré, technique, outil, …), gestion du corpus, peuplement de l'ontologie à partir de textes
Construction de l'ontologie à partir des textes
Fonctionnalités : Module réutilisation / adaptation BD
Fonctionnalités : Module de réutilisation thésaurus / terminologie
Fonctionnalités – critères transverses
Compatibilité (langages informatiques entre eux, exigences utilisateur et fonctionnalités de l'outil)
Homogénéité (stabilité environnement permet modes opératoires quasi automatiques)
Concision (lisibilité, mémorisation, actions minimales)
Organisation de l'affichage (repérage et ciblage des informations dans IHM)

Choix retenus pour DAFOE

Cycle de modélisation d'ontologie à partir de texte

- Permet de situer les convergences et divergences entre partenaires
- Permet de localiser la part du processus prise en charge dans la plateforme

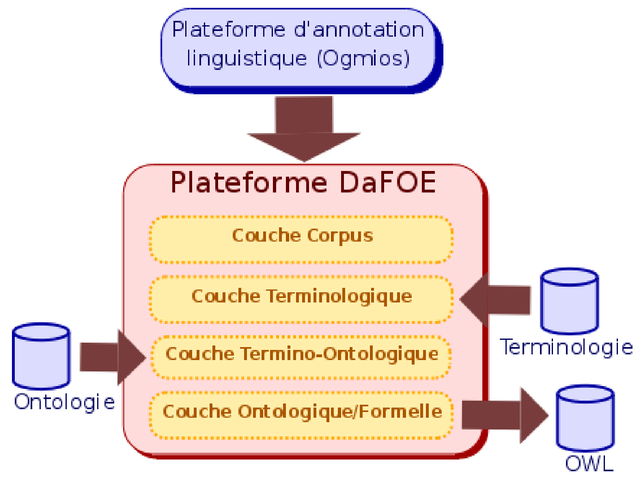


Principes dont s'inspire DAFOE

- Traçabilité des textes aux modèles et inversement (Terminae)
- Différentes couches de modélisation (Terminae)
- Fonctions d'édition d'ontologie (Protégé ou Swoop)
- Exploitation d'un extracteur de termes (Syntex) et extracteur de relations (Caméléon)

Spécificités de la plateforme DAFOE

- Considérer les outils de TAL en dehors de Dafoe
- Favoriser l'utilisation des outils de TAL sous forme de greffons via des interfaces
- Gérer de gros volumes de données
- Gérer des ontologies et des bases de connaissances (instances)
- Combiner réutilisation et modélisation à partir de textes
- Ontologie différentielle (DOE, Archonte)

Propriétés de la plateforme DAFOE

- Flexibilité et paramétrage : des modèles de données
- Apprentissage et personnalisation
- Modèle de données en quatre couches et permettant de traiter des ressources provenant de langues différentes (multilinguisme) et/ou leur présentation dans différentes langues (localisation).
  1. Niveau textuel: le corpus
  2. Niveau terminologique : termes et relations proposés par les outils de TAL
  3. Niveau termino-ontologique : termes et relations désambiguïsées
  4. Niveau formel : concepts et relations décrivant une ontologie