Application BASTRI
Fiches Equipes

GEMO (SR0145VR)

Intégration de données et de connaissances distribuées sur le Web
VERSO (SR0309YR) GEMO (LEO (SR0414TR) , DAHU (SR0226ER))

Statut: Terminée

Responsable : Serge Abiteboul

Mots-clés de "A - Thèmes de recherche en Sciences du numérique - 2024" : *Aucun mot-clé.*

Mots-clés de "B - Autres sciences et domaines d'application - 2024" : *Aucun mot-clé.*

Domaine : Perception, cognition, interaction
Thème : Représentation et traitement des données et des connaissances

Période : 01/01/2003 -> 31/12/2009
Dates d'évaluation :

Etablissement(s) de rattachement : U. PARIS 11 (P.-SUD), CNRS
Laboratoire(s) partenaire(s) : LRI (UMR8623)

CRI : Centre Inria de Saclay
Localisation : Centre de recherche Inria de Saclay
Code structure Inria : 111024-0

Numéro RNSR : 200218379Y
N° de structure Inria: SR0145VR

Présentation

Les informations disponibles en ligne sont de plus en plus complexes, distribuées, hétérogènes, répliquées, multiformes, changeantes et parfois sous forme de services (comme SOAP).

L'objectif de l'équipe-projet est l'étude des problèmes fondamentaux posés aux systèmes modernes de gestion de données et de connaissances, et de développement de solutions novatrices appropriées. Notre but est d'obtenir des systèmes plus ouverts à des informations plus riches, orientés vers le réseau. Le thème central de GEMO est l'intégration de ces informations. Il s'agit de découvrir les informations et services intéressants, comprendre leur contenu, les intégrer et finalement contrôler leurs changements.

On souhaite offrir des environnements puissants et flexibles pour faciliter le déploiement d'applications permettant un accès efficace aux données pertinentes. Pour cela, on s'appuie sur la conception d'entrepôts de données ou de médiateurs offrant un accès uniforme à des sources de données multiples et hétérogènes.

Les problèmes nouveaux que nous considérons demandent souvent de combiner des techniques d'intelligence artificielle (comme la classification) et de bases de données (comme l'indexage).

GEMO est une équipe-projet créée à partir du projet VERSO de l'INRIA Rocquencourt et de membres de l'équipe IASI du Laboratoire de Recherche en Informatique (UMR 8623 CNRS) de l'Université Paris Sud.

Axes de recherche

- Médiation de données XML.
  Nous nous intéressons à l'intégration de données hétérogènes représentées dans le format XML. Le but est de regrouper les éléments d'une collection de documents XML présentant des similarités, et à partir des classes obtenues et de leurs descriptions, générer un schéma global (une ontologie) qui puisse servir de point d'entrée unique pour interroger cette collection.
- Médiation pour le Web sémantique.
  L'objectif du Web sémantique est de tendre vers un Web dont la sémantique des données serait à la fois compréhensible par des utilisateurs humains et appréhendables par des entités informatiques (agents, moteurs de recherche, serveurs d'informations). Les ontologies vont jouer un rôle central pour le marquage sémantique des données

Contact

- **Responsable :** Serge Abiteboul
- **Tél :** 01.72.92.59.20
- **Secrétariat Tél :** 01.74.85.42.25

En savoir plus

- Site de l'équipe
- Site sur inria.fr
- Site du responsable
- Derniers Rapports d'Activité :

Documents sur la structure

- Intranet
- Privés

Décisions

- 3732 (10/12/2002) : création
- 5156 (10/10/2006) : prolongation
- 6072 (19/02/2008) : changement de rattachement
- 6695 (31/03/2009) : cessation du responsable
- 6696 (31/03/2009) : nomination responsable
- 7016 (16/12/2009) : fermeture

Localisation

- **Adresse postale :** Centre de recherche Inria de Saclay Campus de l'École Polytechnique - Bâtiment Alan Turing 1 rue Honoré d'Estienne d'Orves 91120 Palaiseau France
- **Coordonnées GPS :** 48.714, 2.206

Web qui va ouvrir de nombreuses prospectives d'amélioration de la qualité des moteurs de recherche. Dans GEMO, nous nous attaquons à 3 problèmes clés pour le passage à l'échelle du Web sémantique : la médiation entre ontologies, la médiation disribuée (Peer to Peer) entre sources de données, et la médiation entre le Web et ses utilisateurs.

- Entrepots thématiques de données du Web.
  Nous souhaitons développer une approche flexible et générique permettant de spécifier de facon déclarative les données recherchées pour l'entrepot thématique visé, de faciliter leur acquisition à partir du Web, et de structurer les données acquises en vue de leur interrogation.
- Ouverture vers les services du Web.
  En combinant les approches de type entrepot et médiateur, on s'interesse à intégrer également des services du Web (basés sur l'échange de données XML). le but est de découvrir des services intéressant une application particulière et comprendre comment ils peuvent être utilisés. Dans ce but, des travaux portent en particulier sur Acive XML, un modèle basé des documents XML incluant des appels à des services Web.
- Théorie de la modélisation de données :
  Nous considérons les aspects théoriques spécifiques d'une vision de l'informatique centrée sur les données en nous appuyant sur des outils classiques de logique et de complexité pour dégager les spécificités du calcul sur des collections (relations) ou sur des graphes irréguliers (Web).

Relations industrielles et internationales

Projet PICSEL avec France Télécom R&D.
Projet européen DbGlobe sur l'évaluation de requêtes sur le Web.
Projet RNTL E.dot sur un entrepôt dans le domaine du risque alimentaire avec notamment l'INRA.
Industriel : avec Xyleme (start-up issue du projet).