

Intégration des Systèmes d'Information en Biologie

<http://bioinfo.loria.fr/Members/devignes/ISIBIO/ISIBioPresentationFr>

CONTEXTE

A l'ère de la post-génomique, la problématique de l'intégration des sources de données biologiques, énoncée dans la communauté bio-informatique dès 1995, reste d'une actualité frappante.

La prolifération des bases de données et autres ressources (systèmes d'information au sens large), loin de combler les attentes de tous les utilisateurs, suscite encore de nombreuses difficultés. Identifier la ou les ressources pertinentes pour un problème donné n'est pas facile. Les interactions avec ces ressources se font selon des modalités variables. Les résultats obtenus peuvent être discordants suivant les ressources interrogées. La qualité des données est difficile à estimer. Les recherches doivent être actualisées régulièrement à cause des mises à jour.

Les solutions proposées à ce jour prennent en compte différents niveaux d'intégration :

- intégration de l'accès aux ressources (ex : SRS)
- intégration des modèles de données des ressources (entrepôts, médiation)
- intégration de la mise en oeuvre des ressources (workflows, composition de services web)

Les défis à relever concernent l'intégration des **données** proprement dites (ex : détection de discordances), l'estimation de la **qualité** des données et la couverture par les solutions proposées de la **totalité** des sources disponibles.

OBJECTIFS

La réflexion du groupe de travail ISIBio porte particulièrement sur le rôle des métadonnées et des ontologies pour contribuer à l'intégration de systèmes d'informations en biologie, d'une part en introduisant la dimension sémantique, d'autre part en préparant le passage à l'échelle annoncé par les programmes GRID.

L'activité du groupe de travail se manifeste par l'organisation d'un séminaire de portée internationale. Les buts recherchés sont de renforcer les collaborations amorcées, de favoriser de nouvelles interactions et d'accroître la visibilité internationale de la communauté française dans cette thématique.

ORGANISATION - CALENDRIER

Le séminaire « Intégration de Systèmes d'Informations en Biologie » a lieu 2 fois par an. Il est précédé d'une réunion de travail du groupe ISIBio pour renforcer les collaborations en cours.

- 29-30 mars 2005** 1er Séminaire ISIBio ; Institut Curie, Paris
- 4 juillet 2005** ISIBio partenaire de la journée satellite JOBIM : « Ontologies, Grille et Intégration Sémantique pour la Biologie (OGSB) ».
- Automne 2005** 2ème Séminaire ISIBio (en préparation)

THÉMATIQUES

Les thématiques abordées se rattachent aux domaines suivants de l'informatique :

- interopérabilité de ressources hétérogènes
- architectures de médiation
- représentation et gestion des connaissances
- web-services sémantiques
- production et utilisation de métadonnées
- construction et partage d'ontologies

Du côté biologie, quelques domaines d'applications pertinents sont privilégiés afin de mieux saisir les enjeux techniques d'une intégration de ressources et de données qui réponde réellement aux attentes des biologistes (ex : analyse du transcriptome).

Des interactions avec la communauté « Statistiques et analyse de données pour l'étude du transcriptome » et avec la communauté « Grille de Calcul pour la Génomique » s'imposeront.

MOTS CLÉS

Systèmes d'informations biologiques ; intégration ; interopérabilité ; ontologies ; métadonnées ; web services ; transcriptome ; hétérogénéité ; médiation

COMITÉ SCIENTIFIQUE

- Marie-Dominique Devignes (coordinatrice), LORIA, UMR7503, Nancy
- Philippe Bessières, MIG, INRA, Jouy-en-Josas
- Gwennaële Fichant, LMGM, UMR 5100, Toulouse
- Christine Froidevaux, LRI, UMR 8623, Orsay
- Isabelle Mougnot, LIRMM, UMR 5506, Montpellier
- Malika Smail, LORIA, UMR7503, Nancy

GROUPE DE TRAVAIL

En plus des laboratoires des membres du Comité Scientifique, sont concernés les laboratoires suivants :

- ✓ BBE, UMR 5558, Lyon
- ✓ LIF, UMR 6166, Marseille
- ✓ LSIS, UMR 6168 Marseille
- ✓ IGR, UMR 8125, Villejuif
- ✓ LaMI, UMR 8042, Evry
- ✓ Genexpress-LGN, UMR 7091, Villejuif
- ✓ CIRAD, Montpellier
- ✓ LBV, UMR6543, Nice
- ✓ EA-3888, Faculté de médecine, Université Rennes 1

CONFÉRENCIERS INVITÉS

1er séminaire ISIBio : 29-30 mars 2005

- | | |
|---|---|
| Arek Kasprzyk (EBI, Hinxton) | <i>BioMart - a federated query architecture</i> |
| Paul Kersey (EBI, Hinxton) | <i>Assembling Integr8: Adventures with Unidentified Flying Objects</i> |
| Eric Coissac (Université Paris VI et HELIX, Grenoble) : | <i>MicroBI : A pragmatic approach to biological data integration .</i> |
| Anne Morgat (SIB et HELIX, Grenoble) | <i>MicroBI: An application to metabolic data standardization in the Swiss-Prot database</i> |
| Anita Burgun (Université de Rennes) | <i>Dependence relations in Gene Ontology.</i> |

Journée satellite JOBIM : 4 juillet 2005

Ontologies, Grille et Intégration Sémantique pour la Biologie

- | | |
|-------------------------------------|--|
| Susan B. Davidson (U. Pennsylvania) | <i>Data cooperatives: rapid and incremental data sharing in bioinformatics</i> |
|-------------------------------------|--|

ORIENTATIONS

- Intégration des données et détection/correction des incohérences :
 - « des ontologies pour raisonner sur les données ? »
- Estimer la qualité des données et des sources :
 - « des métadonnées pour faire le tri ? »
- Intégration et grille ?