

Un mésocentre thématique: GenOuest

Olivier Collin
20 septembre 2011

INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTEMES ALÉATOIRES



La bio-informatique

- Concepts et techniques nécessaires à l'interprétation informatique de l'information biologique
- Champ de recherche multi-disciplinaire
- Volet recherche (élaboration de nouvelles méthodes) et volet service (analyse de données assurée par/sur les plateformes).

La bio-informatique à l'IRISA/INRIA

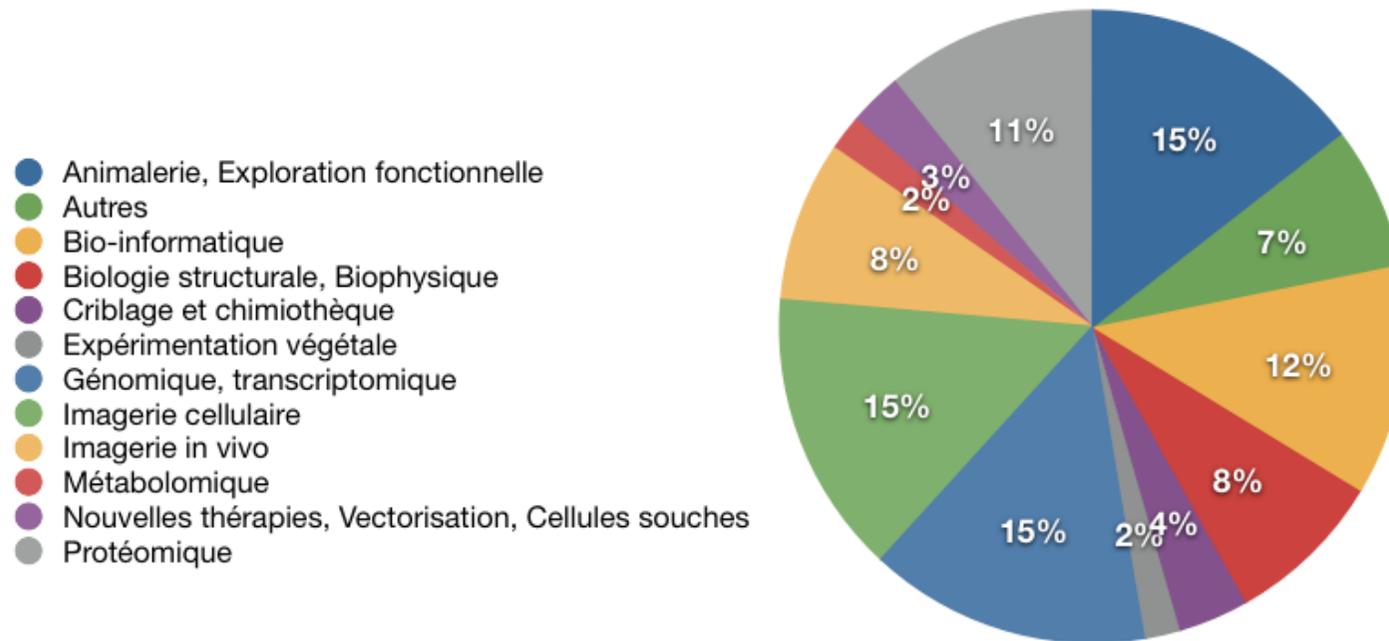
- Symbiose : un centre de ressources en bio-informatique
 - Une équipe de recherche
 - Une plate-forme bio-informatique : GenOuest
- Profiter du savoir-faire et de l'infrastructure d'un centre de recherche en informatique.
- Une association unique en France

Plates-formes en Sciences de la Vie

- GIS IBiSA : Coordination nationale des plates-formes de recherche en sciences du vivant (<http://www.ibisa.net>)
- Définition des plates-formes:
 - Regroupement sur un même site des équipements et des moyens humains destinés à offrir à une communauté d'utilisateurs des ressources technologiques de haut niveau
- Charte (critères) : ouverture, mode de gestion, évolution technologique, formation, évaluation

Plates-formes IBiSA

- 110 plates-formes reconnues (www.ibisa.net au 15-11-2010)
- 13 plates-formes bioinformatiques avec une structure de concertation : ReNaBi



GenOuest

8 ingénieurs
(3 permanents, 5 contractuels)

Cluster (300 cœurs,
80 To stockage)
Banques, logiciels

Regroupement sur un même site des équipements et des moyens humains destinés à offrir à une communauté d'utilisateurs des ressources technologiques de haut niveau

laboratoires INRA, INSERM, CNRS,
IFREMER, Université de Rennes,
Université de Nantes, Ligue contre le
Cancer

2006-2010 Label IBiSA

50 laboratoires – 180 utilisateurs
> 100 000 jobs/mois

Comités animation
ISO9001:2008



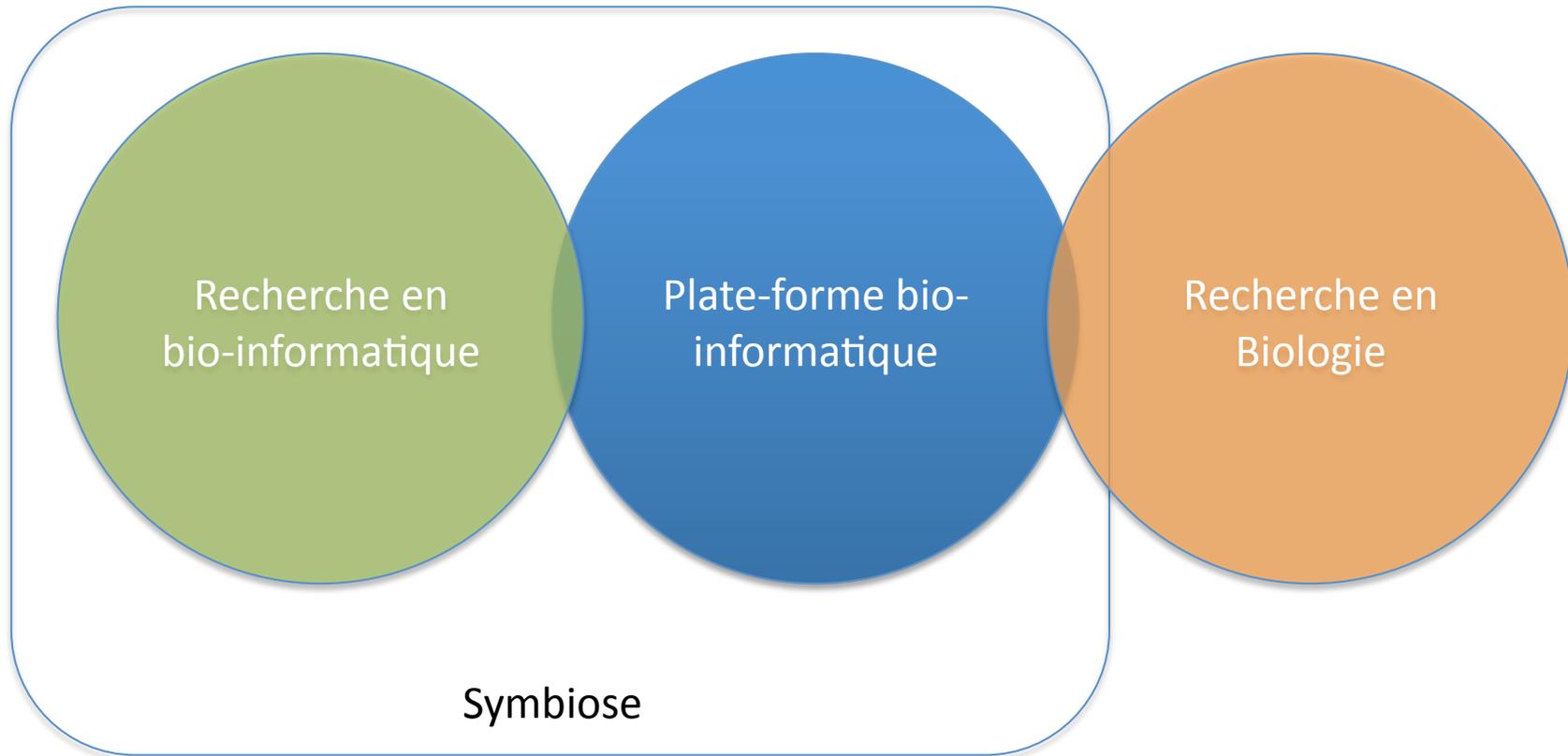
Charte (critères) : ouverture, mode de gestion,
évolution technologique, formation, évaluation

Dossier labellisation

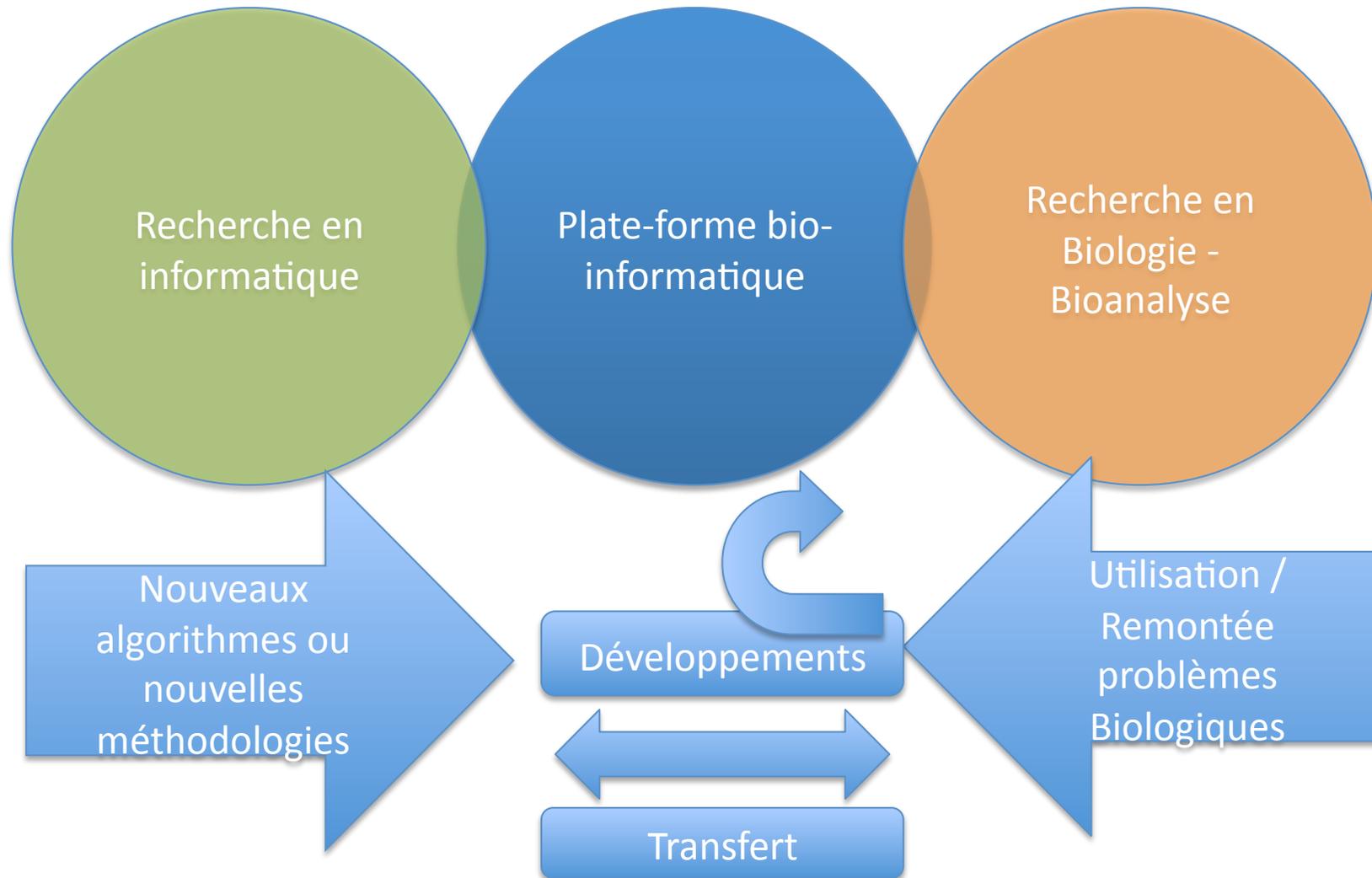
Projets GRISBI, BioMAJ,
MobyNet, Metadata,
SeqCrawler, Opal Toolkit

70 personnes/an

Un lien entre les communautés



Un lien entre les communautés



Services

Un environnement de soutien à la Recherche

Services

Calcul, Développement
Hébergement, Formations
Transfert technologique

Infrastructure

Ressources de calcul, logithèque
Stockage
Banques de données mises à jour

Activités

Transférer...

les résultats de la recherche informatique : Traitement et accélération des traitements des données biologiques, Modélisation des systèmes biologiques

Mettre à disposition...

un environnement de calcul et d'expérimentation, logithèque adaptée

Fédérer et structurer...

la communauté bio-informatique en France grâce à des projets communs d'infrastructure

Développer...

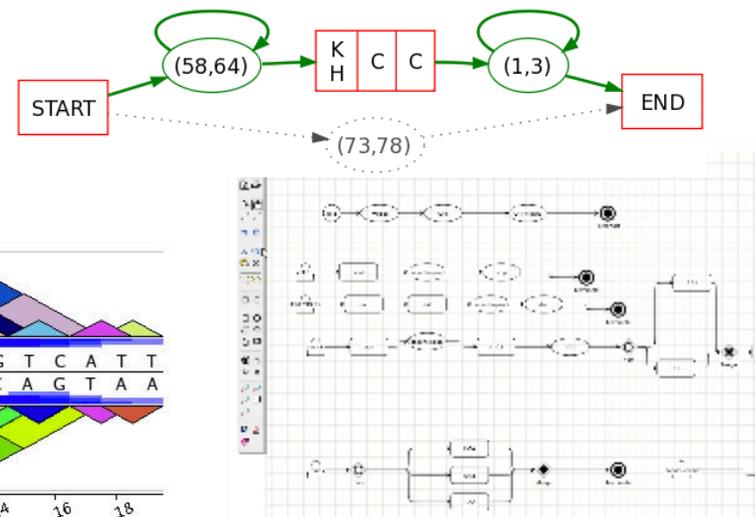
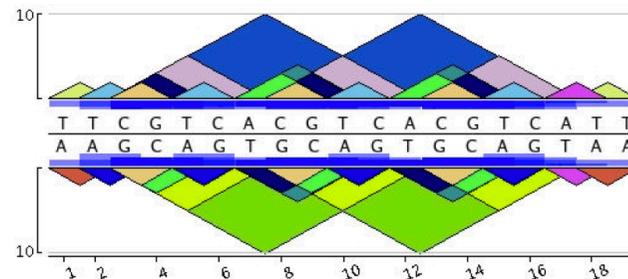
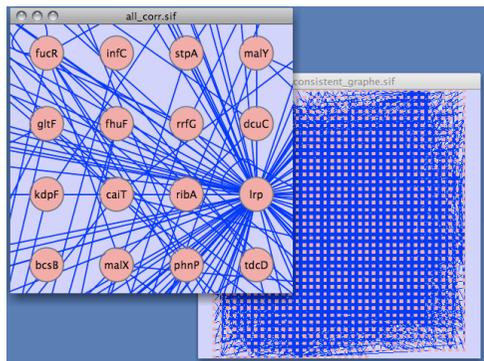
de nouvelles solutions pour la bio-informatique

Transférer

L'ensemble des logiciels développés par l'équipe Symbiose sont mis à disposition de la communauté sur la plate-forme GenOuest .

Exemples : Bioquali, Chavlh, Domain Organizer, GenoFrag, Logol, Pygram, Protomata, Stan, Tuiuiu, Wapam

Accès en ligne de commande, interfaçage web, web service.
Permet aux scientifiques de se focaliser sur le programme



Mettre à disposition



Environnement de calcul et d'expérimentation (~100 logiciels, banques mises à jour)

- Bio-analyse

- Expérimentation/Développement :
Next Generation Sequencing : Alcovna
Classification de protéines : Protomata



Compétition SHREC '10 (SHape REtrieval Contest 2010) 3rd
Eurographics Workshop on 3D Object Retrieval. Catégorie
« Protein Model Classification »



Suite MEME (3 sites dans le monde)

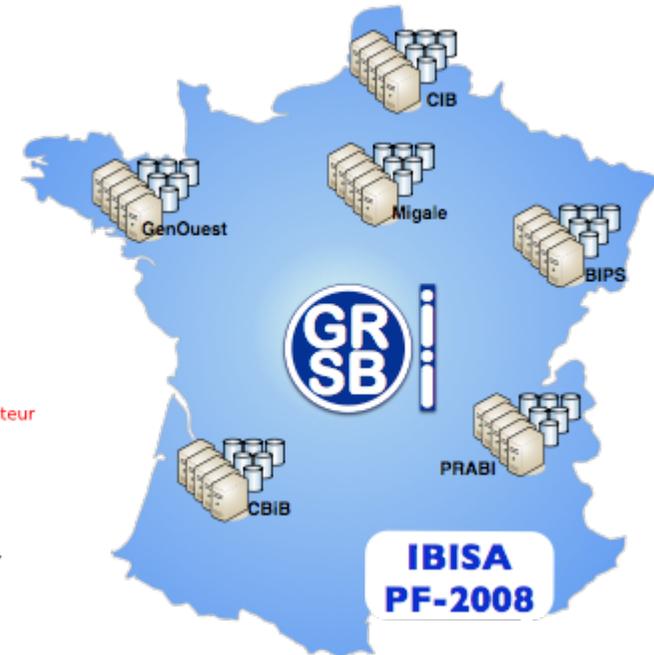
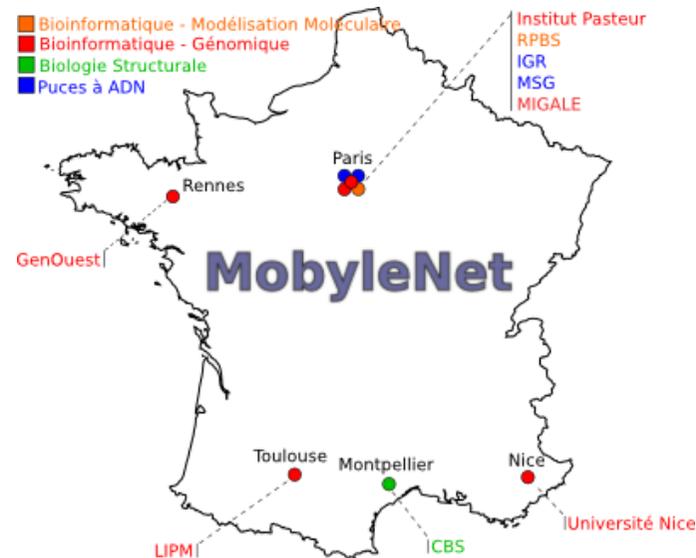
Fédérer

Participer aux actions de structuration de l'infrastructure bio-informatique en France
(ReNaBi)

Exemples :

GRISBI : 6 plates-formes

MobyleNet : 9 plates-formes



Développer



Solution de mise à jour automatisée des banques de séquence

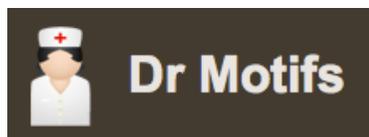
+ 100 téléchargements
+ de 10 plates-formes

BIOINFORMATICS APPLICATIONS NOTE Vol. 24 no. 16 2008, pages 1823–1825
doi:10.1093/bioinformatics/btn325

Databases and ontologies

BioMAJ: a flexible framework for databanks synchronization and processing

Olivier Filangi¹, Yoann Beausse², Anthony Assi¹, Ludovic Legrand³, Jean-Marc Larré²,
Véronique Martin³, Olivier Collin¹, Christophe Caron³, Hugues Leroy¹ and
David Allouche^{2,*}



Nouvelle plate-forme intégrée de découverte et recherche de motifs (également transfert)

<http://drmotifs.genouest.org>

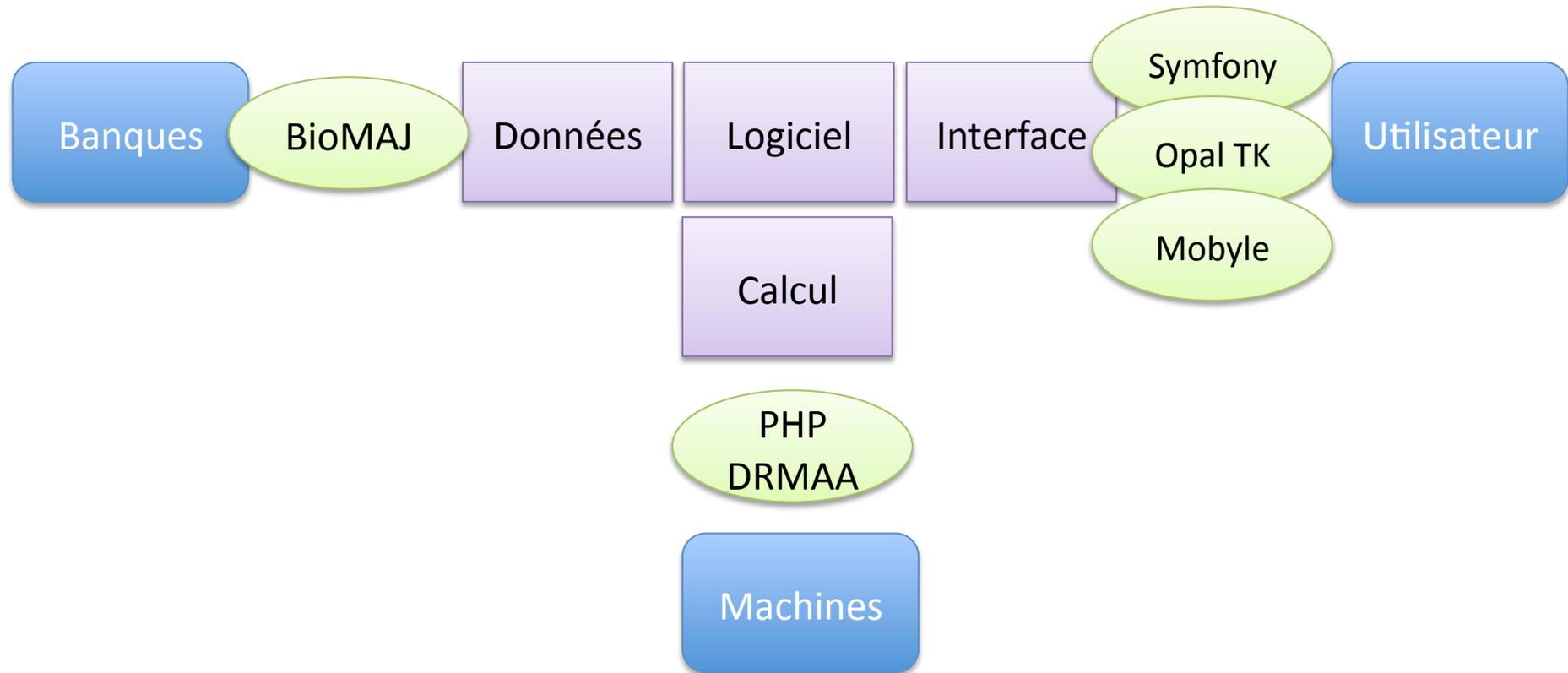


Contributeurs de cet environnement de déploiement de web services

Environnement

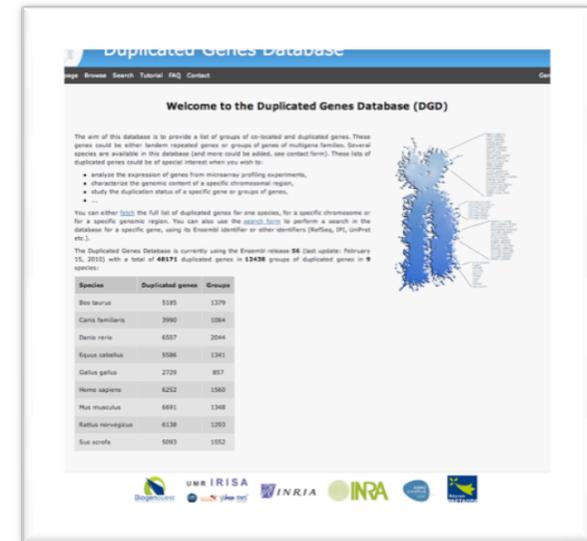
- Briques logicielles
 - BioMAJ : intégration
 - Opal Toolkit : web services
 - Framework Symfony : interfaçage
 - Plugin PHP : soumission DRMAA
 - Mobylye : interfaçage rapide
- Environnement
 - Forge INRIA
 - Git

Environnement



Exemple : DGD

- Duplicate Gene Database v1
- Réunion client : INRA
- Rédaction cahier des charges
- Développements : 1 mois
 - Intégration BioMAJ :
 - Bases de données : données Ensembl! et Xref
 - Interface web (interaction client +++)
 - Web services
- Rédaction article



Utilisateurs

- Local (INRIA-IRISA : informaticien, bioinformaticien)
- Régional (*, bioinformaticien, biologiste)
- National (*, bioinformaticien, biologiste)

Prestation élaborée

Développements
Accompagnement

Prestation de base

Environnement



Autonomie

Positionnement GenOuest

- Laboratoire : interactions avec les équipes INRIA/IRISA
- Université : partenariats avec les laboratoires en Sciences de la Vie
 - Déjà 2 IE mutualisés avec le CAREN et le GERHM – détection des nouvelles problématiques à la source
- National
 - Un des six centres régionaux de ReNaBi : ReNaBi-Grand Ouest
 - Réponse Appel d'Offre Infrastructures Nationales en Biologie Santé
- Européen : projet commun ReNaBi
 - Réponse commune ESFRI/Elixir

Contexte

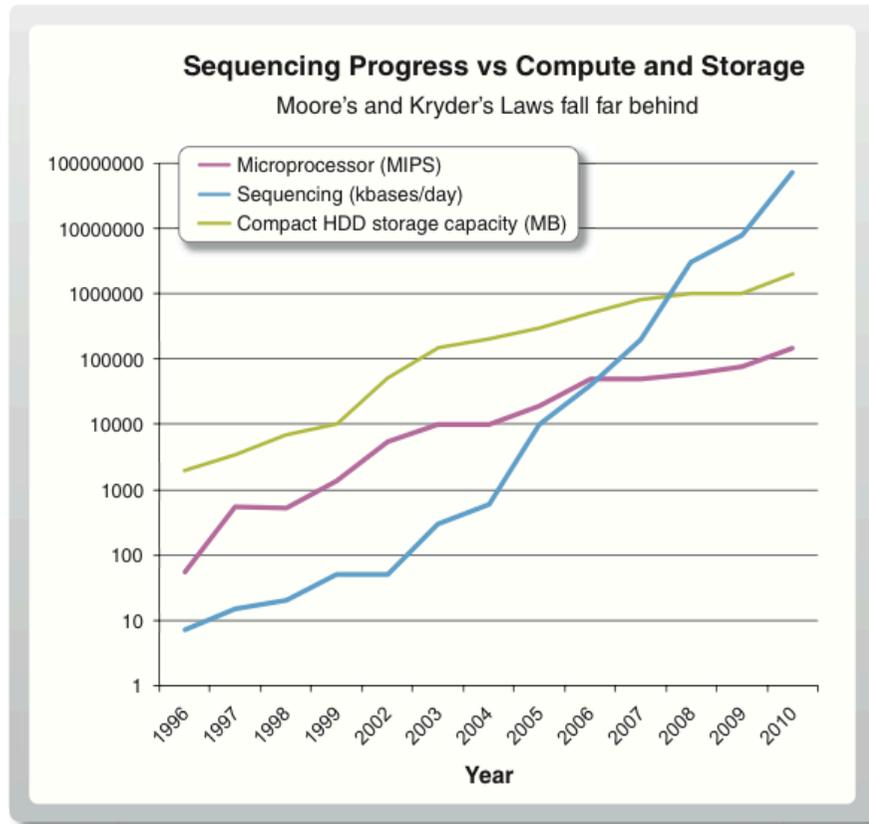


Fig. 1. A doubling of sequencing output every 9 months has outpaced and overtaken performance improvements within the disk storage and high-performance computation fields.

Partenariats forts avec les autres plateformes du réseau Biogenouest

- 1 CDD mutualisé NGS
- 1 CDD mutualisé protéomique

Détection des problèmes « à la source »

- Groupes de réflexion
- Projets communs (ANR, actions incitatives, *ex)

Exemples : projets Université de Rennes 1

IRISA – UMR 6074

<http://apurva.genouest.org/>
<http://bioquali.genouest.org/> 
<http://crispi.genouest.org/> 
<http://gassst.genouest.org/> 

IGDR – UMR 6061

<http://dogs.genouest.org/>
<http://autograph.genouest.org/> 
<http://dogs.genouest.org/rh-server.html>
<http://dog-genetics.genouest.org/>

U835, Upres EA2311

<http://rnplanet.genouest.org/> 

U936

<http://go2pub.genouest.org> 

UMR 6026

<http://mipdb.genouest.org> 
<http://edystrophin.genouest.org/>
<http://cobalt.genouest.org/> 
<http://sorgo.genouest.org/> 
<http://genoweb1.irisa.fr/duals/RASTA-Bacteria/> 

GERHM

<http://mimas.genouest.org> 
<http://www.germonline.org> 
<http://sgv.genouest.org> 
<http://gpsy.genouest.org> 

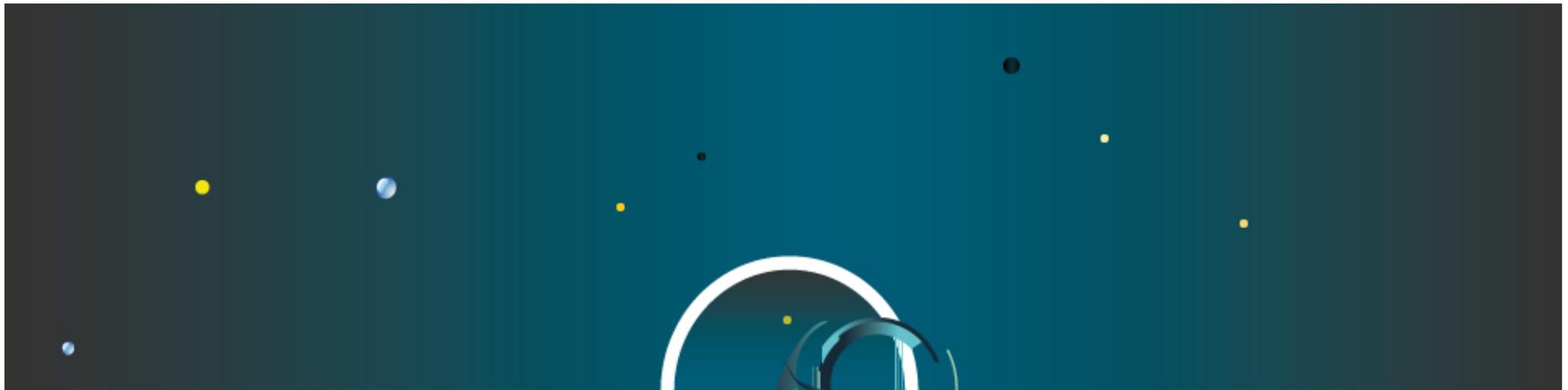
CAREN

<http://seq-454.genouest.org/>
<http://phymycodb.genouest.org/> 
<http://pyroseqman.genouest.org/>

Conclusion

- Travail aux interfaces parfois délicat
- Dualité production/expérimentation
- Investissement en temps (ROI)
- Rôle de conseil vis à vis d'une discipline en pleine mutation

Merci de votre attention
<http://www.genouest.org>



Contrats – RH

Transférer

ANR BioWIC, projet Peapol

Mise à
disposition

ANR PELICAN, ANR ECS, Biogenouest (Région Bretagne)

Fédérer

Projets IBiSA: GRISBI, MobyNet

Développer

ADT INRIA BioMAJ, Projets IBiSA: DrMotifs, Metadata