

Les CATI a INRAE

Centre Automatisé du Traitement de l'Information

- Collectif d'informaticiens qui couvrent un large champ de thématiques

- 3 générations de CATI à INRAE dont la dernière couvre la période 2019 - 2024

- Objectifs des CATI 3G
 - ◆ servir les nouveaux enjeux du numérique
 - ◆ transversalité (inter unités et inter départements) et visibilité renforcée
 - ◆ passage de l'animation à la production

- 18 CATI 3G homologués
 - ◆ collectifs allant de 16 à 150 agents
 - ◆ piloté par 1 ou plus de 10 départements

<https://www.ingenum.inrae.fr/cati>



INRAE

Présentation des CATIs BioS4BioI, BOOM et SysMics

25 juin 2020 / Réunion des bioinformaticiens du centre de Jouy-en-Josas

INRAE

➤ **CATI SysMics**

Sandra Dérozier

Un CATI de 3e génération orienté “recherche”

Objectifs

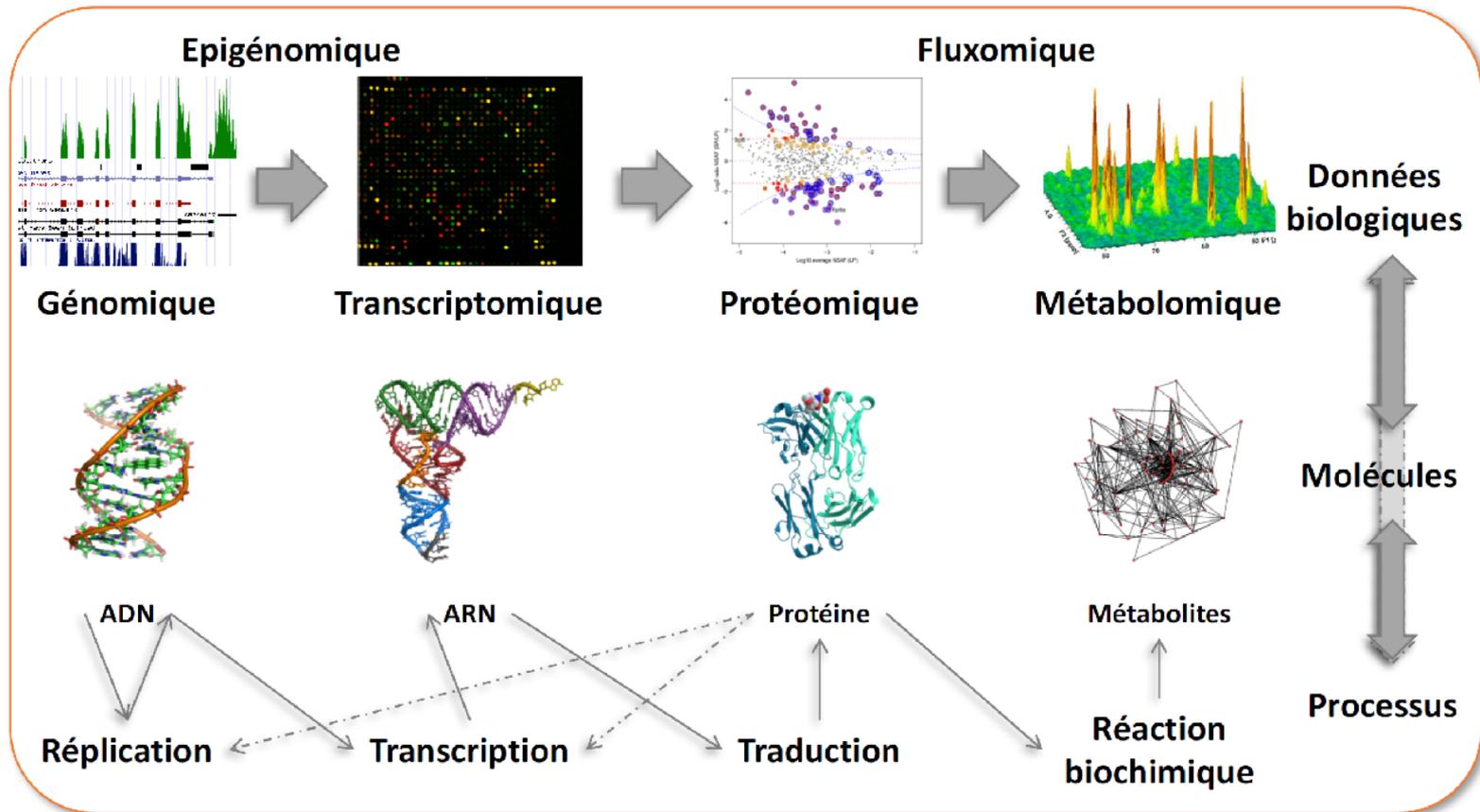
- **développer** un **environnement informatique** intégrant la **notion de processus** (dynamiques) afin de **représenter** un **organisme vivant**,
- **déployer** des **méthodes statistiques** exploitant à la fois des **données “omiques” hétérogènes** (transcriptome, protéome, fluxome, métabolome, phénomène, imageries, ...) et les **processus**.

Missions

- **consolider** et/ou **établir** une preuve de concept de l'**environnement informatique** pour des organismes modèles, *B. subtilis* et *A. thaliana*
- **élaborer** et **déployer** une 1ère **stratégie de transfert** pour un ou des organismes d'intérêt agronomique (maïs ou *R. solanacearum*)
- **constituer** une **communauté interdisciplinaire** autour de l'**intégration systémique** des données et des connaissances biologiques

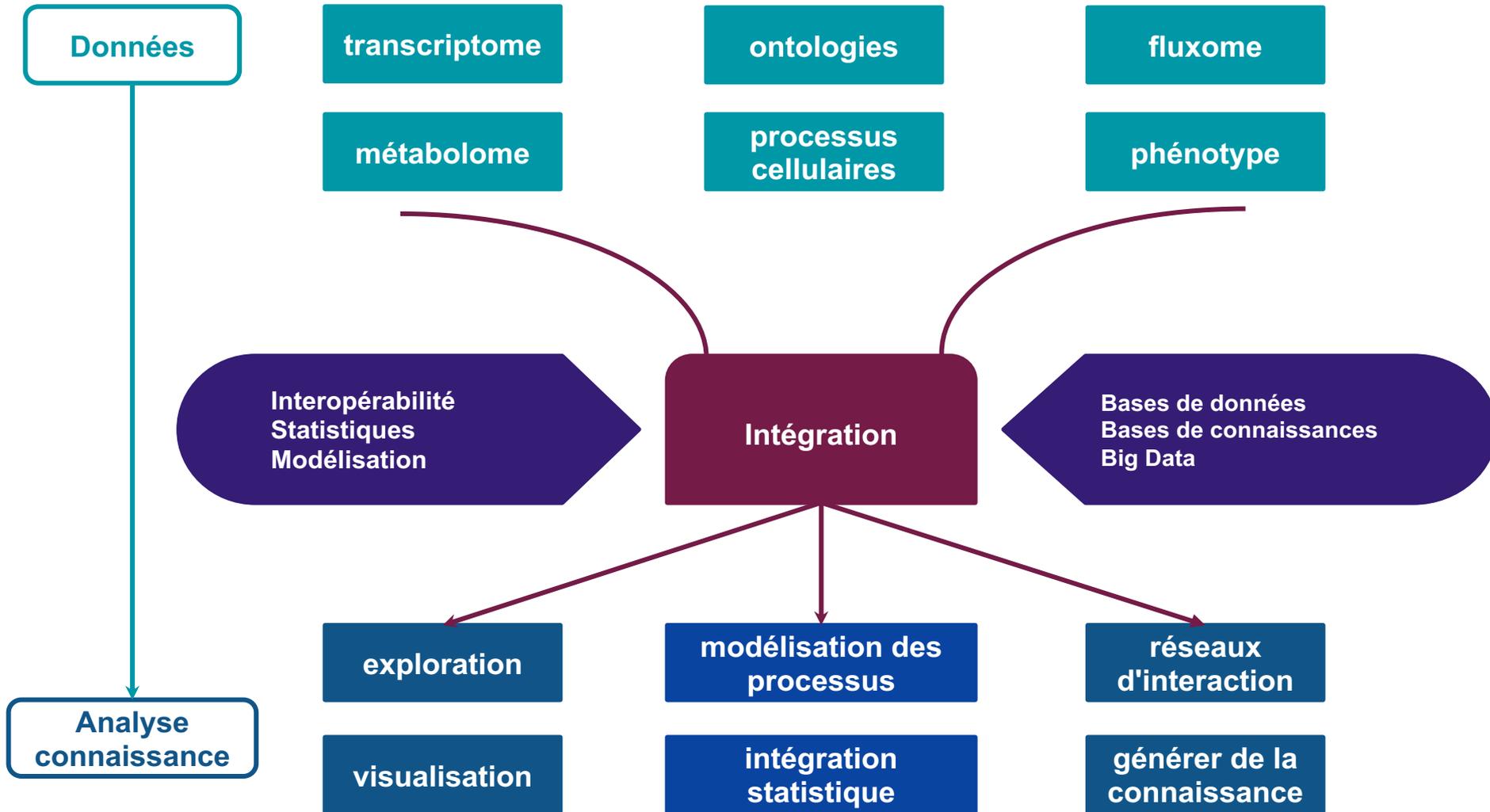


Données hétérogènes et processus biologiques



Un processus biologique : des entités biologiques + un modèle mathématique

Contexte



Les membres

<https://sysmics.cati.inrae.fr/qui-sommes-nous>

Création en 2019 et piloté par **MathNum**, **BAP** et **MICA**



Harry Belcram
GQE



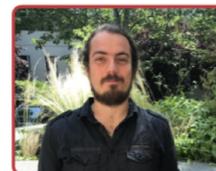
Nadia Bessoltane
IJPB



Eric Biot
IJPB



Véronique Brunaud
IPS2



Guillaume Cerutti
ENS Lyon



Delphine Charif
IJPB



Sandra Dérozier
MaIAGE



Marc Dinh
MaIAGE



Thomas Duigou
MICALIS



Cyprien Guérin
MaIAGE



Cécile Guichard
IPS2



Anne Goelzer
MaIAGE



Olivier Inizan
MaIAGE



Johann Joets
GQE



Christine Paysant-Le Roux
IPS2



Franck Samson
LAMME



Alain Trubuil
MaIAGE

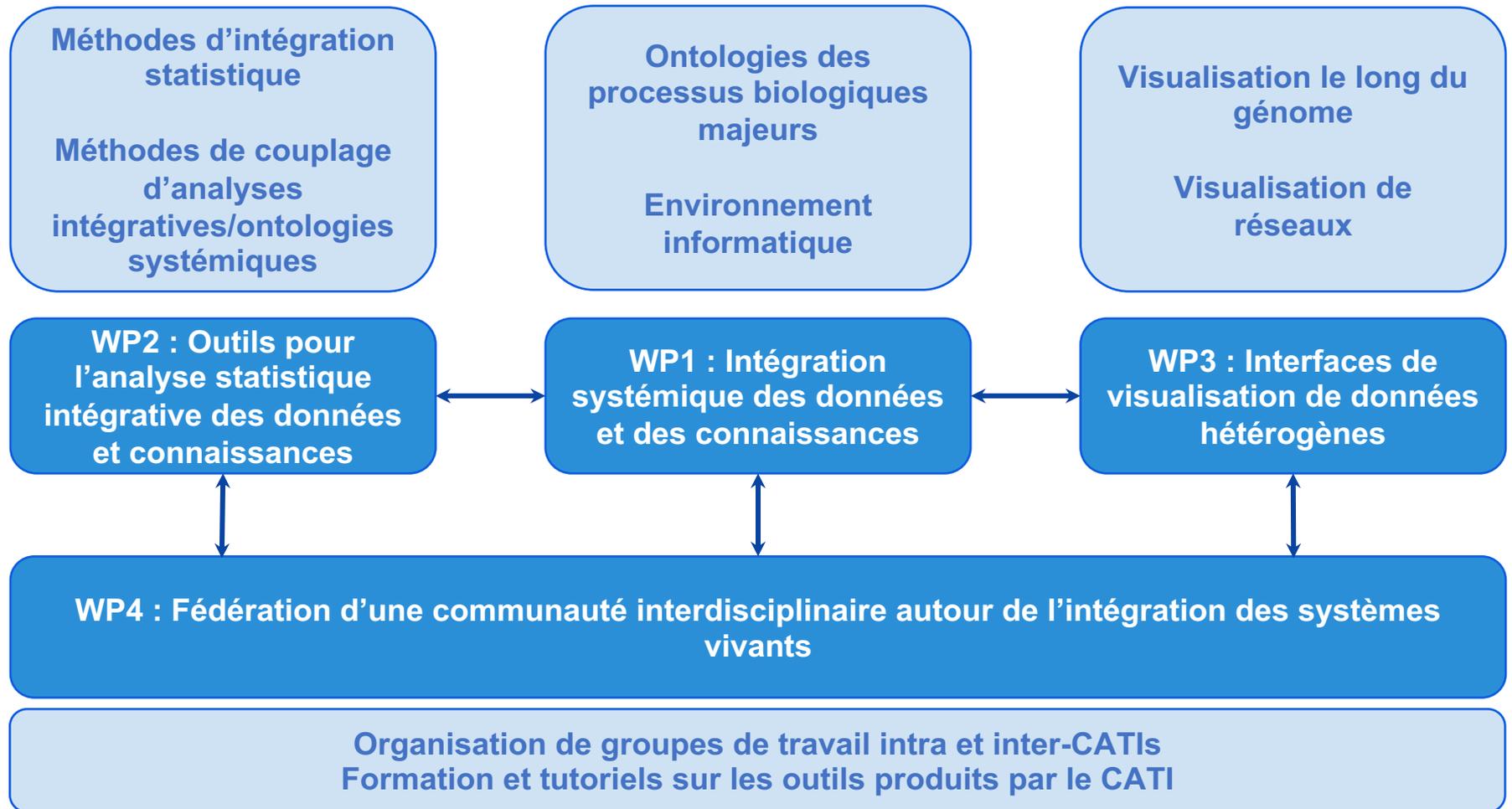


INRAE

Structure du CATI SysMics

<https://sysmics.cati.inrae.fr/workpackages>

Structuration en 4 *workpackages*



INRAE

Présentation des CATIs BioS4Biol, BOOM et SysMics
25 juin 2020 / Réunion des bioinformaticiens du centre de Jouy-en-Josas

Intégration systémique des données et des connaissances

Objectif

Développer un environnement informatique dédié à l'intégration des connaissances et des données "omiques" hétérogènes pour un organisme, en intégrant et unifiant les récentes avancées dans le champ de la biologie des systèmes.

Structuration en **deux entités**

- une ontologie décrivant formellement les processus cellulaires d'un organisme de façon systémique (ainsi que les modèles associés)
- un entrepôt de données hétérogènes structuré par l'ontologie et regroupant la connaissance biologique, des données expérimentales (séquençage, transcriptomique,...), des prédictions bioinformatiques (détection de motifs, phénotypes,...) et des modèles mathématiques associés aux processus cellulaires ainsi qu'à des données plus structurés (réseaux métaboliques et de régulation)



Outils pour l'analyse statistique intégrative de données et connaissances

Objectifs

- produire des outils pour l'analyse intégrative de données hétérogènes
- développer des méthodes d'analyses intégratives avec la prise en compte de connaissances formalisées

Missions

- décrire des données issues des "omiques" et d'imagerie
 - ◆ rédaction de fiches décrivant les données "omiques" et d'imagerie
 - pour une personne d'une autre discipline
 - pour comprendre ce qu'elles représentent/mesurent
 - pour connaître leur nature mathématique (continue/discrète, ...)
 - ⇒ informations utiles pour choisir comment les analyser/les intégrer
- décrire et implémenter des méthodes et outils d'analyse statistiques génériques pour ces données
 - ◆ *package* R et interface RShiny (RFLOMICS : RNA-Seq, protéomique et métabolomique)
- explorer des méthodes et outils d'analyses statistiques intégratives



Interfaces de visualisation de données hétérogènes

Objectif

Développer un ou plusieurs outils permettant de visualiser l'ensemble des données "omiques" hétérogènes disponibles mais également les processus biologiques décrits dans l'environnement structuré proposé à terme dans le *workpackage 1*.

Missions

- **enrichir et adapter** les **outils existants** afin de visualiser les données "omiques" non visualisables actuellement le long du génome
 - ◆ [Genoscapist](#), outil de visualisation de profils quantitatifs (*B. subtilis*)
- **proposer** un outil permettant de **visualiser** les **processus biologiques** (réseaux métaboliques et/ou de régulation) à partir de données extraites du système informatique structuré du *workpackage 1*
 - ◆ [veille](#) sur les outils de visualisation de graphes
- **assurer un lien** entre les différents outils afin de **naviguer** entre les divers types de données ("omiques" et réseaux) ainsi que l'entrepôt de données



Workpackage 4

<https://sysmics.cati.inrae.fr/workpackage4>

Fédération d'une communauté autour de l'intégration des systèmes vivants

Objectifs

Interactions avec les CATIs eMPrEInTE, GREP, BOOM et BARIC sur la représentation, l'organisation et le stockage des données "omiques" hétérogènes.

Actions

- [workshop autour de la description des données "omiques"](#) avec le CATI eMPrEInTE
 - ◆ présentation des données "omiques"
 - ◆ exemples d'intégration multi-"omiques" (intégration statistique, modélisation)

- [présentation du CATI SysMics à l'IJPB](#)
 - ◆ présenter les objectifs du CATI aux biologistes
 - ◆ échanger sur les projets/processus types d'intérêt

- *workshop* autour de l'intégration et de la visualisation de graphes avec les CATIs eMPrEInTE, GREP et BARIC
 - ◆ discussion autour des besoins de chacun
 - ◆ évaluation d'outils



Workshop inter-CATIs SysMics - EMPrEInTE

https://sysmics.cati.inrae.fr/workshop_sysmics_empreinte_2019

Objectifs

- produire un état de l'art sur les connaissances autour des données transcriptomiques, métabolomiques et protéomiques en détaillant les ontologies connues
- échanger sur les analyses statistiques classiquement utilisées dans les différents pipelines d'analyses
- produire un guideline sur les bonnes pratiques de préparation des données "omiques" et faciliter les études nécessitant des approches intégratives



INRAE

Présentation des CATIs BioS4Biol, BOOM et SysMics
25 juin 2020 / Réunion des bioinformaticiens du centre de Jouy-en-Josas

Workshop Work4Graph

CATI impliqués : BARIC, EMPReInTE, GREP et SysMics

Objectifs

- rassembler les membres de 4 CATI autour d'une problématique commune, l'intégration et la visualisation de données sous forme de graphes
- échanger sur les besoins de chaque CATI et partager les expériences de chacun
- évaluer certaines des solutions envisagées afin d'intégrer les données, les interroger et les visualiser
- articuler les futurs flux de données et modèles qui pourront être mis en place entre ces CATI

