



Dispositifs interdisciplinarité IA & science des données en biologie

Mercredi 2 Décembre 2020 - worskhop virtuel



Prospective en Science des Données,
Intelligence Artificielle et Biologie

ORGANISÉ PAR LES CSI DE L'INSB ET DE L'INS2I

Objectif des journées / table-ronde

- **Domaine : Interface entre SD-IA et biologie**
 - SD-IA en biologie moléculaire et cellulaire
 - SD-IA en biologie structurale
 - SD-IA en neurosciences et sciences cognitives
 - SD-IA pour l'utilisation de données humaines pour la recherche en biologie.
- **Objectif : dégager des recommandations aux instituts**
 - Besoins et les ressources de la communauté scientifique
 - Grands axes à renforcer
 - Verrous de tous ordres
 - Investissements nécessaires

Objectif de la présentation: clarifier les dispositifs déjà existant qui pourraient être exploités pour IA&SD@biologie

Organisation globale du CNRS

10 Instituts dont

- **INS2I** = Sciences de l'information et ses interactions
- **INSB** = Sciences biologiques
- (...)

Des missions nationales dont

- **MITI**: Initiatives transverses et interdisciplinarités
- **MiCaDo**: Calcul – Données

Comité national de la recherche scientifique

Conseil scientifique du CNRS

Conseils scientifiques d'instituts (INS2I, INSB...)

41 sections du Comité national dont

- **06** : Sc. de l'information : fondements de l'informatique, calculs, algorithmes, représentations, exploitations
- **07** : Sc. de l'information : traitements, systèmes intégrés matériel-logiciel, robots, commandes, images, contenus, interactions, signaux et langues
- **21**: Organisation, Expression, Evolution des génomes. Bioinformatique et Biologie sys.
- **26** : Cerveau, cognition, comportement
- **28** : Pharmacologie-ingénierie et technologies pour la santé-imagerie biomédicale

4 commissions interdisciplinaires dont

- **CID 51** : Modélisation, et analyse des données et des systèmes biologiques : approches informatiques, mathématiques et physiques

Plans: quelques dispositifs facilitant l'interdisciplinarité

- **A l'échelle du CNRS**
- **Partenariats institutionnels**
- **Soutien aux collaborations**
- **Chercheurs/ses**
- **Infrastructures**
- **Animation**

COP (contrat d'objectif et de performances) CNRS

6 défis sociétaux dont

- L'IA et ses impacts
- Santé et environnement

Priorités scientifiques (extrait)

- Fondements de l'IA: modèles, données et algorithmes
- Systèmes autonomes et interactifs
- Mathématiques pour les sciences du vivant et la médecine
- (...)
- Cerveau humain et fonctions cognitives
- Connaissance des génomes, innovation génétique et environnement
- Vision multi-échelle du fonctionnement des êtres vivants
- Explorer la diversité du vivant

Terrain favorable pour les recherches IA/SD@Biologie

STRUCTURES NATIONALES IMPLIQUEES

Structures transversales au CNRS

- **Recherche:** MITI
- **Calcul/données:** cellule micado
- **Réponses inter-instituts sur actions ciblées**
 - GT santé → CHU
 - Ateliers défis sociétaux COP
 - Modélisation covid

Réflexion interfaces sciences et données

- **Nouveaux domaines de recherche émergents ?**
- **Quelle est la frontière entre recherche et ingénierie en IA/SD expérimentales ?**
- Réflexion avancées pour confronter les points de vue des instituts et des sections.

La problématique IA/SD@Biologie est plutôt en avance de phase

DEUX MODELES DE PARTENARIATS POUR L'INTERDISCIPLINARITE

- Collaboration sur projets ciblés

- Collaboration entre des équipes appartenant à leurs UMR respectives
- Concerne plutôt les thématiques des **sections 6 / 21 (bioinformatique, biologie des systèmes)**

- Collaborations institutionnelles

- Affectation de chercheurs d'un institut dans des unités dépendant d'autres instituts
- UMR avec tutelles partagées
- Concerne plutôt les thématiques des **sections 7 / 26 / 28 (neurosciences...)**

(source: CSI INS2I du 1/12/2020)

Question: quel modèle faire émerger des futures recherches en IA/SD@Biologie ?

COLLABORATIONS ENTRE EQUIPES

Collaborations entre unités INSB et INS2I (> 1/3 unités)

- Génomique
- Biologie des systèmes
- Biologie structurale
- Imagerie
- Neurosciences
- Masses de données

Collaborations INS2I sur thématiques santé (> 1/3 unités)

- Robotique, bio-artefacts, ICM, implants, capteurs
- Données: gestion et intégration de données
- Analyse d'image
- Artefact sociaux, agents virtuels, détection de trouble
- Criblage médicaments: repositionnement, microbiote

(source: enquête GT santé + enquête INS2I, 2020)

COLLABORATIONS ENTRE EQUIPES

Appels de la MITI

- Mastodons
- Imag'In
- Santé numérique
- Projets momentum

Soutien ciblé chercheur

- Osez l'interdisciplinarité (plus de 30 projets)

Bourses 80'prime

- 3 à 4 CD INSB/INS2I par an environ en 2020
- dont 2 SD/IA en 2019 par exemple

AAP instituts

- Jeunes chercheurs
- Projets transversaux des laboratoires

Collaborations entre unités INSB et INS2I (> 1/3 unités)

- Génomique
- Biologie des systèmes
- Biologie structurale
- Imagerie
- Neurosciences
- Masses de données

Collaborations INS2I sur thématiques santé (> 1/3 unités)

- Robotique, bio-artefacts, ICM, implants, capteurs
- Données: gestion et intégration de données
- Analyse d'image
- Artefact sociaux, agents virtuels, détection de trouble
- Criblage médicaments: repositionnement, microbiote

(source: enquête GT santé + enquête INS2I, 2020)

Moyen d'action possibles
→ MITI, Thèses 80'prime

PARTENARIATS INSTITUTIONNELS

Entre organismes: équipes ou unités mixtes

- INS2I : ERL (équipes communes) avec l'Inserm
- INSB : UMR communes avec l'Inserm ou l'Inrae

Entre instituts : UMR communes

Entre équipes CNRS: Labellisation PRIME

- Récent : juillet 2020.
- Collaborations entre deux équipes hébergées par des laboratoires dépendant de deux instituts
- Projet de recherche inscrit dans le temps (2 à 5 ans)
- Validé par le DGDS + soutien financier possible

Unités INS2I & INSB

- 10 unités INS2I/INSB
- 3 unités INS2I/INSIS/INSB

Unités avec d'autres organismes

- 14 unités INSB-INRAE
- 3 unités INS2I-INSERM
- 41 unités INSB-INSERM

Thématiques (analysées par CSI)

- Signal, Image, Capteurs et Visualisation
- IA, Data, Bioinformatique
- Modélisation et simulation

(source: CSI INS2I, labintel)

RECRUTEMENT DE CHERCHEURS

Principe général: *chaque institut recrute et promeut des chercheurs sur la base des propositions des sections et CID*

- **Recrutements via CID 51**
 - Affectations dans une dizaine de laboratoires INS2I depuis 2013.
- **Recrutements croisés**
 - Avant la création de la CID51
- **Recrutements et promotions via les sections**
 - Section 7, section 6
- **Mobilité inter-instituts**
 - INS2I → INSB assez régulier

Chercheurs INS2I affichant des recherches en biologie-santé-handicap

- autour de 75 (30 section 6, 35 section 7)

Unités INS2I hébergeant des chercheurs en biologie

- 33 chercheurs
- 12 unités
- Sections 21, 26, 28

Unités INSB hébergeant des chercheurs en informatique

- 14 chercheurs
- 7 unités
- Section 7++

(source: CSI INS2I, labintel)

Moyen d'action → Recrutements SD/IA via la CID 51 et les sections

SUPPORT ET SOUTIEN, INFRASTRUCTURES

Moyens ingénieurs

- Affectés par les instituts dans des UMR, UMS ou USR
- Sur des missions/besoins nationaux
- Sur des plateformes locales ou des unités
- Actions concertées sur un site à promouvoir.

Institut Français de Bioinformatique

- UMS: tutelle INSB
- Dialogue INS2I pour les demandes de postes
- Projet PIA3 centré sur la SD/IA
- Cloud : CC-IN2P3. Calcul: IDRIS

Centre de calcul national IA : IDRIS (Jean-Zay)

- Jean Zay: accessible à tous gratuitement et très facilement pour les applications IA
- Il suffit de faire la demande : <https://www.edari.fr/>
- Support dédié à IDRIS
- Seule limite : pas de données de santé

IDRIS

- 15 personnes pour le support aux utilisateurs
- 5 personnes pour le support IA

IFB-core

- 14 personnels ITA (dont 2 permanents)
- Action biologie intégrative en lien avec les besoins SD/IA

Moyen d'action

→ IFB (PIA3) et Jean Zay

ANIMATION

GDR

- Structuration des communautés INS2I dans la durée
- Structurations en axe, GT, actions qui sont en lien avec la santé
- Gros GDR et budgets conséquents

Ecoles

- Ecole thématique Masses de données
- Ecole thématique Algo/SB, BioRegul
- Ecole thématique bioinformatique
- Animation de l'IFB sur la biologie intégrative

SD/IA@biologie présent dans au moins 8 GDR

- Bioinformatique moléculaire (BIM)
- Masses de Données, Informations et Connaissances en Sciences (MaDICS, 2015)
- Informatique Géométrique et Graphique, Réalité Virtuelle et Visualisation (IG-RV)
- Informatique Mathématique (IM)
- Information, signal, images, vision (ISIS)
- Modélisation, analyse et conduite des systèmes dynamiques (MACS)
- Robotique
- Fondements formels et algorithmiques de l'Intelligence Artificielle (IA)

Moyen d'action

- (1) Nouvelles actions au sein des GDR
- (2) Croiser les actions des GDR ?

Synthèse

- **A l'échelle du CNRS** : COP, MITI, structures transversales
- **Partenariats institutionnels** : unité mixtes, projets prime CNRS
- **Soutien aux collaborations**: MITI, 80'prime, osez l'interdisciplinarité
- **Chercheurs/ses**: recrutements croisés, via CID, mobilité inter-instituts
- **Infrastructures**: IFB, Jean Zay
- **Animation**: GDR
- *Autres dispositifs à inventer*

TABLE RONDE

Questions scientifiques

- Apports espérés de la SD/IA dans chaque thème
- Points bloquants

Questions organisationnelles

- Verrous structurels ?

Recommandations



www.cnrs.fr