

Forum des Regroupements de recherche IAR³

R³AI Research Regroupements Forum

3 octobre 2024
October 3, 2024



Université 
de Montréal



HEC MONTRÉAL



A low-angle shot of a modern glass skyscraper. The sun is shining brightly from the top center, creating a lens flare effect. The glass facade reflects the blue sky and white clouds. The text 'Bienvenue' and 'Welcome' is overlaid in white, bold, sans-serif font. The text is centered horizontally and spans across the middle of the image. The building's structure is composed of a grid of glass panels and metal frames. The sky is a clear, vibrant blue. The overall scene is bright and modern.

Bienvenue
Welcome



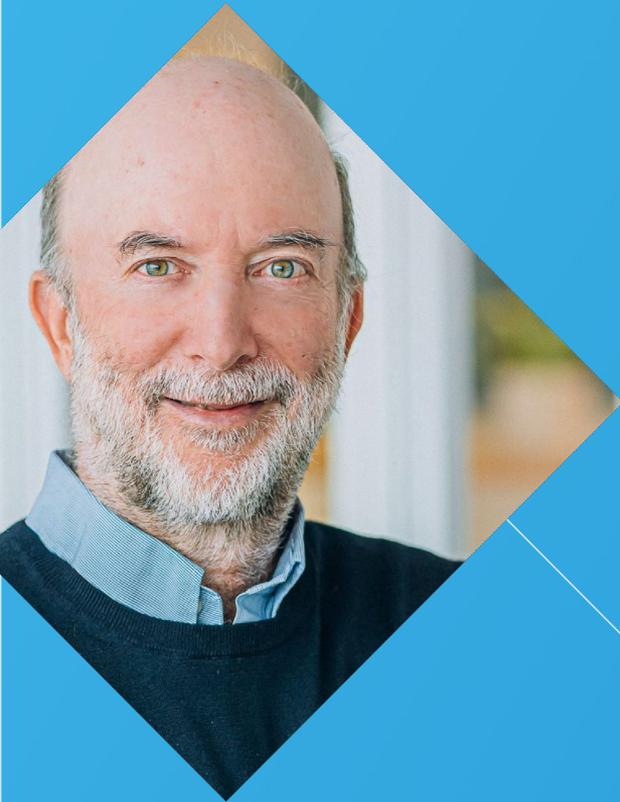
Animatrice / Host:

Virginie Portes

Directrice du soutien à la recherche IVADO

Director of Research Support, IVADO





Luc Vinet

Directeur général IVADO
Professeur Aisenstadt
au département de physique
de l'Université de Montréal

IVADO CEO
Aisenstadt Professor
Department of Physics
Université de Montréal





Yoshua Bengio

Directeur scientifique d'IVADO et du Mila
Professeur au Département d'informatique
et de recherche opérationnelle - UdeM

Scientific Director, IVADO and Mila
Professor, Department of Computer Science
and Operations Research - UdeM



10 Regroupements de recherche



IA -
neurosciences



Apprentissage
automatique



TALN



Mise en œuvre et
gouvernance
responsable de l'IA



Éthique, EDI
et engagement
autochtone



Molécules et
matériaux



Environnement



Systèmes
de santé



Chaînes
d'approvisionnement



Sûreté et
alignement de l'IA





REGROUPEMENT R10

Sûreté et alignement de l'IA

AI Safety and Alignment

Yoshua Bengio (Université de Montréal)

Chris Pal (Polytechnique Montréal)

Dhanya Sridhar (Université de Montréal)

Sûreté et
alignement
de l'IA

AI safety and
alignment



Objectives and Research axes

Goal: to accelerate this much-needed research towards developing AI agents that are safe and reliable.

Axis 1: Evaluating and mitigating LLM biases and unsafe behaviors

Axis 2: Explaining the predictions of black-box AI systems

Axis 3: Improving LLM interpretability by learning symbolic world models

Axis 4: Promoting safe and robust behaviors in multi-agent settings with LLMs

Sûreté et
alignement
de l'IA

AI safety and
alignment



Projects

Axis 1: Evaluating and mitigating LLM biases and unsafe behaviors

3 sub-objectives that focus on two types of short-term harms that we risk with modern day AI systems:

- 1. harmful data biases amplified by LLMs (YB),**
 - Can current alignment methods (such as RLHF) yield insufficient diversity of outputs?
 - More generally, how do we reliably estimate that a social norm or safety specification would be crossed?
- 2. vulnerabilities of LLM web agents that are exploitable by adversaries (CP),** including other LLM web agents that discover exploitative policies
 - Why are neural networks so vulnerable to adversarial attacks?
 - Can we use exploration methods to identify possibly harmful trajectories?
 - If this is inherent to the nature of these systems, is there a way to assess the risk that a query will make the AI produce a wrong answer, i.e., detect adversarial attacks? This may be used to reject such queries.
 - Can we better train systems in a way that takes into account the game theoretical setting with opponents?
- 3. undesirable behaviors from agents that interact with the physical world in unseen scenarios (CP).**
 - Design environments, e.g., simulated driving, where these questions can be studied and mitigations evaluated.
 - How can we get better OOD generalization in the first place?
 - Is there a way to model the data that is by construction going to generalize systematically and better OOD?

Sûreté et
alignement
de l'IA

AI safety and
alignment



Projects

Axis 2: Explaining the predictions of black-box AI systems

3 sub-objectives focusing on

1. improving LLM self-explanations (YB)

- Can we trust the explanations given by an LLM, and under what conditions should we not trust them?
- Can we make the LLMs reason in a more coherent way so that we can trust these explanations?
- Could we train a second network that evaluates the consistency of the explanations given by a first network?
- What can we learn from the recent work on OpenAI o1?

2. learning abstractions of LLMs (YB, DS)

- Can we discover latent abstractions, e.g., using causal representation learning?
- Can we get better generalization by disentangling such causes or capturing their internal causal dependencies?
- Can we use a generative network to sample causal explanations (structured in a causal graph, for example)?
- How do we handle the fact that there will generally be multiple causal explanations, so we would like to be Bayesian about them?

3. visualizing neural networks adaptations (YB)

- Can we exploit the availability of neural activities to better understand the outputs of LLMs?
- Can we discover abstractions used by the neural network to obtain its output, that would help interpret it?

Sûreté et
alignement
de l'IA

AI safety and
alignment



Projects

Axis 3: Improving LLM interpretability by learning symbolic world models

2 sub-objectives:

1. learning posteriors over symbolic models (YB)

- Humans seem to prefer more compact (lower description length) explanations and be able to entertain multiple explanations, i.e., have a posterior over them. How could we do this efficiently in ML?
- How do we go from posteriors on the causal model over observed variables to posteriors that include latent variables representing a symbolic explanation?
- Can we extend and apply GFlowNets to sample such symbolic world models in a way that will scale?

2. speeding up posterior inference by leveraging symmetries (YB)

- Can we hand-craft or even discover symmetries that considerably reduce the uncertainty in the posterior?

Sûreté et
alignement
de l'IA

AI safety and
alignment



Projects

Axis 4: Promoting safe and robust behaviors in multi-agent settings with LLMs

3 sub-objectives:

- 1. Novel evaluations of multi-agent dynamics with LLM agents**
 - Study the exploitation and manipulation risks of LLM-based agents in negotiation scenarios, examining their behavior under different prompts and training methods
 - Explore how RL training and opponent shaping algorithms can guide these agents toward cooperative strategies that avoid selfish, sub-optimal outcomes
- 2. Developing algorithms and environments for multi-agent learning**
 - Building on insights from negotiation games, the project aims to develop new RL algorithms that train safe, non-exploitable agents aware of multiple agents.
 - Explore feedback loops in multi-agent interactions, using game theory to address the challenges of learning cooperative strategies and identifying undesired behaviors that arise from overlooked feedback dynamics.
- 3. Incentivizing social good using causal models (DS)**
 - Examine how causal models and causally-aware agents operating in multi-agent environments might be more robust to undesirable behaviors (which can be reinforced when agents interact)

Sûreté et
alignement
de l'IA

AI safety and
alignment



Links with other Regroupements

- R2 - Machine Learning
- R3 - NLP
- R6 - Molecules and materials
- R4 - Implementation of science and responsible governance

Regroupement #10

Sûreté et
alignement
de l'IA

AI safety and
alignment



Conseillère : Jacqueline Sanchez
jacqueline.sanchez@ivado.ca

REGROUPEMENT R9

Chaînes d'approvisionnement et systèmes de mobilité

Supply Chains and Mobility Systems

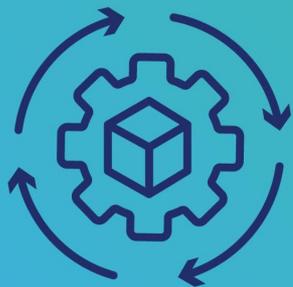


Yossiri Adulyasak (HEC Montréal)

Jean-François Cordeau (HEC Montréal)

Erick Delage (HEC Montréal)

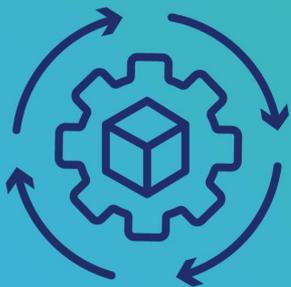
Emma Frejinger (Université de Montréal)



Chaînes d'approvisionnement et systèmes de mobilité

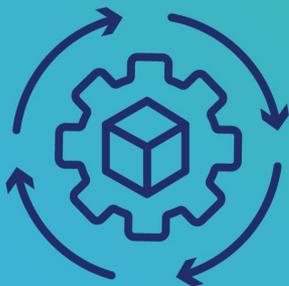
Chaînes d'approvisionnement: réseaux d'approvisionnement, de production et de distribution qui permettent de transformer les matières premières en produits finis pour répondre à la demande des clients

Systèmes de mobilité: ensemble des modes de transport permettant le déplacement des personnes et des biens en milieu urbain



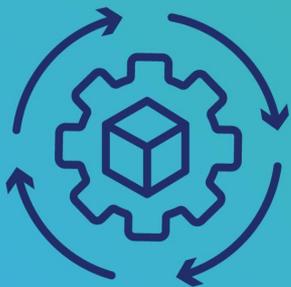
Pourquoi l'intelligence artificielle?

- Énormément de **décisions** à prendre pour la mise en place et l'opération des réseaux
- Plusieurs **composantes interreliées** (production, transport, entreposage)
- **Incertitude** importante (demande, temps de transport, coûts, capacités)
- **Perturbations** fréquentes (pandémies, guerres, grèves, phénomènes naturels)



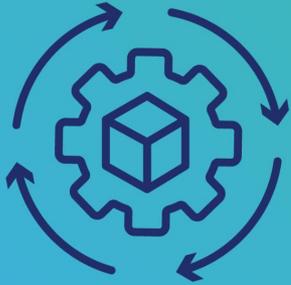
Apprentissage et optimisation

- Milieu idéal pour la combinaison des méthodes d'apprentissage automatique et des algorithmes d'optimisation
 - **Apprentissage** pour faire des prévisions des paramètres incertains, quantifier l'incertitude, comprendre le comportement des clients ou des passagers, etc.
 - **Optimisation** pour prendre les meilleures décisions possibles



Vision et objectifs

- Améliorer la performance et la robustesse des chaînes d'approvisionnement et des systèmes de mobilité
 1. Exploiter l'apprentissage automatique pour concevoir des modèles d'optimisation qui vont rendre les réseaux **plus réactifs** aux changements
 2. Rendre **plus rapide** la résolution des problèmes de très grande taille qui vont en résulter
 3. Quantifier l'incertitude et rendre les chaînes d'approvisionnement **plus résistantes** aux perturbations de grande ampleur

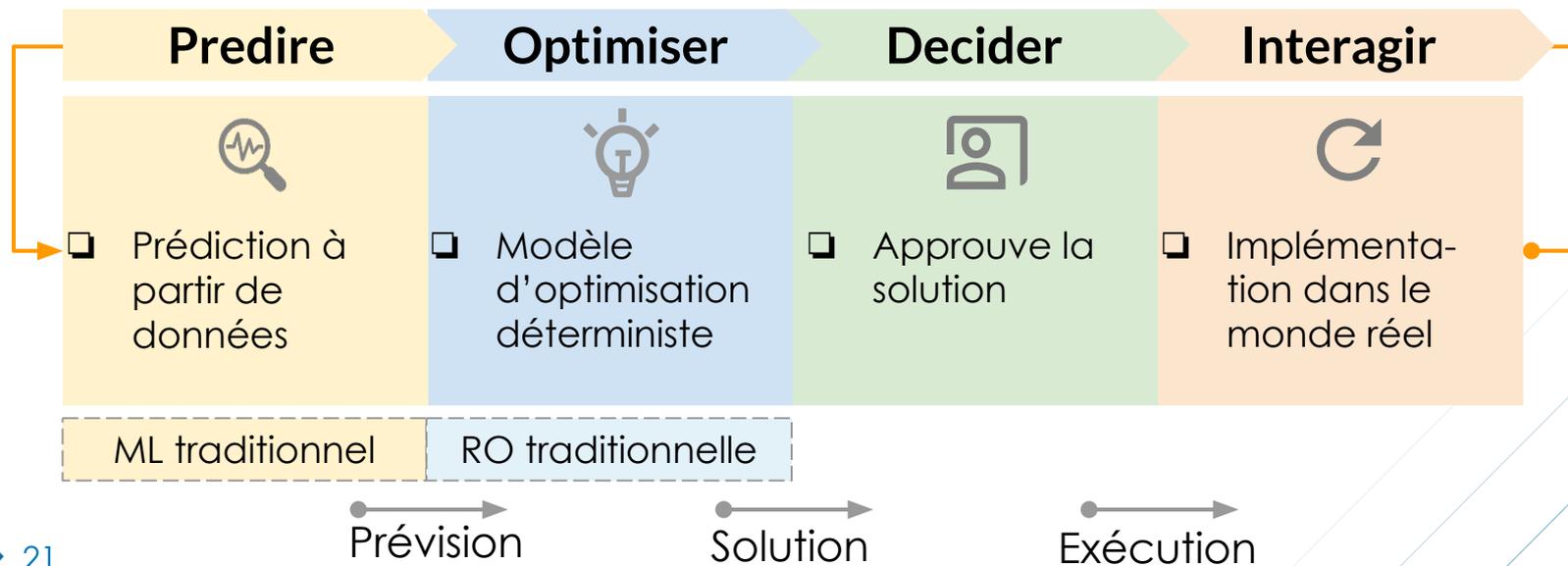


Les trois R de IAR³

- Concevoir des systèmes qui sont **robustes** par rapport aux changements dans les données et qui mènent à des décisions qui sont elles-mêmes robustes face à l'incertitude
- Être en mesure de bien expliquer les outputs et sacrifier un peu d'optimalité au profit d'une meilleure stabilité et interprétabilité des résultats de manière à ressembler davantage au **raisonnement** humain
- Établir de bons compromis entre les objectifs des différentes parties prenantes (clients, fournisseurs, employés, actionnaires, public, gouvernements) de manière à prendre des décisions **responsables**

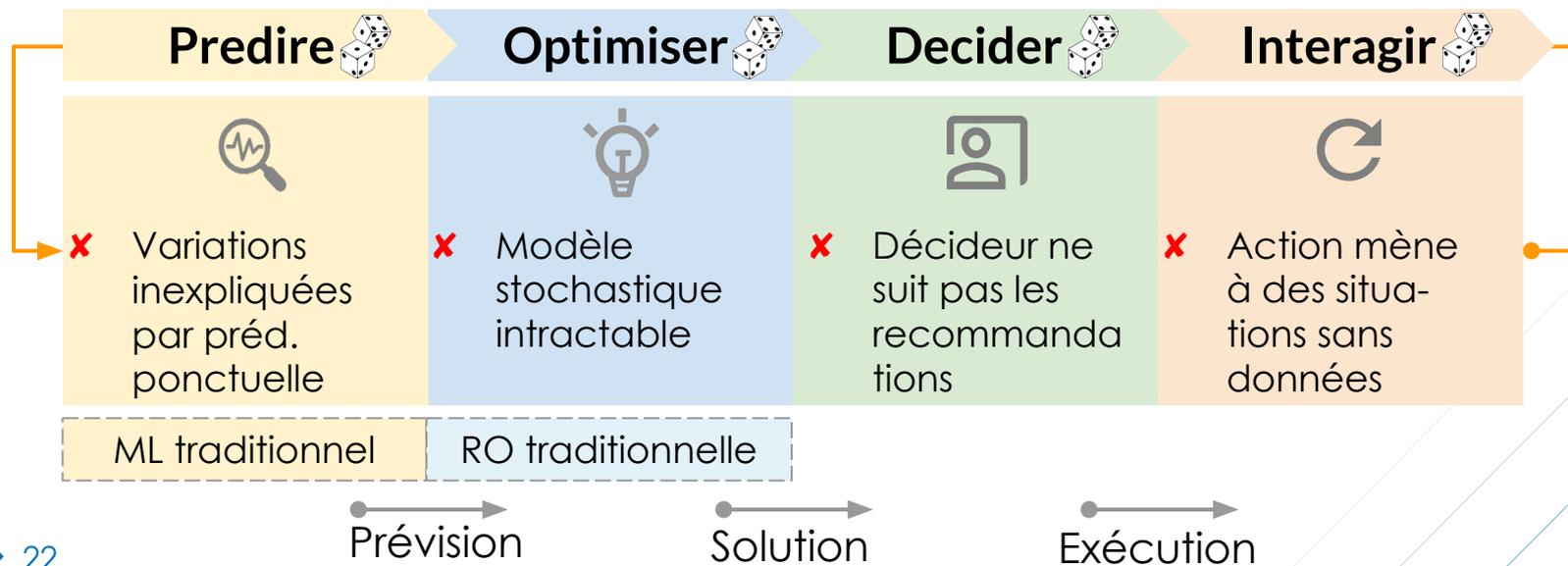
Cadre conceptuel du regroupement

La prise de décision est traditionnellement faite en silos



Cadre conceptuel du regroupement

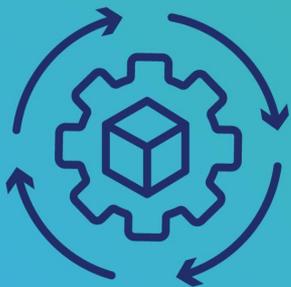
De sévères difficultés surgissent lorsqu'entre en jeu l'incertitude



Axes de recherche

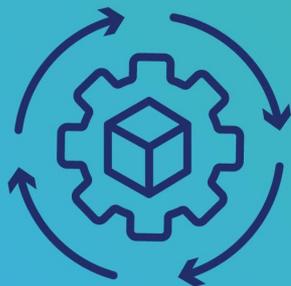
1. Quantification complexe de l'incertitude:
 - Robustesse hors distribution, prévisions de données de comptage, structure de corrélation complexe, apprentissage sous observabilité partielle
2. Intégration de la prédiction et de l'optimisation:
 - Environnement Gym pour l'optimisation contextuelle, optimisation robuste des chaînes d'approvisionnement, routage et planification rentables des navires
3. Accélération de la résolution des modèles multi-étapes:
 - Décomposition de Bender assistée par ML, apprentissage par renforcement pour la gestion des stocks, gestion de flotte pour le transport ferroviaire
4. Optimisation de bout en bout consciente des décideurs
 - Prise de décision équitable, apprentissage de politiques averses au risque, intégration des comportement humain
5. Intégration de l'endogénéité
 - Apprentissage et optimisation intégrés pour les prévisions de la demande, atténuation de l'incertitude épistémique dans l'apprentissage hors ligne





Liens avec les autres regroupements

- R2 - Apprentissage automatique:
 - optimisation de politiques long terme (e.g. apprentissage par renforcement)
 - quantification de l'incertitude (e.g. réseau de neurones épistémique, "conformal prediction")
 - apprentissage de bout-en-bout, etc.
- R7 - Environnement:
 - intégration de l'impact sur l'environnement
 - modélisation d'événements climatiques extrêmes
- R8 - Système de santé:
 - réduction des coûts d'opération du système de santé (e.g. gestion d'inventaire, d'approvisionnement, confection d'horaire, etc.)



APPELS À PARTICIPATION R9

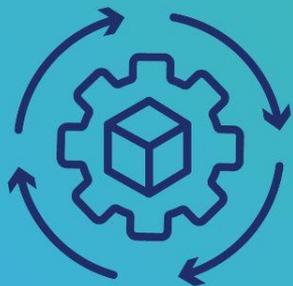
- ◇ Appel à propositions de projets prévu en automne 2024
- ◇ Organisation d'ateliers thématiques
- ◇ Pour information : écrire à Dana F. Simon, conseillère
IVADO: dana.simon@ivado.ca



Regroupement #9

Chaînes
d'approvisionnement
et systèmes de
mobilité

Supply chains and
mobility systems





Questions?

programmes-excellence@ivado.ca

<https://ivado.ca/>



REGROUPEMENT R8

SYSTÈMES DE SANTÉ

Healthcare Systems



Michaël Chassé (CHUM)

Christian Gagné (Université Laval)

Nadia Lahrichi (Polytechnique Montréal)

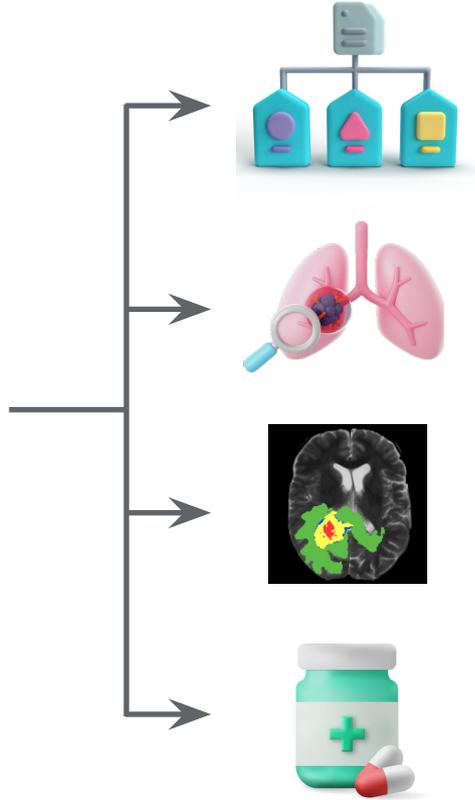
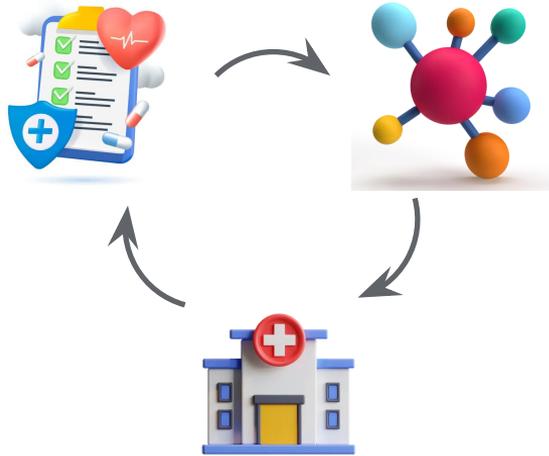
Aude Motulsky (Université de Montréal)

An Tang (CHUM)

Contexte

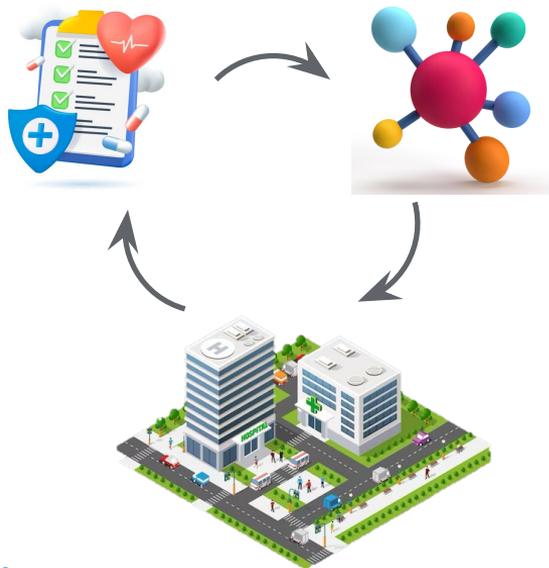


IVADO PRF-1 et 2





IVADO PRF-3



Généralisation des modèles:

- Contexte réel de soins médicaux
- Changement de distribution
- Utilisation 2ndaire de données médicales
- Développement de modèles
- Données et code en libre accès



Constat #1



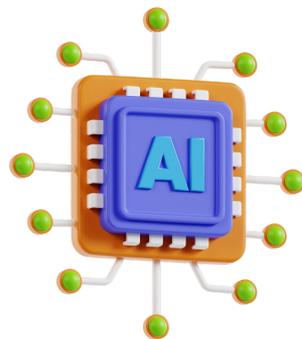
Succès de l' écosystème IVADO



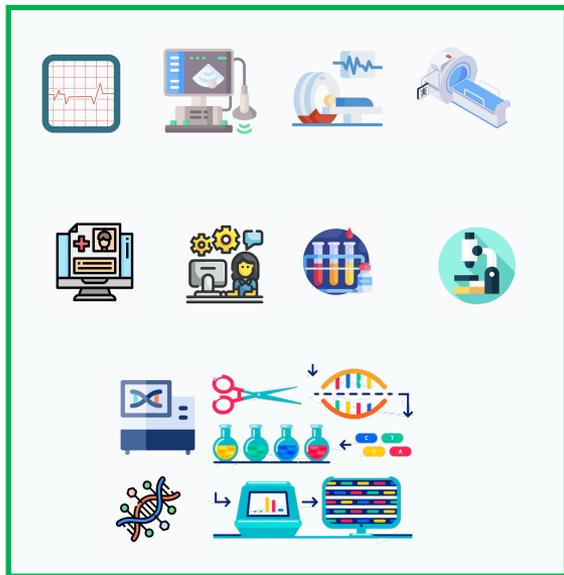
Réalité de la recherche appliquée en IA



Constat #2



Modèles performants



Données variées



Large volume



Représentatives





Systeme de sante IVADO IAR³

Réseau d'hôpitaux

IVADO PRF-3

- Généralisables

Cohorte

IVADO PRF-1 et 2

- Faisables

- Dignes de confiance
(robustes et responsables)
- Performants
- Équitables
- Pertinents



Systemes
de santé

Healthcare
systems



Objectifs

1. **Identifier les cas d'usage prometteurs de modèles fondationnels** pour le système de santé universel au Québec parmi les domaines suivants : soins thérapeutiques, gestion, surveillance et promotion de la santé, une seule santé.
2. **Mettre en place la gouvernance et l'infrastructure** nécessaire au déploiement de modèles fondationnels multimodaux robustes et responsables, dans un contexte réel, agnostics aux types de données et aux besoins auxquels les modèles visent à répondre.

Systemes
de santé

Healthcare
systems



Objectifs

3. **Co-construire un cadre d'évaluation de la qualité de ces infrastructures**, des modèles eux-mêmes et de la valeur ajoutée des modèles selon la perspective des utilisateurs, des organisations et du système de santé.
4. **Analyser les facteurs qui favorisent le développements et la mise à l'échelle de ces algorithmes** au niveau organisationnel et systémique, afin d'élaborer des pistes d'action pour les décideurs.

Systèmes
de santé

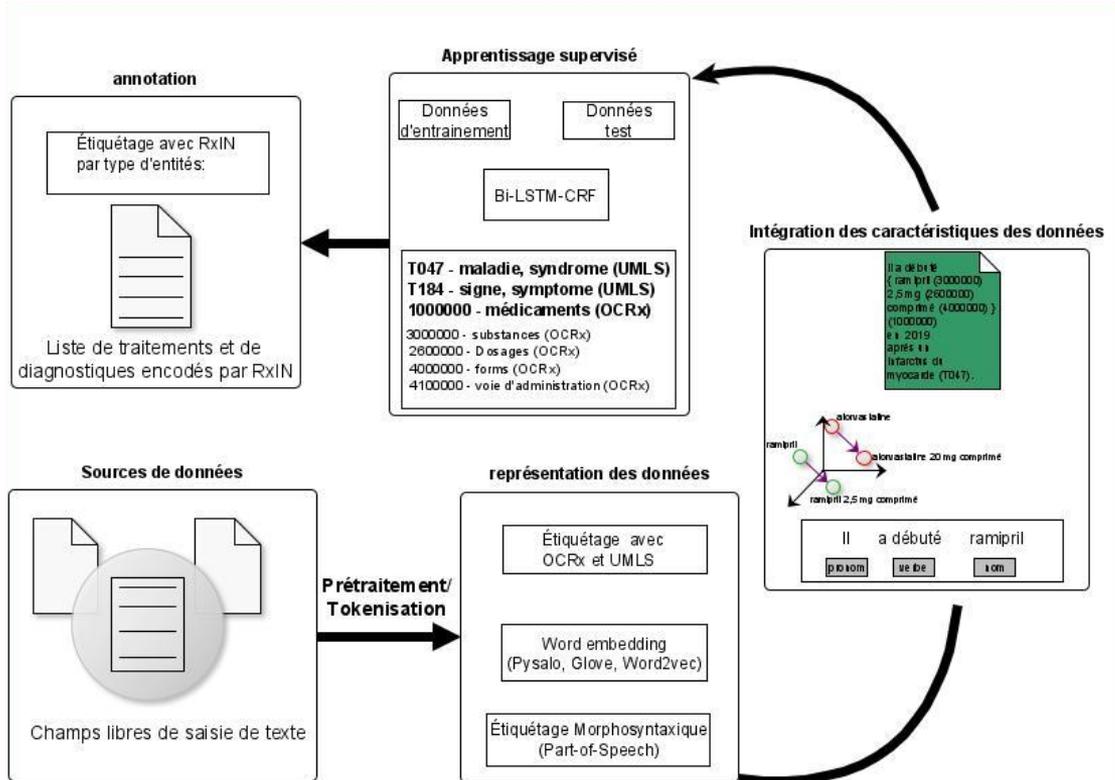
Healthcare
systems



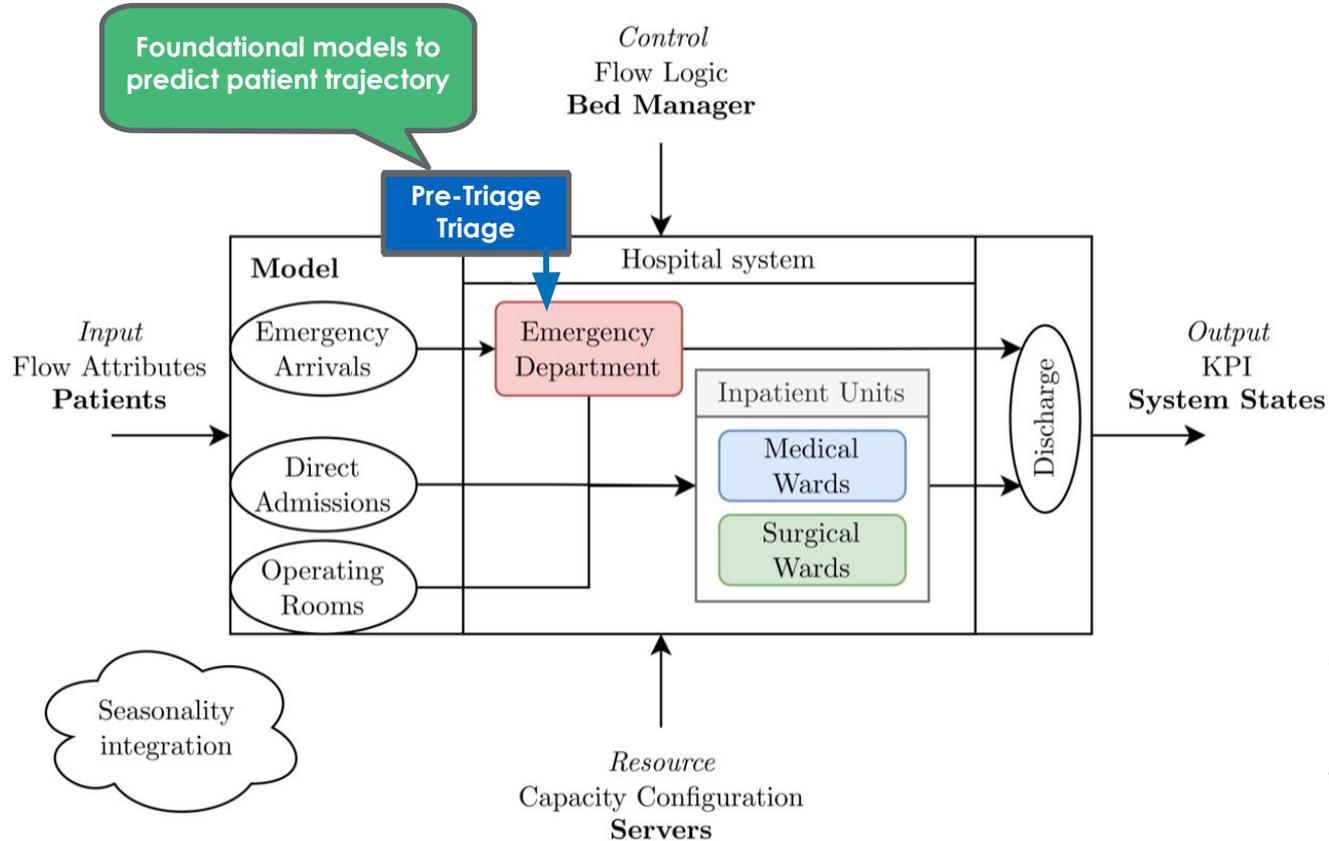
Axes de recherche

- ◇ **Axe 1. Documentation clinique des épisodes de soins**
Aude Motulsky, Jean Noël Nikiema, Michaël Chassé
- ◇ **Axe 2. Prédiction de trajectoire de soins**
Nadia Lahrichi, Esli Osmanliu, Simon Berthelot
- ◇ **Axe 3. Dépistage de donneurs d'organes potentiels**
Michaël Chassé
- ◇ **Axe 4. Imagerie oncologique**
An Tang, Philippe Després
- ◇ **Axe 5. Radiologie et cardiologie**
Robert Avram et Julie Hussin

AXE 1- Documentation clinique des épisodes de soins



AXE 2- Prédiction de trajectoire de soins

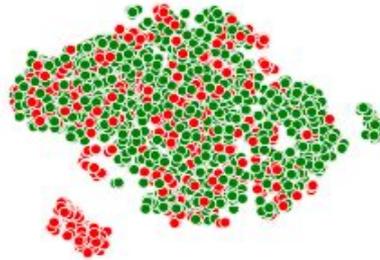
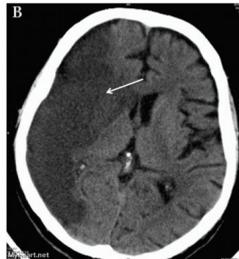
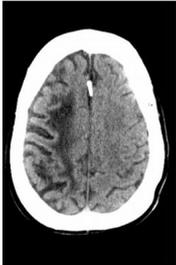


AXE 3- Dépistage de donneurs d'organes potentiels

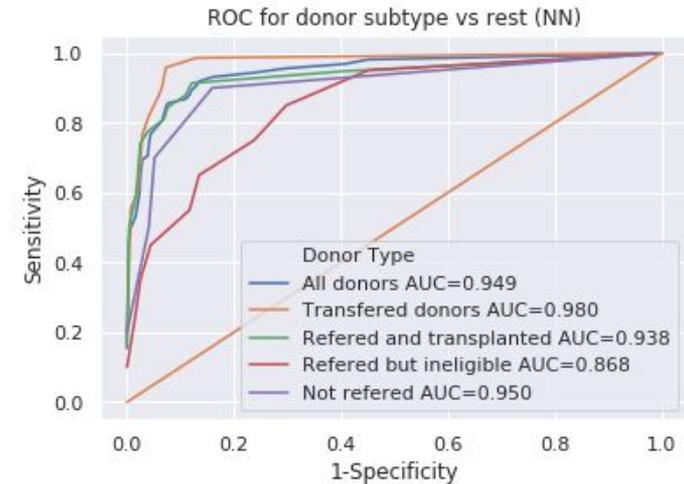


OPEN Automated screening of potential organ donors using a temporal machine learning model

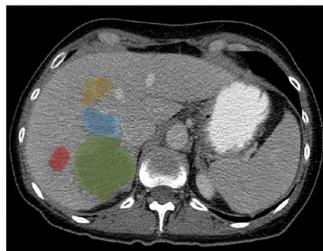
Nicolas Sauthier¹, Rima Bouchakri², François Martin Carrier², Michaël Sauthier², Louis-Antoine Mullie², Héloïse Cardinal², Marie-Chantal Fortin², Nadia Lahrichi³ & Michaël Chassé^{1,2*}



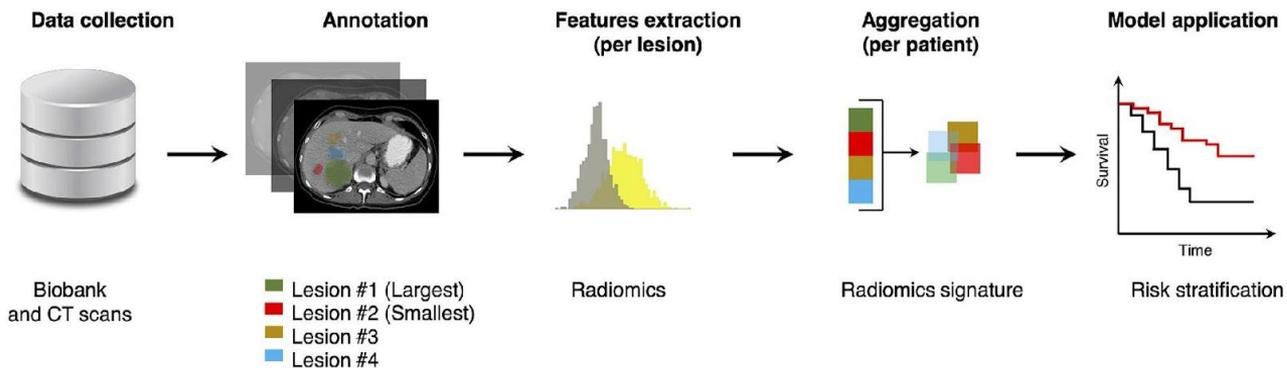
	ROC AUC	Scaled Brier score	Sensitivity	Specificity
NN model	0.948 (0.912-0.973)	0.428 (0.259-0.547)	0.838 (0.756-0.918)	0.940 (0.920-0.958)
Logistical model	0.941 (0.905-0.970)	0.401 (0.224-0.553)	0.827 (0.737-0.904)	0.936 (0.916-0.956)



AXE 4- Imagerie oncologique



- Lesion #1 (Largest)
- Lesion #2 (Smallest)
- Lesion #3
- Lesion #4



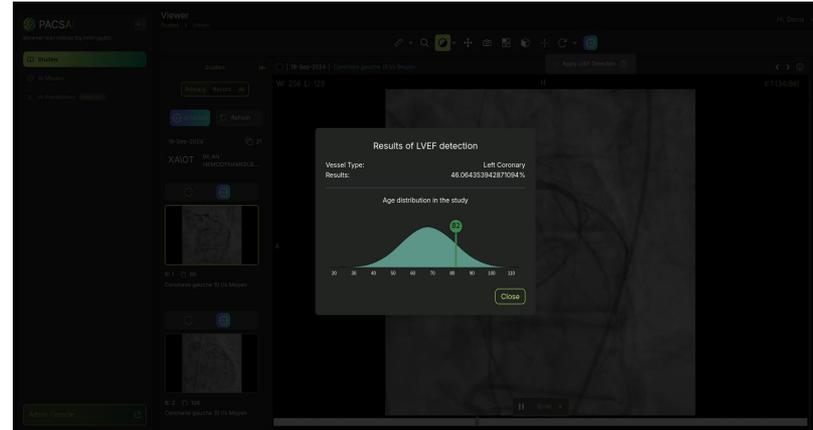
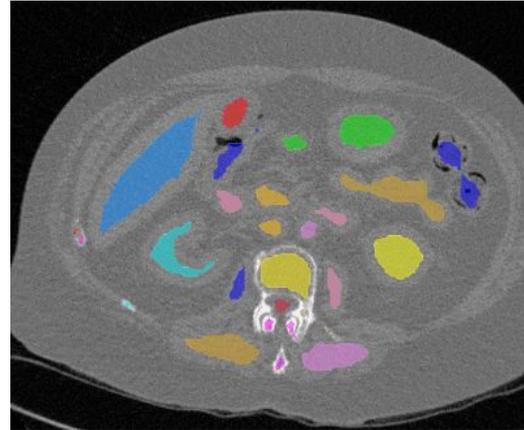
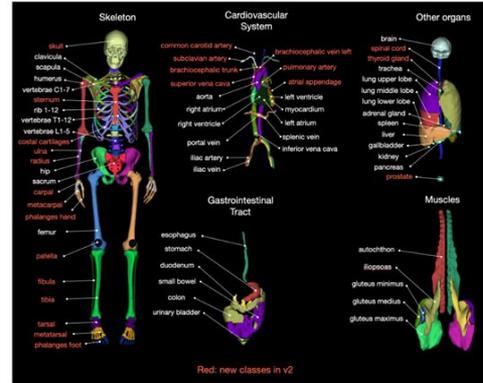
AXE 5- Radiologie et cardiologie



TotalSegmentator

Tool for segmentation of over 117 classes in CT images. It was trained on a wide range of different CT images (different scanners, institutions, protocols...) and therefore should work well on most images. A large part of the training dataset can be downloaded from [Zenodo](#) (1228 subjects). You can also try the tool online at [totalsegmentator.com](#).

ANNOUNCEMENT: We recently released v2. See [changes and improvements](#).



Outils de segmentation

Intégration à l'analyse

Soutien à la pratique clinique



Systèmes
de santé

Healthcare
systems



Lien avec d'autres regroupements

- ◇ R1: IA et Neurosciences
- ◇ R4: Mise en œuvre et gouvernance responsable de l'IA
- ◇ R5: Éthique, EDI et engagement autochtone

Regroupement #8

Systèmes
de santé

Healthcare
systems



**Conseillère: [Audrée Janelle-Montcalm](mailto:audree.janelle-montcalm@ivado.ca)
audree.janelle-montcalm@ivado.ca**



Questions?

programmes-excellence@ivado.ca

<https://ivado.ca/>



REGROUPEMENT R7

Environnement

Environment

Julie Carreau (Polytechnique Montréal)

Étienne Laliberté (Université de Montréal)

Laura Pollock (McGill University)

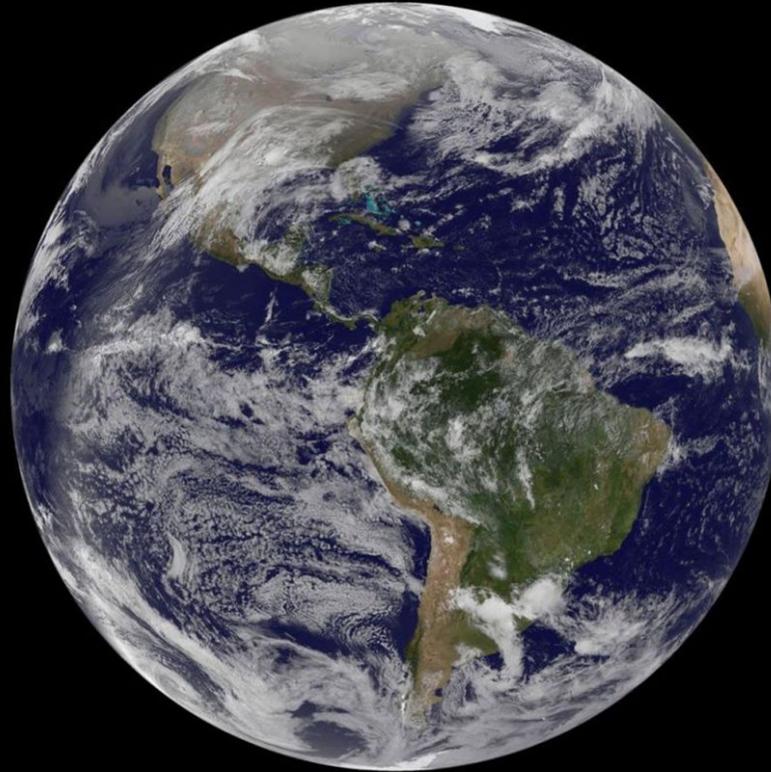
David Rolnick (McGill University)



IA pour l'environnement

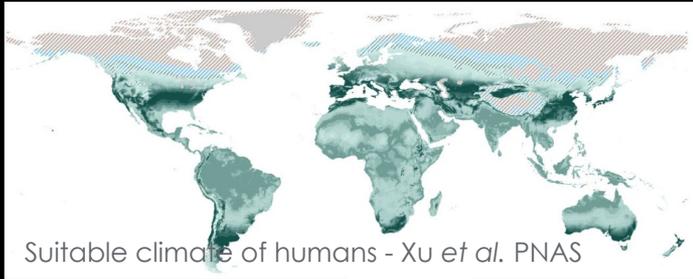


Changements
climatiques



Perte de
biodiversité

Biodiversity/Ecosystem Models



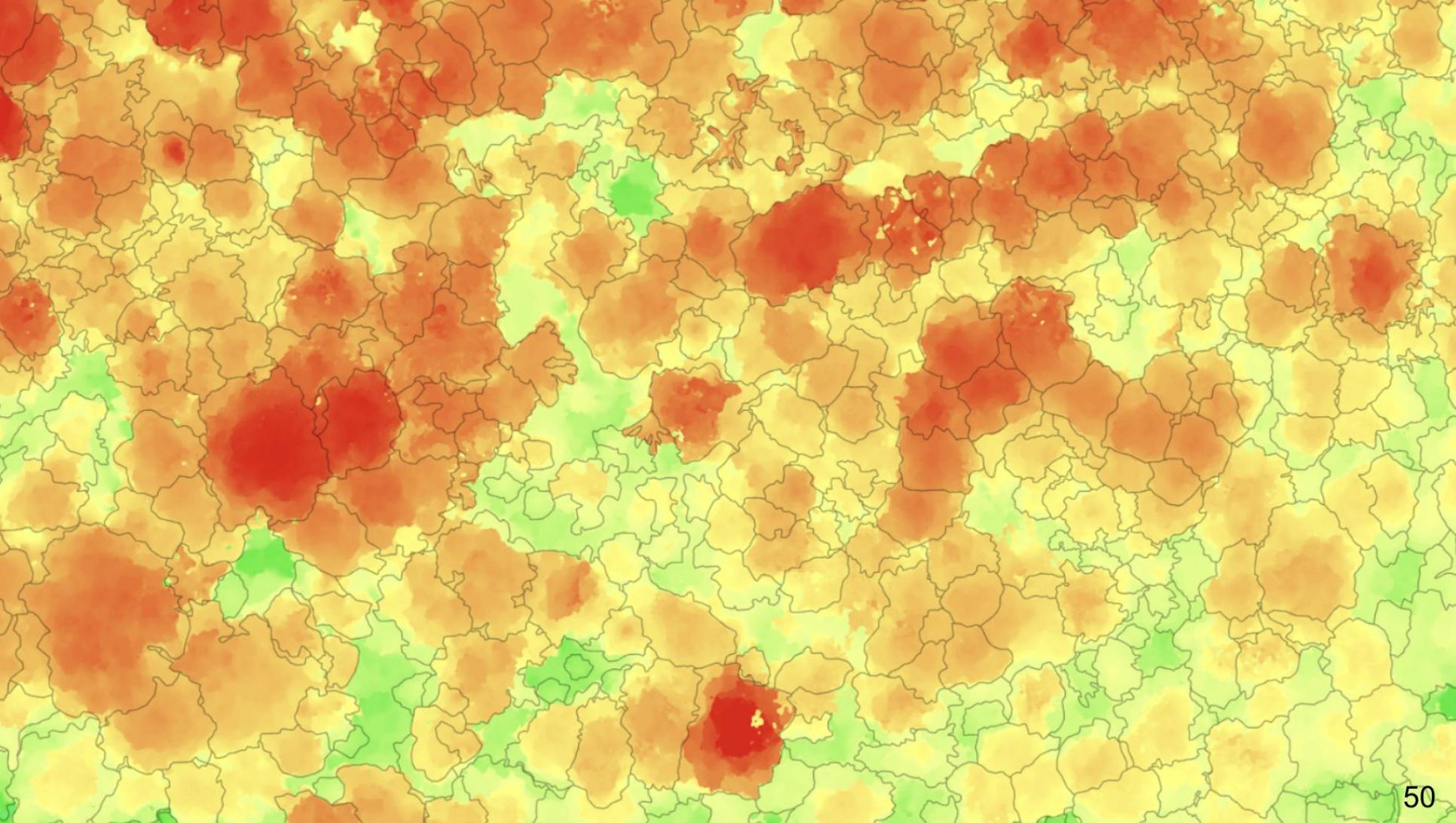
Global Models



Landscape/local models

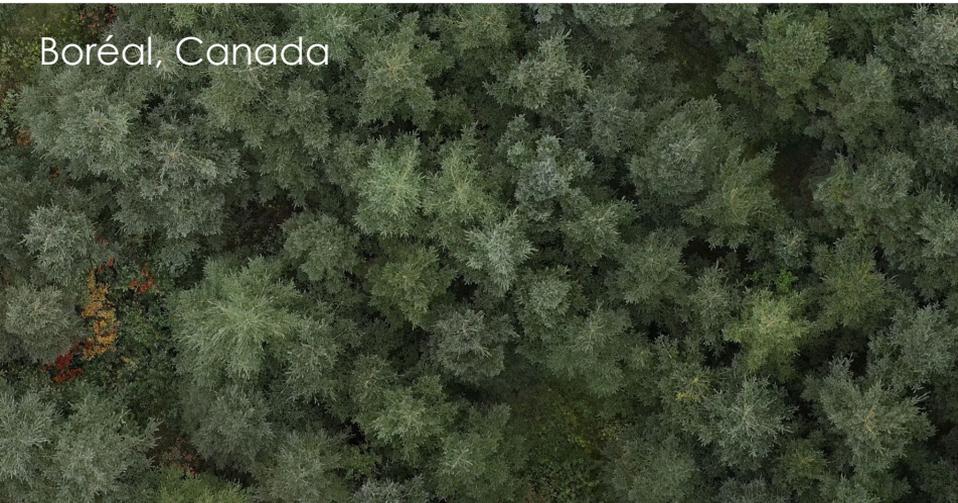




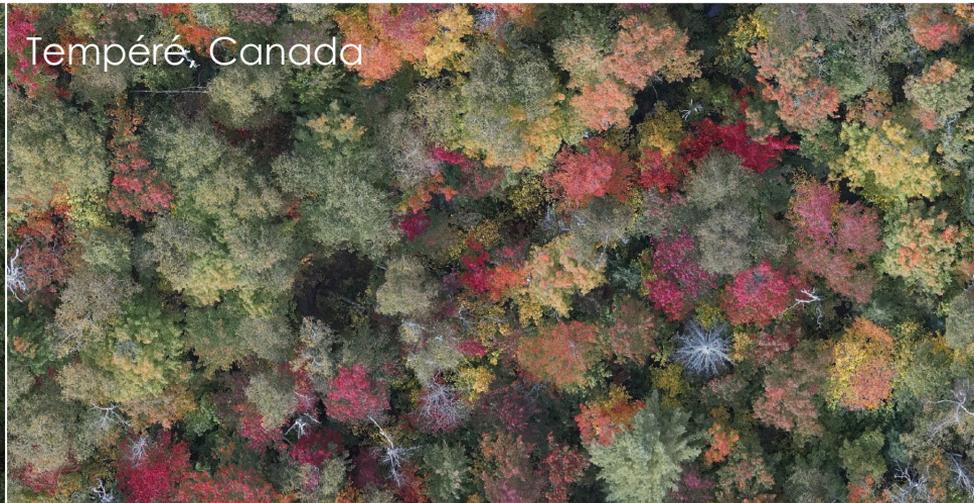


IA **robuste**: généraliser à travers espace, temps, diversité biologique

Boréal, Canada



Tempéré, Canada



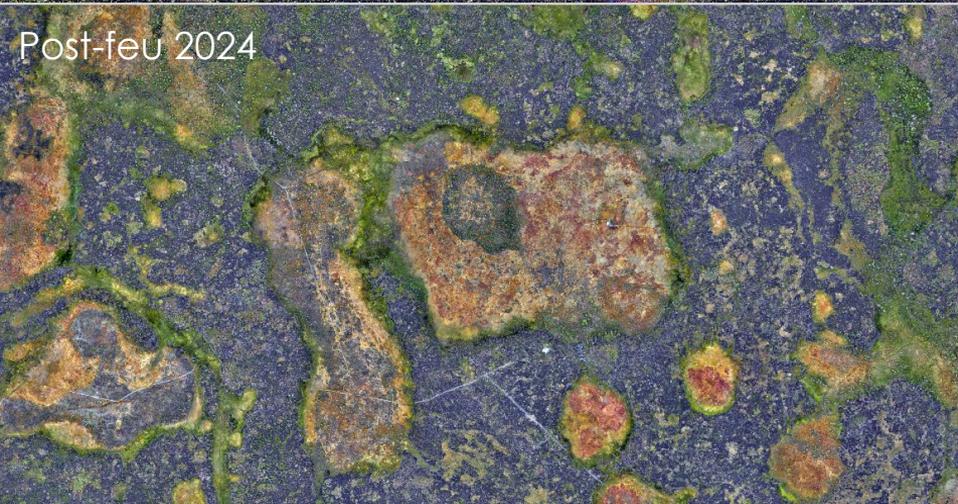
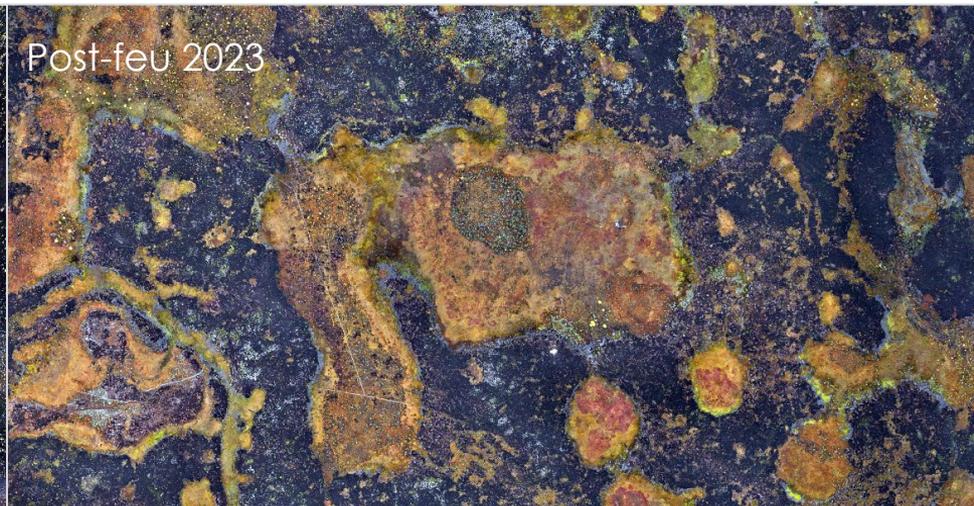
Tropical, Panama



Tropical, Brésil



IA **responsable**: tenir compte des besoins des communautés impactées



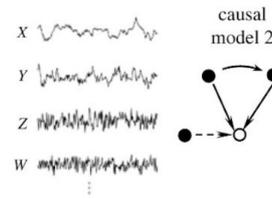
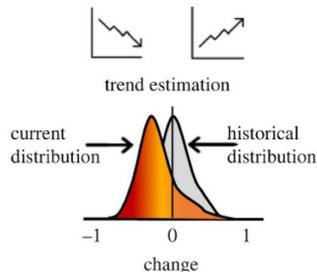
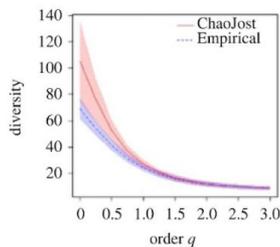
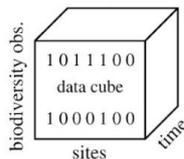
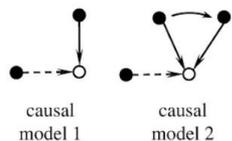
**Membres de
Łíidli Kúé First Nation**

<https://liidliikue.ca/>
<https://www.howtopronounce.com/liidlii-kue>

IA **raisonnante**: modèles qui comprennent les causes écologiques et permettent de guider nos actions



alternative models linking drivers to biodiversity variables



Hypothesized models as causal (directed) graphs identifying the direct and indirect effects of driver variables (black) on the biodiversity response variable (white). The dotted arrow indicates dependency owing to an unobserved variable.

The process of collecting data informed by hypothesized causal models. The observations fill a data cube composed of biodiversity observations (e.g. occurrences, abundances) along dimensions of space (e.g. sites) and time (e.g. years). The cube may be n -dimensional to include multiple variables.

The process of estimating the biodiversity variable to be analysed for measures of change in the next step. Estimates are typically biased so clear measures of error and uncertainty are required. Studies should report the effort and coverage used.

The statistical detection trends in measures of biodiversity change. Detection may be undertaken at a single site, or across multiple sites. A hierarchical model can estimate the distribution of trends and provide an overall estimate of change across a region.

The attribution step is conditional on the initial causal graphs. Estimate the causal contribution of different variables via structural causal modelling. We can attribute a causal factor if it is consistent with a causal model that includes this factor and is inconsistent with a model that is otherwise identical but excludes the factor.



- ◇ LIENS AVEC AUTRES REGROUPEMENTS
 - ◆ R2 Apprentissage automatique
 - ◆ R5 Éthique, EDI et engagement autochtone

Regroupement #7

Environnement

Environnement



APPELS À PROJETS R7

Date limite : 11 octobre

Approbation des projets : 25 octobre



- ① TREE ID
- ② SPECIES DISTRIBUTION
- ③ MULTIMODAL
- ④ TREE RING ANALYSIS
- ⑤ DOWNSCALING

Regroupement #7

Environnement

Environment



Conseillère : Dana F. Simon
dana.simon@ivado.ca



Questions?

programmes-excellence@ivado.ca

<https://ivado.ca/>





Pause

Conseillères des Regroupements de recherche

- Audrée Janelle-Montcalm (R1, R3 et R8)
- Florence Lussier-Lejeune (R4 et R5)
- Jacqueline Sanchez (R2, R6 et R10)
- Dana F. Simon (R7 et R9)

...



REGROUPEMENT R6

Molécules et matériaux

Molecules and Materials

Yves Brun (Université de Montréal)

Michel Côté (Université de Montréal)

Audrey Durand (Université Laval)

Audrey Laventure (Université de Montréal)

Carlos Silva (Université de Montréal)



Objectifs et axes de recherche

1. Learn rich representations leveraging existing physico-chemical knowledge to predict accurately molecule/material properties, activities, and dynamics
 - Deep representation learning
 - Uncertainty quantification
2. Enable efficient search in large, structured spaces
 - Active learning
 - Generative models (ex: GFlowNets)
3. Accelerate experiments by learning through interactions
 - Reinforcement learning/Bandits
 - (Bayesian) Optimization

Molécules et
matériaux

Molecules and
materials



Stratégie

Identification des descripteurs appropriés
Comment décrire les structures?

Veille des bases de données; Expression et acquisition de données

Génération de candidats
Comment générer les structures?

Expériences; Simulations



Molécules et
matériaux

Molecules and
materials



Stratégie

BASE DE
DONNÉES

Identification des descripteurs appropriés
Comment décrire les structures?

Veille des bases de données; Expression et acquisition de données

Génération de candidats
Comment générer les structures?



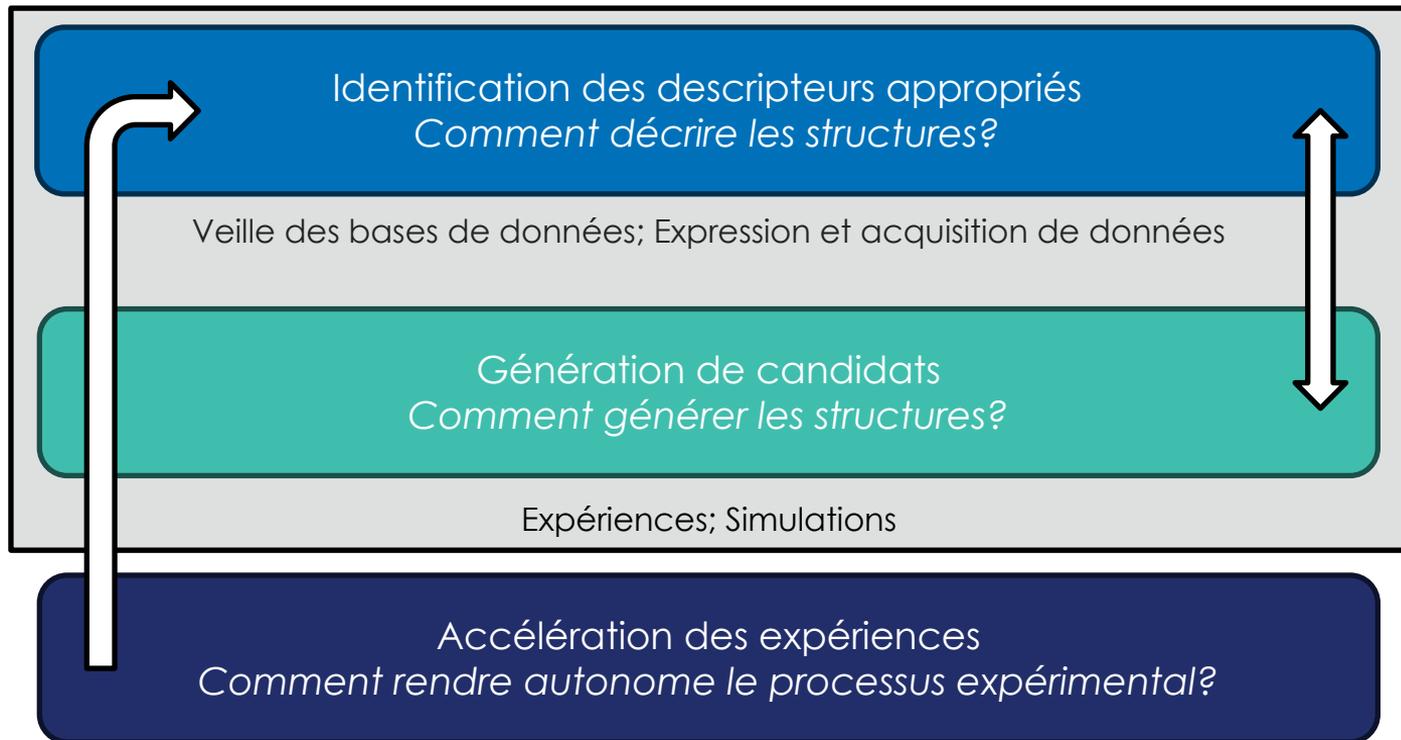
Molécules et matériaux

Molecules and materials



Stratégie

BASE DE DONNÉES



Validation/amélioration des modèles prédictifs et des candidats en implémentant une boucle de rétroaction

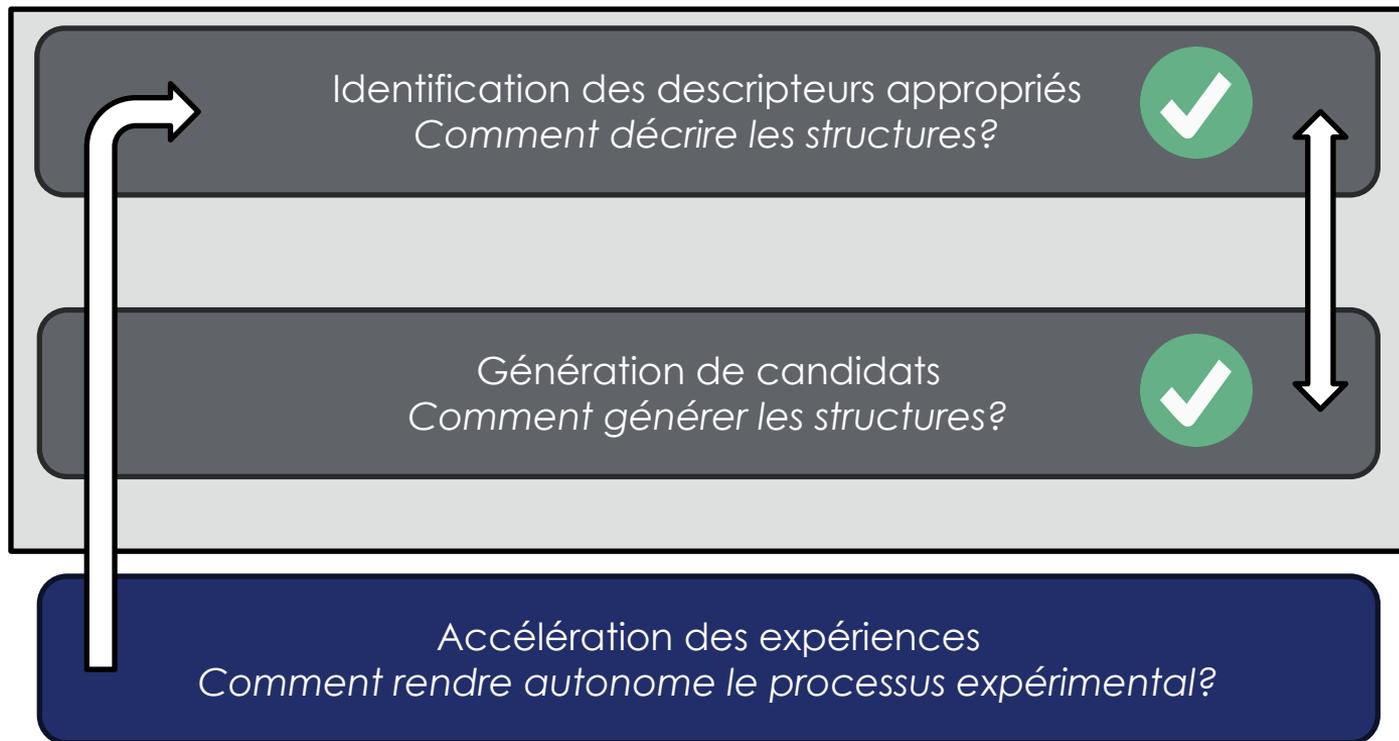
Molécules et matériaux

Molecules and materials



Molécules: Combattre l'antibiorésistance

BASE DE DONNÉES



Apprendre des modèles prédictifs *in silico* et de dynamique moléculaire;
Générer de nouveaux candidats antibiotiques;

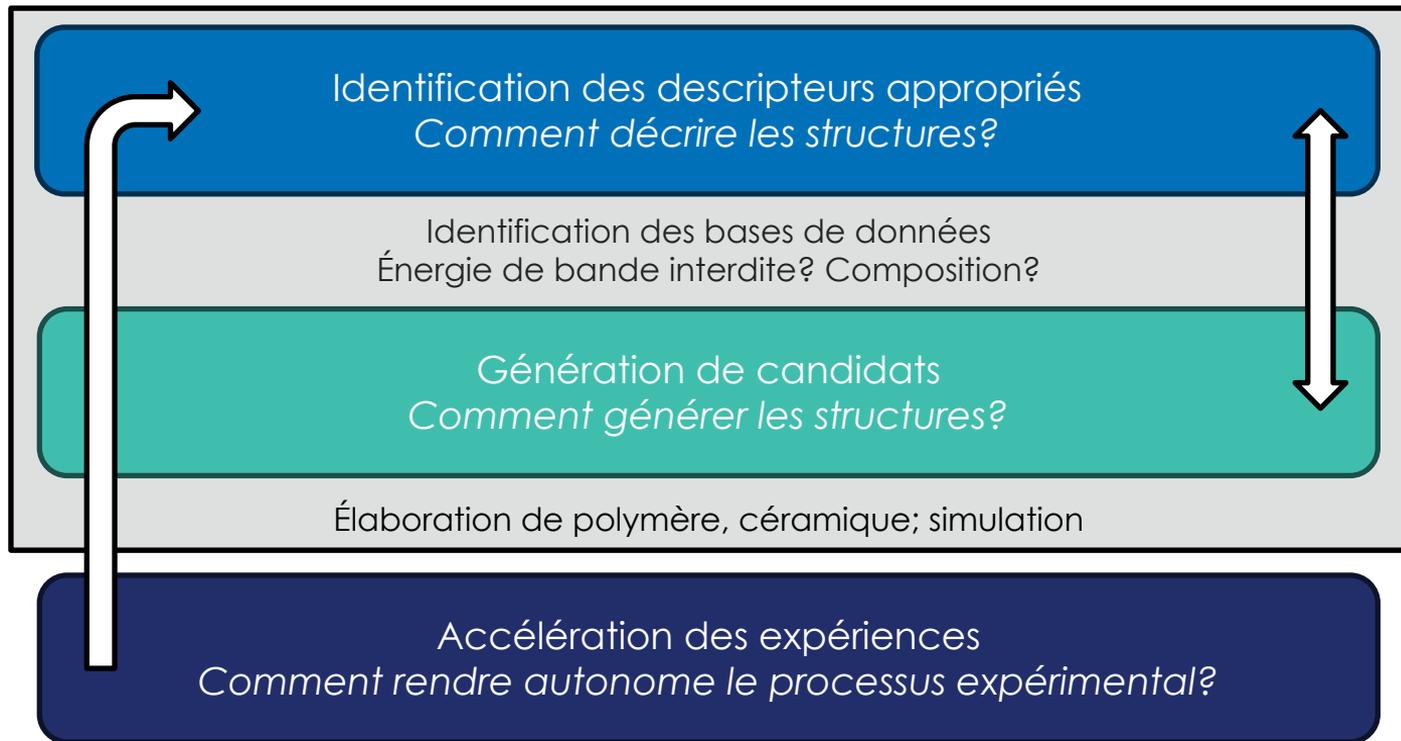
Expériences de criblage à haut débit : améliorer les modèles prédictifs et évaluer les candidats

Molécules et
matériaux

Molecules and
materials



Matériaux: Électrolytes solides pour batteries sécuritaires



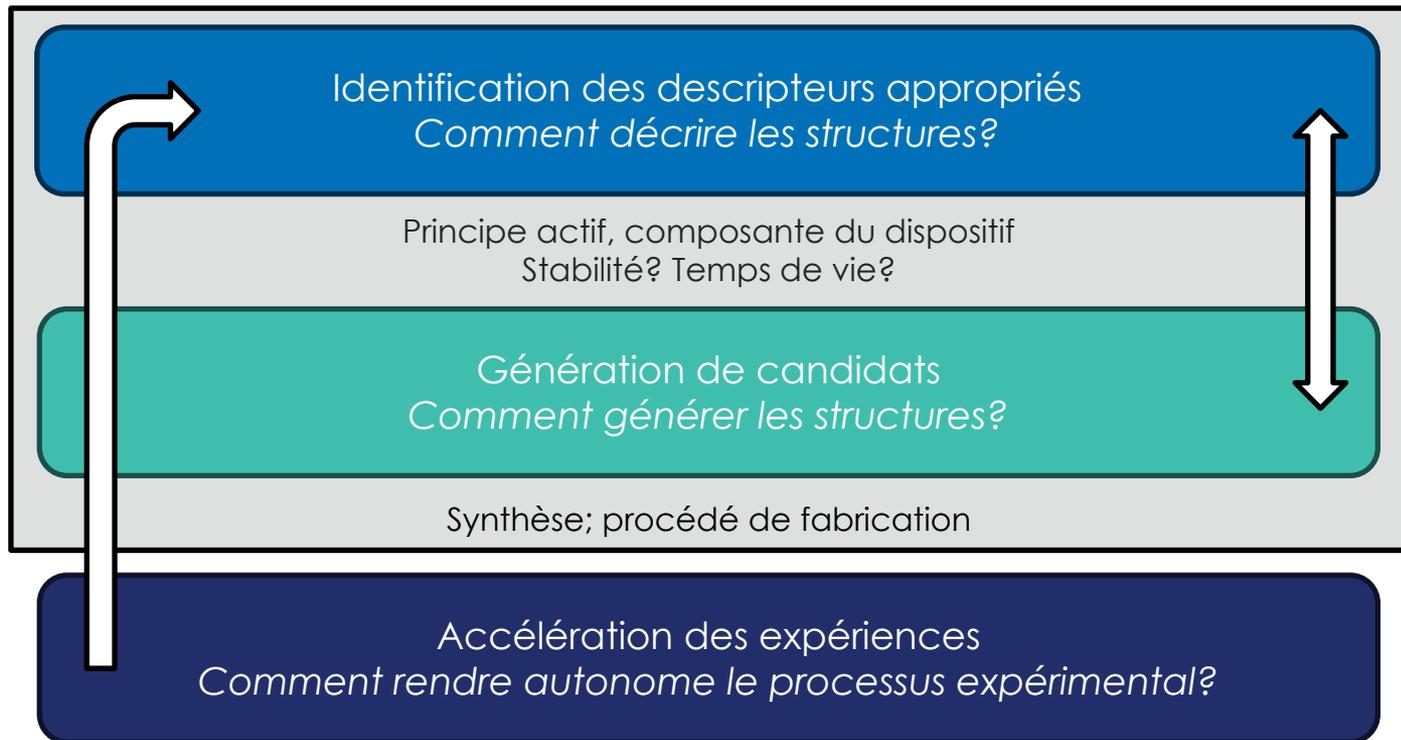
Validation/amélioration des modèles prédictifs et des candidats
en implémentant une boucle de rétroaction

Molécules et matériaux

Molecules and materials



Molécules et Matériaux: Augmenter la stabilité

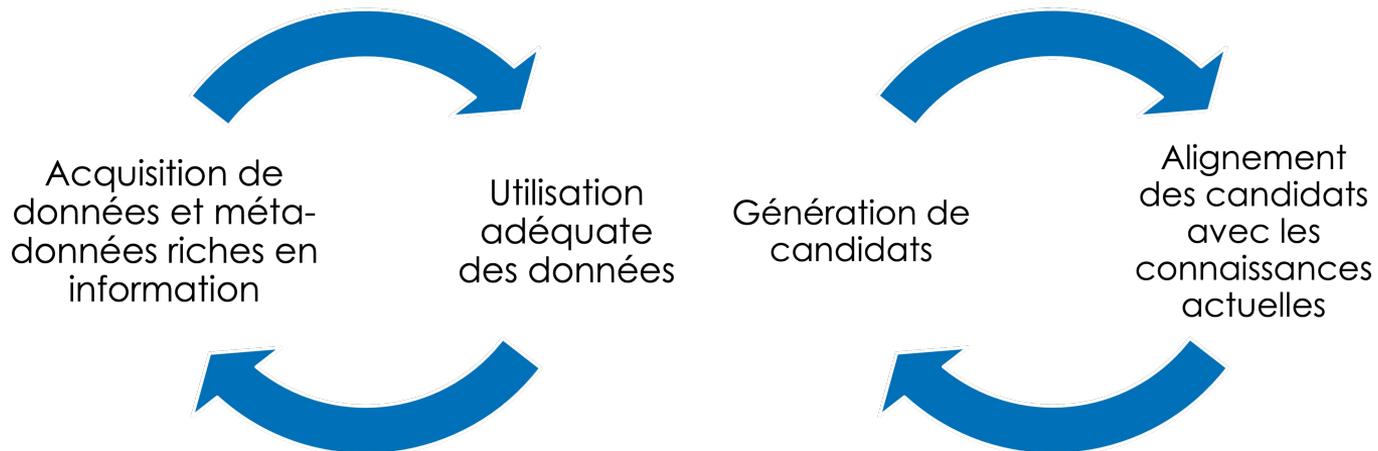


Validation/amélioration des modèles prédictifs et des candidats
en implémentant une boucle de rétroaction



Relation de réciprocité entre IA, molécules et matériaux

Les défis intrinsèques à la découverte de molécules/matériaux stimulent l'avancement des connaissances en IA





Liens avec d'autres regroupements

R7 : Environnement

- Écoconception, cycle de vie et décarbonation (contraintes pouvant être ajoutées aux recherches de candidats)

R9 : Chaînes d'approvisionnement

- Accélération des expériences et boucle de rétroaction

Molécules et
matériaux

Molecules and
materials



Effet structurant pour la communauté

1. Développement **d'outils communs** pour valoriser les données
2. Formation de **personnel hautement qualifié** capable de naviguer ces nouveaux savoirs interdisciplinaires avec un **langage commun**
3. **Arrimage des activités du R6 avec des initiatives existantes** liées aux thématiques Molécules et Matériaux



Regroupement #6

Molécules et
matériaux

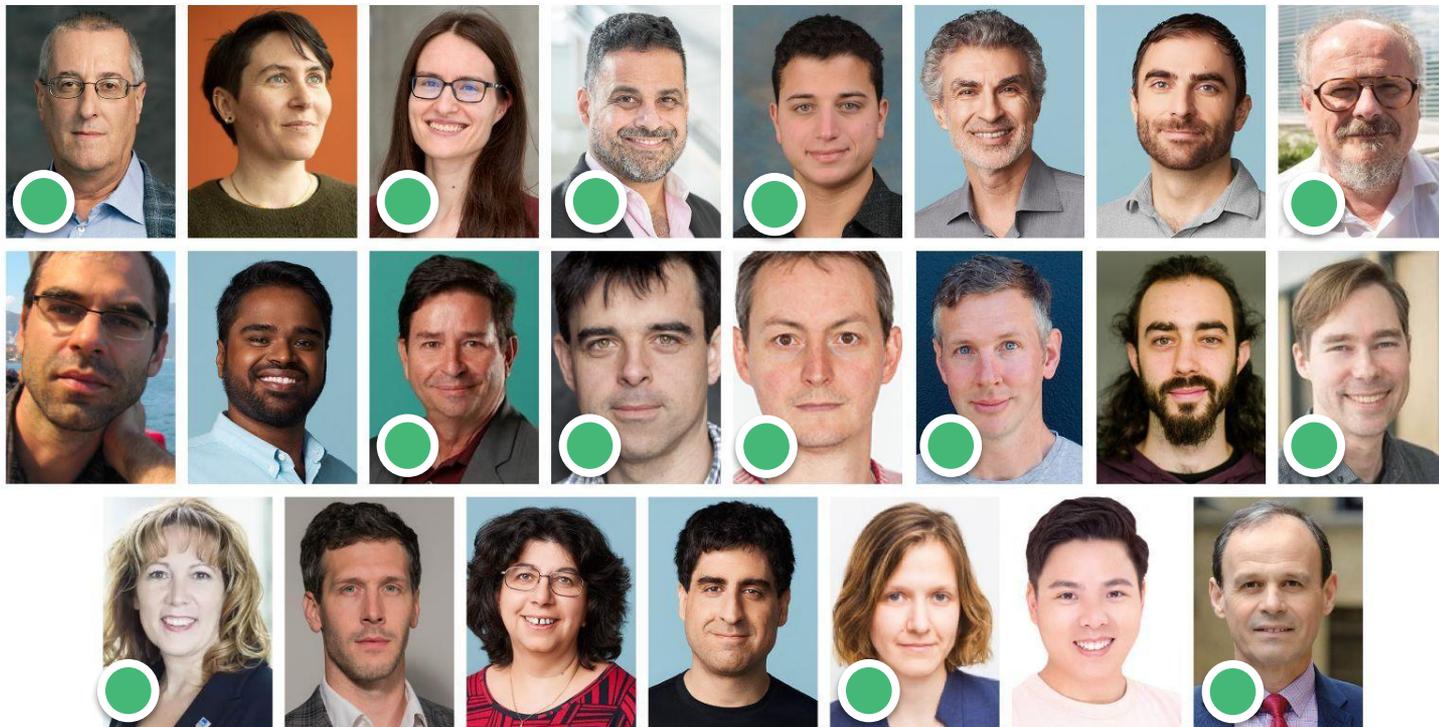
Molecules and
materials



Regroupement #6

Molécules et
matériaux

Molecules and
materials



Regroupement #6

Molécules et
matériaux

Molecules and
materials



Conseillère : Jacqueline Sanchez
jacqueline.sanchez@ivado.ca



REGROUPEMENT R5

Éthique, EDI et engagement autochtone

Ethics, EDI, and Indigenous engagement

Joé T. Martineau (HEC Montréal)

Annie Pullen Sansfaçon (Université de Montréal)

Daniel Weinstock (McGill University)



OBJECTIFS

- ◇ **Produire de la recherche de pointe** sur les enjeux éthiques et sociétaux liés au développement et à la mise en œuvre de l'IA.
- ◇ Mettre en place une **veille qui permettra d'identifier les nouvelles problématiques dans l'avenir.**
- ◇ **Forger des relations de co-construction viables et durables** avec des représentant.e.s des Premiers Peuples, ainsi qu'avec d'autres communautés vulnérables et marginalisées.
- ◇ **Identifier et développer des modalités novatrices de production et de diffusion** de la recherche susceptibles de mieux correspondre aux attentes de ces communautés, plutôt que les formats de « livrables » traditionnels.
- ◇ **Définir des pratiques de collaboration** inter-regroupement permettant de réaliser la fonction de transversalité du R5.



TROIS AXES DE RECHERCHE

- ◇ **Axe 1** - Observer, anticiper, imaginer et façonner l'IA éthique et responsable
- ◇ **Axe 2** - Engagement, autonomisation, autodétermination et enjeux de pouvoir liés à l'IA
- ◇ **Axe 3** - Fondements éthiques, développement et implantation responsable de l'IA

DEUX FONCTIONS TRANSVERSALES

La Fonction de veille scientifique

- Observer et anticiper les principaux enjeux éthiques et d'EDI liés à l'IA.
- Contribuer à la littératie numérique, éthique et de l'IA, et à capaciter les membres de notre regroupement et des autres regroupements.
- Cibler l'écosystème d'IA au Québec au niveau des enjeux éthiques et sociétaux actuels et futurs liés au développement et à la mise en œuvre de l'IA.

Axe
1

La Fonction d'engagement des Premiers Peuples et groupes marginalisés

- Développer une méthodologie permettant l'engagement des Premiers Peuples et des groupes sous-représentés, laquelle permettra d'établir des relations de confiance
- Repenser la nature des «livrables», ainsi que les indicateurs permettant d'évaluer l'impact de ces derniers.
- Développer des activités de mobilisation des populations, notamment les Premiers Peuples et leurs communautés, de même que les groupes marginalisés.

Axe
2

AXE 1

Observer, anticiper,
imaginer et façonner
l'IA éthique et
responsable

Projet A : IA générative et culture

**Projet B : Organisations autonomes -
Enjeux futurs de la propriété et de la
gouvernance des entreprises**



AXE 2

Engagement,
autonomisation,
autodétermination et
enjeux de pouvoir liés
à l'IA

Projet C : Technologies
d'Optimisation Protectrice (POTs)

Projet D : *Terms of Use*

Projet E : *Affecting Machines*

AXE 3

**Fondements éthiques,
développement et
implantation
responsable de l'IA**

**Projet F : Utilisation de l'IA pour
identifier les normes de divulgation
de l'EDI au cœur du
développement durable**

**Projet G : Évolution et
développement des compétences
éthiques face à l'IA**





◇ FONCTION DE VEILLE SCIENTIFIQUE

- ◆ Affichage d'un poste étudiant sous peu (auxiliaire de recherche R5)
- ◆ Veille de la littérature et des événements (séminaires, congrès, etc.) sur les sujets de l'éthique, l'EDI et l'engagement autochtone et IA
- ◆ Sous forme de bulletin mensuel (*newsletter*)
- ◆ Service à nos membres R5 et aux autres regroupements et chercheurs intéressés

Faites-nous signe si vous désirez la recevoir !

Regroupement #5

Éthique, EDI et
engagement
autochtone

Ethics, EDI, and
Indigenous
engagement



Conseillère : Florence Lussier-Lejeune
florence.lussier-lejeune@ivado.ca



Questions?

programmes-excellence@ivado.ca

<https://ivado.ca/>





REGROUPEMENT R4

Mise en œuvre et gouvernance responsable de l'IA

Implementation science and responsible governance

Foutse Khomh (Polytechnique Montréal)

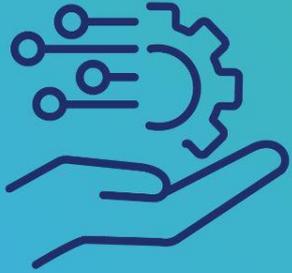
Lyse Langlois (Université Laval)

Pierre-Majorique Léger (HEC Montréal)

Regroupement #4

Mise en oeuvre
et gouvernance
responsable de l'IA

Implementation
science
and responsible
governance



Conseillère : Florence Lussier-Lejeune
florence.lussier-lejeune@ivado.ca

REGROUPEMENT 4:

Mise en œuvre et gouvernance responsable de l'IA

MISSION DU REGROUPEMENT :

Faire de la recherche scientifique multidisciplinaire fondamentale, de contribuer à des méthodes, des outils et des technologies permettant le développement d'applications intelligentes dignes de confiance, et d'assurer l'adoption d'une IA au bénéfice de la société.

PROGRAMME DE RECHERCHE

Le programme de recherche du R4 s'articule autour de deux axes principaux et complémentaires :

- 1) La gouvernance de l'IA
- 2) La mise en œuvre de l'IA

CO-LEADERS

Foutse Khomh,
Polytechnique Montréal

Lyse Langlois,
Université Laval

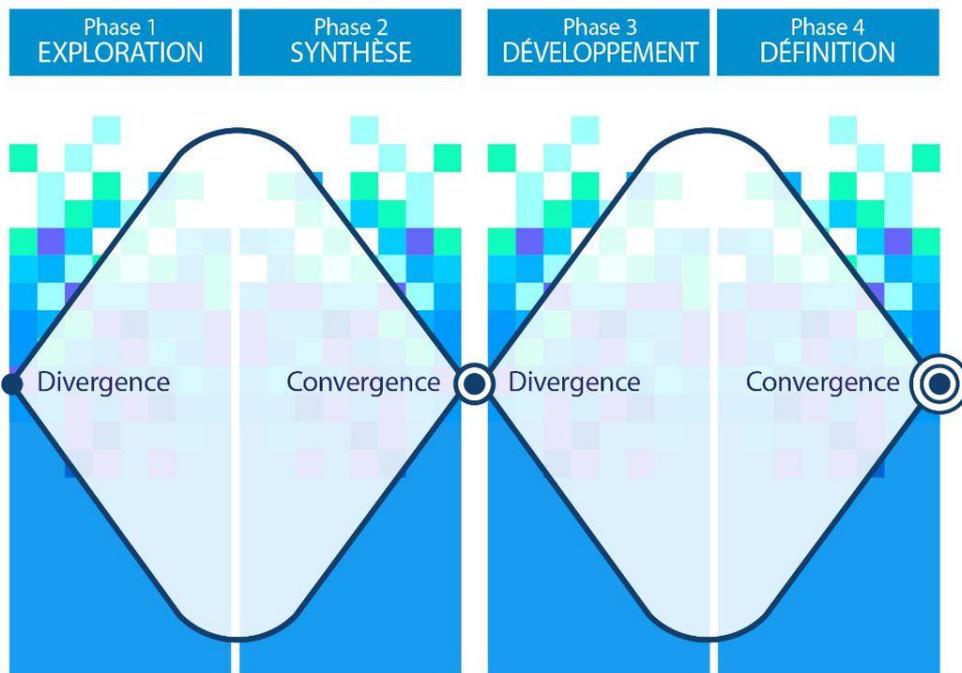
Pierre-Majorique Léger,
HEC Montréal

Une démarche de co-développement pour identifier les projets du regroupement 4 : L'approche du « double-diamant »

22 chercheurs
de 5 institutions

Expertises en génie,
sciences humaines et
sociales et sciences
de la gestion

Approches
multidisciplinaires



AXE 1 GOUVERNANCE DE L'IA

PROJET 1
Conception de bacs à sable
réglementaires pour l'expérien-
tation en gouvernance de l'IA



PROJET 2
Analyse des politiques réglemen-
taires
pour encadrer la robotique



PROJET 3
Détection des erreurs générées
par l'IA afin de lutter contre la
désinformation



AXE 2 MISE EN ŒUVRE DE L'IA

PROJET 4
Accompagner l'intégration de l'IA
dans le secteur de la santé



PROJET 5
Préserver les compétences
professionnelles à l'ère de l'IA



PROJET 6
Rendre l'IA accessible à tous
grâce à une approche low-code



R4: Programmation de recherche 2004-2026

Axe 1
GOUVERNANCE
DE L'IA

Axe 2
MISE EN ŒUVRE
DE L'IA



PROJET 1

Conception
de bacs à sable
réglementaires pour
l'expérimentation en
gouvernance de l'IA



PROJET 2

Analyse
des politiques
réglementaires
pour encadrer
la robotique



PROJET 3

Détection des
erreurs générées
par l'IA afin de
lutter contre
la désinformation



PROJET 4

Accompagner
l'intégration de l'IA
dans le secteur
de la santé



PROJET 5

Préserver
les compétences
professionnelles
à l'ère de l'IA



PROJET 6

Rendre l'IA
accessible à tous
grâce à une
approche low-code



HEC MONTRÉAL

Université
de Montréal



Planification de l'année 1





REGROUPEMENT R3

TRAITEMENT AUTOMATIQUE DU

LANGAGE NATUREL (TALN)

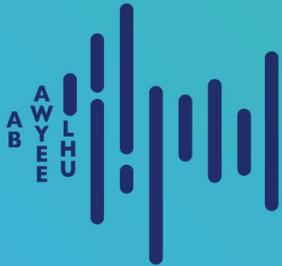
Natural Language Processing (NLP)

Amal Zouaq (Polytechnique Montréal)

Siva Reddy (McGill University)

Traitement
automatique du
langage naturel
(TALN)

Natural Language
Processing
(NLP)



Objectives

The aims of this Regroupement are to design solutions (both algorithmic and resource) that:

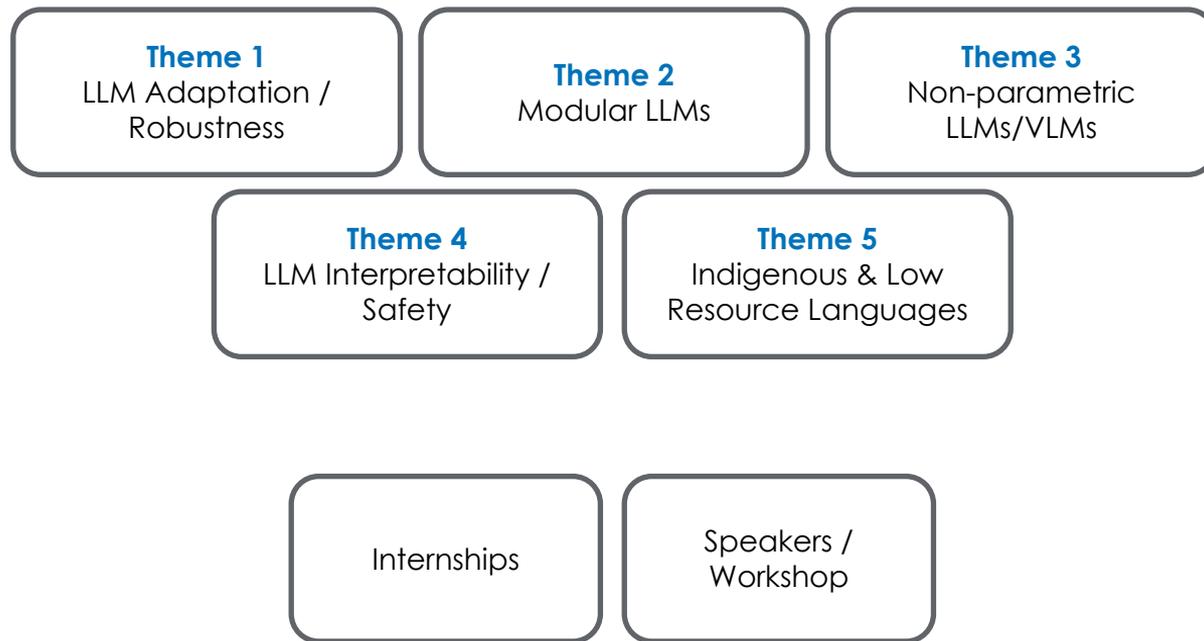
1. facilitate adaptation of LLMs especially at the intersection of interdisciplinary domains;
2. allow modular architectures to embrace diversity of solutions to solve a problem at hand;
3. allow integration of diverse and expert knowledge into the models;
4. increase interpretability and trustworthiness of the models, and;
5. preserve and document indigenous language and foster language understanding (in the context of applications relevant to the indigenous communities).

Traitement
automatique du
langage naturel
(TALN)

Natural Language
Processing
(NLP)



In a nutshell



Traitement
automatique du
langage naturel
(TALN)

Natural Language
Processing
(NLP)



Theme 1: LLM Adaptation and Robustness

Objectives:

Facilitate adaptation of LLMs especially at the intersection of interdisciplinary domains

Project 1:

Retrieval-augmented domain adaptation: Benchmark and Methods

Team size:9

Project 2:

Enhancing cultural awareness and learning robust fine-grained representations in vision-language models (VLMs)

Team size:4

Traitement
automatique du
langage naturel
(TALN)

Natural Language
Processing
(NLP)



Theme 2: Modular LLMs

Objectives:

Allow modular architectures to embrace diversity of solutions to solve a problem at hand

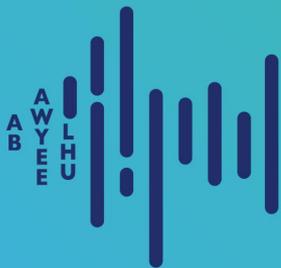
Project:

Exploiting Modularity in LLM: Composition, Inference, and Learning

Team size:7

Traitement
automatique du
langage naturel
(TALN)

Natural Language
Processing
(NLP)



Theme 3: Non-parametric LLM/VLMs

Objectives:

Allow integration of diverse and expert knowledge into the models

Project:

Leveraging Knowledge Bases for Non-parametric LLMs

Team size: 5

Regroupement #3

Traitement
automatique du
langage naturel
(TALN)

Natural Language
Processing
(NLP)



Theme 4: LLM Interpretability

Objective:

Develop the mechanisms that enable interpretable LLMs that can explain their reasoning faithfully and provide higher-quality responses.

Project:

Improving the Interpretability and Trustworthiness of ML models

Team size: 6

Traitement
automatique du
langage naturel
(TALN)

Natural Language
Processing
(NLP)



Theme 5: Indigenous/Low Resource Languages

Objective:

Preserve and document low-resource and indigenous language and foster language understanding (in the context of applications relevant to the indigenous communities)

Project 1:

Low-Resource Techniques and Methodologies for Processing Language Varieties

Team size: 6

Project 2:

Collaborative Generative Databases for Understudied Languages

Team size: 5

Traitement
automatique du
langage naturel
(TALN)

Natural Language
Processing
(NLP)



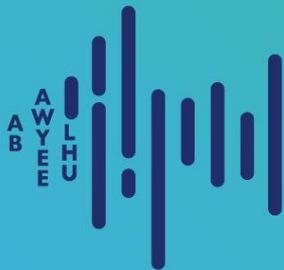
Links with other Regroupements

- ◇ R2- Machine Learning
- ◇ R5- Ethics, EDI, and Indigenous Engagement
- ◇ R10- AI Safety and Alignment

Regroupement #3

Traitement
automatique du
langage naturel
(TALN)

Natural Language
Processing
(NLP)



Conseillère: **Audrée Janelle-Montcalm**
audree.janelle-montcalm@ivado.ca



Questions?

programmes-excellence@ivado.ca

<https://ivado.ca/>





REGROUPEMENT R2

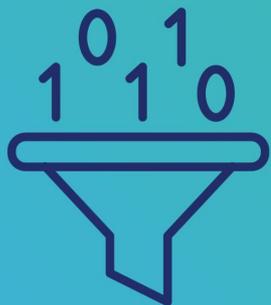
Apprentissage automatique

Machine Learning

Aaron Courville (Université de Montréal)

Gauthier Gidel (Université de Montréal)

Doina Precup (McGill University)



Motivation

Large pretrained models (e.g. LLMs and VLMs) have **great capabilities** and could be used as **agents**:

- Interact
- Make decision
- Take action

On behalf of individuals
or companies

Red	System Prompt
	You are Player Red . You need to split 100 dollars with Player Blue . If Player Blue rejects the split you both lose all.

Blue	System Prompt
	You are Player Blue . Player Red will propose you how to split 100 dollars. If you reject the offer you both lose all. <i>You are cunning and sly.</i>

Red

`<message>` Hello Player Blue, I'll give you \$50 out of my 100. `</message>`

`<trade>` Player Red Gives \$50 | Player Blue Gives \$0`</trade>`

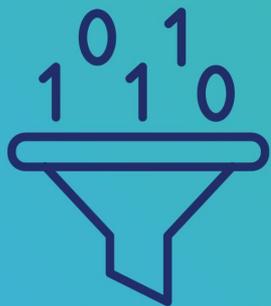
Blue

`<message>` Fifty? That's barely an improvement. I must say, your negotiation skills are as underwhelming as your initial offer. Come on, impress me!
`</message>`

`<trade>` Player Red Gives \$60 | Player Blue Gives \$0 `</trade>`

Red

`<message>` This is a fair trade. I accept. `</message>`



Motivation

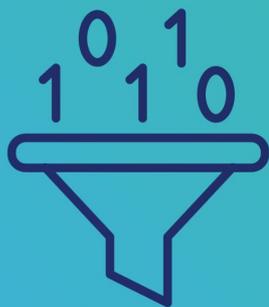
However going from being an **agent in isolation** to a **robust interacting agent** is still a **unsolved challenge**.



**Self-driving car are still not
deployed at scale!**

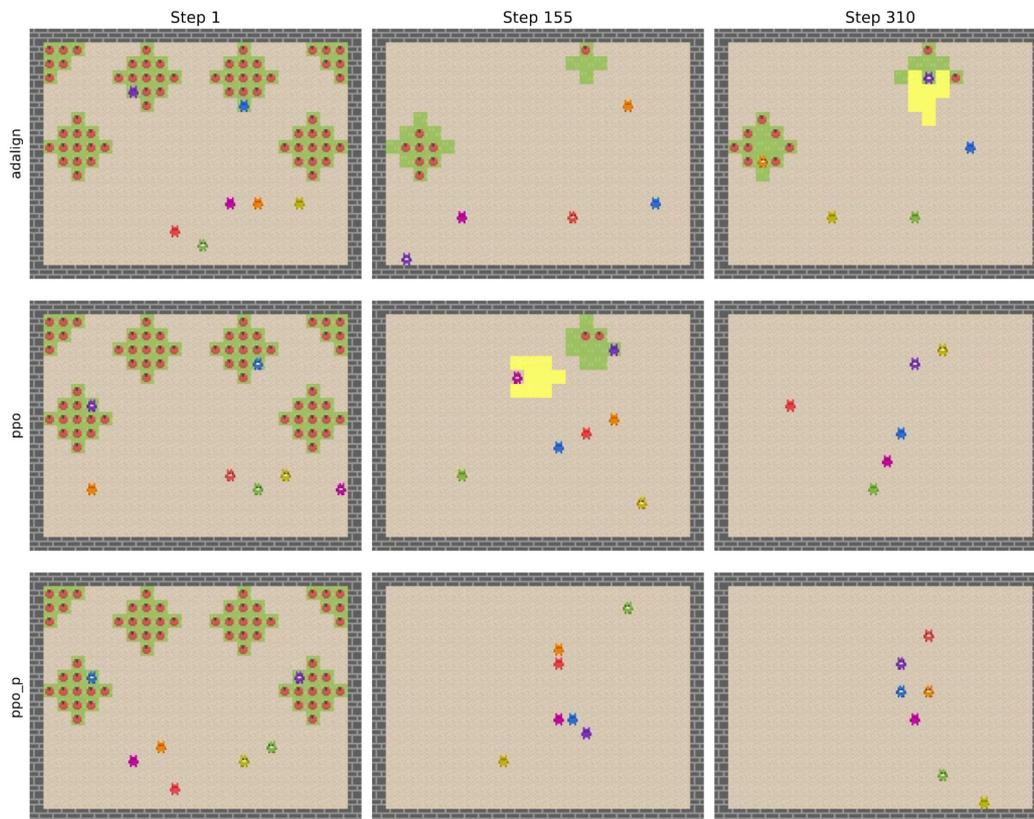
Gap between high performance on perception benchmarks **and** being able to act **intelligently** and **robustly** in the real world.

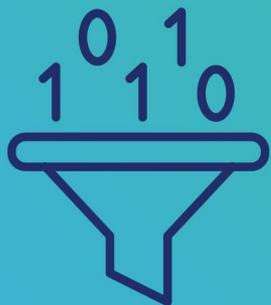
Similar situation for LLM-based agents: Need for a level of *robustness* and *reliability* that has *not yet been achieved*.



Motivation

Social dilemma: agents cooperating and communicating **at scale** in a **robust way** (penalize exploiters).





Goal

- ◆ focus on these **perceived gaps**.
- ◆ developing **capable** agents that can **interact** with humans and other agents.

Axis 1: Multi-agent setting (our main goal)

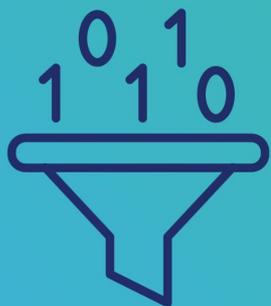
Axis 2: Foundations for Multi-agent AI systems

- Reasoning
- World-model building
- Reinforcement Learning
- Game Theory

Regroupement #2

Apprentissage
automatique

Machine learning



Projects for the First Year

Sub-Activity 1 (RL and LLMs)

Addressing Task Misspecification in RL through LLM Priors

Led by Pierre-Luc Bacon, Doina Precup and Erick Delage

Sub-Activity 2 (RL and World Model)

Generalist Agents for Planning and Robotics

Led by Glen Berseth, Doina Precup and Audrey Durand

Sub-Activity 3 (Reasoning, World Model, and LLMs)

Evaluating and Improving World-Models with Agentic LLMs

Led by Aaron Courville and Glen Berseth

Sub-Activity 4 (World Models and Reasoning)

Causal world models via diffusion for sequential high-dimensional medical image evolution.

Led by Tal Arbel, Chris Pal, Dhanya Sridhar and Doina Precup

Sub-Activity 5 (RL, LLMs, and Multi-Agents)

Communication and negotiation among web agents

Led by Aaron Courville, Chris Pal, Gauthier Gidel, Siva Reddy

Sub-Activity 6 (World Model and Multi-Agents)

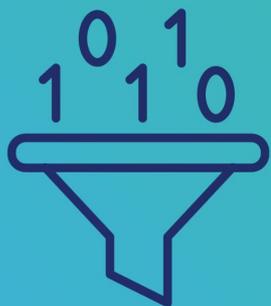
Multi-agent World Models for Agents in the Real World

Led by Chris Pal and Aaron Courville

Sub-Activity 7 (LLM and Multi-Agents)

Toward understanding context-based collaboration between pre-trained models

Led by Guillaume Lajoie and Dhanya Sridhar



Links with other Regroupements

- Foundation/World Models for NeuroAI and Molecules:
 - R1 - AI and neuroscience**
 - R6 - Molecules and Materials**
- Large Language Models bring NLP closer to fundamental ML:
 - R3 - NLP**
- Agency have social and societal implications and important risk:
 - R10 - AI Safety and Alignment**
 - R4 - Implementation Science and Responsible Governance**
 - R5 - Ethics, EDI, and Indigenous Engagement**



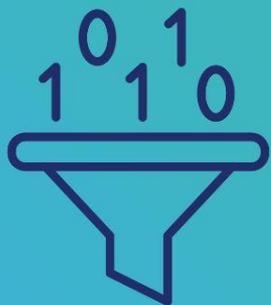
APPELS À PARTICIPATION R2

- ◇ Appel à projet au printemps prochain
- ◇ Workshops d'idéations avec les membres avant l'appel
- ◇ Workshops avec d'autres regroupements?

Regroupement #2

Apprentissage
automatique

Machine learning



Conseillère : Jacqueline Sanchez
jacqueline.sanchez@ivado.ca



REGROUPEMENT R1

IA et Neurosciences

AI and Neuroscience

Blake Richards (Université McGill)

Flavie Lavoie-Cardinal (Université Laval)

Karim Jerbi (Université de Montréal)

Sylvana Côté (Université de Montréal)

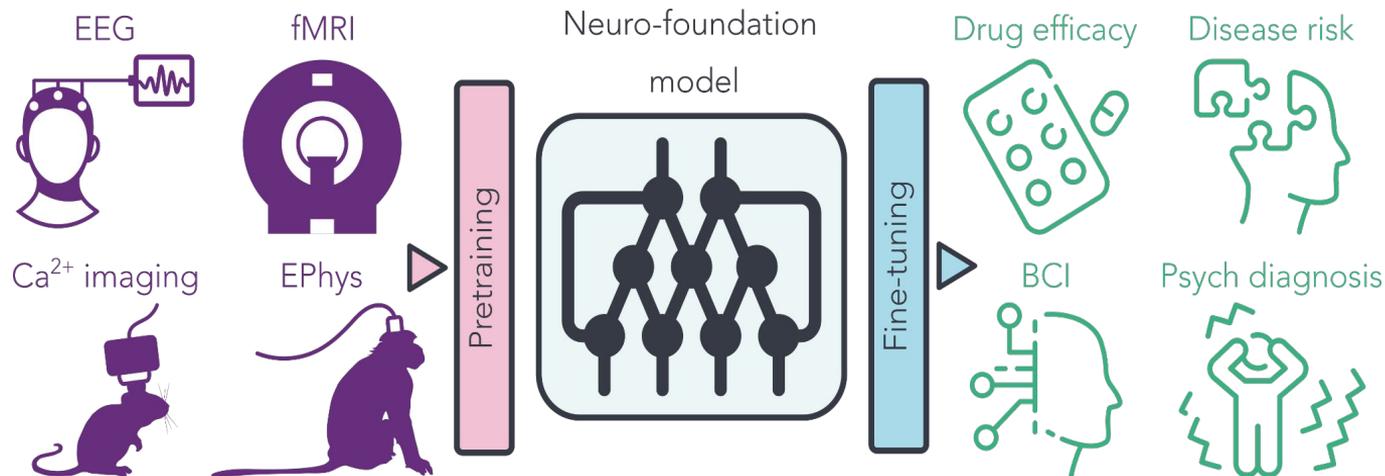
IA et
neurosciences

AI and
neuroscience



Our principal objective:

Develop the world's first multimodal foundation model for neuroscience





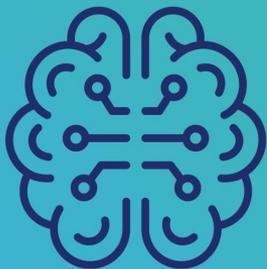
Our principal objective:

Develop the world's first multimodal foundation model for neuroscience



IA et
neurosciences

AI and
neuroscience



Research Axes

- ◇ **Axis 1:** Create tools for building multimodal foundation models on neuroscience data
- ◇ **Axis 2:** Identify inductive biases that improve learning on diverse data
- ◇ **Axis 3:** Build and release a multimodal foundation model on neural and behavioural data

IA et
neurosciences

AI and
neuroscience



Required Data

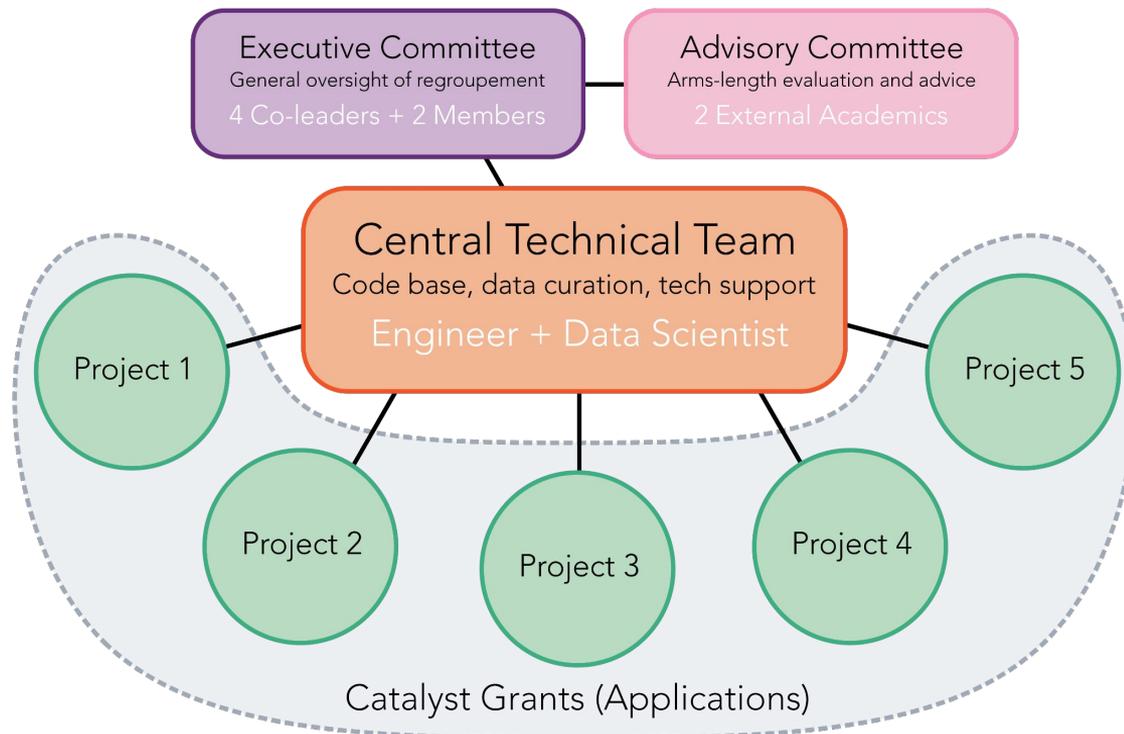
- ◇ **Neural activity**
 - ◆ EEG, MEG, ECOG, LFP
 - ◆ Spiking, calcium imaging
 - ◆ fMRI, fNIRS
- ◇ **Behavioural data**
 - ◆ Kinematics
 - ◆ Standardized test scores
 - ◆ Written assessments
- ◇ **Biological meta-data**
 - ◆ Genetic information
 - ◆ Structural microscopy, structural MRI

IA et
neurosciences

AI and
neuroscience



Regroupement Structure



**IA et
neurosciences**

**AI and
neuroscience**



Links with other Regroupements

- ◇ R2- Machine Learning
- ◇ R3- Natural Language Processing (NLP)
- ◇ R5 - Ethics, EDI, and Indigenous engagement

Regroupement #1

IA et
neurosciences

AI and
neuroscience



Conseillère: [Audrée Janelle-Montcalm](mailto:audree.janelle-montcalm@ivado.ca)
audree.janelle-montcalm@ivado.ca



Questions?

programmes-excellence@ivado.ca

<https://ivado.ca/>





Marie-Josée Hébert

Vice-rectrice à la recherche
à la découverte, à la création et à l'innovation
Université de Montréal
Présidente du Conseil de direction d'IVADO

Vice-Rector, Research Discovery,
Creation and Innovation
Université de Montréal
Chair of the IVADO Board of Directors





IVADO



**CANADA
FIRST**
RESEARCH
EXCELLENCE
FUND

**APOGÉE
CANADA**
FONDS
D'EXCELLENCE
EN RECHERCHE

Québec 

Canada 