

# Enquête sur l'Intelligence Artificielle dans le Département MathNum : Référentiel, Positionnement et Dynamique

CS MathNum, 11/09/2025

F. Garcia & V. Guigue

# IA encore et toujours



- Etude prospective sur l'IA et son impact dans le processus de recherche et d'innovation (cabinet Arthur D. Little, 2024)
- Note usage assistant IAG, DispSO (06/2024)
- Rapport Aghion-Bouverot sur l'IA (03/2024)
- IA générative à l'INRAE, P. Mauguin, STSforum (10/2023)
- Démarrage Mission IA MathNum (04/2023), veille science des données - focus IA
- DigitBio, animation l'IA en sciences du vivant (2022-23)
- Animation sur l'IA, AG MathNum (05/2022)
- Intelligence artificielle et santé animale, INRAE Productions Animales (2020)
- Dossier de direction EPMRA « Intelligence Artificielle pour la recherche agronomique » (Klein et al., 2019)
- Rencontre Mission C. Villani (Garcia, 09/2017), Rapport Villani (2018)
- Rapport FrancelIA « INRA et Intelligence Artificielle - Note d'opportunité » (06/2017)

# Calendrier



- Mission confiée mi-2023 à F. Garcia (MIAT) et V. Guigue (MIA Paris-Saclay)
- Enquête via entretiens (2 à 3 heures) avec toutes les unités MathNum entre mars et septembre 2024
- Première restitution en octobre lors de l'AG MathNum 2024
- Finalisation du rapport début 2025

# Objectifs de l'enquête



- **Proposer un référentiel** décrivant les nombreuses facettes de l'IA et permettant de positionner les recherches menées et les compétences présentes au sein des différentes unités du département.
- **Décrire les directions actuellement prises** par les unités et leurs équipes, entre leur expertise historique et la prise en compte des techniques émergentes, en particulier du deep learning et de l'IA générative.
- **Questionner le ressenti** des unités par rapport à cette vague actuelle de l'IA, et des transformations qu'elle entraîne à la fois dans les questions et dans les pratiques de recherche.

# Entretiens avec les unités



Environ 2h, 1 à 5 personnes

Semi-directifs

Grille d'entretien élaborée en amont

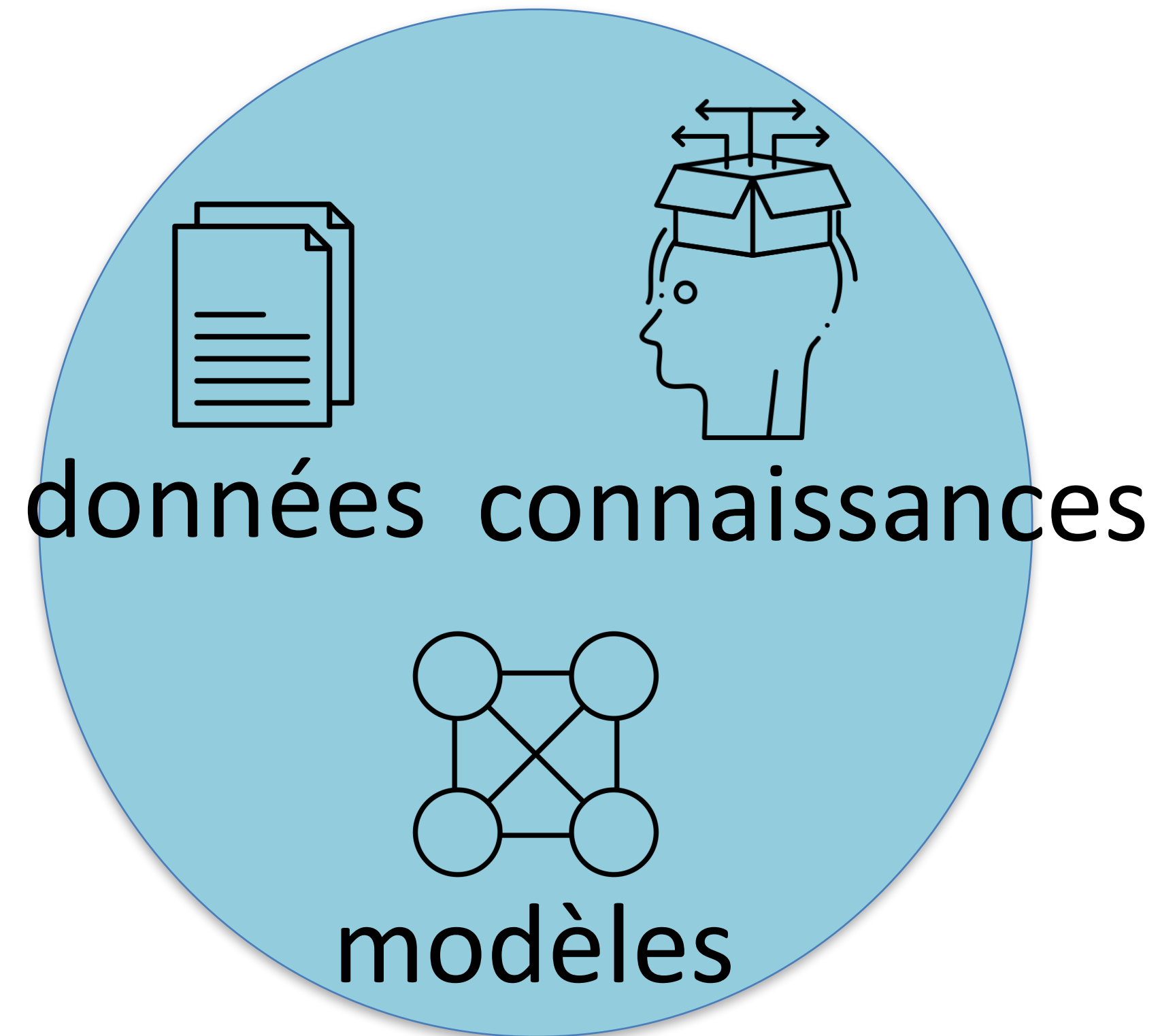
Unité	entretien
OPAALE	16/07
MalaGE	12/06
MIA-PS	19/06
LISC	14/05
TSCF	17/05
BIOSP	29/02
MIAT	29/05
ITAP	22/04
MISTEA	23/05
TETIS	03/09



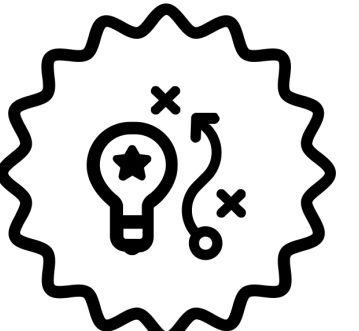
# Brève histoire de l'IA

# Intelligence Artificielle, pour une simulation de l'intelligence humaine

  
capteurs,  
robots



  
humains

  
ordinateurs

## Machine Learning

Neural Networks, Statistical learning

Reinforcement Learning

Generative AI

*From 60s and  
then from 90s*

## Knowledge representation and management

Ontologies

Semantic Web

Knowledge modeling

*From 70s*

## Artificial Intelligence

Modeling

Observation

Action

Diagnosis

Recommendation

## AI decision and planning

*From 60s*

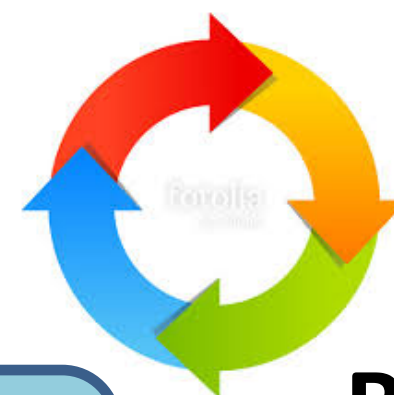
## AI modeling, sim & optimization

CSP

Multi-agent simulation

Evolutionary algorithms

*From 80s*

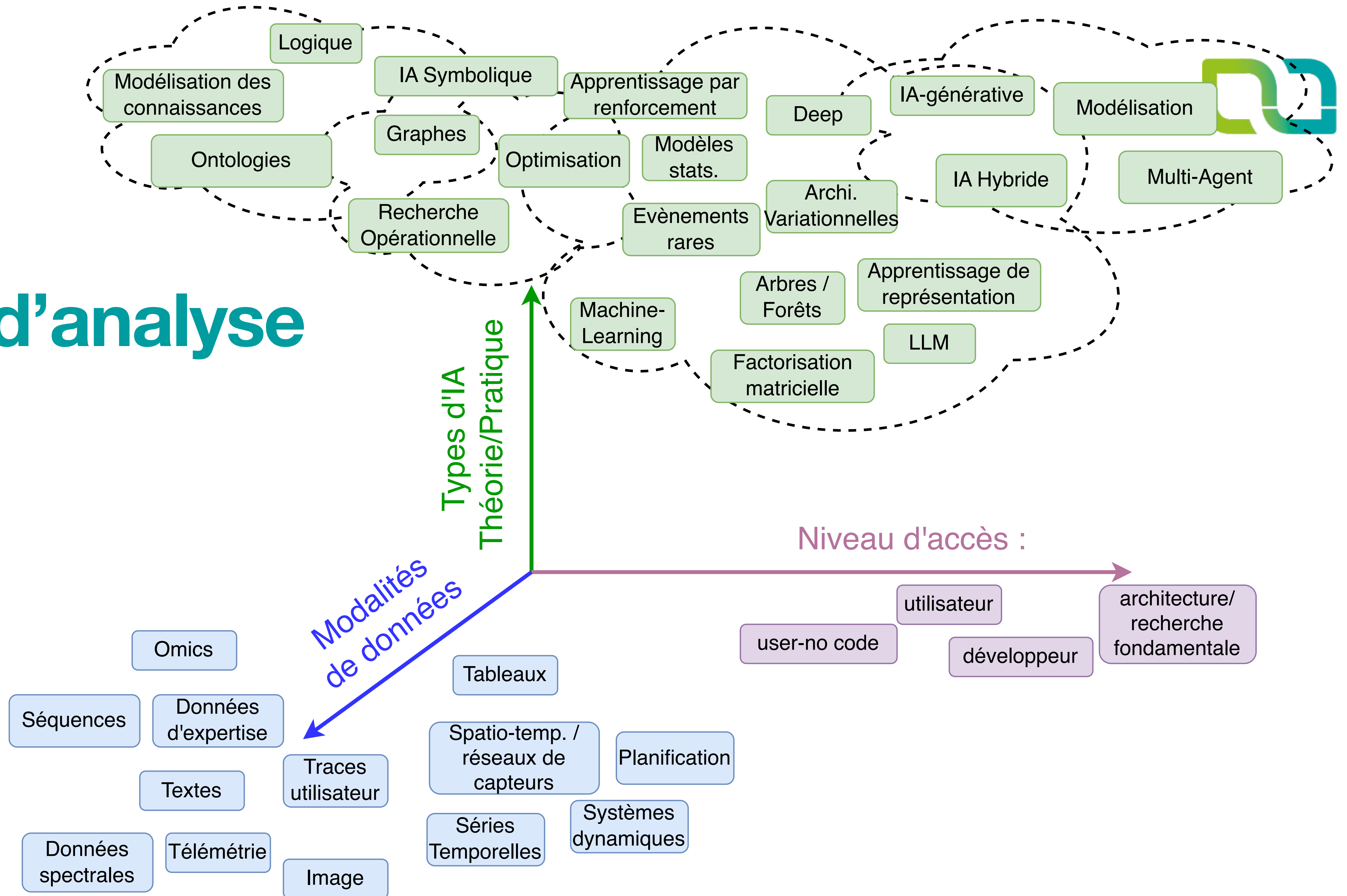






# Éléments d'un référentiel

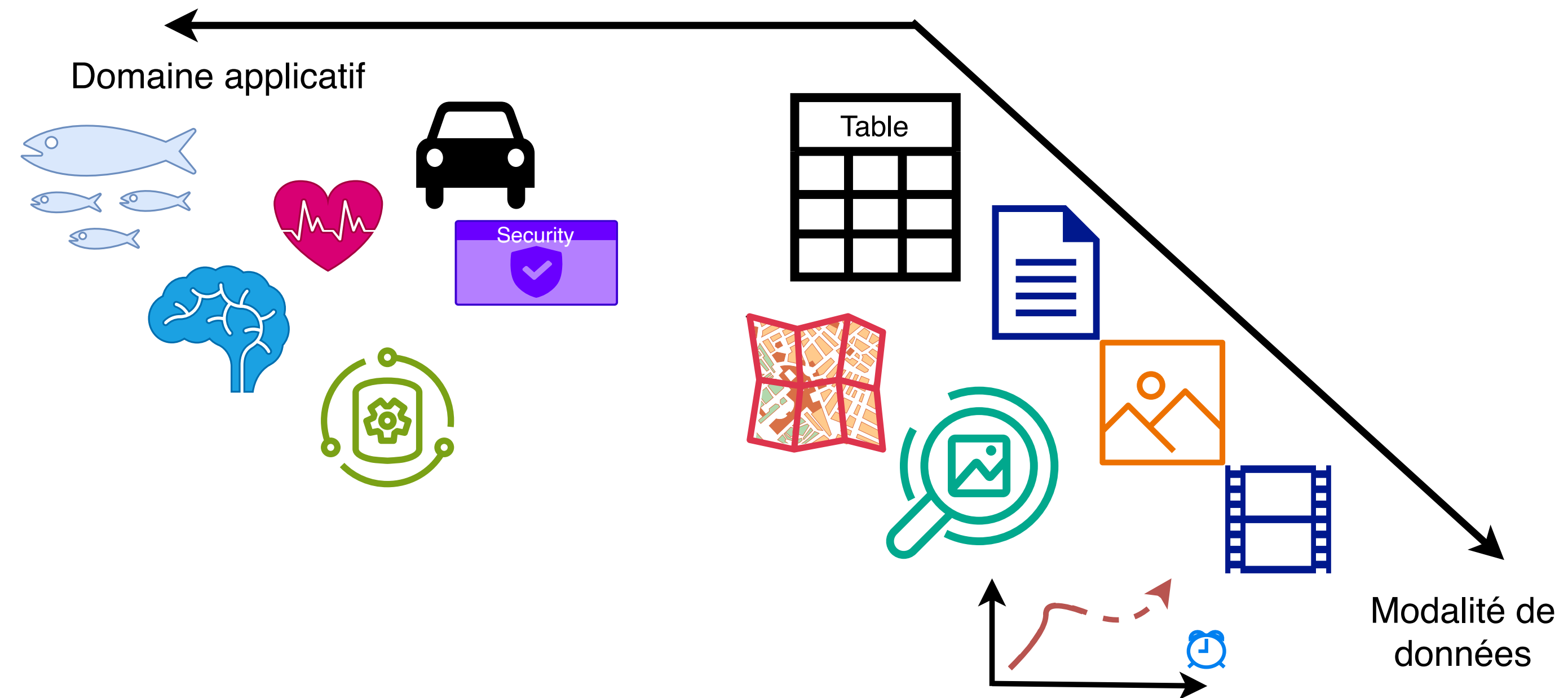
# Grille d'analyse



# AXE 1: Modalités de données



- (Tab) Données tabulaires
- (ST) Séries temporelles, systèmes dynamiques, capteurs
- (SpT) Données spatiales et séries spatio-temporelles
- (PM) Paroles et musique
- (Vis) Image/vision
- (Vid) Vidéo
- (Tel) Télémétrie
- (Spm) Spectrométrie
- (TAL) Texte
- (Exp) Données d'expertise
- (OmQ) Données omiques
- (Rec) Données d'interaction & recommandation
- (Enq) Données d'enquête
- (Par) Données de paramétrage IA, Processus séquentiels ou combinatoires



# AXE 2 : Approches en IA



- (IAS) Logique et IA symbolique
- (KB) Ingénierie des connaissances et ontologies
- (RI) Indexation et Recherche d'Information
- (CTR) Automatique et contrôle (MDP, ...)
- (Seq) Approches Séquentielles: Chaines de Markov, CRF, RNN
- (GR) Graphes: détection communauté, voyageur de commerce...
- (ML1) Modélisation statistique
- (ML2) Modèles linéaires, noyaux, régularisation
- (ML3) Arbres, Forêts, Boosting, ...
- (ML4) Non supervisé, Clustering
- (DL1) MLP, embeddings
- (DL2) Archi: CNN, RNN, GNN, Loss: contrastive learning, ...
- (DL3) Modèles pré-entraînés, YOLO, ResNet, Modèles de langue
- (DL4) IA Générative: GAN, VAE, mod. de diffusion; Augm. de données
- (DL5) Neural ODE, PINNs
- (RL1) MDP, Q-learning, modèles à états discrets, Bandits
- (RL2) Deep-learning pour le renforcement (PPO, DQN, AC)
- (RL3) Modèles pré-entraînés: AlphaFold
- (OPT) Optimisation continue (Gradient, BFGS, ...)
- (RO) Planification, Optimisation Discrète (SAT, Algo. Gén., Prog. Dyn.)
- (SIM1) Approches mécanistes (EDP), modélisation hybride
- (SIM2) Systèmes multi-agents, modélisation informatique

# Axe 3 : Niveaux d'accès à l'IA



- Théoricien amont (optimisation, travail sur les infrastructures de calcul, ...)
- Théoricien de l'IA (théorèmes proba/stats, démonstration formelle sur graphes, ...)
- Exploitation des outils + preuves & garanties formelles / développement d'architectures
- Exploitation pratique des outils disponibles sur l'étagère
- Simple utilisateur de chatGPT : est-on un acteur de la vague IA?

Maturité des outils = Nouveaux niveaux d'accès

Continuum des profils / pas de clivage franc

# Axe 4 : Dynamique



Tenter de détecter les sujets émergents ou en disparition dans les équipes

- Thématique en régression
- En progression







# **Dynamiques et ressentis des unités**



# Perception de la vague IA



**Interrogation sur ce qu'est aujourd'hui l'IA**, avec la conscience que c'est un peu devenu un terme fourre-tout.

L'IA n'est pas forcément un terme utilisé dans les labos, on va parler de **Machine-Learning** ou de **Deep-Learning**.

Les travaux en **IA « classique »** sont encore présents dans certains labos, mais on en parle moins.

Le DL et l'**IA générative** ont émergé de manière impressionnante en très peu de temps

Certains perçoivent déjà un infléchissement des NN « classiques », et un boom de tout ce qui est relatif à l'**hybridation**, à la **multi-modalité**

# Les domaines de MathNum largement impactés



Le développement de l'IA sous forme ML/DL touche **tous les domaines** de MathNum,

En particulier sous forme **d'hybridation** des méthodes (statistique, physique, IA)

Les **statistiques** apportent des choses (explicabilité, tests, ...) aux modèles **Deep-Learning** boîtes noires et peuvent être rapprochées de l'IA générative

# Une dynamique en cours vers l'IA ML/DL

La **disponibilité des données**, un facteur important, problématique.

Mieux prendre en compte cette dimension données -> **nouvelles perspectives**.

Dynamique de plusieurs collègues vers un **positionnement méthode plus IA**, au sens ML/DL

Exploitation pratique **d'outils** au **développement** de méthodes

De manière générale, cette orientation vers le ML/DL *peut signifier pour certaines unités un abandon de travaux en IA plus classique*, plus symbolique

A l'inverse, occasion pour certains de développer ou de relégitimer certains travaux

# La promesse de l'IA générative



## Problème de **définition** de l'IA générative

Curiosité des unités de Mathnum vers les méthodes d'IA générative, qui semblent prometteuses (VAE, data-augmentation, ...)

L'**hybridation** entre ces techniques d'IA générative et les autres démarches plus traditionnelles de modélisation reste un sujet ouvert.

IAG également perçue comme pouvant faire évoluer l'IA vers une discipline moins informatique et plus math appli.

Comme pour le ML puis le DL: **barrière d'accès** / filtre. Tout le monde n'y va pas

# Compétences et de moyens humains



Le **besoin de compétences est manifeste**, et il y a le sentiment d'avoir de moins en moins de **temps pour se mettre à niveau**, en particulier via la biblio.

Tout va très vite et que cela est difficile de suivre ces fronts de recherche.

L'arrivée de l'IA s'accompagne de besoins importants en ingénierie informatique, ce qui peut entraîner des "**frottements**" **vis à vis du rôle des chercheurs informaticiens**.

Question de l'évolution des besoins en ingénierie, en particulier au niveau des plateformes au sein des unités du département.

# Recrutements



Des demandes de poste de chercheurs ou d'ingénieurs, souvent justifiés par le manque de temps pour se former

Recrutements parfois jugés comme difficiles (vivier, niveau)

Des choix variés quant au positionnements de ces profils, entre utilisateurs et experts développeurs, entre méthodologistes et thématiciens

Les nouveaux recrutés apportent cette compétence ML/DL, et permettent de faire le pont avec les méthodes originales de l'unité.

Parfois des inquiétude sur l'avenir des ces chercheurs recrutés pour faire du DL dans des labos de non spécialistes.



# Formation, réseaux d'animation



La formation individuelle (MOOC, ...) est la première réponse à ce besoin de montée en compétences

La formation organisée à l'échelle des équipes ou de l'unité  
Besoin d'espaces de discussion

Expertise / formation de quelques-uns, + diffusion en interne.

Difficulté de suivre l'état de l'art / vitesse

# Publications



ML/DL maintenant très présents dans les publications de tous les domaines applicatifs

=> Conférences disciplinaires ou thématiques traditionnelles se sont donc largement ouvertes au ML et au DL.

Publication de travaux à l'interface dans des revues thématiques qui acceptent les travaux méthodologiques en ML et DL.

Sur les aspects plus théoriques en ML et DL + autres aspects de l'IA,  
=> Explosion des travaux soumis / aléa du processus de soumission, travail de rédaction de plus en plus important.

La profusion des publications rend difficile de suivre l'état de l'art, de faire la part des choses entre les papiers qui comptent vraiment et les autres.



# Moyens de calcul



Associés aux nouvelles méthodes d'IA, des besoins importants en calcul apparaissent au sein des unités. Les solutions retenues pour y répondre sont variées, et passent au départ par des clusters locaux, et un ou quelques GPU.

Le coût important des GPU est néanmoins un obstacle à leur large développement dans les unités ou au sein des plateformes, pour des utilisations non régulières.

On observe alors pour quelques unités le passage sur des solutions nationales distantes, plus puissantes et plus sûres.

# L'impact de ChatGPT



Arrivée de ChatGPT : nombreux usages, que ce soit pour l'écriture ou le codage.

Pour de nombreuses personnes, c'est un réel plus, qui fait gagner du temps.

Souvent les étudiants / doctorants premiers expérimentateurs, mais sans que cela soit véritablement organisé au niveau des unités.

Interrogations sur la place de ces outils en recherche  
Attentes / Craintes de ne pas être en capacité de les exploiter au mieux.