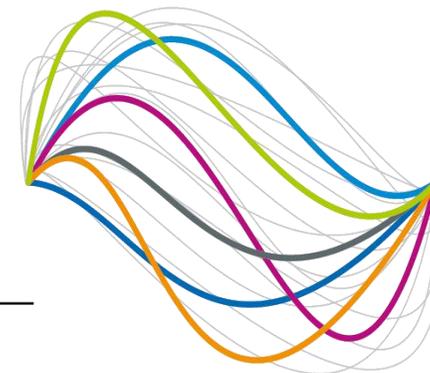


Printemps de la donnée 2022

Datacenter Régional de Bourgogne- Franche-Comté

Un datacenter et un centre de calcul au
service du calcul haute performance et de l'IA

UBFC



UNIVERSITÉ
BOURGOGNE FRANCHE-COMTÉ

RÉGION
BOURGOGNE
FRANCHE
COMTÉ

UNIVERSITÉ de
FRANCHE-COMTÉ



Didier Rebeix – Université de Bourgogne

Kamel Mazouzi – Université de Franche-Comté

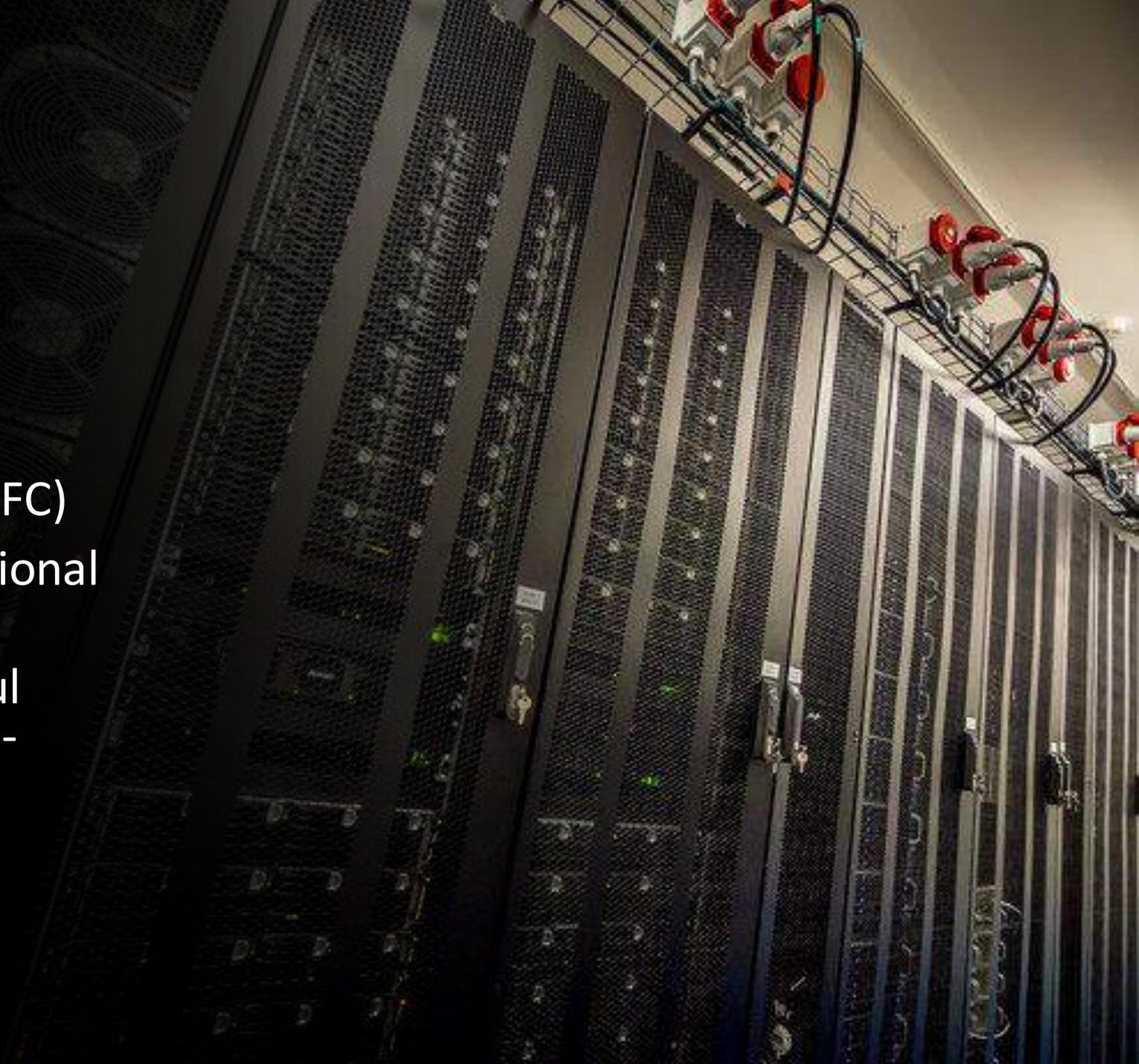
Patrick Senet – Lab. Interdisciplinaire Carnot de Bourgogne (ICB)

Fabien Picaud – Lab. de Nanomédecine, Imagerie et Thérapeutiques (LINIT)



Sommaire

- Présentation de Université de Bourgogne-Franche-Comté (UBFC)
- Présentation du datacenter régional de l'UBFC
- Présentation du centre de calcul régional de Bourgogne-Franche-Comté (MesoBFC)
- Exemples d'applications
- Remerciements





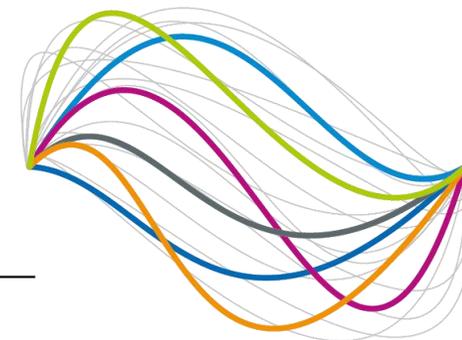
Sommaire

- Présentation de Université de Bourgogne-Franche-Comté (UBFC)
 - Présentation du datacenter régional de l'UBFC
 - Présentation du centre de calcul régional de Bourgogne-Franche-Comté (MesoBFC)
 - Exemples d'applications
 - Remerciements
- 

Présentation de l'université de Bourgogne-Franche-Comté (UBFC)

UBFC

UNIVERSITÉ
BOURGOGNE FRANCHE-COMTÉ



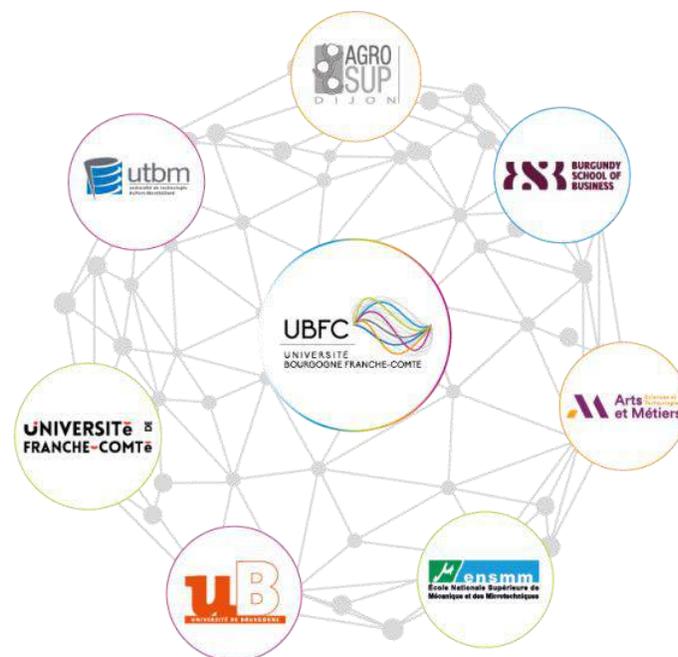
8 800
personnels
dont
2 200 enseignants-chercheurs



60 000
étudiants



7 établissements membres
implantés sur 13 sites



51

laboratoires

Dans une multitude de domaines.
Certains d'entre eux
comprennent les instituts de recherche nationaux



250

chercheurs permanents

(CNRS, INRA, INSERM, INRAP,
EFS, BRGM, IFSTTAR)



Sommaire

- Présentation de Université de Bourgogne-Franche-Comté (UBFC)
 - Présentation du datacenter régional de l'UBFC
 - Présentation du centre de calcul régional de Bourgogne-Franche-Comté (MesoBFC)
 - Exemples d'applications
 - Remerciements
- 

Présentation du Datacenter Régional UBFC (DCR-UBFC)

- Construction 2015
- Niveau Tier 3 : 99,982 % de disponibilité
- Label MESRI – Datacenter regional BFC
- Label European Code of Conduct for datacenters
- Projet d'extension en cours
- Superficie : 600 m²
- Onduleurs + groupe électrogène
- Récupération des calories informatiques
- Contrôle d'accès jusqu'à la baie informatique
- Vidéosurveillance
- Détection/extinction incendie





Présentation du Datacenter Régional UBFC (DCR-UBFC)

Bénéfices :

- Mutualisation des infrastructures
- Sécurité physique
- Sécurité électrique
- Sécurité de refroidissement

Offre de service :

- Hébergement sec
- Serveurs virtuels
- Calcul haute performance
- Ressources pour l'IA
- Expertise scientifique





Présentation du Datacenter Régional UBFC (DCR-UBFC)

Les baies informatiques

- Confinement d'allée chaude pour améliorer la récupération des calories.



Présentation du
Datacenter Régional
UBFC (DCR-UBFC)

Groupe électrogène

+

Onduleurs

=

Résilience en cas de
panne EDF



Présentation du Datacenter Régional UBFC (DCR-UBFC)

Pompe à chaleur :

- Récupération des calories IT
- Injection dans le réseau de chaleur du campus
- Chauffage et eau chaude sanitaire pour le campus





Sommaire

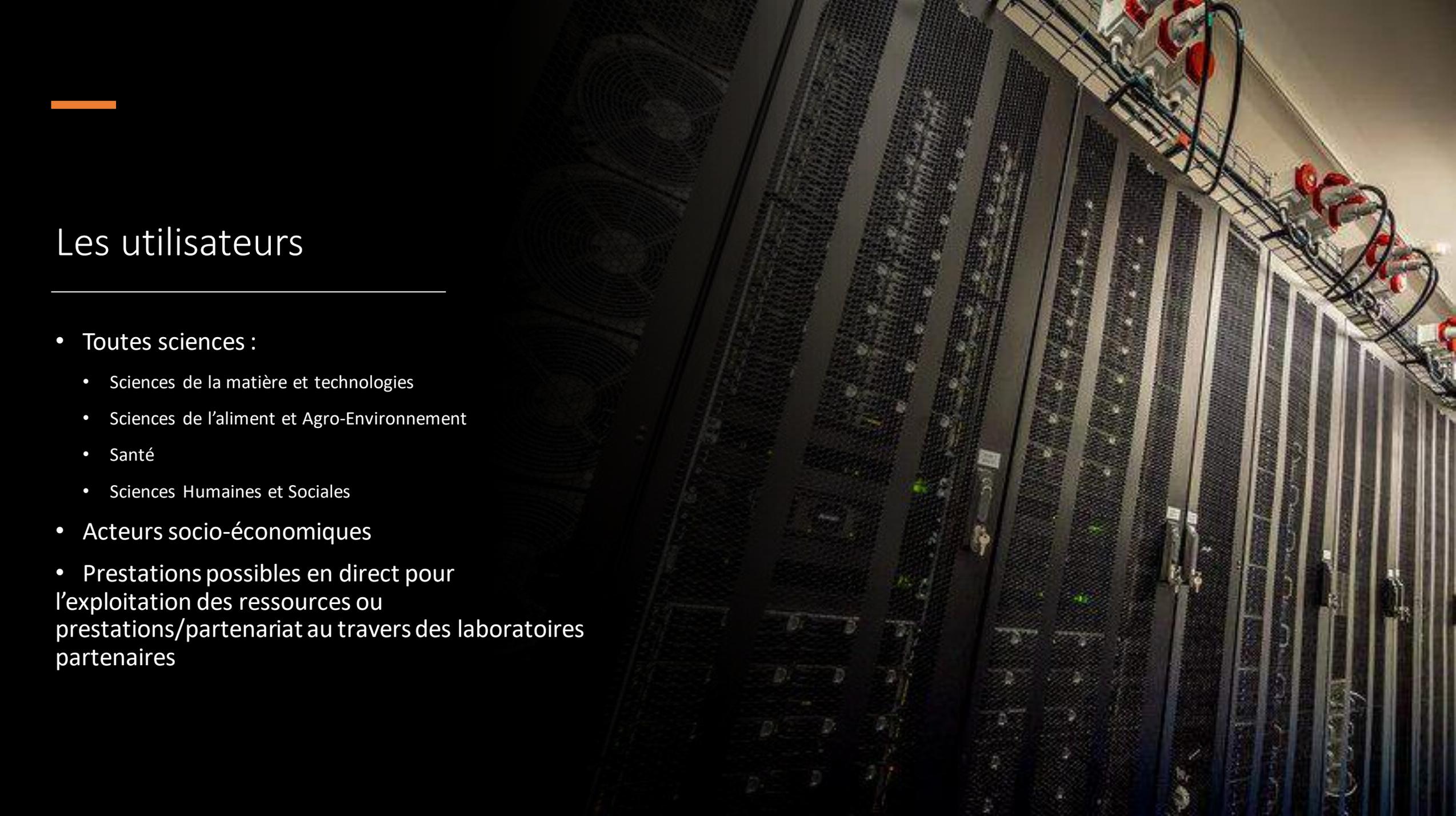
- Présentation de Université de Bourgogne-Franche-Comté (UBFC)
 - Présentation du datacenter régional de l'UBFC
 - Présentation du centre de calcul régional de Bourgogne-Franche-Comté (MesoBFC)
 - Exemples d'applications
 - Remerciements
- 



Présentation du centre de calcul régional de Bourgogne-Franche-Comté (MesoBFC)

Missions

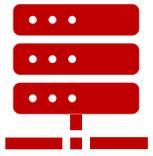
- Plateforme scientifique pluridisciplinaire.
- Fédérer les ressources pour le calcul intensif et l'IA au niveau régional.
- Expertise en simulations numériques, traitements de données massives à haute performance et le développement de l'Intelligence Artificielle.
- A la disposition des acteurs de la recherche et du monde socio-économique régional.
- Formation et accompagnement des utilisateurs.



Les utilisateurs

- Toutes sciences :
 - Sciences de la matière et technologies
 - Sciences de l'aliment et Agro-Environnement
 - Santé
 - Sciences Humaines et Sociales
- Acteurs socio-économiques
- Prestations possibles en direct pour l'exploitation des ressources ou prestations/partenariat au travers des laboratoires partenaires

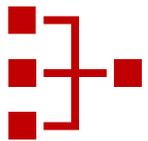
Puissance **957 TéraFlop/s**



12 300 coeurs de
calcul généralistes (Intel, AMD)



3.5 Péta octet
stockage de données



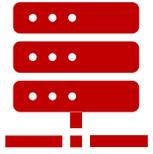
Réseaux très haut
débit 100 Gb/s



Plus de 100 logiciels
scientifiques optimisés par les
équipes de MesoBFC



Ressources pour I.A



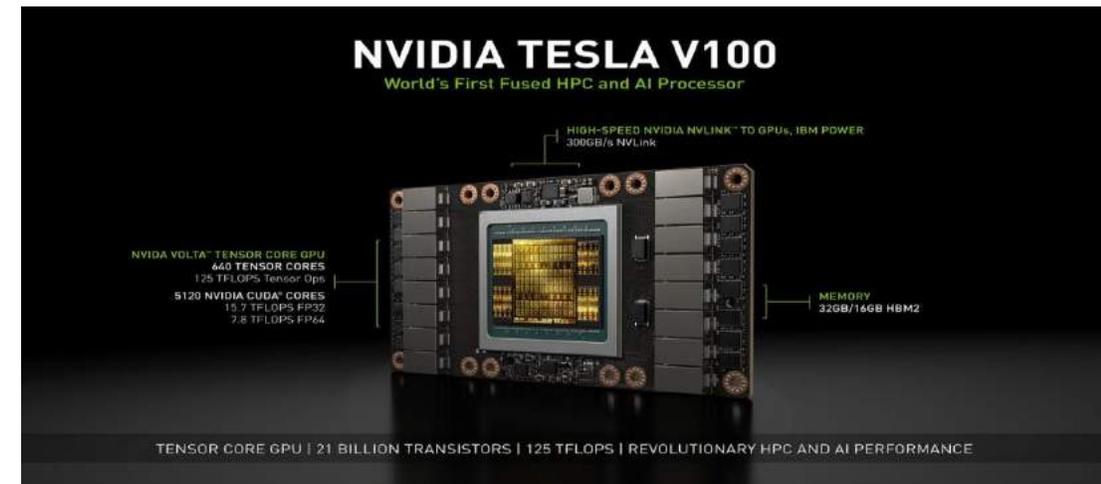
34 GPU nVIDIA
V100/A100



Tensorflow, PyTorch,
Caffe, Keras, Scikit-
learn, ...



Docker & Singularity



Ressources pour les données : 3.5 Peta octet



Espace de travail (Work, Scratch)



Espace de stockage (BDD, Cloud)



Espace d'archivage (Bande, DVD)

Classification selon :

la localité des données par rapport au CPU

la durée de vie de la donnée

le système de fichiers utilisé



Sommaire

- Présentation de Université de Bourgogne-Franche-Comté (UBFC)
 - Présentation du datacenter régional de l'UBFC
 - Présentation du centre de calcul régional de Bourgogne-Franche-Comté (MesoBFC)
 - Exemples d'applications
 - Remerciements
- 

CLASSIFICATION DE DOSSIERS PATIENTS (FEMTO & HNFC)

Codes CIM10

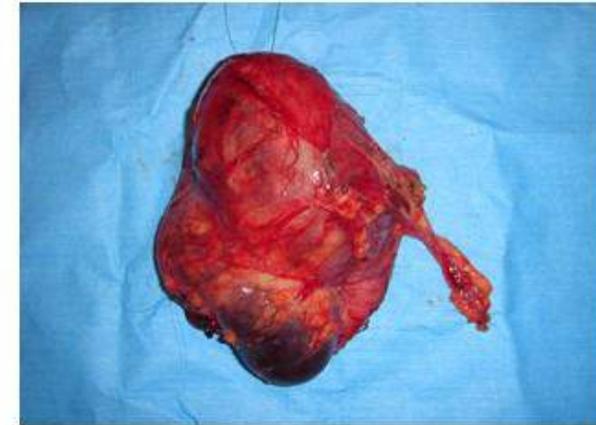
- Classification internationale obligatoire pour chaque dossier patient
- 12 000 codes.
- Ex : **U07.10** = *COVID-19, forme respiratoire, virus identifié*
- Classification manuelle, très coûteuse en temps et en personne.
- Automatisation grâce à l'IA et au Traitement Automatique des Langues (NLP)
- Utilisation des derniers modèles de Deep Learning : BERT



Segmentation Automatique de reins tumoraux chez l'enfant par Intelligence Artificielle Distribuée

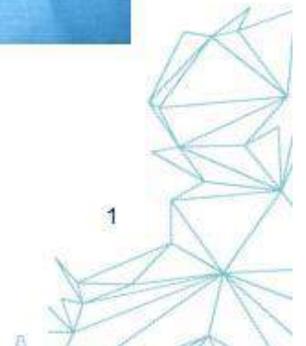
Néphroblastome :

- La tumeur rénale la plus fréquente chez l'enfant
- Segmentations guidées manuellement : sources d'erreur humaine et chronophages
- Planification pré-opératoire de l'intervention (anticipation des risques opératoires)
- Choix du geste chirurgical (néphrectomie totale versus néphrectomie partielle)

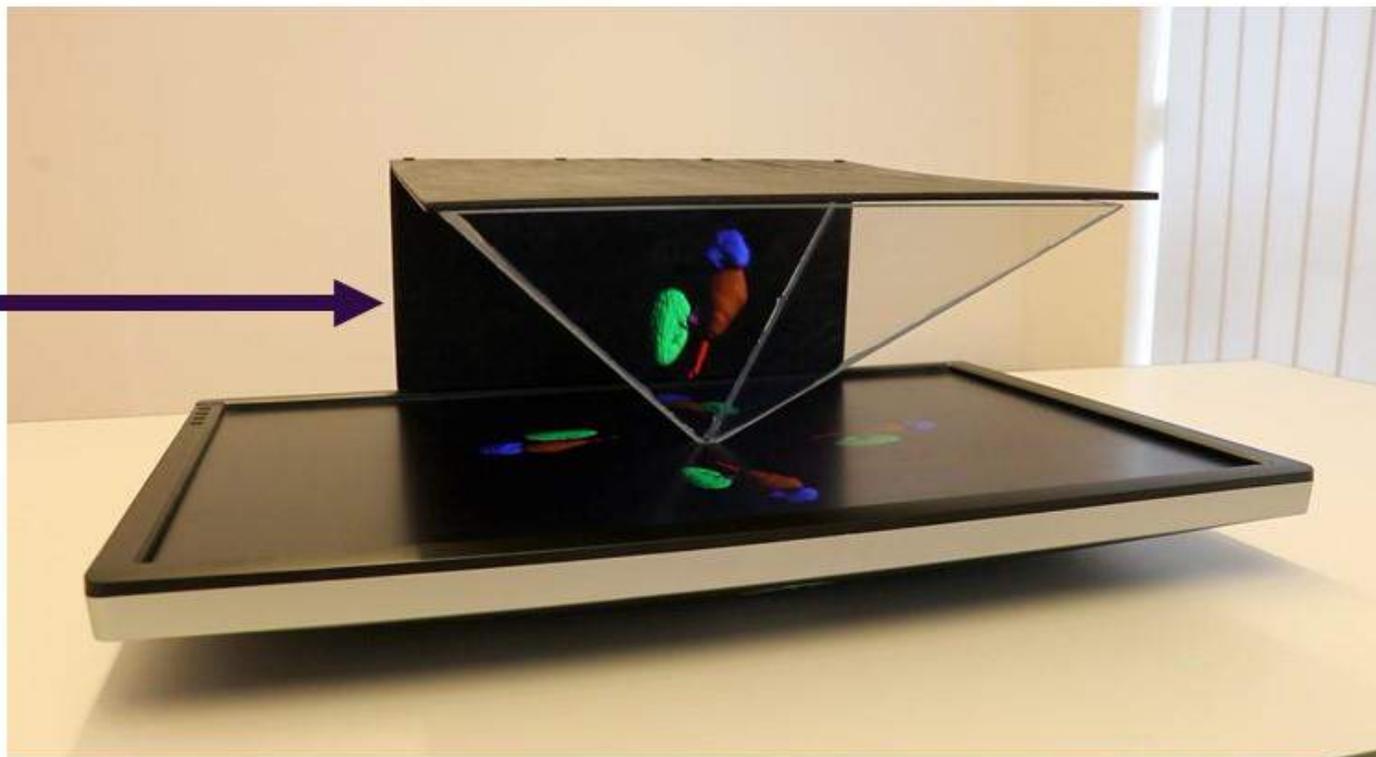
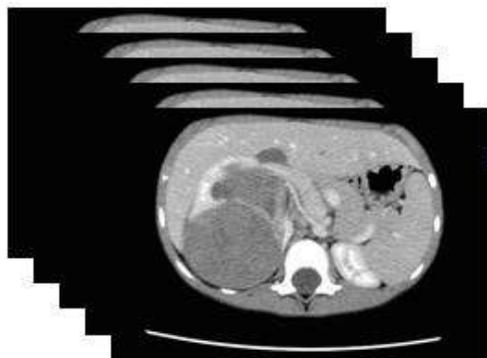


16 & 17 juin 2021

Salon de l'IA - Belfort



Résultats



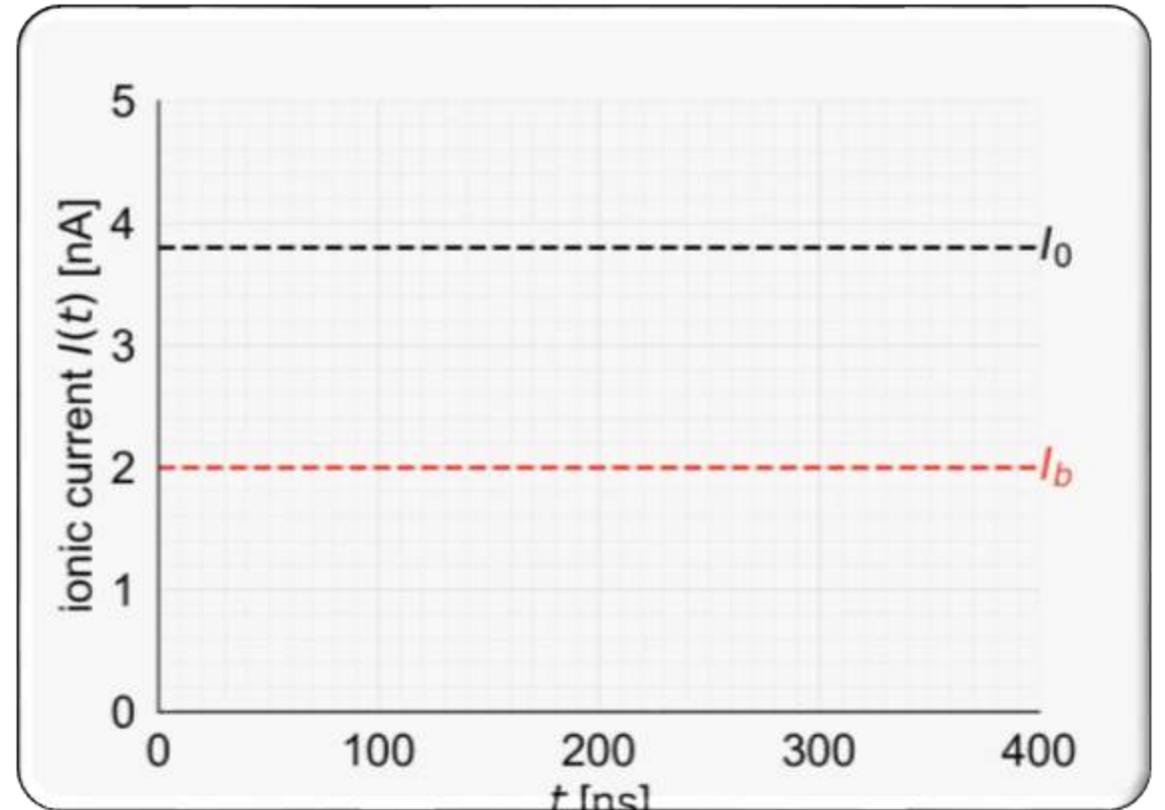
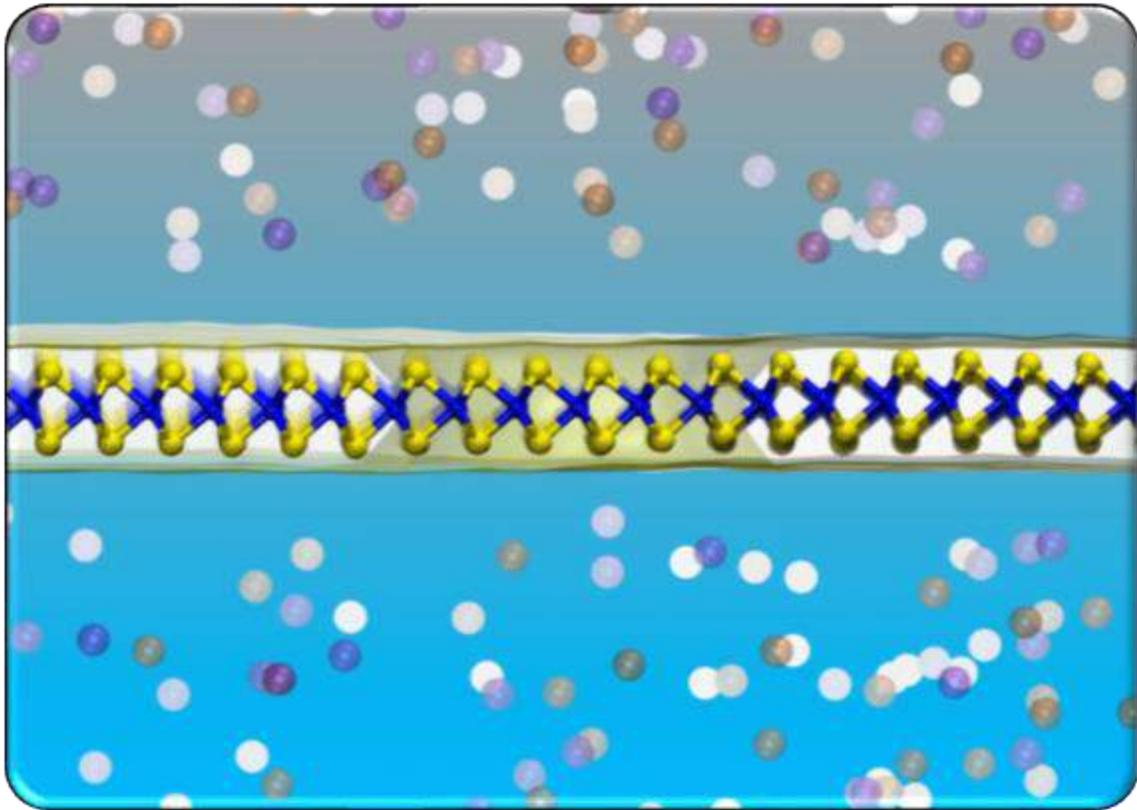
Structure anatomique	Précision inter-opérateur (Dice)	Précision plate-forme SAIAD (Dice)
Rein pathologique	0,86	0,85
Tumeur	0,95	0,94
Globale	0,91	0,90

16 & 17 juin 2021

Salon de l'IA - Belfort

SEPIA (2021-2023)

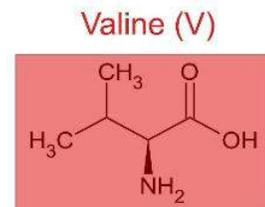
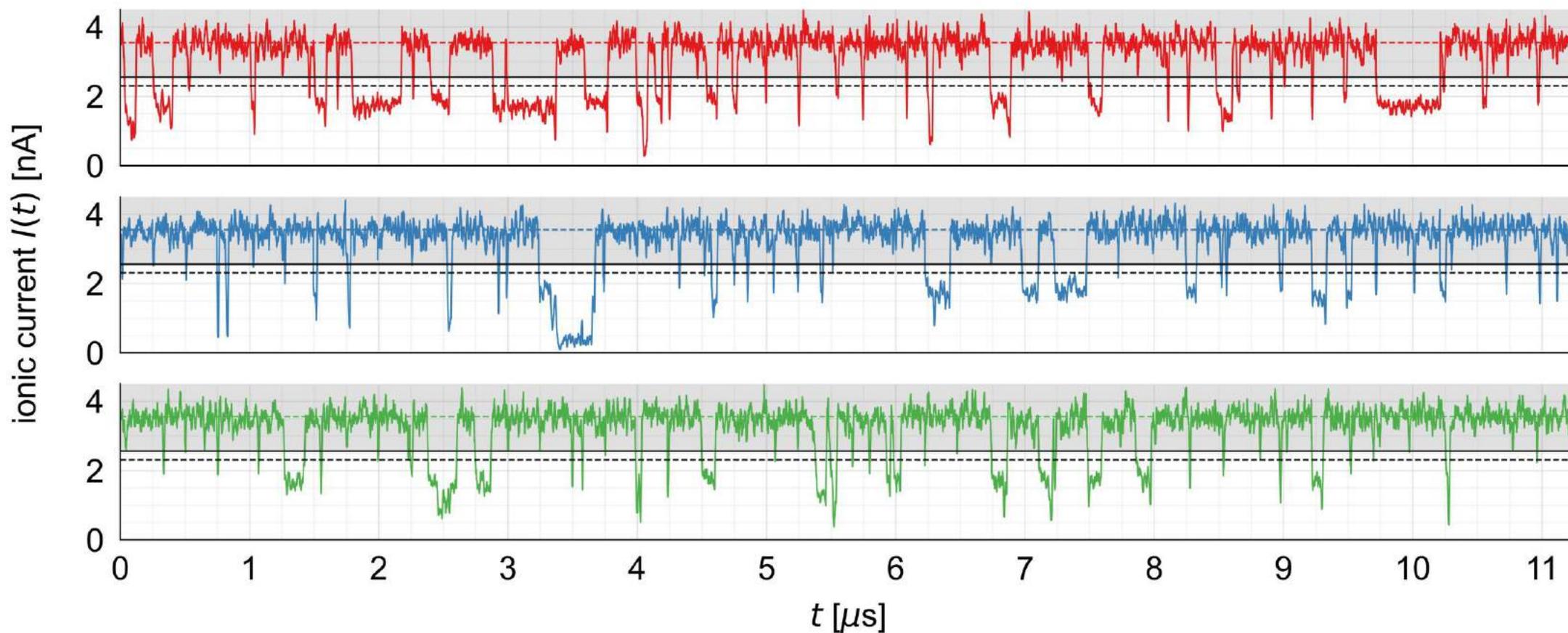
Données: signal de courant extrait de simulations de dynamique moléculaire du nano-séquenceur de protéines dans son environnement réel.



SEPIA (2021-2023)

Big data: Plusieurs centaines de simulations générées

(1 simulation GROMACS MPI: 30h wallclock time – 224 coeurs - AMD 7302 3 GHz)



Une architecture à haute performance pour la collecte, le stockage et l'analyse des données des réseaux sociaux

Annabelle Gillet & Éric Leclercq & Nadine Cullot

LIB - EA 7534 Université de Bourgogne Franche-Comté
Équipe Science des Données

SIA 2021



Calculer des indicateurs en temps réel

Production d'indicateurs en temps réel, qui correspondent à des besoins bien identifiés (ex. : calculer des tendances, détecter des évènements).

De nouveaux indicateurs peuvent être ajoutés en fonction des découvertes dans les analyses exploratoires.

Effectuer des analyses exploratoires

Pour fouiller les données lors de situations ou besoins spécifiques.

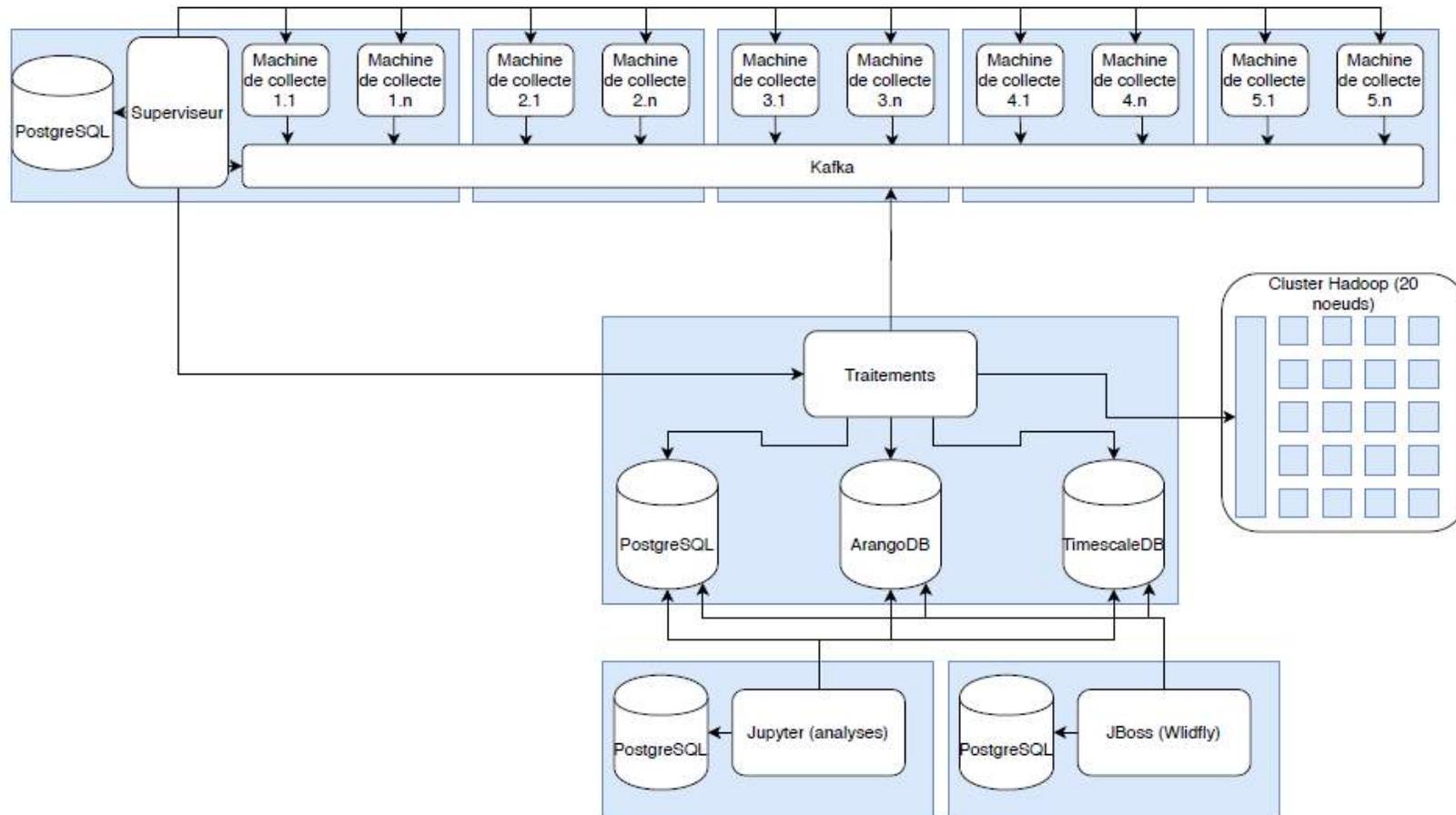
Les analyses peuvent être guidées par des anomalies détectées dans les indicateurs en temps réel.



ISITE-BFC (ANR-15-IDEX-0003), responsable : Gilles Brachotte

Le but du projet Cocktail est de créer un observatoire en temps réel des tendances, des singularités et des signaux faibles dans les discours à propos de l'alimentaire et de la santé sur Twitter.

Architecture physique (déployée au DC UBFC)



Depuis le début de la collecte (Octobre 2019) : 11To de données brutes,
450M de tweets

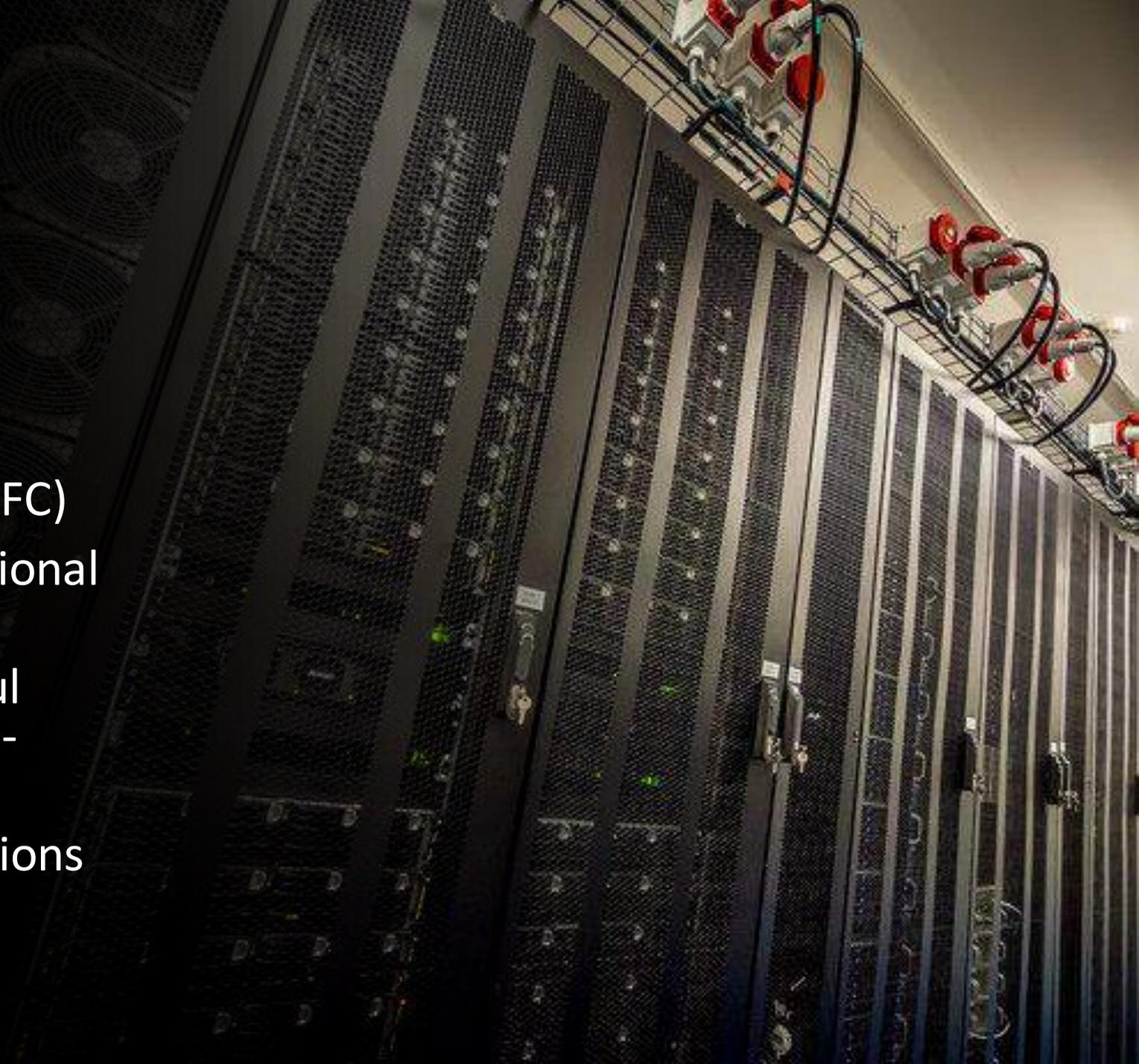
- Architecture en production depuis 2019, capable d'absorber 6 000 tweets/s
- Transfert de compétences avec les deux entreprises régionales partenaire du projet Cocktail
- 3 thèses en cours
- Des études réalisées sur l'incendie de l'usine Lubrizol, les vaccins et les discours complotistes, le mouvement vegan, etc.
- Plus d'une dizaine de publications en informatique

Contact : Éric Leclercq (eric.leclercq@u-bourgogne.fr),
responsable de l'équipe sciences des données du Laboratoire
d'Informatique de Bourgogne,
responsable du Master 2 Bases de Données et Intelligence Artificielle

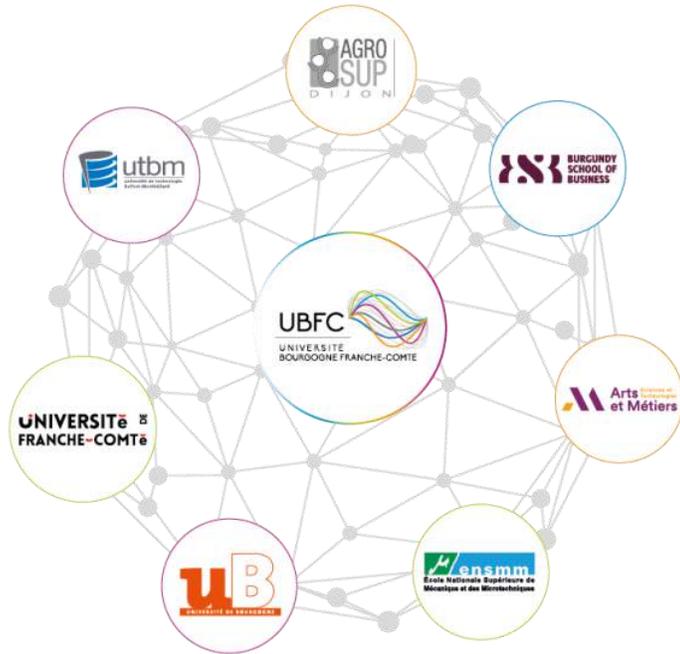


Sommaire

- Présentation de Université de Bourgogne-Franche-Comté (UBFC)
- Présentation du datacenter régional de l'UBFC
- Présentation du centre de calcul régional de Bourgogne-Franche-Comté (MesoBFC)
- IA : Démonstrateurs et applications
- Remerciements



Remerciements



RÉGION BOURGOGNE FRANCHE COMTÉ

