

# Reproductibilité de la vision par ordinateur basée sur l'apprentissage machine en bio-imagerie

david.rousseau@univ-angers.fr

# Reproductibilité de la vision par ordinateur basée sur l'apprentissage machine en bio-imagerie

david.rousseau@univ-angers.fr

Midjourney prompt : **“De Gaulle calling for reproducible AI, black &white“**

Midjourney prompt : “De Gaulle calling for reproducible AI, black &white“

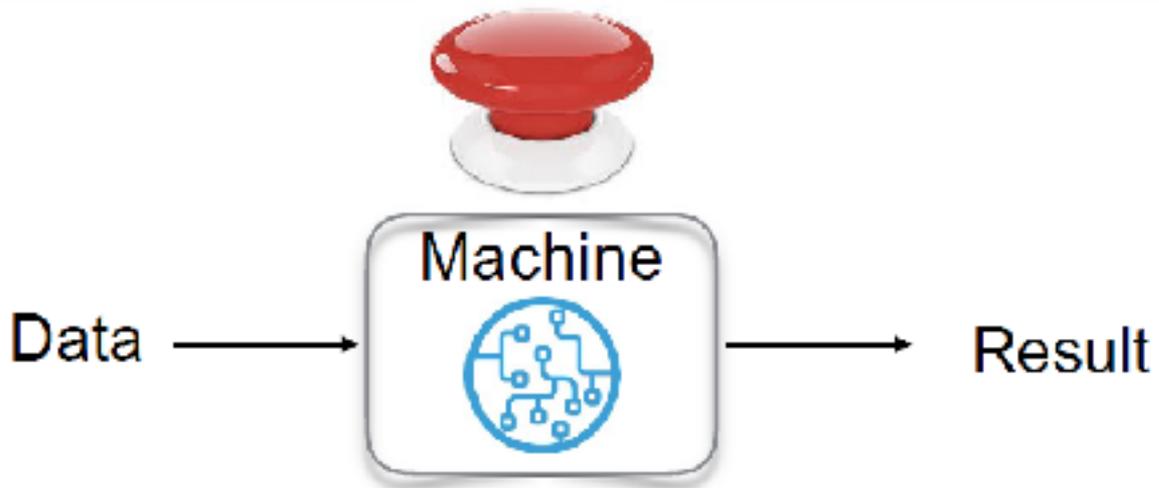


Midjourney prompt : “De Gaulle calling for reproducible AI, black &white“



# Machine learning

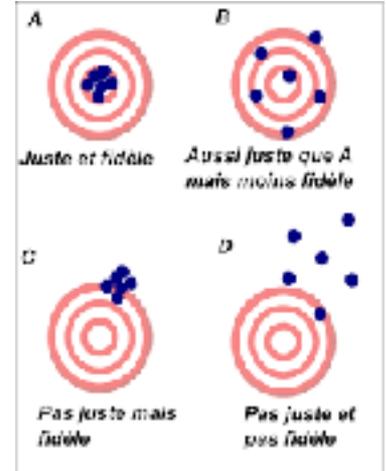
Apprentissage automatique ... ne signifie pas reproductible



# Reproductibilité ?

la capacité d'une expérience à être répétée par un autre expérimentateur ou une autre expérimentatrice. La reproductibilité est l'un des principaux aspects des processus scientifiques.

Deux notions : justesse et fidélité



<https://github.com/royerlab/napari-chatgpt>

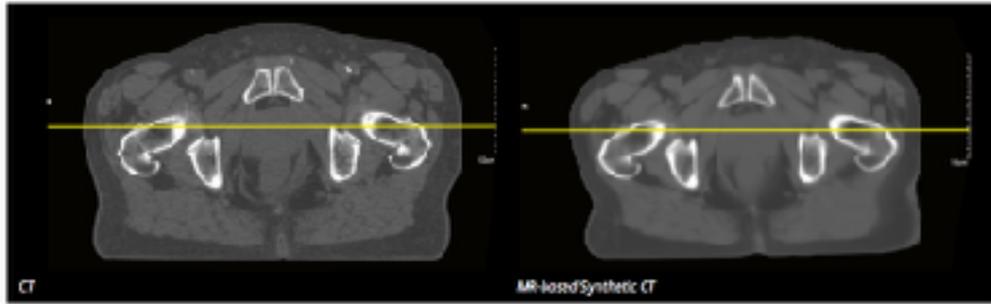
The screenshot shows the GitHub repository page for `napari-chatgpt` by user `haeleinhuepf` (Robert Haase). The repository description reads: "Home of *Omega*, a napari-aware autonomous LLM-based agent specialized in image processing and analysis." The repository has 217 stars and 21 forks. The languages section shows: CSS (0.3%), Python (14.1%), JavaScript (0.5%), and HTML (0.1%).

Below the repository information, there are three panels:

- Logo:** A circular logo with a stylized eye in the center, surrounded by various icons representing different tools and technologies. Below the logo is the text "NAPARI - CHATGPT".
- Image Processing:** A screenshot of the napari interface showing a 3D visualization of green, glowing, spherical structures, likely representing biological data.
- Chat Interface:** A screenshot of a chat window titled "Omega" with a text input field and a "Send" button.

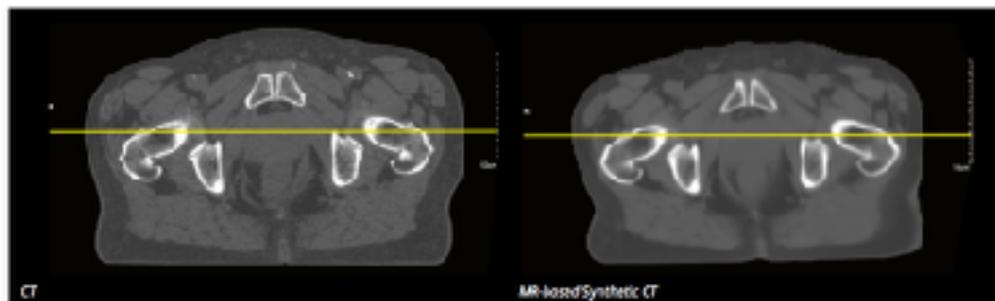
Royer, L. A. (2024). Omega—harnessing the power of large language models for bioimage analysis. *Nature Methods*, 1-3.

Plus dangereux ?



**SIEMENS**

# Plus dangereux ?



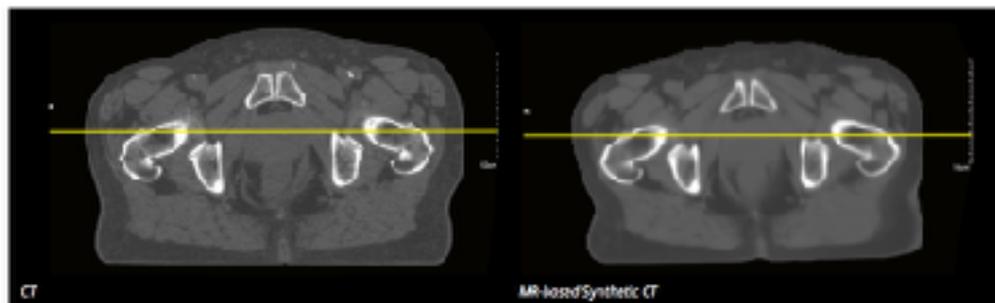
SIEMENS

Mean dose difference [Gy]

Patient	Prescription dose	PTV	GTV	Bladder	CTV/ovary
P1	120cc 4 Gy	-0.18	-0.14	0	0
P2	450cc 2 Gy	0.02	0.08	0.17	0.62
P3	280cc 1.8 Gy	0.83	0.85	0	1.37
P4	120cc 4 Gy	-0.04	-0.04	0.51	0.46
P5 (Plan 1)	11cc 18 Gy	0.04	0.04	0	0
P5 (Plan 2)	11cc 18 Gy	0.75	0.89	0	0
<b>Median</b>		0.02	0.06	0	0.21
<b>95th Percentile</b>		0.75	0.89	0.17	1.10
<b>MEAN</b>		0.83	0.85	0.52	1.37

Table 1: Differences in dose distributions of planning CT and Synthetic CT in the PTV, GTV, and two organs at risk (ovary, the bladder and uterus), for breast cancer patients. P5 has two different plans for two different target volumes.

# Plus dangereux ?



# SIEMENS

Quelle base d'entraînement ?

Mean dose difference [D]

Patient	Prescription dose	PTV	GTV	Residuals	CI Deviation
P1	120cc 4 Gy	-0.18	-0.14	0	0
P2	450cc 2 Gy	0.02	0.08	0.17	0.42
P3	280cc 1.8 Gy	0.83	0.85	0	1.37
P4	120cc 4 Gy	-0.04	-0.04	0.51	0.46
P5 (Plan 1)	11cc 18 Gy	0.04	0.04	0	0
P5 (Plan 2)	11cc 18 Gy	0.75	0.89	0	0
<b>Median</b>		<b>0.02</b>	<b>0.06</b>	<b>0</b>	<b>0.21</b>
<b>95th Percentile</b>		<b>0.75</b>	<b>0.89</b>	<b>0.17</b>	<b>1.10</b>
<b>MAX</b>		<b>0.83</b>	<b>0.85</b>	<b>0.51</b>	<b>1.37</b>

Table 1: Differences in dose distributions of planning CT and Synthetic CT in the PTV, GTV, and two organs at risk (the brainstem and larynx), for three breast patients. P1 has no residual dose plans for use of breast augmentation.



# Avoiding a replication crisis in deep-learning-based bioimage analysis

Deep learning algorithms are powerful tools for analyzing, restoring and transforming bioimaging data. One promise of deep learning is parameter-free one-click image analysis with expert-level performance in a fraction of the time previously required. However, as with most emerging technologies, the potential for inappropriate use is raising concerns among the research community. In this Comment, we discuss key concepts that we believe are important for researchers to consider when using deep learning for their microscopy studies. We describe how results obtained using deep learning can be validated and propose what should, in our view, be considered when choosing a suitable tool. We also suggest what aspects of a deep learning analysis should be reported in publications to ensure reproducibility. We hope this perspective will foster further discussion among developers, image analysis specialists, users and journal editors to define adequate guidelines and ensure the appropriate use of this transformative technology.

Romain F. Laine, Ignacio Arganda-Carreras, Ricardo Henriques and Guillaume Jacquemet

# IA de confiance (EU AI Act)

Pour répondre à ces enjeux, l'IA de confiance doit être déclinée sur 8 piliers

Ces 8 principes sont, telles assignées par l'IA Act, mais également celles qui sont des enjeux majeurs pour les organisations.

## Robustesse

Définir et repousser les limites de validité du modèle

## Performance

Répondre aux besoins métiers de manière optimale

## Qualité de données

Maîtriser la qualité des données en entrée du modèle

## Responsabilité

Déployer une gouvernance transparente et explicite

## IA DE CONFIANCE

## Frugalité

Mesurer et minimiser l'impact environnemental

## Équité

Maîtriser les biais et garantir l'équité des résultats

## Contrôle des dérives

Contrôler l'impact opérationnel du modèle en production

## Explicabilité

Expliquer les prédictions aux utilisateurs finaux

Calendrier de déploiement de l'IA Act



<https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

# Qui parle ?

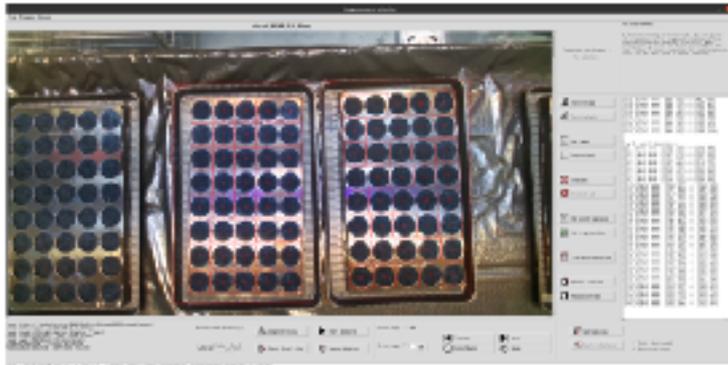
David R. (responsable équipe ImHorPhen Angers IRHS - IA dans ISC Phenome)

# Qui parle ?

David R. (responsable équipe ImHorPhen Angers IRHS - IA dans ISC Phenome)



PHENOGRID (cinétique) : émergence, symptômes foliaires, cycle circadien, germination ...



Techno-providers



Instituts Techniques



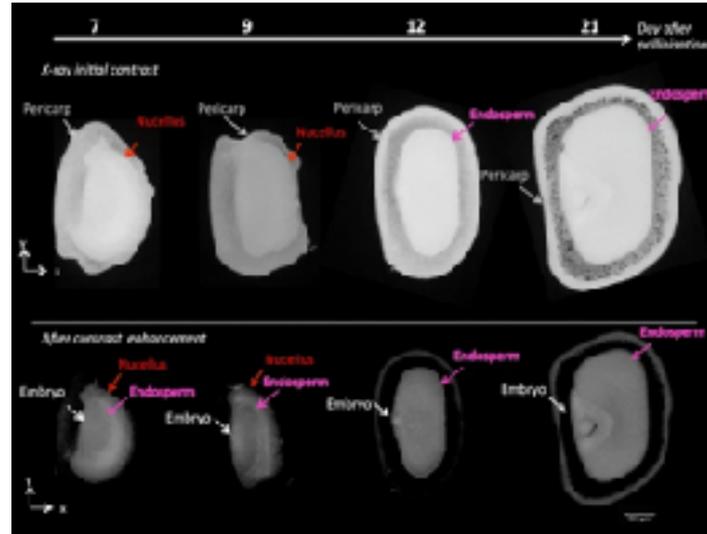
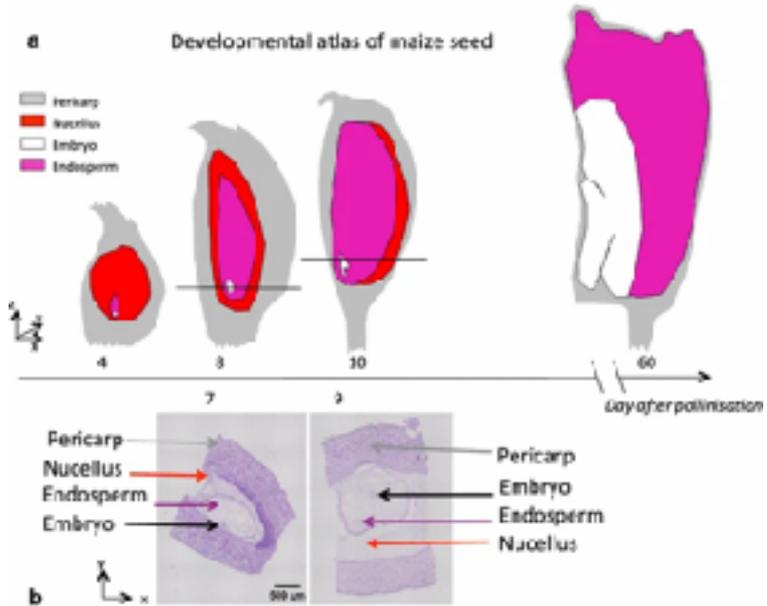
Agro-entreprises



# Qui parle ?

David R. (responsable équipe ImHorPhen Angers IRHS - IA dans ISC Phenome)

Rousseau, D., Widiez, T., Di Tommaso, S., Rositi, H., Adrien, J., Maire, E., ... & Rogowsky, P. (2015). Fast virtual histology using X-ray in-line phase tomography: application to the 3D anatomy of maize developing seeds. *Plant Methods*, 11, 1-10.



# Qui parle ?

David R. (responsable équipe ImHorPhen Angers IRHS - IA dans ISC Phenome)

Growth of plants

Day 1 : Camera's installation



Day 10 : plant has grown more than expected



Heads of the plants rotated



Individuals in the scene



# Qui parle ?

David R. (responsable équipe ImHorPhen Angers IRHS - IA dans ISC Phenome)

## Work done and results in T2.2 Trouble shooting

Technical problem related to the sensor



Change the position of the camera during the emergence

Image day 3



Image in day 4



Weather : the fog, sun, water



Camera is not on top view

# Road map ?

Les sources de variabilité en apprentissage machine

- Algorithmes
- Données
- Annotations
- Composants librairies et matériels
- Procédures de création

Illustrations en imagerie du vivant

Bonnes pratiques

# Apprentissage machine et vision par ordinateur kezako ?

## Computer Vision and Machine Learning

### Trainig process

The machine is shown different foods to find a tomato



Image Features

Training

Result

Tomato ✓  
Apple ✗  
Peach ✗

Images with tomatoes already highlighted and labeled

Training Labels

Tomato ✓

### Image Recongition process

Now we show a tomato to the machine



Image Features

Learned Model

Prediction

Tomato ✓

# Source de variabilité dans les algorithmes

De nombreux algorithmes incluent une part d'aléatoire dans leur fonctionnement propre

# Source de variabilité dans les algorithmes

De nombreux algorithmes incluent une part d'aléatoire dans leur fonctionnement propre

## K-Means

- Iterative method to cluster your data
1. Initialization: pick  $k$  random cluster centroids
  2. Based on the centres, assign all points to a cluster
  3. Based on the assignments, recalculate the centroids
  4. Repeat 2-3 until convergence



# Source de variabilité dans les algorithmes

De nombreux algorithmes incluent une part d'aléatoire dans leur fonctionnement propre

K-Means

**Random Forest**

**Stochastic Gradient Descent**

# Bonne pratique

Lancer plusieurs entraînements

Calculer la valeur moyenne et écart type

Estime la variabilité liée à l'algorithme lui-même

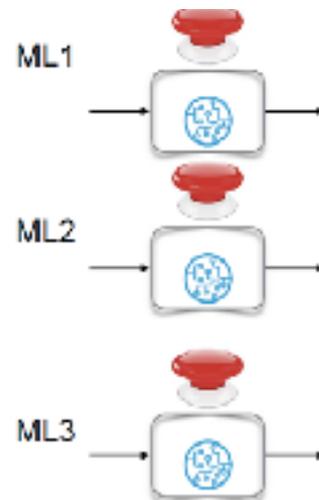


# Conséquence d'être piloté par la donnée

Comme les algo sont pilotés par la donnée

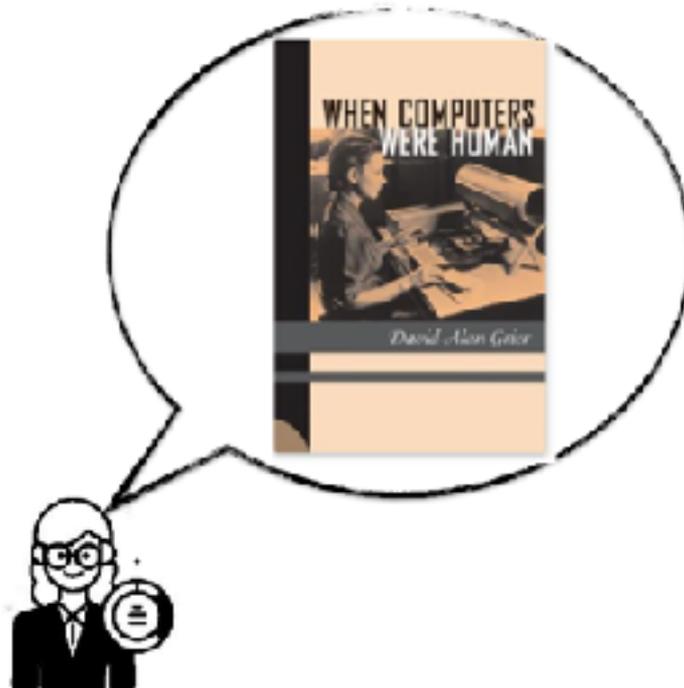
Important de ne pas se limiter à un algorithme

Comparer plusieurs machines ... surtout si on prétend en avoir un qui est le "meilleur"



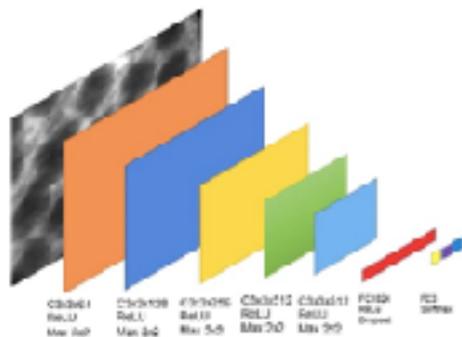
# Conséquence d'être piloté par la donnée

Une très ancienne bonne pratique



# Exemple de présentation de résultats

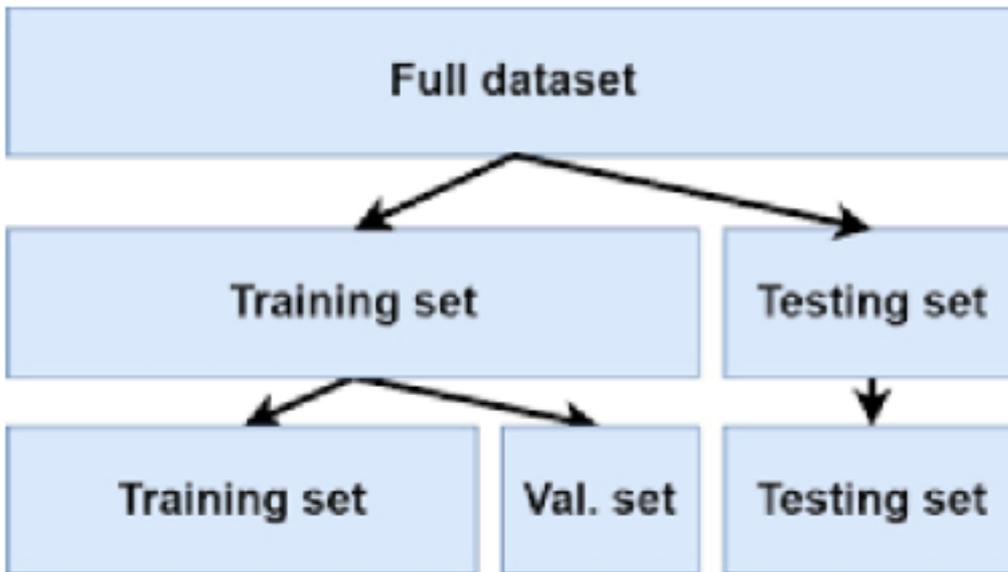
Left			Right			
Classifiers	Transfer Learning	Accuracy		True Cancer	True Inflammation	True Healthy
Proposed CNN architecture	—	96.93% ± 0.13	Predicted Cancer	13994	0	0
LBP features + linear SVM	—	97.7% ± 0.33	Predicted Inflammation	0	4032	0
VGG16 + linear SVM	X	86.6% ± 0.4	Predicted Healthy	0	6	1849
VGG16	X	82.12% ± 4.1				
ResNet50	X	79.94% ± 4.6				
DenseNet	X	78.01% ± 3.3				
VGG16	—	78.46% ± 1.27				



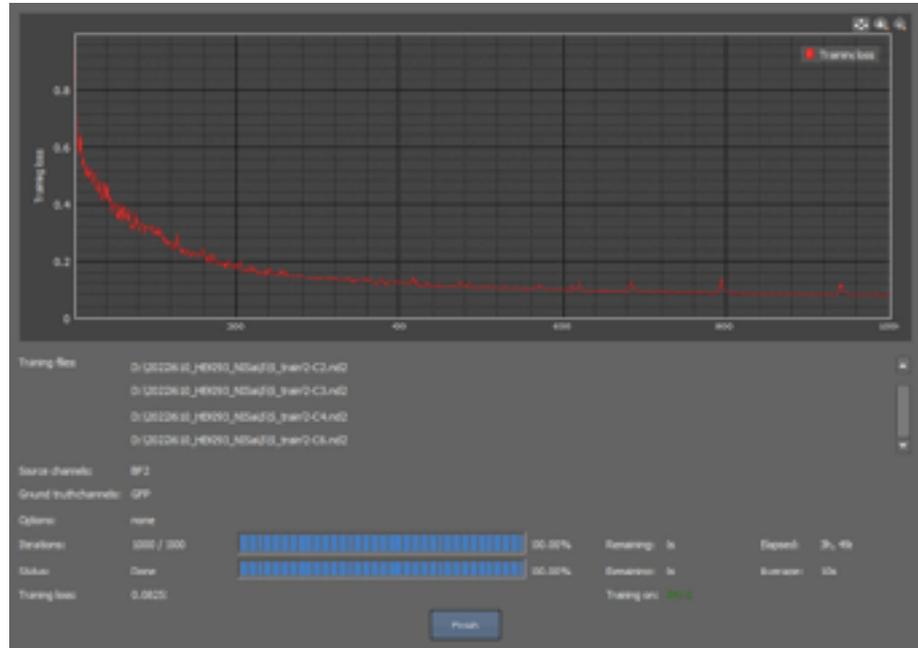
Rasti, F., Wolf, C., Dorez, H., Sablong, R., Moussata, D., Samiei, S., & Rousseau, D. (2019). Machine Learning-Based Classification of the Health State of Mice Colon in Cancer Study from Confocal Laser Endomicroscopy. *Scientific reports*, 9(1), 1-11.

# Source de variabilité dans les données

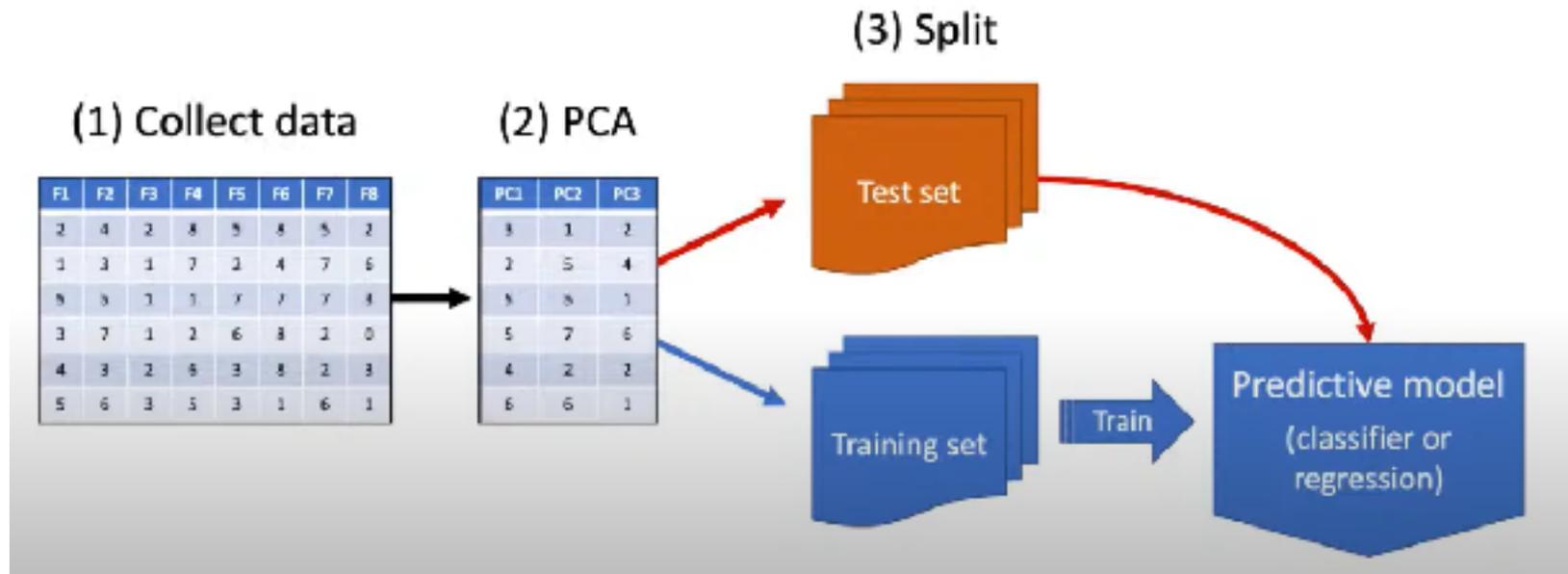
Les jeux de données doivent être arbitrairement séparés pour établir le jeux d'apprentissage



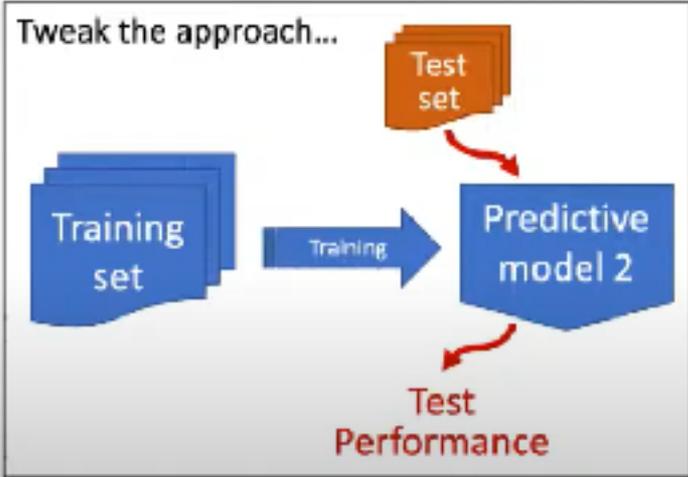
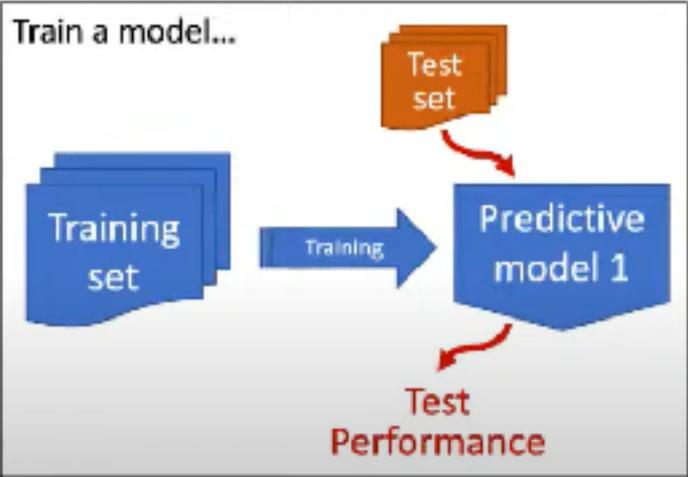
# Exemple 1 No split



## Exemple 2



# Exemple 3



# Bonne pratique

Réaliser un K-Fold cross validation

Calculer l'écart type et la moyenne

Estimer l'incertitude liée au data set lui-même



# Biais et Data



Thermomètre : 60%



Arme : 61%

Il n'existe pas un biais mais de nombreux types de biais statistiques:



## Types de biais [\[ modifier \]](#) [\[ modifier le code \]](#)

- **bias effet-centre**
- **bias de vérification** (work-up bias)
- **bias d'autosélection**, estimé à 27 % des travaux d'écologie et pseudoréplications<sup>1</sup>
- **bias de sélection** : les personnes sondées ne sont pas représentatives de l'**Échantillon biaisé**.
- **bias de mesure** : les techniques de mesures sont incorrectes
- **bias de notification**
- **bias de publication** : les données sont davantage diffusées si elles sont positives
- **bias de confirmation**
- **bias de confusion**
- **bias d'un estimateur**
- **bias de suivi** (appelé aussi **biais de réalisation**) : lorsque les données ne sont pas représentatives de la population
- **bias d'attrition** : retrait de certains patients de l'analyse
- **bias d'évaluation** : lorsque la mesure du critère de jugement est biaisée
- **bias d'indication** : cas particulier de biais de confusion, lorsque les données sont biaisées en fonction de la présence ou non d'une maladie
- **bias d'information** : lorsque les données sont biaisées en fonction de la qualité de l'information
- **bias d'interprétation** : erreur dans le mode d'analyse des données
- **bias de spectre** : en médecine, un test diagnostique peut offrir un spectre de sensibilité différent de celui de la population
- **erreur écologique**
- **problème d'agrégation spatiale**

Quelques biais courants dans les projets IA:

**Biais de sélection** : le jeu de données ne reflète pas les réalités de l'environnement dans lequel le modèle fonctionnera.

**Biais de mesure** : lorsque les données recueillies pour l'entraînement diffèrent de celles recueillies dans le monde réel

**Biais d'association** : Ce biais se produit lorsque les données d'un modèle d'apprentissage machine renforcent et/ou multiplient un biais culturel

# Bonne pratique

Go FAIR

Méta données un moyen d'identifier des biais

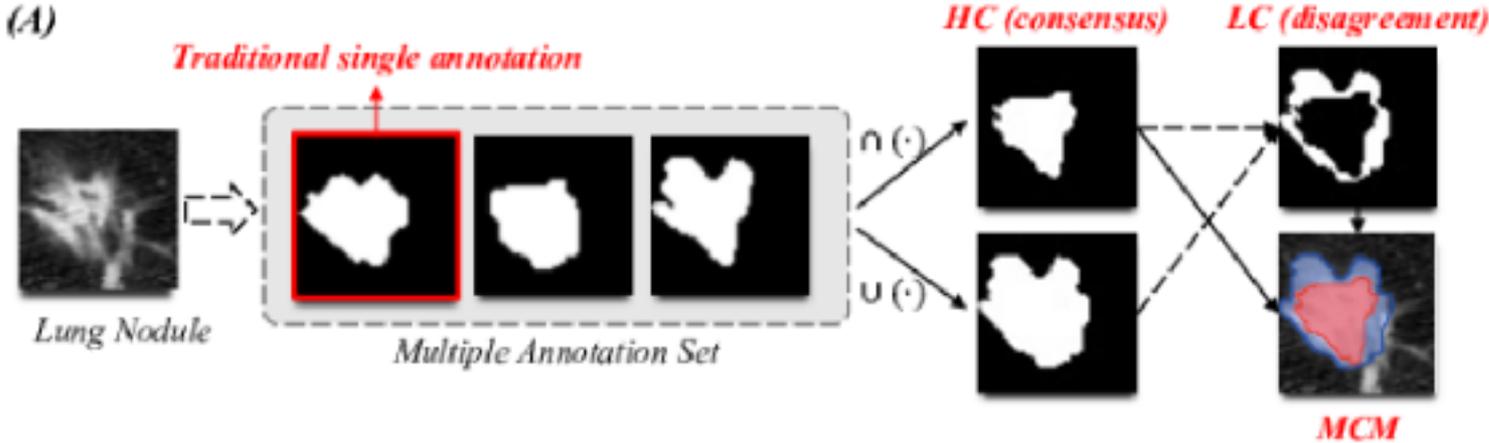
Ouvrir les données et annotations



# Source de variabilité dans les annotations

Les humains qui fournissent des annotations peuvent être en désaccord

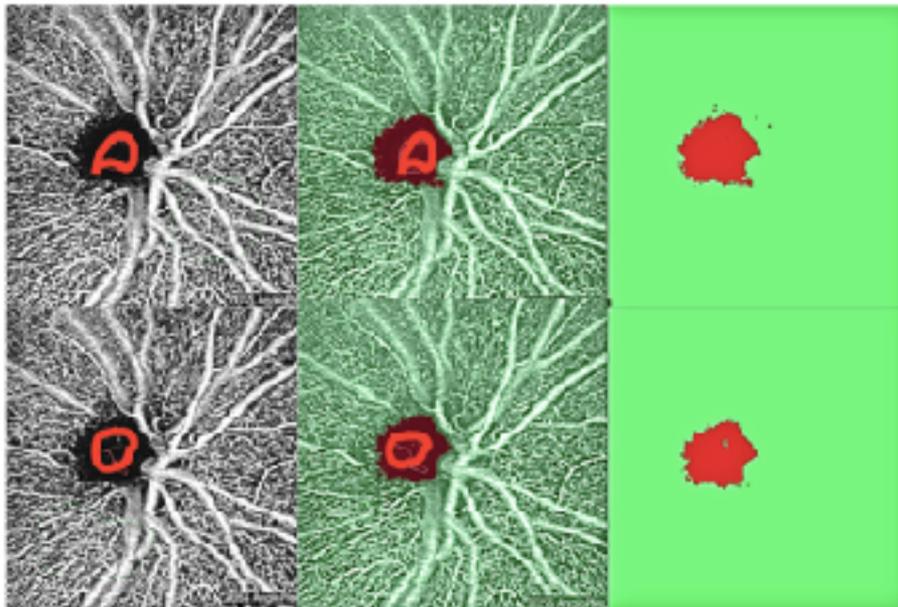
Exemple diagnostic médical



# Source de variabilité dans les annotations

Les humains qui fournissent des annotations peuvent être en désaccord

Exemple diagnostic médical



# Bonne pratique

Enregistrer cette variabilité

L'utiliser à l'entraînement ou l'analyse



# Source de variabilité Librairie

[Learn](#) ▾[Ecosystem](#) ▾[Edge](#) ▾[Docs](#) ▾[Blogs &](#)[News](#) ▾[About](#) ▾[Become a](#)[Member](#)

2.3 ▾

[Community](#) [ + ][Developer Notes](#) [ + ][Language Bindings](#) [ + ][Python API](#) [ - ][torch](#)[torch.nn](#)[Docs](#) > [Reproducibility](#) [Shortcuts](#)

## Reproducibility

Completely reproducible results are not guaranteed across PyTorch releases, individual commits, or different platforms. Furthermore, results may not be reproducible between CPU and GPU executions, even when using identical seeds.

However, there are some steps you can take to limit the number of sources of nondeterministic behavior for a specific platform, device, and PyTorch release. First, you can control sources of randomness that can cause multiple executions of your application to behave differently. Second, you can configure PyTorch to avoid using nondeterministic algorithms for some operations, so that multiple calls to those operations, given the same inputs, will produce the same result.

### Reproducibility

- + [Controlling sources of randomness](#)
- + [Avoiding nondeterministic algorithms](#)
- [DataLoader](#)

# Source de variabilité liée au Hardware

TABLE IV  
CNN ACCURACY AND MEMORY SUBSYSTEM DYNAMIC ENERGY CONSUMPTION AT DIFFERENT VOLTAGE LEVELS. 850mV REPRESENTS THE MAXIMUM ACCURACY AT NOMINAL VOLTAGE.

	850 mV		750 mV		700 mV		650 mV		600 mV	
	Accuracy (%)	Energy ( $\mu$ J)								
FP_32	99.8	976.4	33.8	889.2	-	851.6	-	818.4	-	789.0
FXP_8_16	99.8	511.9	99.8	482.9	99.4	470.4	95.3	459.3	76.6	449.6
FXP_4_32	95.0	634.3	95.0	615.0	94.9	606.7	93.7	599.3	89.6	592.9
FXP_4_8	92.7	378.3	92.7	359.0	92.7	350.8	90.9	343.4	86.2	336.9

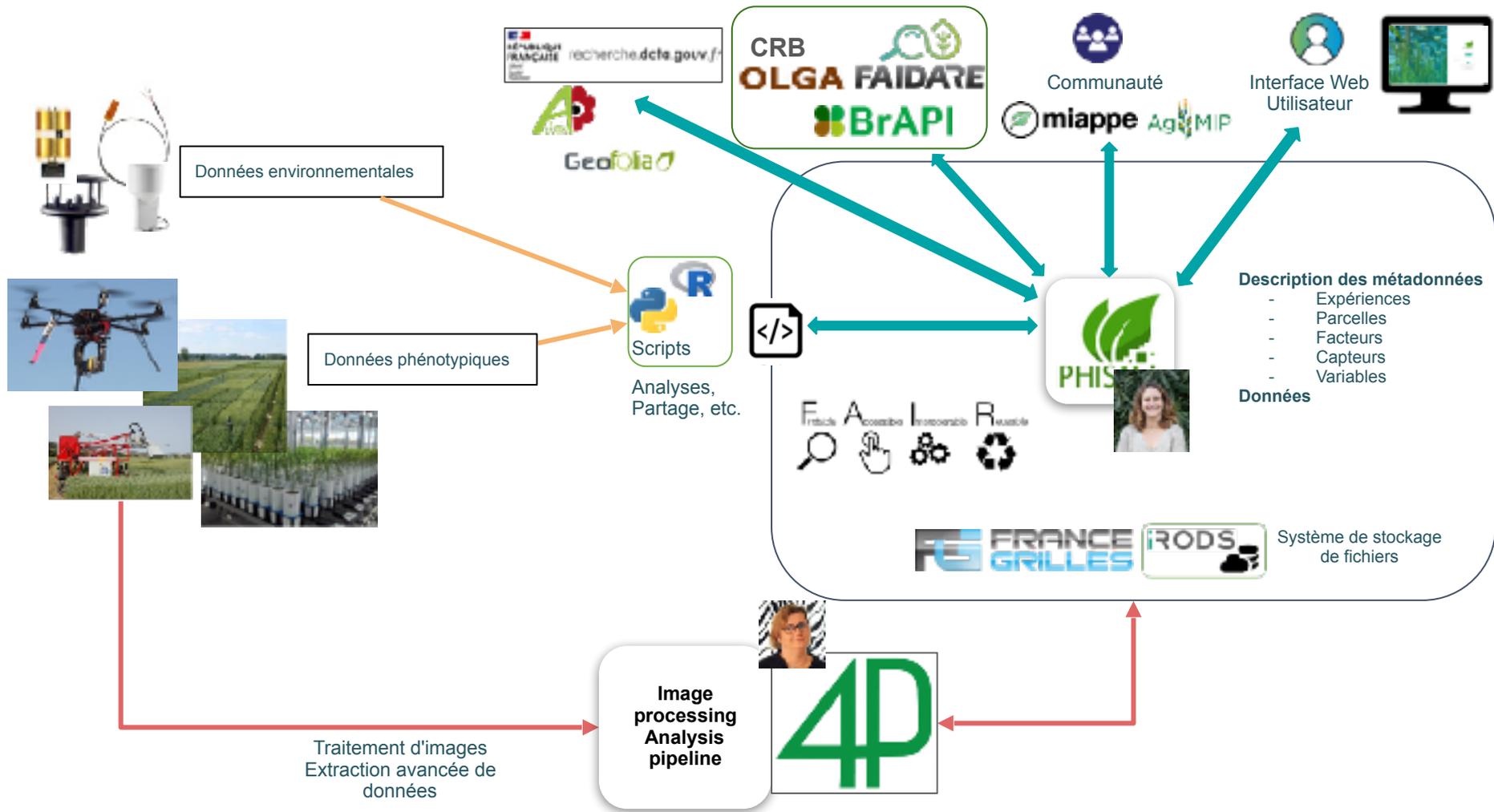
Denkinger, Benoît W., et al. "Impact of memory voltage scaling on accuracy and resilience of deep learning based edge devices." *IEEE Design & Test* 37.2 (2019): 84-92.

Source de variabilité lié au processus de création

# Révolution et reproduction



# Au sein de l'infrastructure PHENOME



# Comment sont produits ces outils ?

- Réponse A : des Geeks créent le modèle à partir d'un petit jeu de données

# Comment sont produits ces outils ?

- Réponse A : des Geeks créent le modèle à partir d'un petit jeu de données
- Réponse B : des Geeks + early adopters réalisent un prototype puis industrialisation par d'autres genres de Geeks

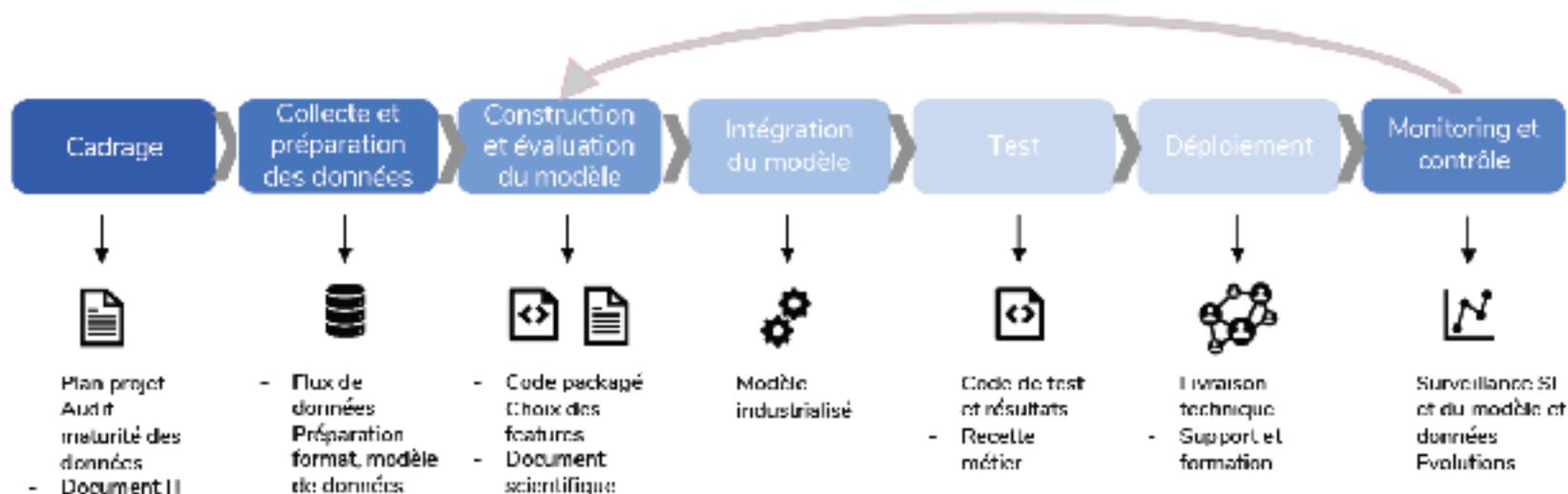
# Comment sont produits ces outils ?

- Réponse A : des Geeks créent le modèle à partir d'un petit jeu de données
- Réponse B : des Geeks + early adopters réalisent un prototype puis industrialisation par d'autres genres de Geeks
- Réponse C : des utilisateurs finaux définissent leurs besoins, des Geeks créent le produit final

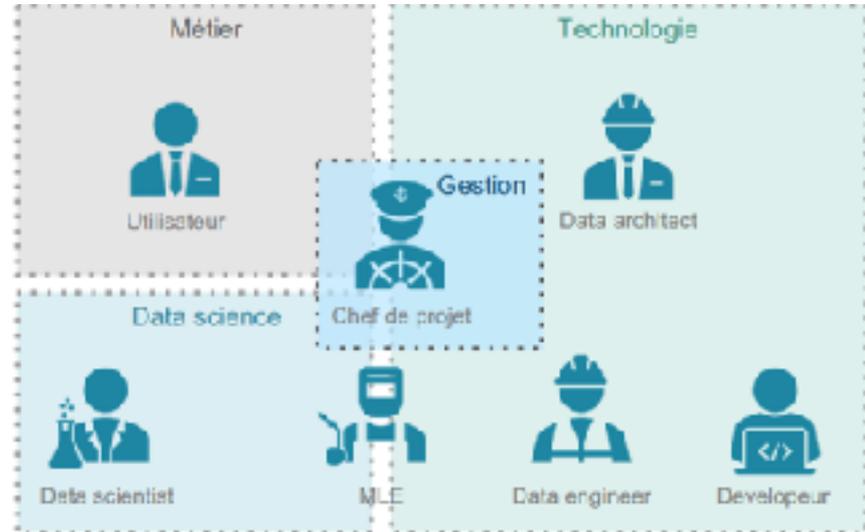
# Comment sont produits ces outils ?

- Réponse A : des Geeks créent le modèle à partir d'un petit jeu de données
- Réponse B : des Geeks + early adopters réalisent un prototype puis industrialisation par d'autres genres de Geeks
- Réponse C : des utilisateurs finaux définissent leurs besoins, des Geeks créent le produit final
- Réponse D : des utilisateurs finaux définissent leurs besoins, des Geeks créent un prototype qui s'améliore avec les contributions d'early adopters en lien avec des représentants des utilisateurs finaux, puis l'outil est industrialisé pour tous les utilisateurs par d'autres genres de Geeks.

# Un processus itératif...



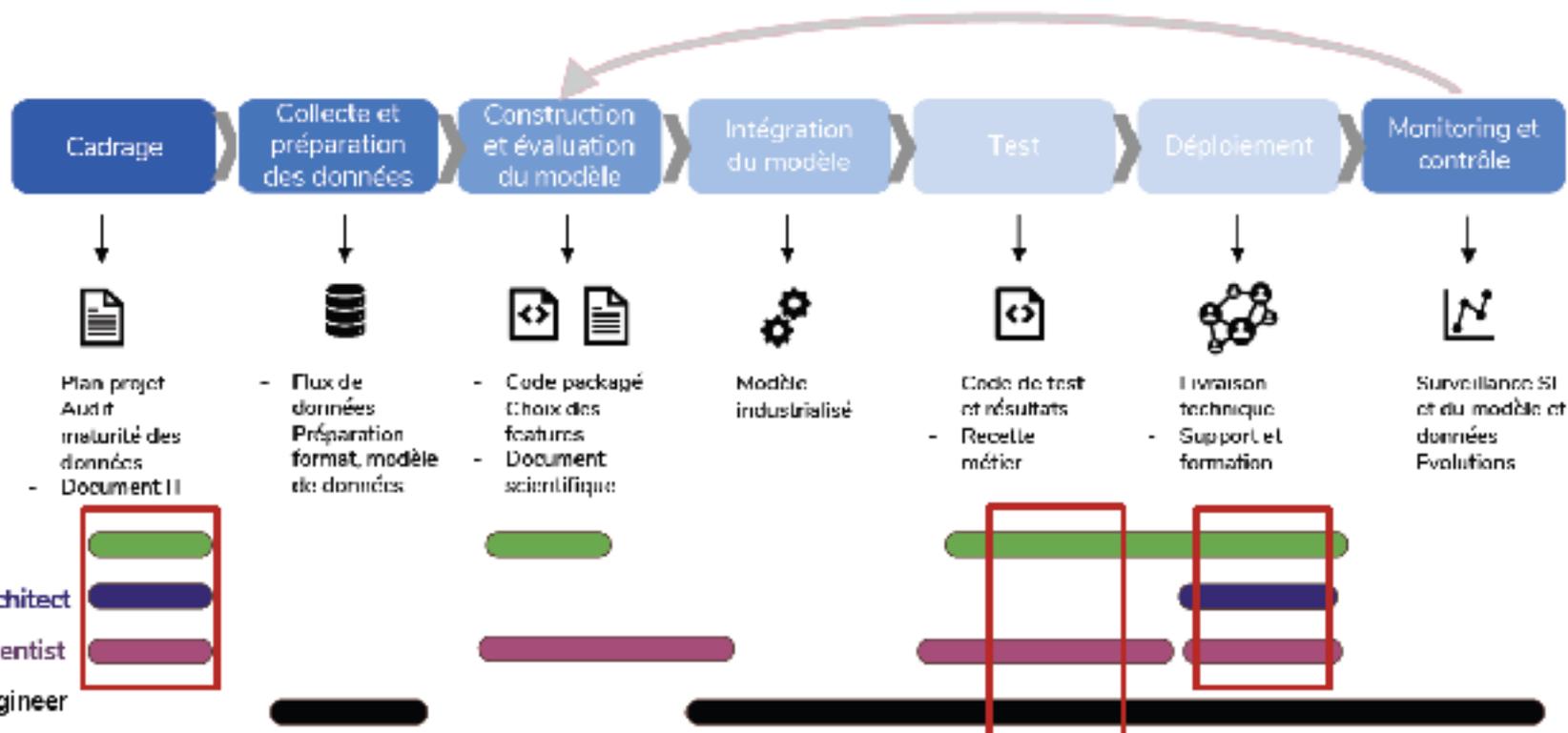
... avec de multiples rôles



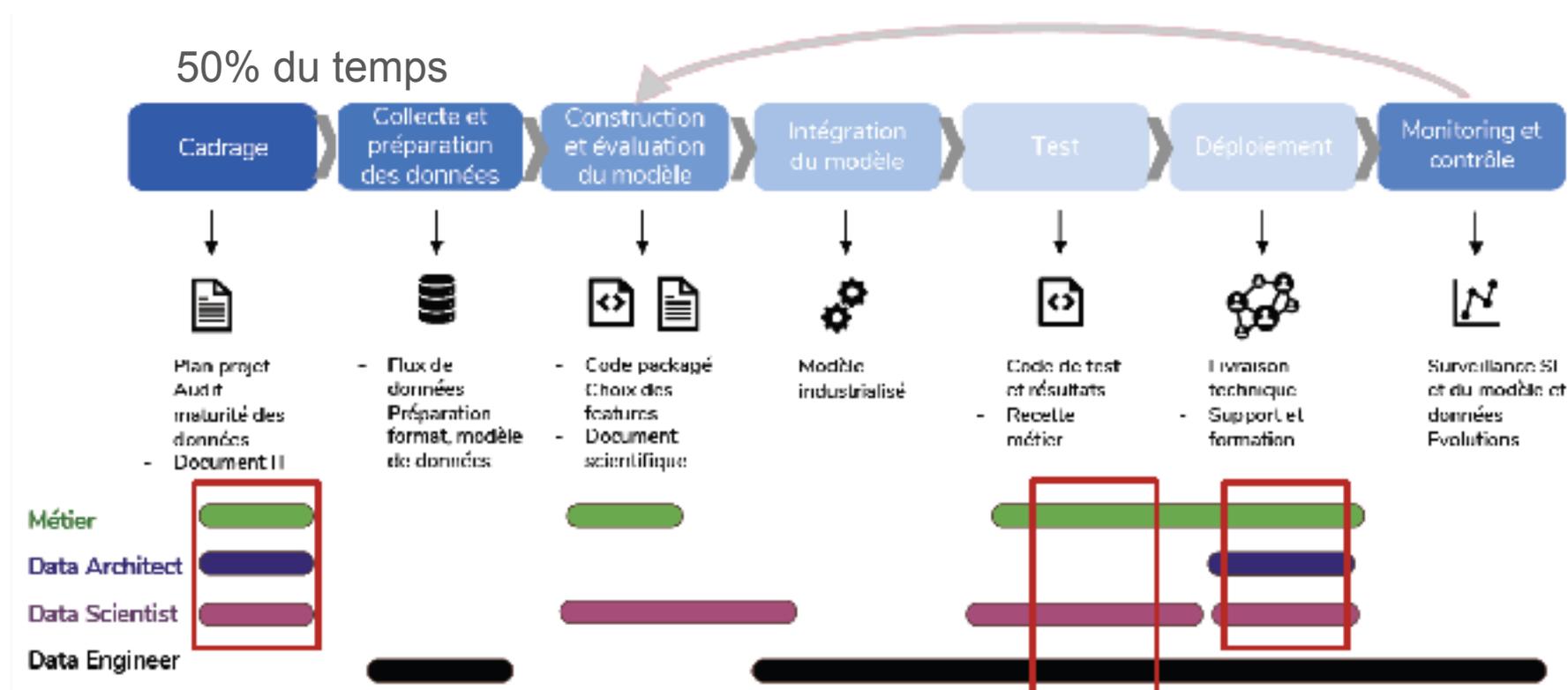
# Rôle de chacun

- Chef de projet data: Pilote les projets Data en respectant les objectifs et les contraintes.
- Data analyst : Exploite la donnée préparée pour mener des analyses et répondre à des problématiques métiers
- Data scientist: Construit des algorithmes pour répondre aux besoins des utilisateurs en extrayant de connaissances à partir de données structurées et non structurées.
- Data engineer: Construit et optimise les infrastructures et met en place un pipeline de données.
- Développeur : Conçoit, développe et déploie des applications spécialisées et assure l'expérience utilisateur par le débogage, la maintenance, la mise à niveau et le versionnage.
- Data architect : Conçoit et crée les infrastructures de données pour répondre aux besoins de l'entreprise.
- Utilisateur : définit son besoin et participe à l'évaluation itérative de l'outil

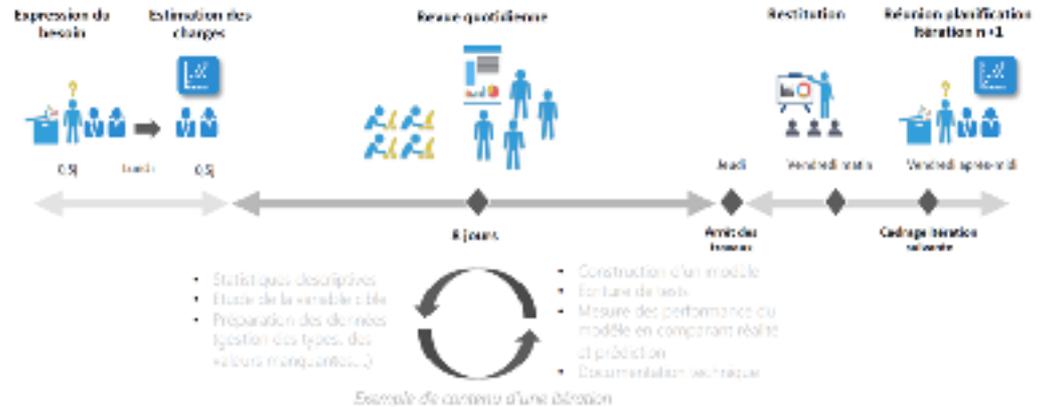
# Une contribution de chacun tout au long du processus



# Une contribution de chacun tout au long du processus

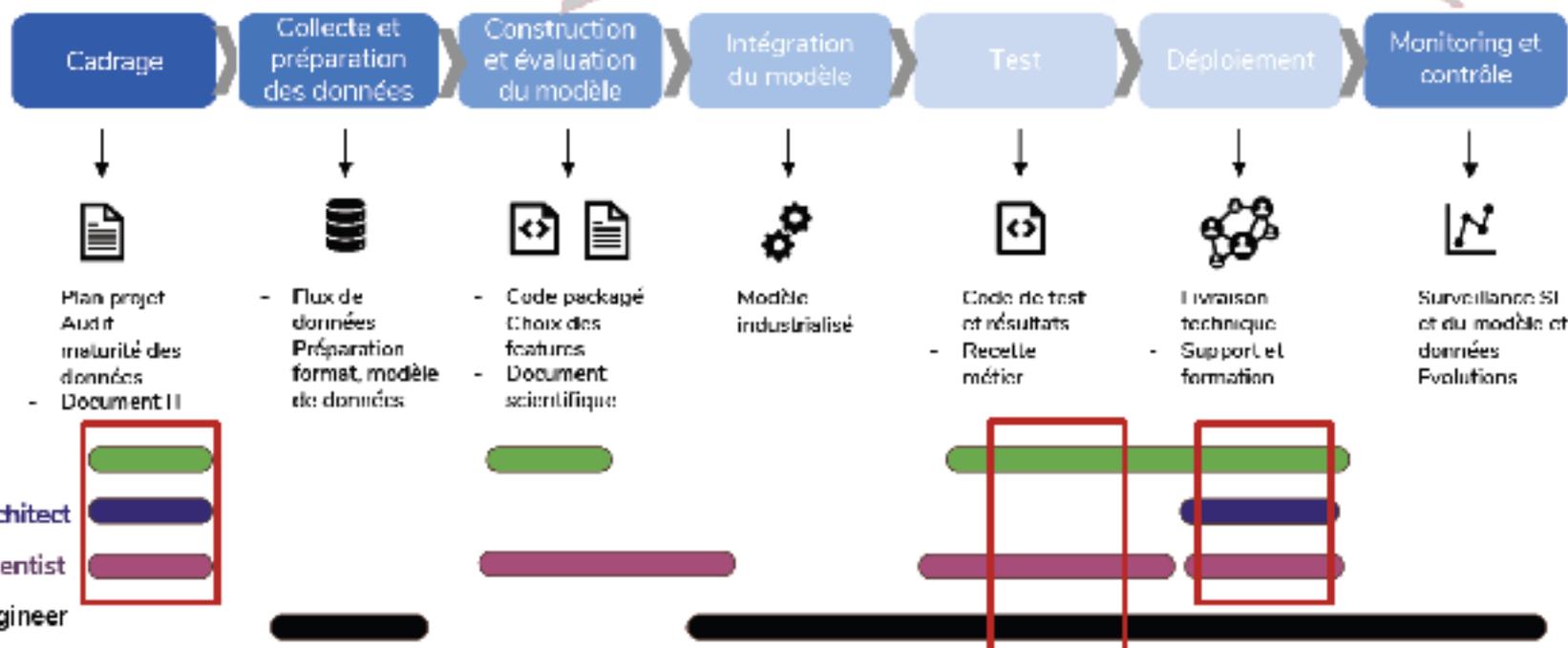


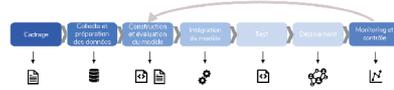
# Un oeil sur l'industrie... l'approche agile



# Explicitons ce schéma étape par étape

50% du temps





# Cadrage (métier, data scientist, data architecte)



## MÉTIER

- **Compréhension métier :**
  - Quel est le **contexte métier** ?
  - Quel(s) **objectif(s) stratégique(s)** le cas d'utilisation sert-il ?
  - Quel **périmètre métier** est concerné ?
- **Impact métier :**
  - Quel est l'impact du cas d'usage sur le **processus métier existant** ?
  - Quels sont les **KPIs de performance** à atteindre ?
  - Quelle est la **date de mise en production** prévue ?
- **Contraintes et standards métier**
  - Y a-t-il des **standards métier** à respecter ?
  - Y a-t-il des **contraintes métier** à suivre ?



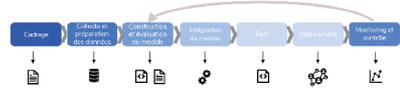
## FONCTIONNEL

- **Utilisateurs :**
  - Qui sont les **utilisateurs finaux** ?
  - Combien sont-ils ?
  - Quel est leur **environnement de travail** ?
- **Processus fonctionnel :**
  - Quel est le **périmètre fonctionnel** ?
- **Systèmes en adhérence :**
  - Quels sont les **systèmes en adhérence** à prendre en compte dans ce projet ?
- **Features, Fonctions & Services :**
  - Dans quelle **interface utilisateur** doit-on afficher l'output ?
  - Quelle est la **fréquence de mise à jour** ?



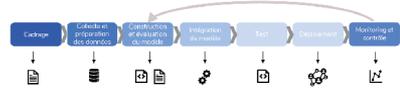
## DATA

- **Données d'entrée :**
  - Quelles sont les **sources de données** utilisés ?
  - Quel est le **volume de données** ? Combien d'**historique** ?
- **Reporting & Analyses :**
  - Y a-t-il un **besoin de reporting/BI** ?
- **Conformité & Sécurité :**
  - Des **données personnelles** sont-elles utilisées ?
- **Data Processing :**
  - Quel est le **niveau de qualité de données** estimé ? (besoin de preprocessing)



# Préparation de la donnée (data ing.)

- Collecte/Acquisition
- Séparation des données test et entraînement
- Exploration : évaluer la qualité des données
- Nettoyage : éliminer données inutiles ou de mauvaise qualité, prétraiter
- Transformation : pour permettre l'analyse ultérieure
- Méta données : Go FAIR at least with MIAPPE (voir talk d'Isabelle)



# Construction et évaluation du modèle (métier + data ing.)

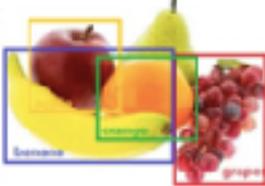
- Annotation des données d'entraînement

# L'annotation kézako ?

## Image Level

<b>Scene classification</b>	<b>Attributes</b>
	
Is this a field, a park, a restaurant?	Which leaf has disease?

## Object Level

<b>Object classification</b>	<b>Object detection</b>	<b>Classification + Localization</b>
		
Is it a tomato or an apple?	Where are the objects?	Where are the apples?

## Pixel Level

<b>Instance segmentation</b>	<b>Semantic segmentation</b>	<b>Object parts</b>
		
What is shown in every pixel? (Identify each entity)	Which pixels belong to the plants?	Where is the flowers? Where is the stem?

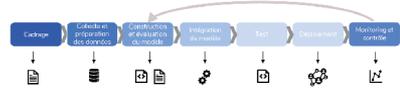
# L'annotation un processus en tant que tel



1 staff (Data scientist)  
2 to 10 hours  
to design model

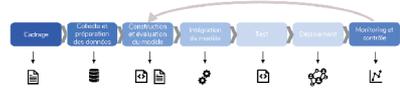
2 to 3 peoples  
5 to 15 hours  
for a first model

Crowd sourcing or  
outsourcing 50 to 100h  
for a strong model



# Construction et évaluation du modèle (métier + data ing.)

- Annotation des données d'entraînement
- Séparation des données en 2 (train & validation)



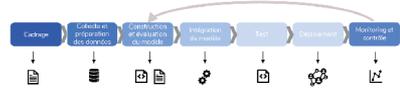
# Construction et évaluation du modèle (métier + data ing.)

- Annotation des données d'entraînement
- Séparation des données en 2 (train & validation)
- Entraînement d'un modèle sur ressource de calcul (cloud ou local)



[This image](#) by [GIFPublic\\_MH](#) is licensed under [CC-BY 3.0](#)





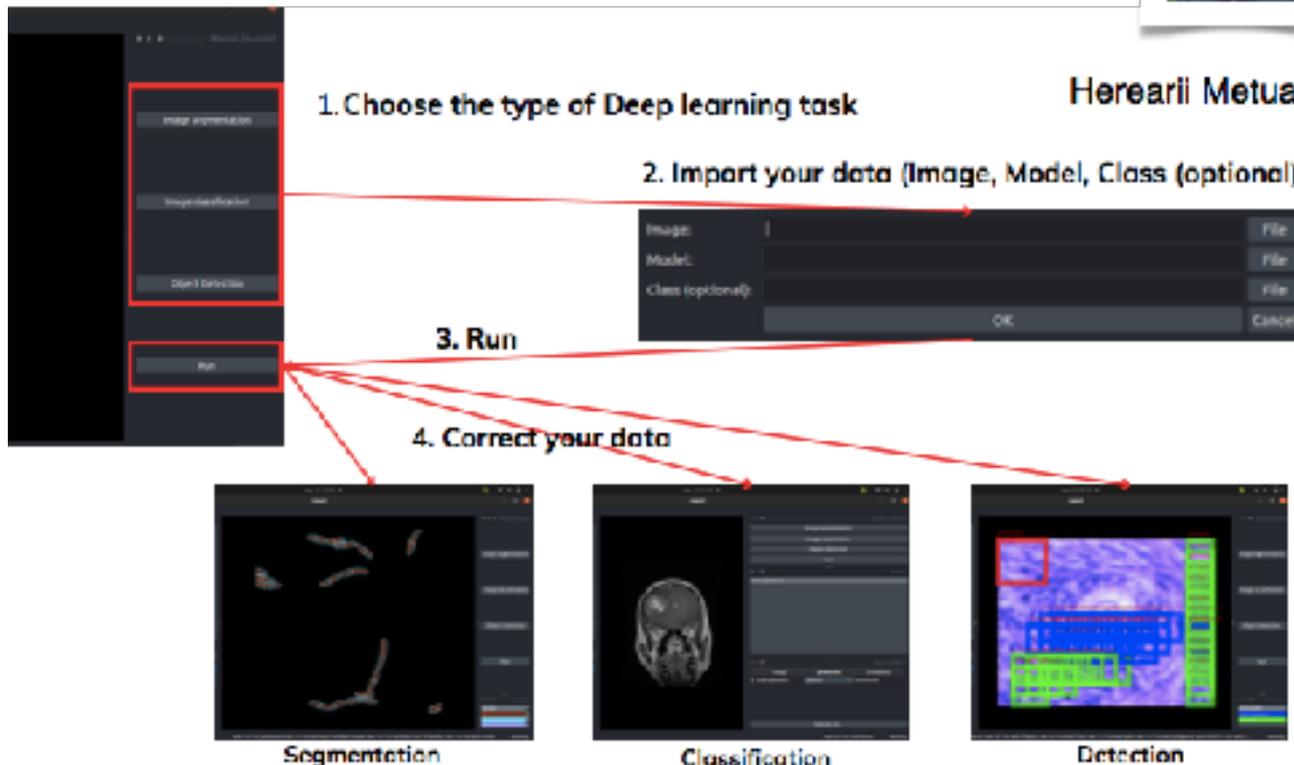
# Construction et évaluation du modèle (métier + data ing.)

- Annotation des données d'entraînement
- Séparation des données en 2 (train & validation)
- Entraînement d'un modèle sur ressource de calcul (cloud ou local)
- Evaluation du modèle sur données de validation et test
- POC partagée avec utilisateurs finaux via une interface minimale (par exemple NAPARI- MANINI dans phénomène)

# Manini Plugin Napari



Herearii Metuarea



# Téléchargement Manini



Home / Posts

## Manini MAchIne Inference & correction

An user-friendly plugin that enables to annotate images from a pre-trained model (segmentation, classification, detection) given by an user.

Download

Documentation

Description Activity [View](#)

Categories: [Machine Learning](#) [AI](#) [Software](#) [Tools](#) [Open Source](#) [Python](#) [Deep Learning](#) [Computer Vision](#) [Image Processing](#) [Data Science](#)

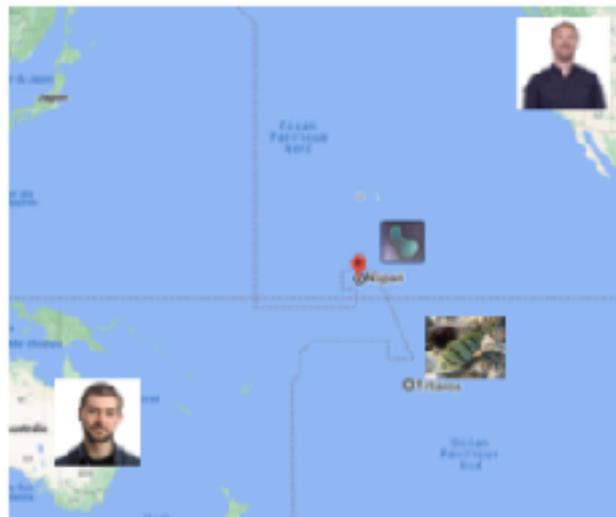
Manini (MAchIne Inference & Correction) is thought as a tool to boost the collaborative contribution of end-users to the assessment of deep learning model during their testing phase. It is a user-friendly plugin that enables to manually correct the result of an inference of deep learning model by an end-user. The plugin covers the following informational tasks: segmentation, classification and object detection.

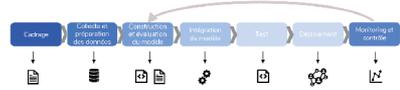
### White paper

Manini: From user to user. David Rousseau. Toward more collaborative deep learning project management in cloud



White paper  
Installation  
Description  
License  
Issues

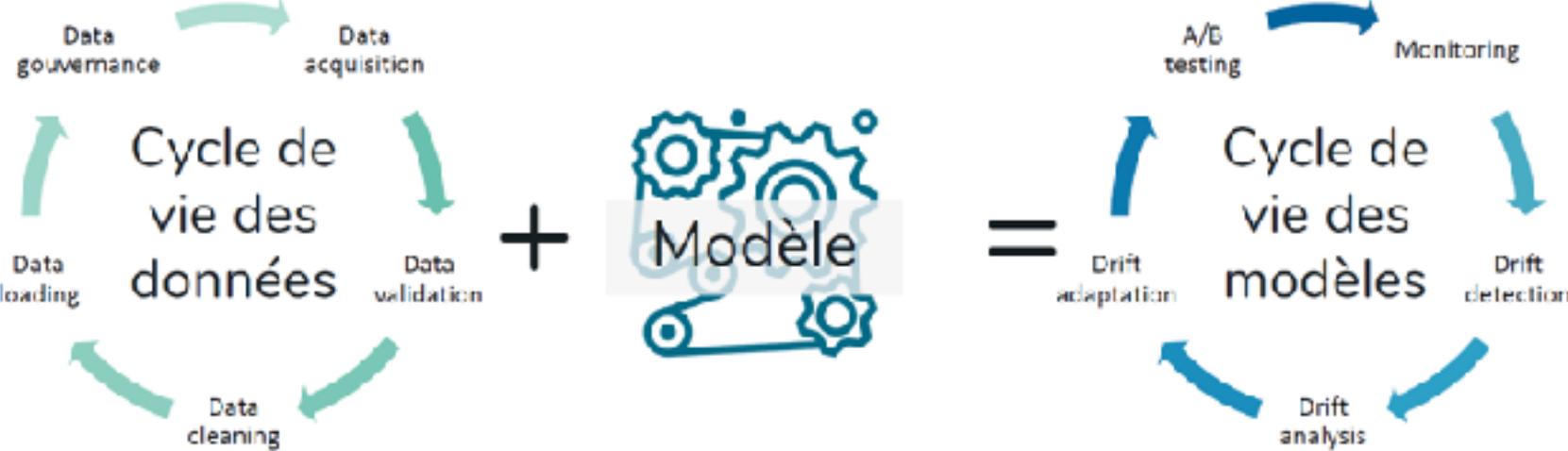


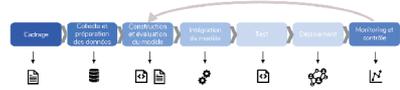


# Construction et évaluation du modèle (métier + data ing.)

- Annotation des données d'entraînement
- Séparation des données en 2 (train & validation)
- Entraînement d'un modèle sur ressource de calcul (cloud ou local)
- Evaluation du modèle sur données de validation et test
- POC partagé avec utilisateurs finaux via une interface minimale (par exemple NAPARI- MANINI dans phénomène)
- Synthèse des performances qualitatives et quantitatives
- Evaluation de quantité et qualité nécessaire de données supplémentaires si les performances ne sont pas encore satisfaisantes

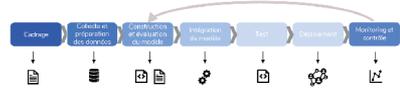
# Cycle des modèles





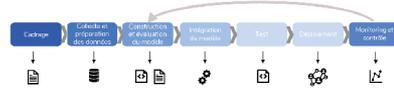
# Intégration du modèle (data ing.)

- Export du modèle dans un format générique via un dépôt mutualisé (exemple MLFlow)
- Portage du modèle dans un Docker
- Installation dans un environnement exécutable sur Cloud avec interface utilisateur définie lors du cadrage (sous 4P dans Phénome)



# Test du modèle (métier, data scientist, data ingé.)

- Test sur de gros volumes de données non annotées
- Analyse des performances et retour utilisateur



# Déploiement (métier, data scientist, data ing., data architecte)

- Ouverture de l'accessibilité à tous
- Supports de formation (tutoriel, fiche d'utilisation, exemple de données jouets permettant d'évaluer les limites de l'algorithme)

**BioImage.IO** — Upload — Documentation — About

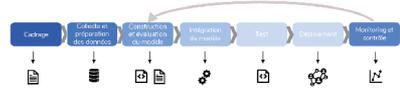
## BioImage Model Zoo

Advanced AI models in one-click

- Integrated with Fiji, Ilastik, ImJoy and more
- Try model instantly with BioEngine
- Contribute your models via GitHub
- Link models to datasets and applications

[Explore the Zoo](#)

Community Partners



# Monitoring et contrôle (Data ing.)

- Contrôle qualité des données en entrée et en sortie

# Spécificités liées à la gestion de projets Data IA

- Spécificités méthodologiques : une approche progressive et itérative
- Spécificité organisationnelle : de la nécessité de mobiliser des acteurs variés et nouveaux
- Spécificités structurelles et réglementaires - Gouvernance des données
- Spécificités technologiques : un écosystème en constante évolution
- Spécificité juridique et éthique : pour une utilisation maîtrisée de la donnée

# Conclusion générale sur IA et reproductibilité

- Le risque de biais lié aux données d'entraînement
- L'incertitude liée aux aspects aléatoires des algorithmes, librairies

# Conclusion générale sur IA et reproductibilité

- Le risque de biais lié aux données d'entraînement
- L'incertitude liée aux aspects aléatoires des algorithmes, librairies
- La consommation d'énergie



L'IA permet de combattre le réchauffement climatique en :

- Détectant plus tôt les zones de dérives climatiques
- Optimisant la gestion des ressources énergétiques renouvelables
- Prévoyant l'amplitude des catastrophes naturelles

Impact positif en allant chercher des cas d'usages environnementaux



L'IA accélère le réchauffement climatique en :

- S'appuyant sur des sources de données toujours plus volumineuses
- S'appuyant sur des calculs de plus en plus lourds
- Etant déployée sur de plus en plus d'appareils

Impact négatif si on ne prend pas conscience de la dette environnementale créée par l'IA en général.

# Check list des bonnes pratiques ?

## The Machine Learning Reproducibility Checklist (v0.0, Apr 7 2020)

For all models and algorithms presented, check if you include:

- A clear description of the mathematical setting, algorithm, and/or model.
- A clear explanation of any assumptions.
- An analysis of the complexity (time, space, sample size) of any algorithms.

For any theoretical claim, check if you include:

- A clear statement of the claim.
- A complete proof of the claim.

For all datasets used, check if you include:

- The relevant statistics, such as number of examples.
- The details of train / validation / test splits.
- An explanation of any data that were excluded, and all pre-processing step.
- A link to a downloadable version of the dataset or simulation environment.
- For new data collected, a complete description of the data collection process, such as instructions to annotators and methods for quality control.

For all shared code related to this work, check if you include:

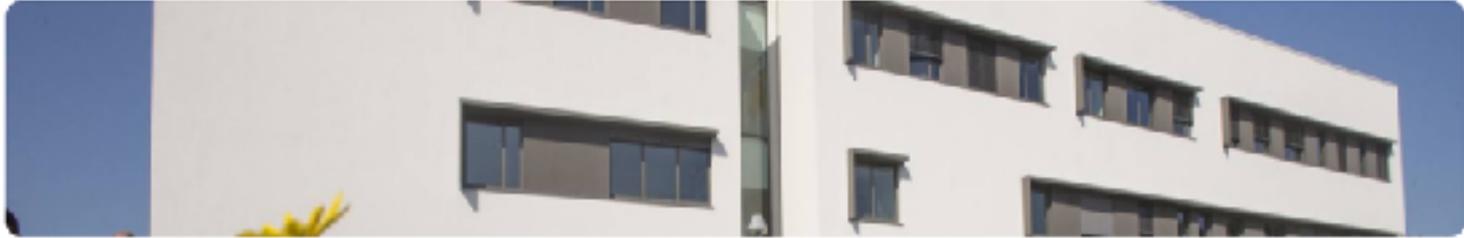
- Specifications of dependencies.
- Training code.
- Evaluation code.
- (Pre-)trained model(s).
- README file includes table of results accompanied by precise command to run to produce these results.

Utilisé par la conférence NEURIPS  
depuis 2021 pour le process de review

# Pour aller plus loin



Rechercher



## ImHorPhen Bio imaging research group

@imhorphenbioimagingresearch95 407 abonnés 205 vidéos

ImHorPhen is a bioimaging, research group, headed by Prof. David Rousseau, located in An... >

[www.universite-nantes.fr/irs/Recherche/Imagerie-pour-Horticulture-et-Pl...](http://www.universite-nantes.fr/irs/Recherche/Imagerie-pour-Horticulture-et-Pl...) et 1 autre lien



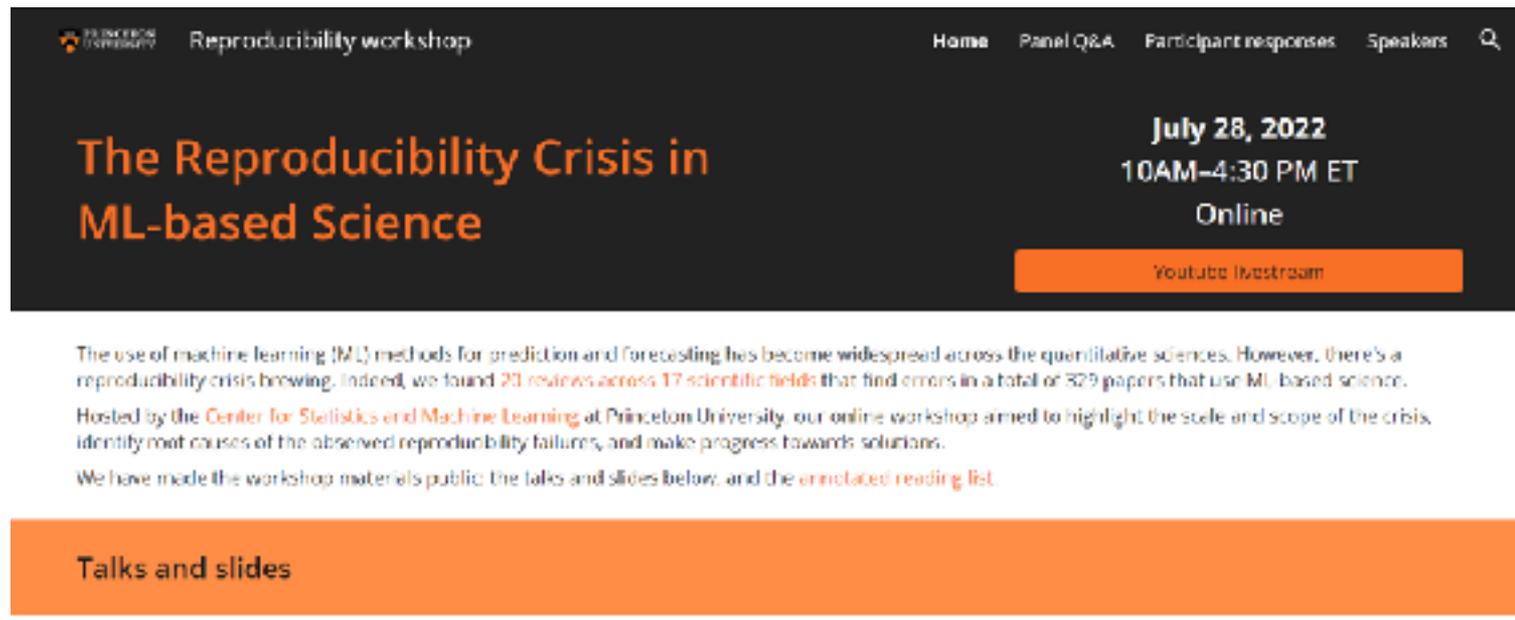
Abonné

... par exemple comment rédiger un matériel et méthode incluant de l'IA

<https://www.youtube.com/watch?v=hbaKWqj8MZ8>

# Biblio point de départ

<https://reproducible.cs.princeton.edu/>



The screenshot shows the top section of a website for a workshop. At the top left is the Princeton University logo and the text 'Reproducibility workshop'. To the right are navigation links: 'Home', 'Panel Q&A', 'Participant responses', and 'Speakers', followed by a search icon. The main content area has a dark background with orange text. The title 'The Reproducibility Crisis in ML-based Science' is on the left. On the right, the date 'July 28, 2022' and time '10AM-4:30 PM ET' are listed, with 'Online' below. An orange button labeled 'Youtube livestream' is positioned below the time. Below this dark section, on a white background, is a paragraph of text: 'The use of machine learning (ML) methods for prediction and forecasting has become widespread across the quantitative sciences. However, there's a reproducibility crisis brewing. Indeed, we found 20 reviews across 17 scientific fields that find errors in a total of 329 papers that use ML-based science. Hosted by the Center for Statistics and Machine Learning at Princeton University, our online workshop aimed to highlight the scale and scope of the crisis, identify root causes of the observed reproducibility failures, and make progress towards solutions. We have made the workshop materials public: the talks and slides below, and the annotated reading list.' At the bottom of the screenshot is an orange bar with the text 'Talks and slides' in white.

PRINCETON Reproducibility workshop

Home Panel Q&A Participant responses Speakers

**The Reproducibility Crisis in ML-based Science**

July 28, 2022  
10AM-4:30 PM ET  
Online

Youtube livestream

The use of machine learning (ML) methods for prediction and forecasting has become widespread across the quantitative sciences. However, there's a reproducibility crisis brewing. Indeed, we found 20 reviews across 17 scientific fields that find errors in a total of 329 papers that use ML-based science. Hosted by the Center for Statistics and Machine Learning at Princeton University, our online workshop aimed to highlight the scale and scope of the crisis, identify root causes of the observed reproducibility failures, and make progress towards solutions. We have made the workshop materials public: the talks and slides below, and the annotated reading list.

Talks and slides

**Merci pour votre attention**

[david.rousseau@univ-angers.fr](mailto:david.rousseau@univ-angers.fr)

