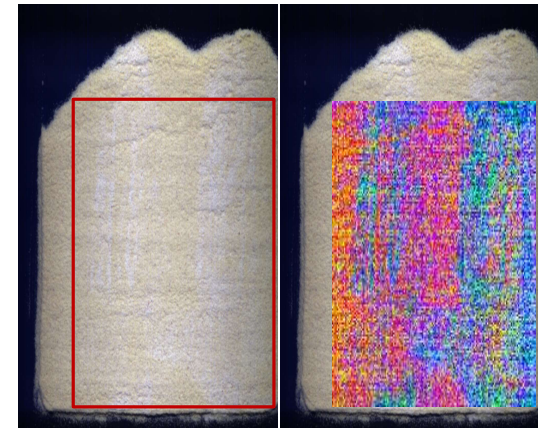


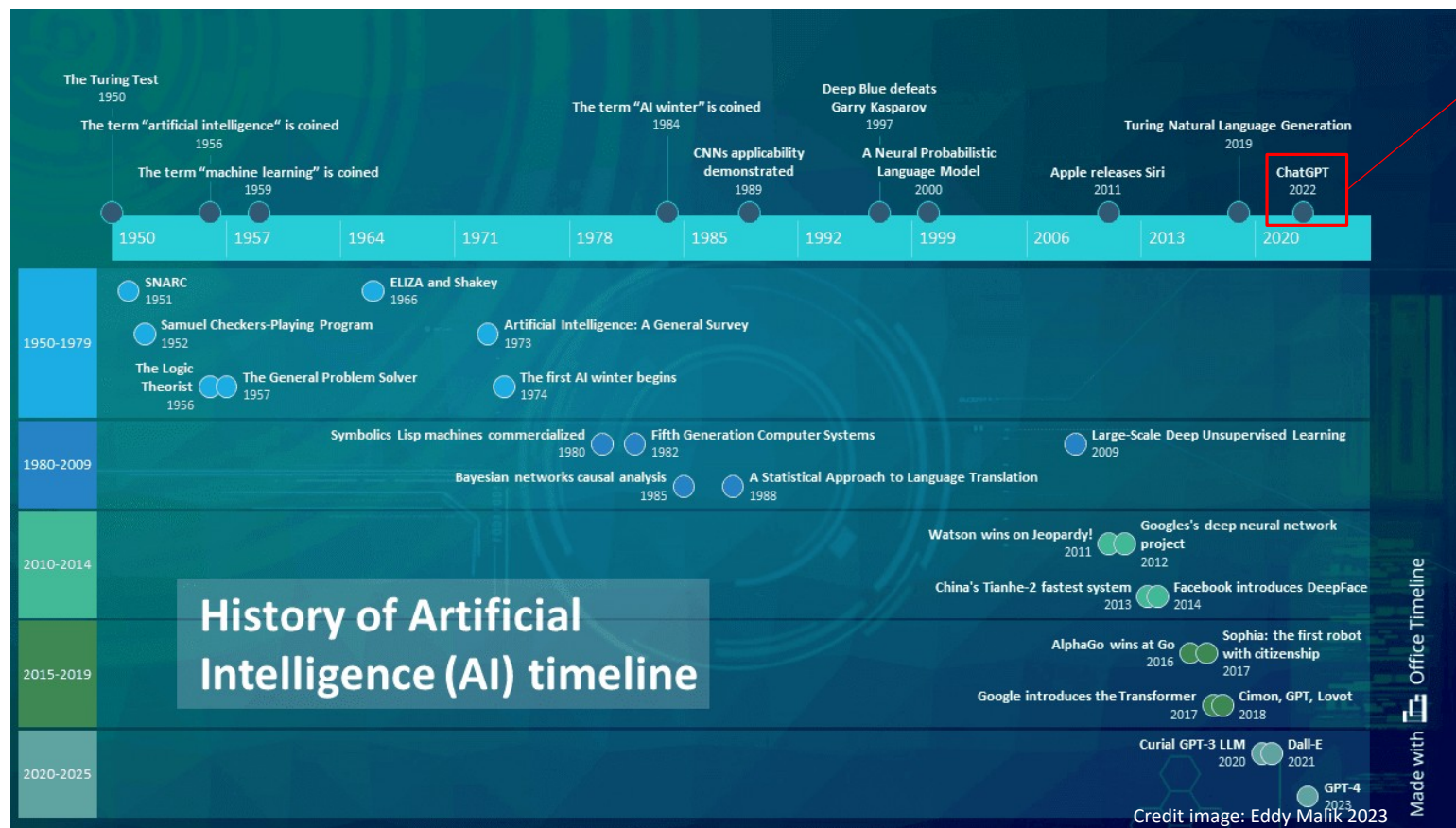
L'Intelligence Artificielle au service de la recherche agronomique



La place de l'Intelligence Artificielle (IA) dans notre travail de conseil et d'encadrement
Namur, 13/10/2025

Antoine Deryck
Chercheur et doctorant au CRA-W

Contexte



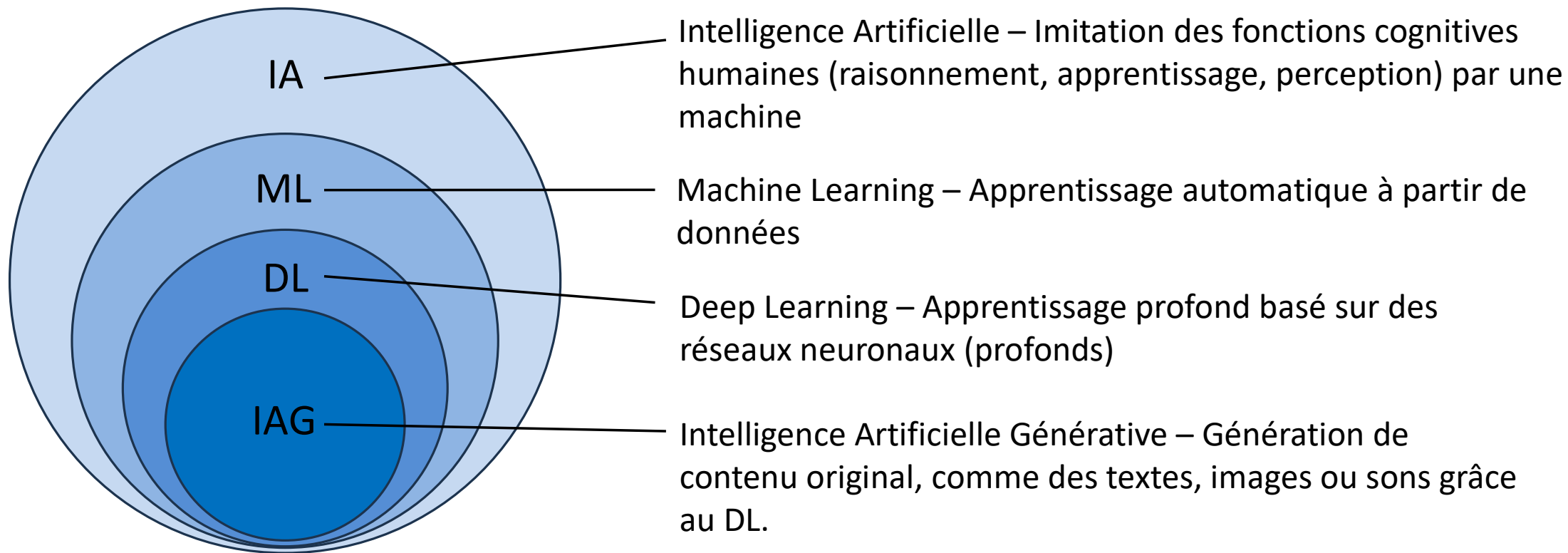
30 novembre 2022 –
Lancement ChatGPT

Gain de popularité,
l'IA devient connue
du grand public

Pourtant, IA existe
depuis bien
longtemps

Pourquoi ce gain ?

Qu'est-ce que l'intelligence artificielle ?



ChatGPT = 1ère interface permettant d'interagir avec une IAG qui a été rendue accessible au grand public

L'IA au CRA-W

Deux niveaux d'actions

Développement d'algorithmes pour résoudre diverses problématiques

- *Comptage automatique*
- *Qualispectra*
- *InnovEau*
- *Sagriwasent/Belcam*
- *Cartofoor*

...

Utilisation des IAG pour soutenir les activités de recherche

- *Recherche d'outils IAG pertinents*
- *Vérification de bases de données*

...

Comptage automatique de larves

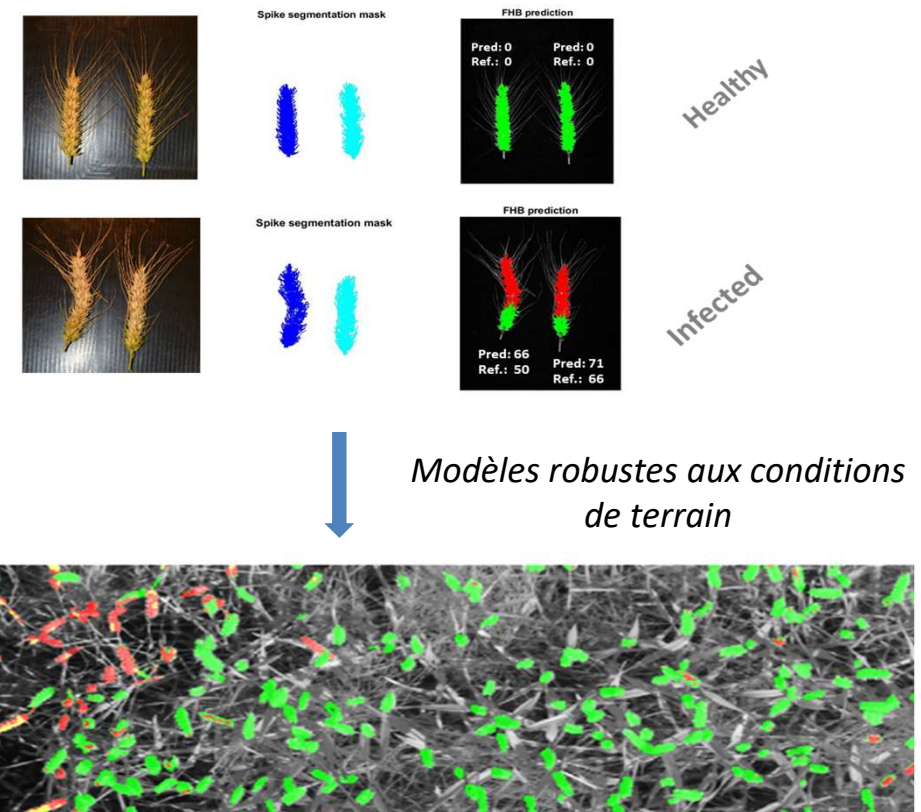
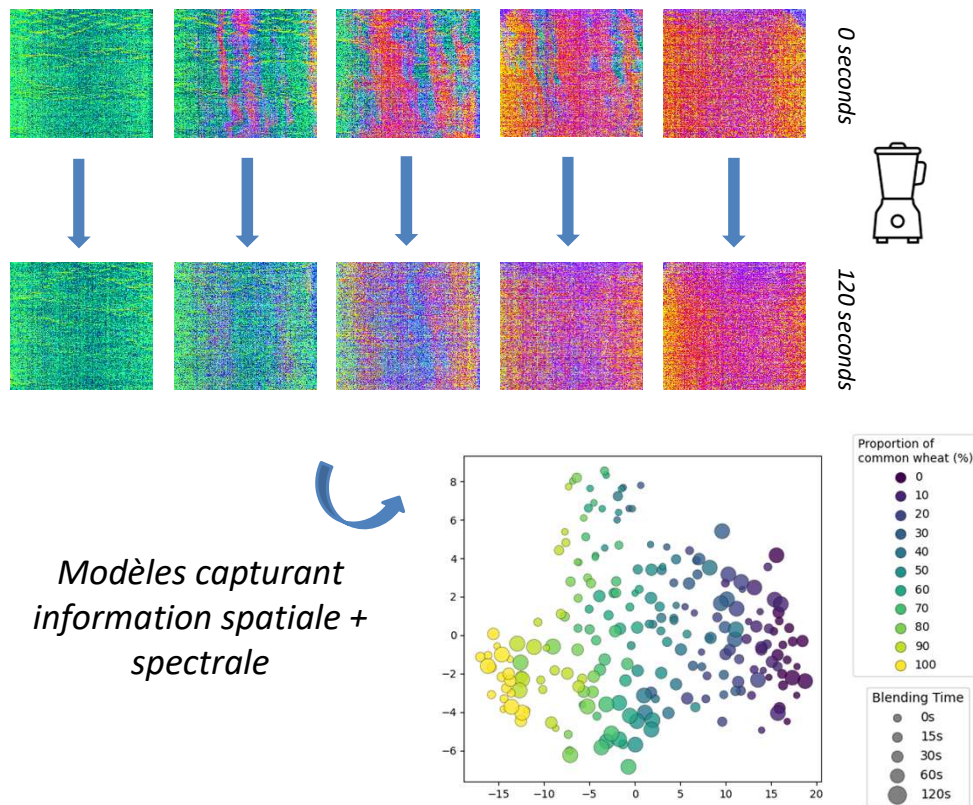
- Contexte: Etude du CRA-W sur *Sitodipolis mosellana* et *Contarinia tritici* (Cecidomyiidae), ravageurs du froment, blé dur, seigle et triticales



- Développement d'une plateforme simple d'utilisation permettant de compter le nombre de larves en quelques secondes (implémentée rapidement grâce aux IAG)

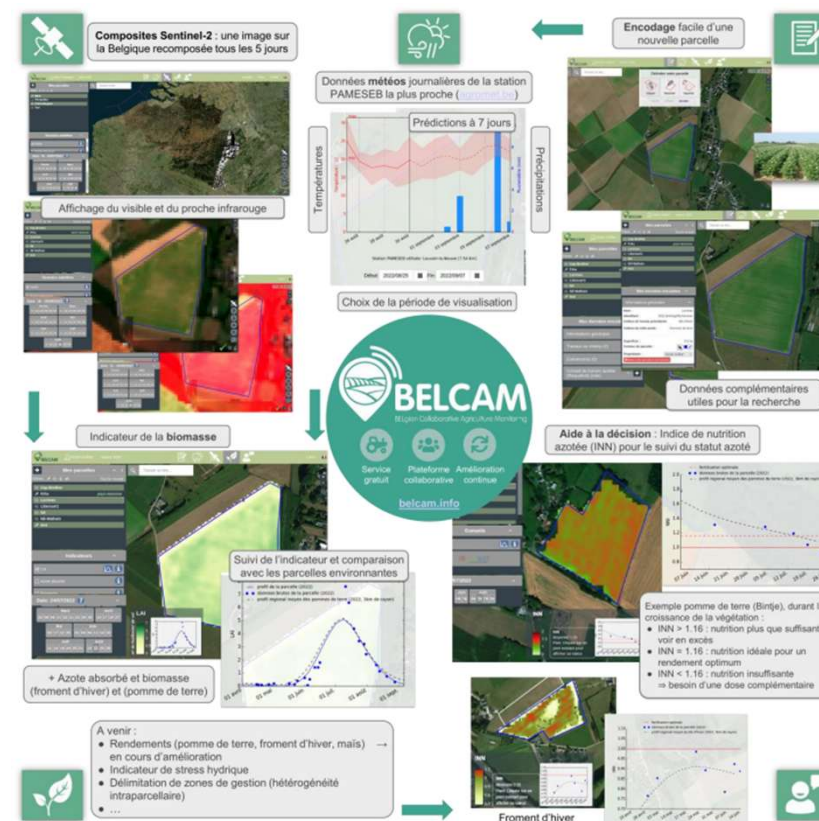
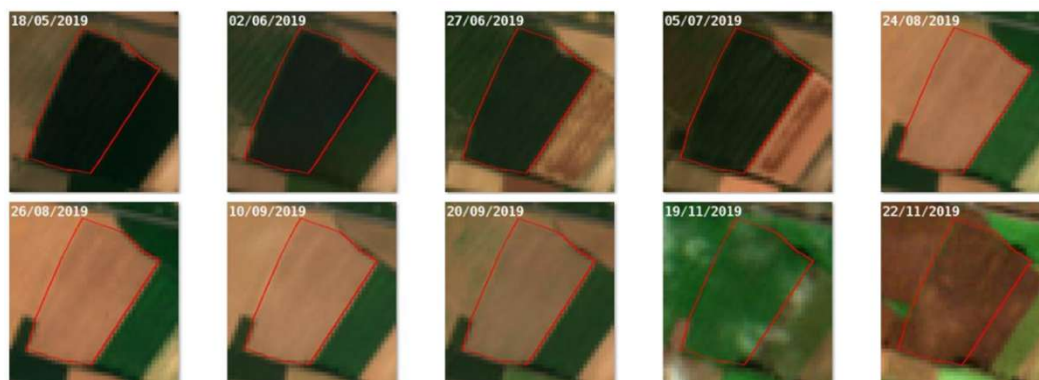
Qualispectra

- Traitement d'images hyperspectrales grâce à des algorithmes de Deep Learning

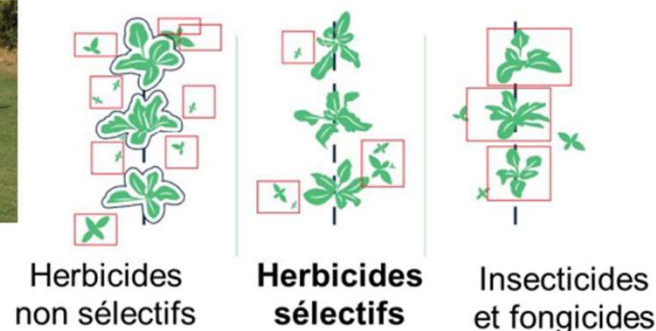


Sagriwasent II / Belcam

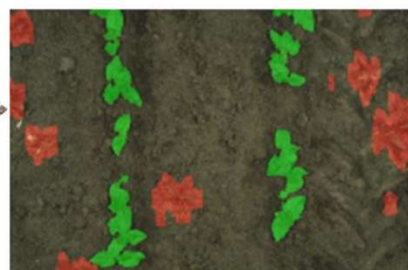
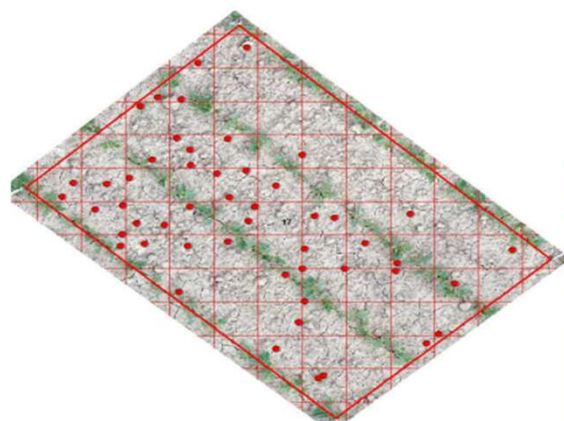
- Suivi continu de l'agriculture Wallonne à l'aide d'outils innovants incluant télédétection et nouvelles technologies



- Réduction de l'utilisation des herbicides grâce à l'agriculture de précision
- Détection automatique des adventices via réseaux de neurones + "spot-spraying"

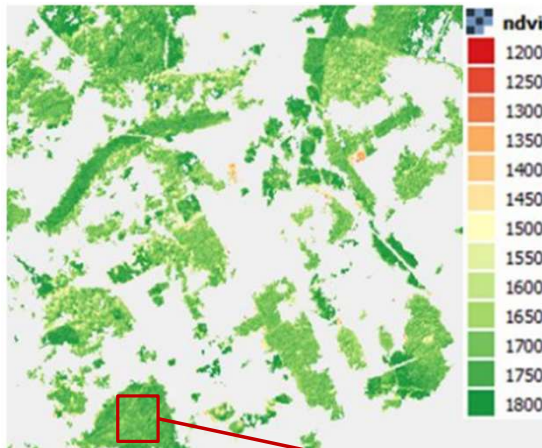


- Gains théoriques importants
- Efficacité et détection variable, relativement bas pour des capteurs embarqués (13% - 62% d'adventices détectées)
- Coûts élevés
- Prometteur, développements nécessaires pour entraîner IA (autres cultures et adventices)



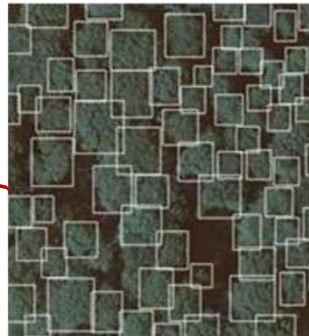
Cartofor

- Utilisation d'images à haute résolution spatiale pour le suivi de l'état de santé de forêts
- Identification des houppiers via réseaux de neurones

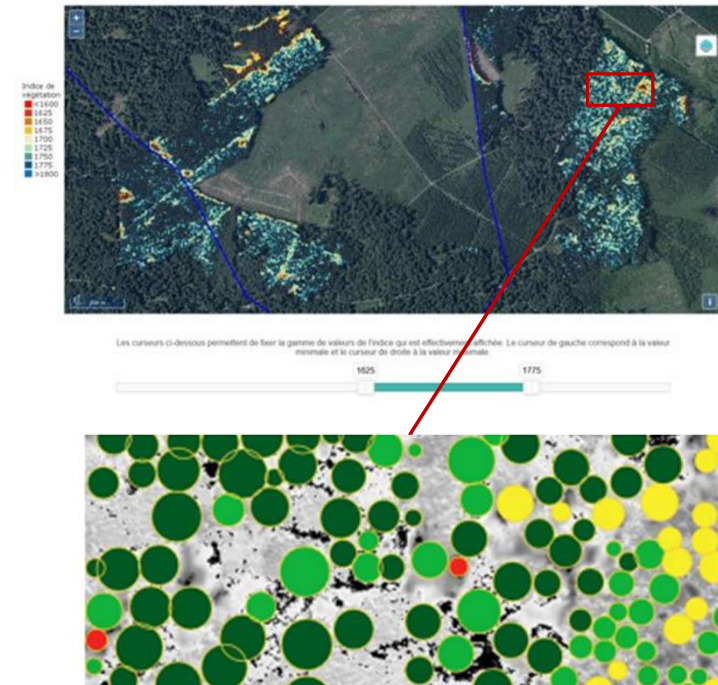


Calcul du NDVI

Identification des
houppiers par CNN



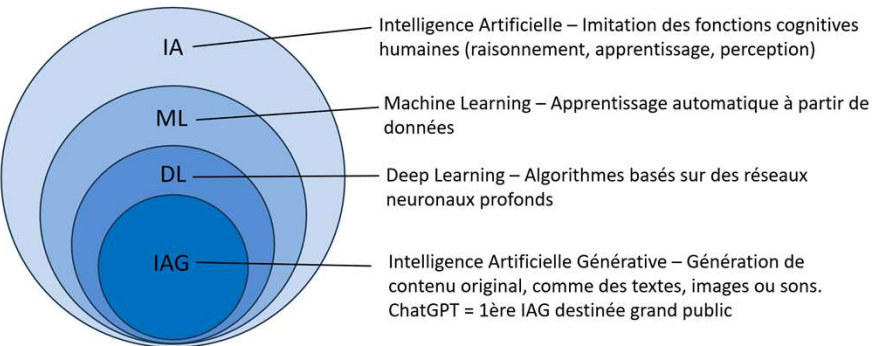
Suivi NDVI à l'échelle de l'arbre et de la parcelle



Comité Intelligence Artificielle Générative

- Comment utiliser les outils d'Intelligence Artificielle Générative dans la recherche ?

Qu'est-ce que l'intelligence artificielle ?



Guidelines

Travailler au CRA-W

Informatique
Formulaires informatiques
Logiciels
Matériels et téléphonie
Intelligence artificielle



1

Transparence

Transparence
de l'utilisation de l'IA, lors
de la création de contenu

Cela implique l'obligation de signaler explicitement lorsque le contenu a été généré ou assisté par IA, afin de ne pas induire le public en erreur quant à la nature humaine ou artificielle de la production de contenu.

Il est conseillé d'être aussi précis que possible quant à la part du travail réalisée par l'IA.

Cette pratique vise à préserver la confiance et l'intégrité de nos communications et à garantir que le public reste informé de la manière dont les technologies émergentes sont utilisées à son service.

3

Vérification

Vérification
de la justesse des résultats
générés

Les employés doivent s'assurer que les informations, les données ou tout autre contenu généré par l'IA, sont exacts, à jour et issus de sources fiables. Avant de publier ou d'utiliser le contenu généré, une revue critique doit être effectuée pour valider sa pertinence et sa validité.

Cela peut impliquer de croiser les informations avec des sources authentiques ou de consulter des experts dans le domaine concerné.

Cette étape est cruciale pour maintenir la confiance du public dans l'intégrité et la qualité de l'information que nous diffusons.

2

Responsabilité

Responsabilité
du contenu créé par l'IA
générative.

Cela signifie que chaque employé utilisant de telles technologies doit s'assurer que le contenu généré est conforme à nos standards éthiques, légaux et professionnels. En cas de diffusion d'informations inexactes, trompeuses ou préjudiciables, il incombe à l'employé de corriger ou de retirer le contenu et d'assumer les conséquences de ces erreurs.

La responsabilité inclut également l'usage prudent de l'IA générative pour ne pas compromettre la qualité et la fiabilité de nos services.

4

Respect

Respect
des données confidentielles
et du RGPD

Les employés doivent veiller à ne pas utiliser de données sensibles ou protégées sans l'autorisation explicite des personnes concernées ou sans respecter les cadres légaux applicables.

Cela inclut le respect de la propriété intellectuelle et le fait de s'abstenir d'utiliser des œuvres protégées par le droit d'auteur sans licence appropriée.

La protection de la vie privée et des droits d'auteur est non seulement une obligation légale mais aussi une question de respect et d'intégrité professionnelle.

OPINION

**ChatGPT for Research and Publication:
A Step-by-Step Guide**

Son S. Brown, MD, MS, DM, DMSc, MBA
DOI: 10.5833/1531-6719-28-6-576

Introduction

This commentary provides a concise step-by-step guide on using ChatGPT, an advanced natural language processing (NLP) model, for research and publication purposes. The guide assesses crucial aspects, including data preprocessing, fine-tuning techniques, prompt engineering, and ethical considerations. By addressing challenges related to biases, interpretability, and plagiarism, this commentary offers insights and recommendations for the responsible and ethical use of ChatGPT. The guide empowers researchers to ethically integrate ChatGPT effectively into their workflows, enhancing productivity and improving the quality of their scientific publications. Through clear instructions and guidelines, researchers can tap into the transformative potential of ChatGPT, driving scientific progress in a language-driven world.

In recent years, the field of NLP has witnessed remarkable advancements, bringing us closer to the realm of human-like language generation. Among these advancements, ChatGPT, based on the groundbreaking GPT-3.5 architecture developed by OpenAI, stands as an impressive language model capable of generating coherent and contextually relevant text responses. With its ability to understand and respond to user inputs, ChatGPT has opened up new possibilities for various applications, including research and publication.

The traditional process of conducting research and publishing scientific papers has been significantly transformed by the emergence of ChatGPT. Researchers and authors can now leverage the power of this sophisticated language model to streamline and enhance their workflow, leading to improved efficiency and higher-quality publications. However, using ChatGPT effectively in the research and publication domain requires a comprehensive understanding of its capabilities, limitations, and best practices.

In this commentary, I offer my thoughts for a step-by-step guide for researchers and authors who seek to harness the potential of ChatGPT in their research endeavors and publication efforts. By exploring various aspects, such as data preprocessing, fine-tuning techniques, prompt engineering, and ethical considerations, this guide will equip researchers with the necessary knowledge to harness the full potential of ChatGPT in their scientific pursuits. Moreover, this commentary will delve into the challenges associated with using ChatGPT for research and publication, including biases, interpretability, and concerns regarding plagiarism. By addressing these challenges directly, I aim to provide researchers with valuable insights and recommendations to navigate these important issues and ensure the responsible and ethical use of ChatGPT in a research tool.

The significance of my guide lies in its potential to bridge the gap between the rapid progress of language models like ChatGPT and the research and publication process. By elucidating the intricacies of integrating ChatGPT into scientific workflows, researchers will be empowered to leverage this advanced technology effectively, thereby enhancing the overall quality and impact of their research output.

In the following sections, I present a comprehensive overview of the steps involved in using ChatGPT for research and publication.

Step 1: Title and Title Page Creation by ChatGPT

ChatGPT can be a valuable tool in generating titles for research papers. Its ability to understand and generate human-like text allows it to analyze and synthesize information provided by researchers to craft concise and impactful titles. By leveraging its vast knowledge base and language capabilities, ChatGPT can assist in capturing the essence of a research paper, conveying the main focus and contributions succinctly. Researchers can collaborate with ChatGPT by providing relevant information, such as the subject, objectives, methodology, and key findings of their study. ChatGPT can then generate multiple title options, offering different perspectives and angles that researchers can consider. This collaboration with ChatGPT can save time and stimulate creativity, helping researchers refine their titles to accurately represent their work and engage potential readers. ChatGPT can then be used to create the entire title page and then can also customize based on each journal's recommendations.

FORMATION / WORKSHOP

Maîtrisez l'IA générative et transformez votre organisation – Mars 2025

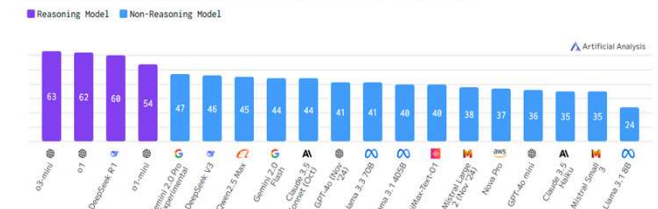
Programmes liés : DigitalWakons4.ai

Intelligence artificielle (IA/IA) Secteur du numérique Sensibilisation et formation voir plus

Cet atelier pratique a pour objectif de familiariser les participants avec l'utilisation efficace des dernières versions d'intelligences artificielles génératives grand public (Generative AI).

Artificial Analysis Intelligence Index by Type

Artificial Analysis Intelligence Index (Version 2, released Feb '25) by Type: Higher is better



Objectifs: Veille active des nouveaux outils, encadrer et proposer l'utilisation, trouver des applications potentielles pour le CRA-W

Comité Intelligence Artificielle Générative

Exemple d'output: Proposition d'outils de soutien pour la recherche bibliographique



Elicit Recent Library Alerts

Q. How can autoencoder be used for dimensionality reduction of agronomical and food spectral data?

Sort: Most relevant Search Filters + Column Export UPGRADE Save all Get full-texts?

Paper	10 found	Summary
Evaluation of an autoencoder as a feature extraction tool for near-infrared spectroscopic discriminant analysis. Seun Jo, Woosuk Sohy, Hyeseon Lee, Hyeon Chung 		Autoencoders can be used for dimensionality reduction of agronomical and food spectral data by extracting sample-descriptive features, improving discrimination accuracy, especially when spectral features are indistinct.
Auto-encoder based dimensionality reduction Yasi Wang, H. Yao, Sicheng Zhao 		Auto-encoders can learn unique features for dimensionality reduction , demonstrated through experiments on synthesized and real datasets like MNIST and Olivetti face datasets.
Neural networks for dimensionality reduction of fluorescence spectra and prediction of drinking water disinfection by-products. Nicolas M. Peleato, R. Legge, R. Andrews 		Autoencoders can be used for dimensionality reduction of fluorescence data, improving predictability of DBPs by mitigating overfitting and providing optimal prediction accuracies.
Comparing AutoEncoder Variants for Real-Time Denoising of Hyperspectral X-Ray Nicola Bonettini, C. A. Gonano, Paolo Bestagini, Marco Marcon, B. Garavelli, and 1 more 		
Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging		Segmented autoencoders (S-SAEs) can be used for dimensionality reduction by dividing the spectral domain into smaller segments, reducing complexity and improving feature extraction and classification accuracy

Great, I'll look for studies that apply autoencoders for dimensionality reduction of spectroscopic or hyperspectral data specifically in the food or agronomical sector. I'll prioritize those that compare autoencoders with traditional methods like PCA, but I'll also include relevant studies that use autoencoders without comparison. I'll get back to you soon with a detailed summary.

Research completed in 14m 37 sources

Autoencoder-Based Dimensionality Reduction for Hyperspectral and Spectroscopic Data in Food and Agriculture

Autoencoder-Based Dimensionality Reduction for Spectral Data in Food and Agronomy

Background: High-Dimensional Spectral Data and Reduction Techniques

Spectroscopic and hyperspectral data are inherently high-dimensional, often comprising hundreds of wavelength bands or spectral features per sample. This high dimensionality poses challenges for data storage, processing speed, and model overfitting. To address this, dimensionality reduction techniques are applied to distill essential information into a lower-dimensional feature space. Traditional methods include both linear and nonlinear approaches: for example, Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and t-Distributed Stochastic Neighbor Embedding (t-SNE) are commonly used to reduce spectral data dimensions . PCA finds orthogonal directions of maximal variance (unsupervised), while LDA (supervised) seeks projections that best separate known classes . Nonlinear manifold methods like t-SNE are often used for visualizing complex spectral differences by preserving local structures in a low-dimensional plot . Each method retains different aspects of the data, and the choice can impact what features of the spectra are emphasized.

Autoencoders, a class of unsupervised deep learning models, have emerged as powerful tools for nonlinear dimensionality reduction. An autoencoder (AE) is a neural network trained to encode the input data into a compressed latent representation and then decode it to reconstruct the original input. By training on reconstruction, the network learns a low-dimensional encoding that captures the most salient features of the data. Unlike PCA (which is linear), autoencoders can capture complex nonlinear spectral patterns through their multiple layers and activation functions. They typically perform a data-driven feature extraction and can be viewed as a nonlinear generalization of PCA. Recent hyperspectral imaging research highlights that

Obtenir un rapport détaillé

+ Rechercher

Sources:

- Li et al., *Remote Sens.* 14(18):4579 (2022) – hyperspectral feature extraction methods overview
- Hang et al., *Remote Sens.* 11(7):864 (2019) – unsupervised spectral feature learning with autoencoders (AE vs PCA results)
- Munyendo et al., *Int. J. Food Sci. Tech.* 58(3):1284 (2023) – coffee adulteration detection with NIR and deep autoencoder
- Vasafi et al., *J. Food Eng.* 299:110510 (2021) – autoencoder for milk process anomaly detection (Raman spectroscopy)
- Shen et al., *Appl. Sci.* 10(11):3769 (2020) – sparse autoencoder for soluble solids in plums (NIR)
- Yang et al., (2020) – deep features (SAE) for peach SSC prediction (cited in).
- Liu et al., *Comput. Electron. Agric.* 198:107007 (2022) – SSC-AE method for strawberry defect detection
- Luo et al., *Food Control* 126:108078 (2021) – t-SNE for vibrational spectroscopy visualization in agro-food

Comité Intelligence Artificielle Générative

NotebookLM



The screenshot displays the NotebookLM interface with a document titled "Hyperspectral unmixing for raman spectroscopy via physics constrained autoencoders". The interface is divided into three main sections: Sources, Discussion, and Studio.

Sources: The document text is visible, discussing hyperspectral unmixing using autoencoder neural networks. It mentions that a set of Raman measurements $X = \{x_i\}$ are collected from a sample of interest, where each x_i is a spectrum, viewed as mixtures of a set of Raman spectral signatures M (endmembers), related to the chemical species present, and their respective fractional abundances a_i . The document aims to infer the set of unknown underlying endmembers M and fractional abundances a_i from the set of measurements X , thereby obtaining qualitative and quantitative compositional information about the sample. It also mentions that the decoder D learns endmember signatures during training and the encoder E learns to derive the corresponding fractional abundances. Encoders can accommodate different concepts from representation learning, such as convolutional layers and attention, to improve feature extraction and provide more accurate and robust unmixing. Decoders can be structured to model different linear and nonlinear mixing models. Physics-inspired constraints are indicated in red in C and D.

Discussion: A discussion thread titled "How to set constraints in autoencoders for spectral unmixing?" is shown. It discusses the importance of constraints in the autoencoders (AE) for spectral unmixing (SU). The constraints are crucial because the underlying physical models, such as the Linear Mixing Model (LMM), require that the proportions of materials (called abundances) satisfy two main conditions: 1. The sum of abundances must be equal to 1 (ASC). 2. The abundances must be non-negative (ANC). The discussion also mentions that these constraints are integrated directly into the architecture of the AE, generally in the encoder and the decoder, to guide the learning process. It lists two fundamental constraints in spectral unmixing: 1. **La Contrainte de Somme à Un (ASC)**: stipule que la somme des abundances de tous les constituants dans un pixel doit être égale à un ($\sum \alpha_n = 1$). 2. **La Contrainte de Non-Négativité des Abondances (ANC)**: garantit que toutes les abundances sont non-négatives ($\alpha_n \geq 0$). It further explains that when using autoencoders for SU, the decoder is often limited to a single layer so that its weights represent the endmembers (pure spectra), and a non-negativity constraint on the endmembers is also applied. It concludes by stating that these constraints are defined and implemented in the architectures of autoencoders for SU: 1. **Contraintes sur les Abondances (Espace Latent / Sortie de l'Encodeur)**. It notes that the fractional abundances are represented by the output of the last layer of the encoder.

Studio: The Studio section shows a list of sources and a note titled "Autoencoders et Décomposition Spectrale". It also includes a button to "Ajouter une note".

Detective



- Base de données GenEdit – Modifications génétiques ciblées – Recense résultats de publications scientifiques

→ Vérification automatique des entrées via IA

A	B	C	D	E	F	G	H	I	J	K	L	M
kingdom	Species	Country	Trait Category	Trait	gene	Method	Technique de Modif	type	Year	Bibtex	References	Status
Plants	Rapeseed	China	Modified cor	Modified fatt	FAD2	CRISPR	SDN1 / CRISPR/Cas		202	\cite{huangM Huang, H., Ci	Unknown	
Plants	Maize	US	Herbicide tol	Herbicide tol	AHAS	ODM (Oligo-I	ODM (Oligo-Directed Muta		2000	\cite{zhuTarg Zhu, T., Peter	Unknown	
Plants	Maize	US	Herbicide tol	Herbicide tol	AHAS	ODM (Oligo-I	ODM (Oligo-Directed Muta		2000	\cite{zhuEngi Zhu, T., Mett	Unknown	
Plants	Rice	Japan	NC	The mutation	ALS	CRISPR	CRISPR/Cas		2004	\cite{okuzaki Okuzaki, A., t	Unknown	
Plants	Maize	Germany		Herbicide res	PAT		ZFN		2006	\cite{schorna Schornack, S	Unknown	
Plants	Maize	US	Herbicide tol	Herbicide tol	IPK1	ZFN	ZFN	Gene replace	2009	\cite{shuklaP Shukla, V.K.;	Unknown	
Plants	Maize	China	NC	Heritable tar	LG1	Meganucleas	Meganucleas	Gene knocko	2010	\cite{gaoHeri Gao, H.; Smit	Unknown	
Plants	Soybean	US	NC	Heritable tra	DCL4b	ZFN	ZFN	Gene knocko	2011	\cite{curtinTa Curtin, S.J.; Z	Unknown	
Animals	Poultry	US	Patent	Product qual	six BBL gene	Meganucleas	Meganucleas	SDN1	2011	https://www.	Unknown	
Animals	Cattle		Modified cor	Prevision :re	BLG	ZFN		Knock out	2011	\cite{yuHighl Yu, S., Luo, J.	Unknown	
Animals	Pig									\cite{yangGe Yang, D., Yan	Unknown	



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	kingdom	Species	Country	Trait Category	Trait	gene	Method	Technique de Modif	type	Year	Bibtex	References	Status
2	Plants	Rapeseed	China	Modified cor	Modified fatt	FAD2	CRISPR	SDN1 / CRISPR/Cas		2020	\cite{huangM Huang, H., Ci	Unknown	
3	Plants	Maize	US	Herbicide tol	Herbicide tol	AHAS	ODM (Oligo-	ODM (Oligo-Directed Muta		1999	\cite{zhuTarg Zhu, T., Peter	Unknown	
4	Plants	Maize	US	Herbicide tol	Herbicide tol	AHAS	ODM (Oligo-	ODM (Oligo-Directed Muta		2000	\cite{zhuEngi Zhu, T., Mett	Unknown	
5	Plants	Rice	Japan	NC	The mutator	ALS	ODM (Oligo-	ODM (Oligo-Directed Muta		2004	\cite{okuzaki Okuzaki, A.,	Unknown	
6	Plants	Maize	Germany		Herbicide re	PAT		ZFN		2006	\cite{schorna Schornack, S	Unknown	
7	Plants	Maize	US	Herbicide tol	Herbicide tol	IPK1	ZFN	ZFN	Gene replace	2009	\cite{shuklaP Shukla, V.K.;	Unknown	
8	Plants	Maize	US	NC	Heritable tar	LG1	Meganuclea	Meganuclea	Gene knocko	2010	\cite{gaoHeri Gao, H.; Smit	Unknown	
9	Plants	Soybean	US	NC	Heritable tra	DCL4b	ZFN	ZFN	Gene knocko	2011	\cite{curtinTa Curtin, S.J.; Z	Unknown	
10	Plants	Tobacco	US	Patent	Product qual	six BBL gene	Meganuclea	Meganucleas	SDN1	2011	https://www.	Unknown	
11	Animals	Cattle		Modified cor	Prevision :re	BLG	ZFN		Knock out	2011	\cite{yuHighl Yu, S., Luo, J.	Unknown	
12	Animals	Pig	China	Biomedical	Ppar-? heter	PPAR-y	ZFN	somatic cell nuclear trans		2011	\cite{yangGe Yang, D., Yan	Unknown	

- Suite: Entrée automatique des informations via IA pour de nouveaux articles + double-vérification

Conclusion

- L'IA est déjà bien intégrée dans plusieurs projets du CRA-W
 - Les outils évoluent vite : besoin d'une veille et d'adaptations continues
 - Potentiel immense pour la recherche agronomique
 - Communication entre utilisateurs d'IA essentielle afin de mutualiser les efforts
 - Réflexion en cours pour un outil dédié aux applications ayant recours au Deep Learning.
- Comment favoriser la communication autour de ces outils au sein de nos services ?

Merci pour votre attention

Antoine Deryck

Chercheur Junior et Doctorant

Département Connaissance et valorisation des produits

Unité Qualité et Authentification des produits

a.deryck@cra.wallonie.be | <https://www.cra.wallonie.be/fr/antoine-deryck>



Embedded and Artificially intelligent VISion Engineering (EAVISE)

antoine.deryck@kuleuven.be

