Hackathon SDD 2025
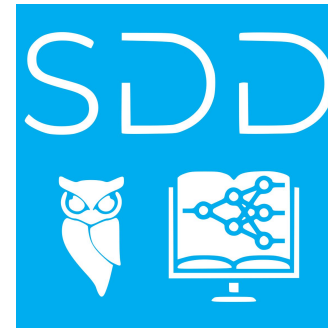
Hack on what?

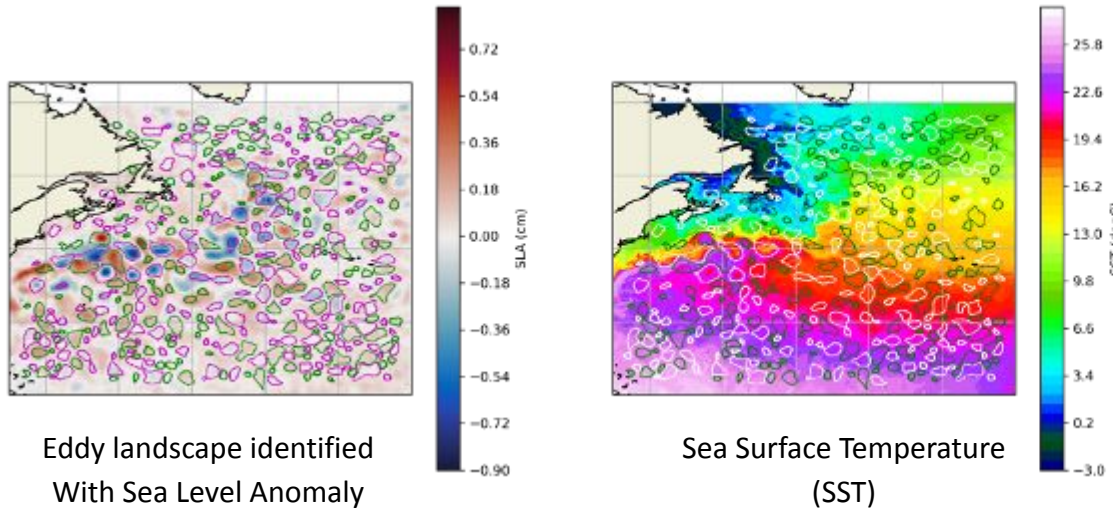1. **Ocean Eddy Identification and Tracking**

2. **IoT – Overall Equipment Effectiveness Prediction using watsonx**

3. **Intelligent Text Extraction**

Subject 1: Ocean Eddy Identification and Tracking

Responsable : Daria Botvynko, Anass el Aouni, Simon van Gennip

Eddy landscape identified
With Sea Level Anomaly

Sea Surface Temperature
(SST)

Oceanic Eddies are vortices of the order of ~10km in horizontal scale, whose signature is clearly visible in satellite products and well reproduced in ocean models. Sea Level Anomalies (SLA) are used for detecting such object by means of numerical techniques, yet such approach lacks accuracy in detection, namely because Sea Level Anomaly contains errors. Eddy signature is also visible in other variables such as Sea surface Temperature (SST) that do not suffer from such limitation.

The objective here is to **develop a Deep Learning approach to identify eddies using SST, SLA, the ocean velocity field and to follow them in time.** For this purpose students will have a dataset consisting of SLA, SST, and ocean currents velocity images which are slightly distorted relative to reality, together with labels (the eddy contours, see figure) obtained directly from reality.

Subject 2:

Responsable : Elsa Charpentier, Serge Bonnaud, Eric Daubie

The subject focuses on exploring the **watsonx platform** to accelerate Data Science projects. Participants will gain hands-on experience across the full lifecycle of a Data Science project, including **data preparation**, **data visualization**, **training Machine Learning (ML) models**, and **deploying and monitoring these models**.

The challenge centers on an **IoT** use case: predicting drops in the **Overall Equipment Effectiveness (OEE)** of industrial machines. OEE is a key metric for measuring the performance of production lines, accounting for **production speed** and the **quality of items** produced. Using realistic simulated datasets, participants will identify the features influencing OEE, select the best Machine Learning model, deploy it as an API, and establish lifecycle governance for the project.

Leveraging **AutoAI**—a graphical AutoML tool within Watsonx—participants will develop an AI model to predict OEE and evaluate it using standard metrics such as **accuracy**, **F1-score**, **precision**, and **recall**. The model will be deployed and tested on the **Watsonx platform**, ensuring a practical, low-code, end-to-end Data Science workflow.

AVAILABILITY   PERFORMANCE   QUALITY

= OEE

Subject 3: Intelligent Text Extraction

Responsable : Kinza KASMI, Mathis CHENET

Context :

Our computer vision projects predominantly center on **Optical Character Recognition** (OCR). Clients come to us with vast collections of documents seeking to transform their analog data into digital assets. Our mission is to read and convert these documents into structured databases, enabling querying and analytics capabilities.

The hackathon :

The main challenge of OCR lies in accurately extracting specific fields from documents that often have **poor quality or artifacts**. For this hackathon, we've designed a challenge that mirrors these real-world scenarios: participants will work with a dataset of **PDF images** containing annotated fields, particularly focusing on a "references" field that we've deliberately varied in quality. This hackathon tests both technical skills and practical problem-solving abilities that match our daily production work.

Objective :

Develop an OCR solution capable of :

1. Accurately **detecting** specific text bounding boxes.
2. Correctly **reading** the text from these areas.

SIAPARTNERS

Hack when?

Monday 8h00 :
    Kickoff presentation

Tuesday 12h :
    Board games

Wednesday 12h:
    Pizza 🍕 🍕 🍕

Wednesday 16h00:
    Amphi 3
    Wrap-up presentation

Day schedules depend on subject, check with partner

Hack where?

☕ **61.101**

MERCATOR OCEAN INTERNATIONAL **61.102**

IBM **61.103**

SIA PARTNERS **61.104**

Dedicated Slack channel
(#hackathon1 - #hackaton3)

Extra rooms:
    Monday: 61.105
    Tuesday: 61.105
    Wednesday: 61.105 (morning)
                  61.112 (afternoon)

🏅🏅🏅🏅🏅🏅🏅🏅🏅🏅
🏅 Hacking on site is the best way to win!
🏅
🏅
🏅🏅🏅🏅🏅🏅🏅🏅🏅🏅

Hack how?

Goals for this hackathon:

📚🧠 Practice knowledge gained in SDD

🤷‍♀️🤷‍♂️ Cover parts of SDD material you're not sure on

👩‍💻👨‍💻 Work with your classmates in teams of 3 to 5

🌍🛰️ See examples of real-world data science

🥳🥳 Have fun!

This hackathon will be a success if you:

📣🙋‍♀️ **Communicate**

Make sure everyone is on the same page about your objectives and plan.

🙋‍♀️🙋‍♂️ **Ask for help**

Partners will be available throughout the hackathon to answer questions. Dennis is also available for general questions.

🤝🙏 **Help each other**

That's what teammates are for!

🤖💻 **Hack**

From Pandas to PyTorch, show off your SDD skills!

Hack how?

Evaluation is based on:

Effort

Try your hardest during these three days to get the best results

Communication

How well you work as a team and coordinate with the partners

Planning

You only have 3 days! How will you use them? Plan well, including what task each member of your team will do.

Results

Push your results to the maximum!