SPE International Aberdeen Section — Solutions. People. Energy.℠

Fiber-Optic Sensing for Field Development Asset Integrity & Optimization Workshop

24–25th March 2026
Ardoe House Hotel,
Aberdeen, UK

Distributed Acoustic Sensing and Artificial Intelligence: Challenges, Constraints, and SSL-Based Mitigation Strategies.

Camille HUYNH

FEBUS Optics

C. Huynh[1,2], J. Rimpot[1,3], C. Hibert[1,3], A. Turquet[4], T. Stangeland[4], J-P. Malet[1,3], V. Lanticq[2]

[1] ITES, CNRS UMR 7063, Université de Strasbourg – 5 rue René Descartes, 67084 Strasbourg, FRANCE

[2] FEBUS Optics – 2 av. du Président Pierre Angot, 64000 Pau , FRANCE

[3] EOST, CNRS UAR 830, Université de Strasbourg – 5 rue René Descartes, 67084 Strasbourg , FRANCE

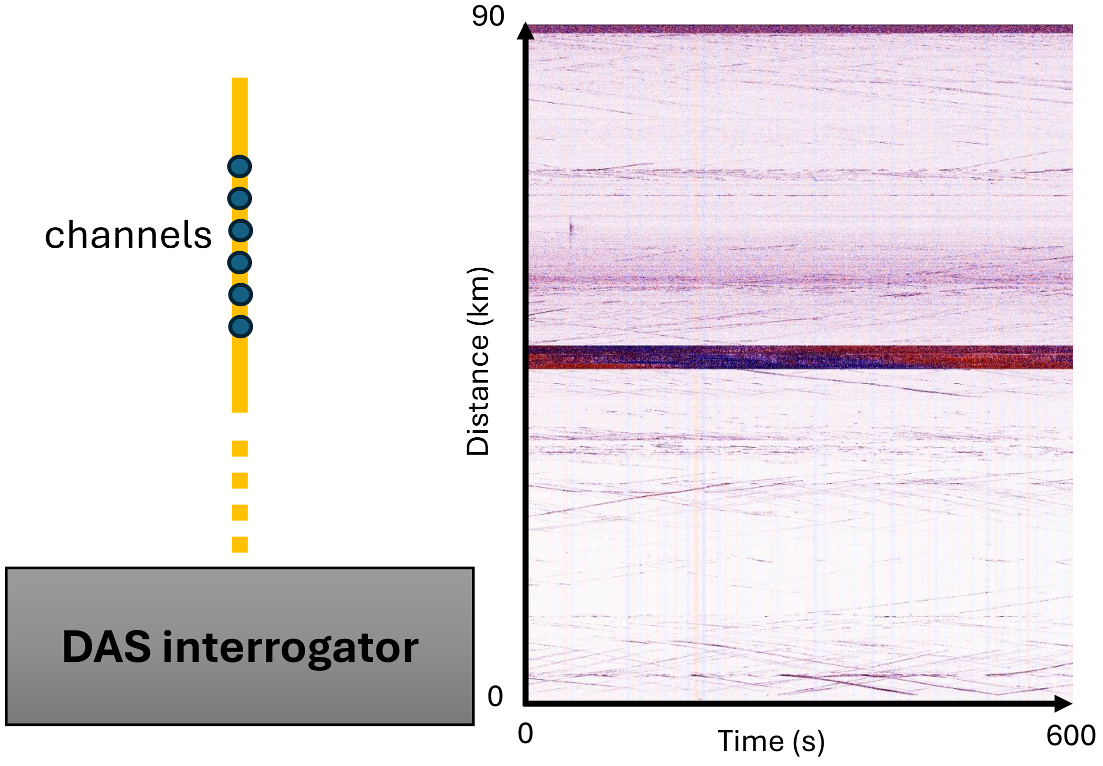
[4] NORSAR – Gunnar Randers vei 15, 2007 Kjeller, NORWAY

FEBUS optics

The reference in distributed sensing

INTRODUCTION

Distributed Acoustic Sensing – Data characteristics



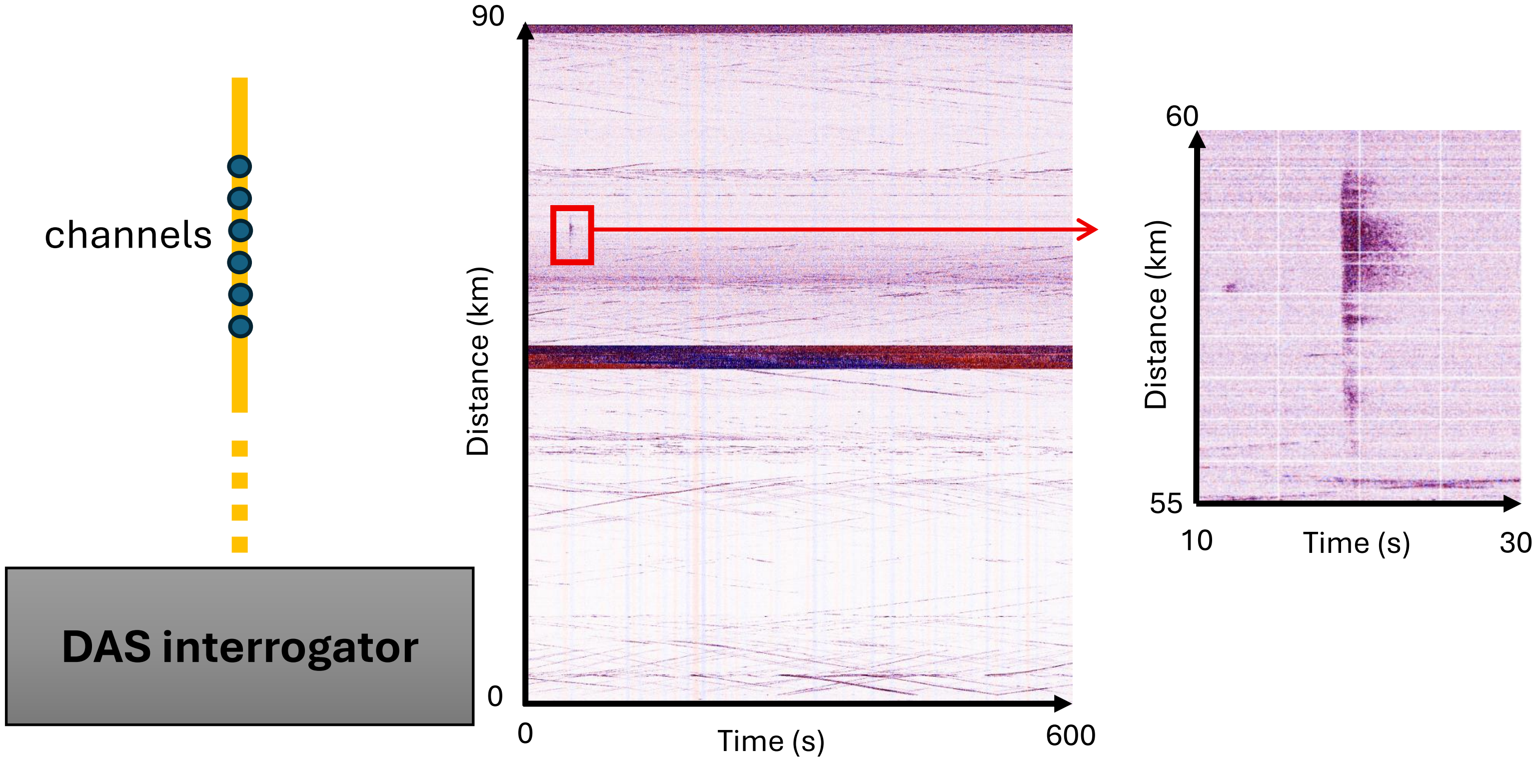
- Why is DAS used?
  - Distributed measurement
  - High distance range (several hundreds km)
  - High spatial (~meter) and temporal (~kHz) resolution

+ Very good to study localized known event (in distance and in time).

− High data volume to manage when used with no apriori about the expected events.

INTRODUCTION

Distributed Acoustic Sensing – Data characteristics



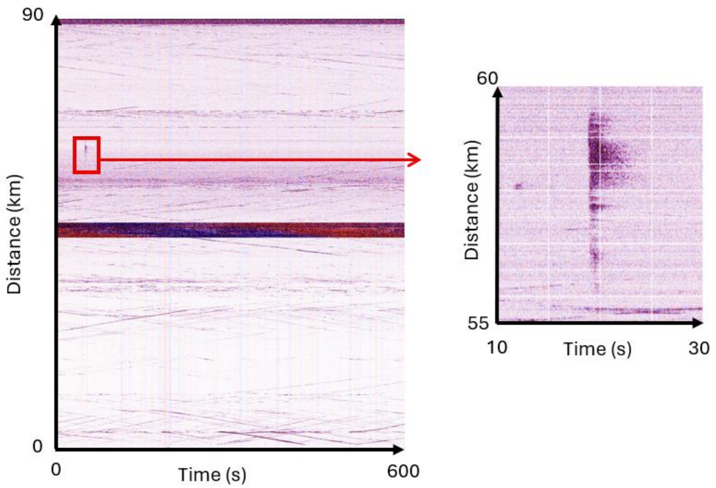
- Why is DAS used?
  - Distributed measurement
  - High distance range (several hundreds km)
  - High spatial (~meter) and temporal (~kHz) resolution

+ Very good to study localized known event (in distance and in time).

– High data volume to manage when used with no apriori about the expected events.

INTRODUCTION

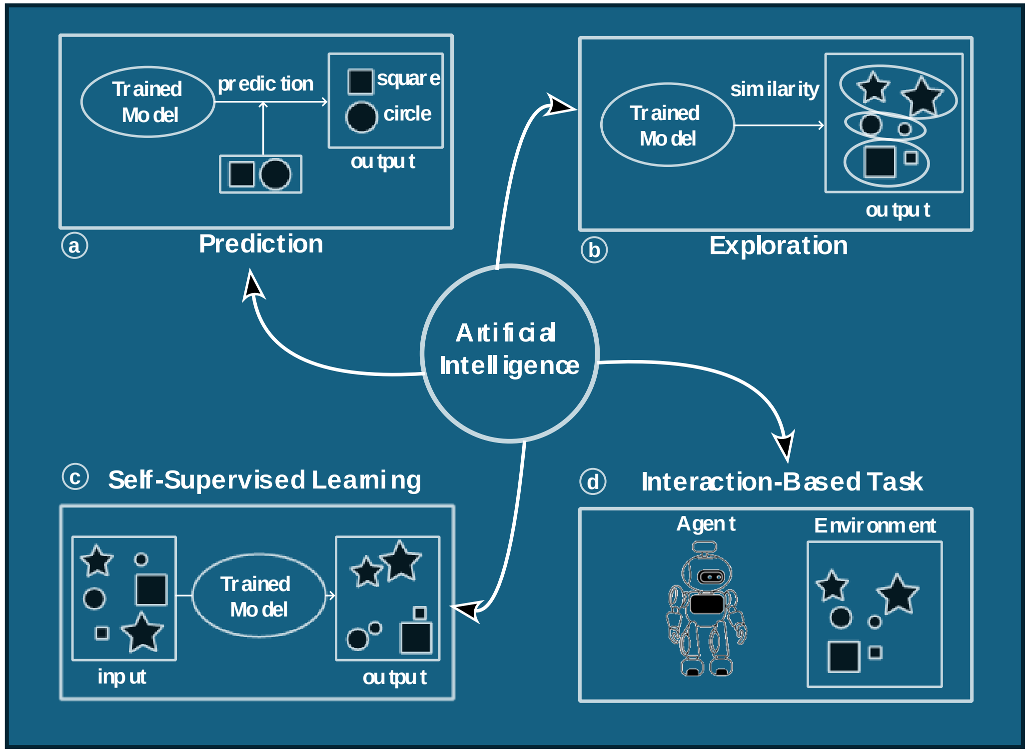
Distributed Acoustic Sensing – Current application limitations



- Challenges
  - C1: **High data volume**: High spatial (~meter) and temporal (~kHz) resolution, acquisition range up to hundreds km.
  - C2: **Extremely few known events, data mainly dominated by noise:** Methods must detect anomalies and treat rare and frequent events equally.
  - C3: **High diversity of signals inside the same event class, environmental behaviour influence :** Automated separation of different signal regimes requires learning informative representations from raw signals.

INTRODUCTION

Insight about artificial intelligence (AI) world



Prediction: Learn from labeled event.
e.g. Detect earthquakes using a model trained on a reference catalog.

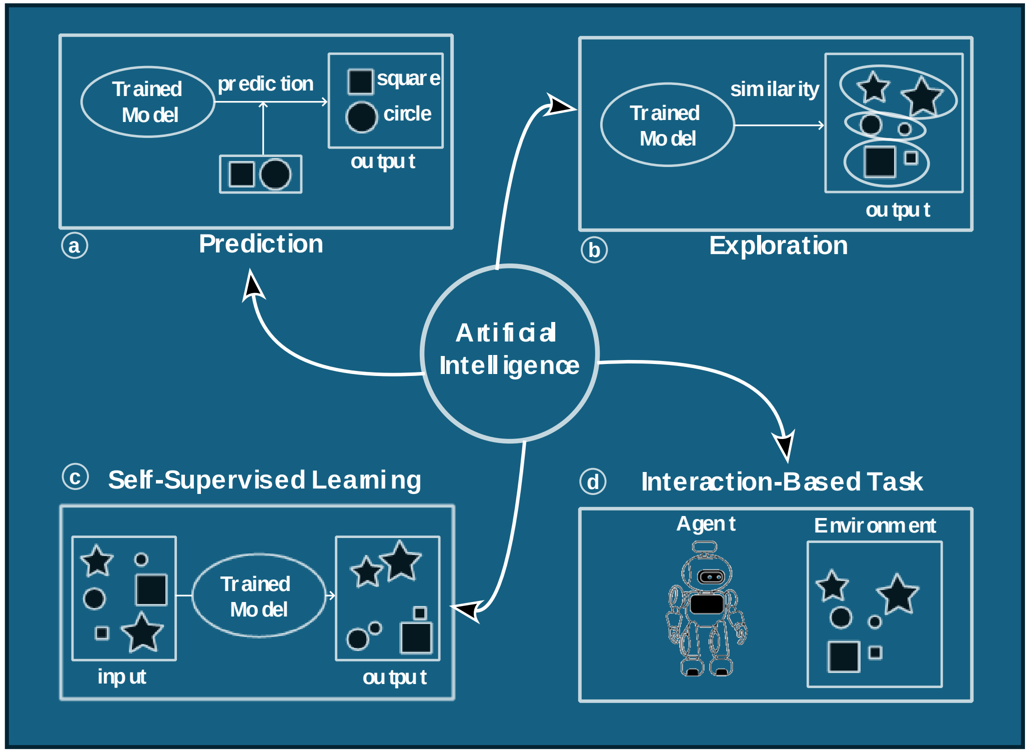
Exploration: Group signal without labels
e.g. Cluster propagation behaviour

SSL: Learn structure from unlabeled data.
e.g. Try to predicting neighboring DAS traces in space or time.

Reinforcement Learning: Learn with feedback.
e.g. Identify the most informative DAS channels

INTRODUCTION

Insight about artificial intelligence (AI) world



Prediction: Learn from labeled event.
e.g. Detect earthquakes using a model trained on a reference catalog.

Exploration: Group signal without labels
e.g. Cluster propagation behaviour → useful for C2

SSL: Learn structure from unlabeled data.
e.g. Try to predicting neighboring DAS traces in space or time. → useful for C3

Reinforcement Learning: Learn with feedback.
e.g. Identify the most informative DAS channels

INTRODUCTION

Objectives of the study, targeted scope

- Objective: Cluster vibrational events without any assumptions, and within a reasonable timeframe compared to a purely visual annotation.
- Our methodology
  - Self-Supervised and Unsupervised Learning for dataset representation, exploration and assisted labelization (*Huynh et al. 2025*).

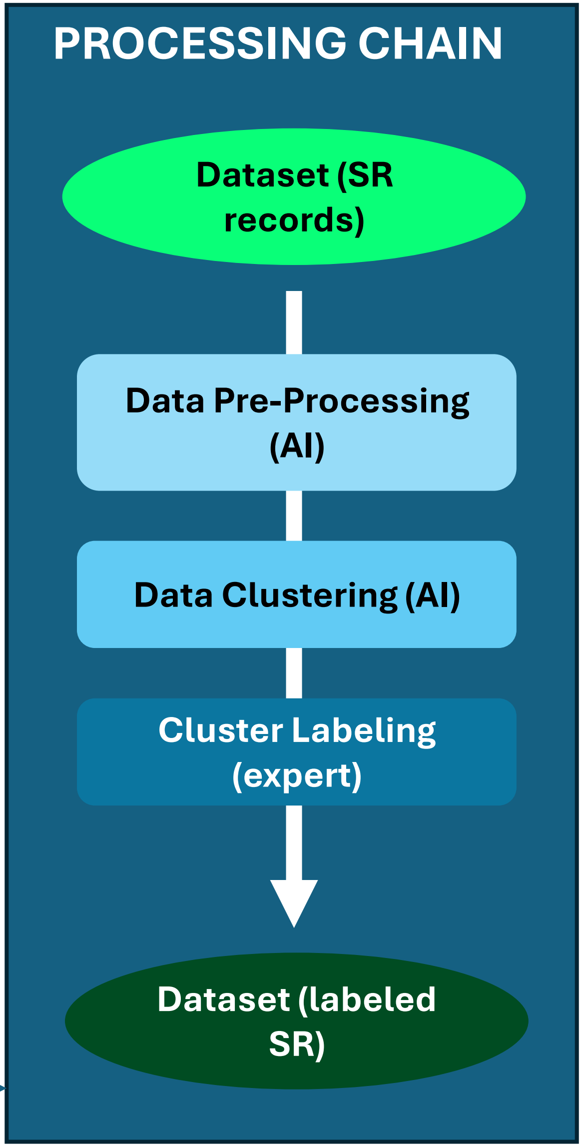
- Instrument: Distributed Acoustic Sensing (FEBUS A1-R)
- 2 datasets:
  - Pyrenees
  - Viella

INTRODUCTION

Objectives of the study, targeted scope

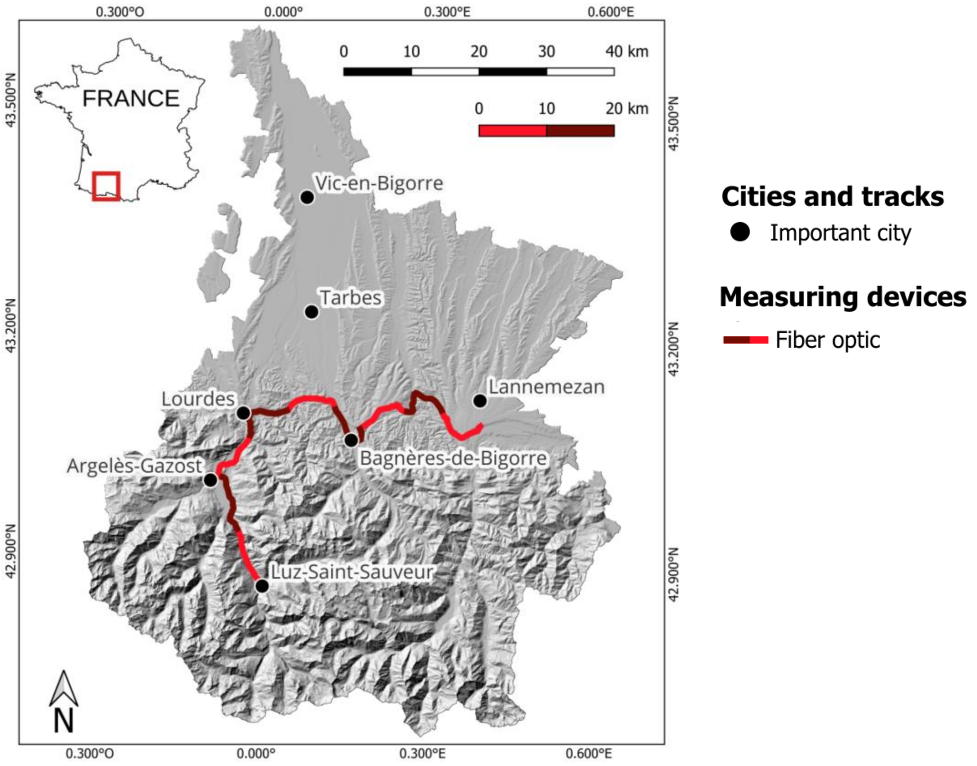
- Objective: Cluster vibrational events without any assumptions, and within a reasonable timeframe compared to a purely visual annotation.
- Our methodology
    - Self-Supervised and Unsupervised Learning for dataset representation, exploration and assisted labelization (*Huynh et al. 2025*).
- Instrument: Distributed Acoustic Sensing (FEBUS A1-R)
- 2 datasets:
    - Pyrenees
    - Viella

Outline of the presentation



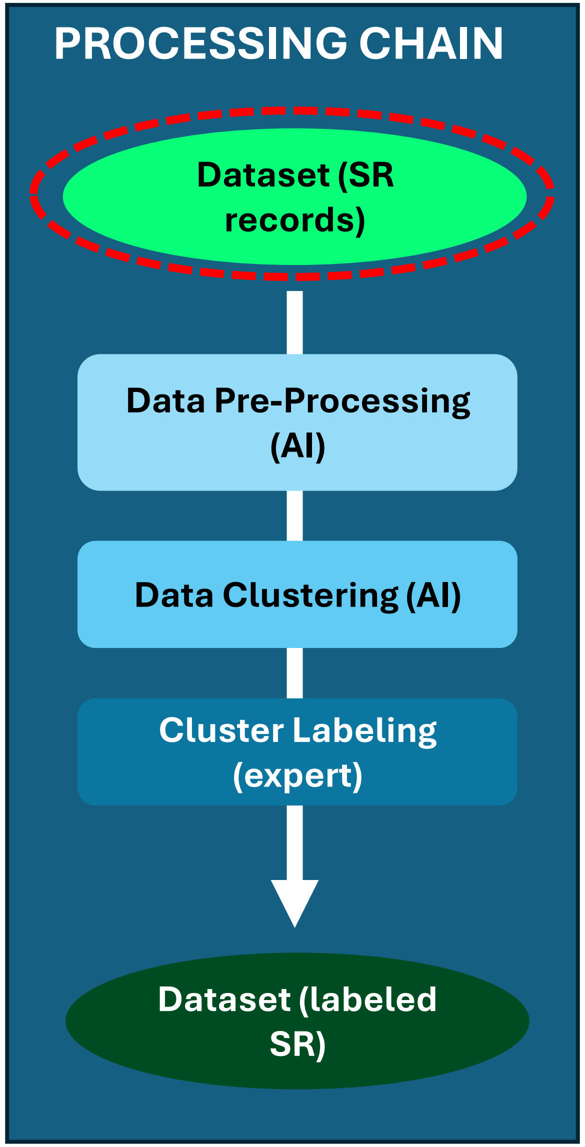
DATASET

Data Acquisition Site 1: Pyrenees



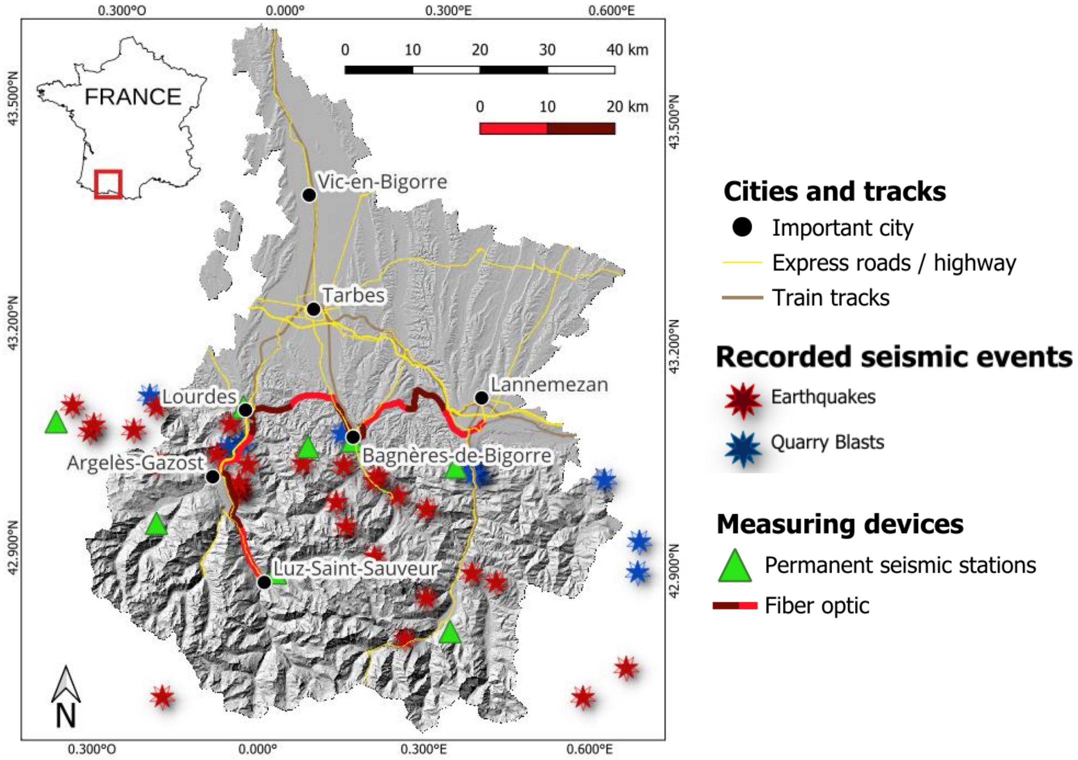
Location: Pyrenees (SW France)

- **Fiber length:** 91 km.
- **Amount of spatial virtual sensors:** 19k channels.
- **Sampling frequency:** 250 Hz.
- **Data Volume:** 19x 10-min files



DATASET

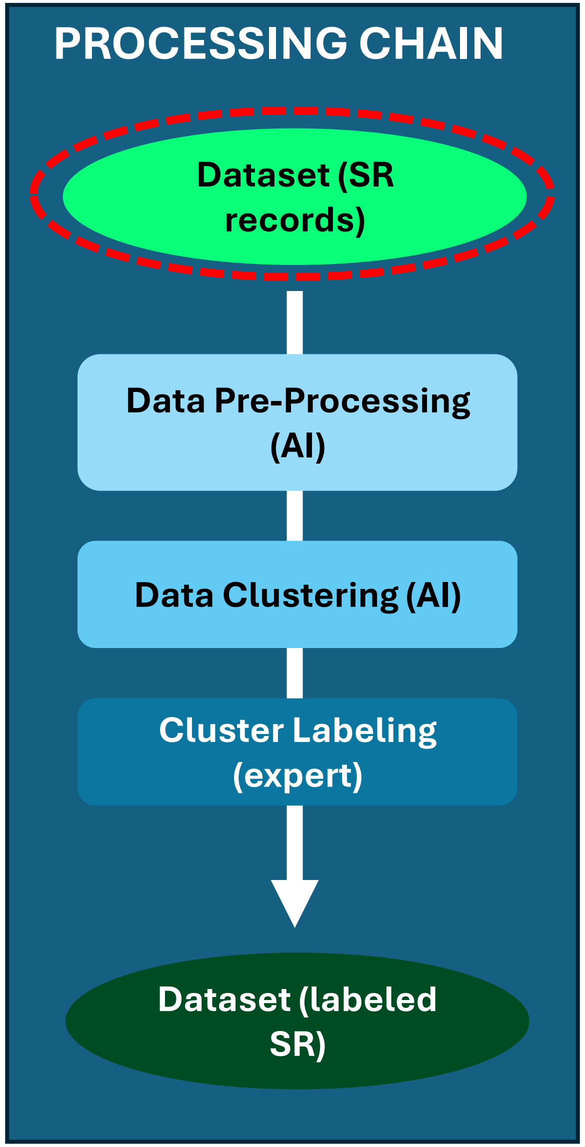
Data Acquisition Site 1: Pyrenees



Location: Pyrenees (SW France)

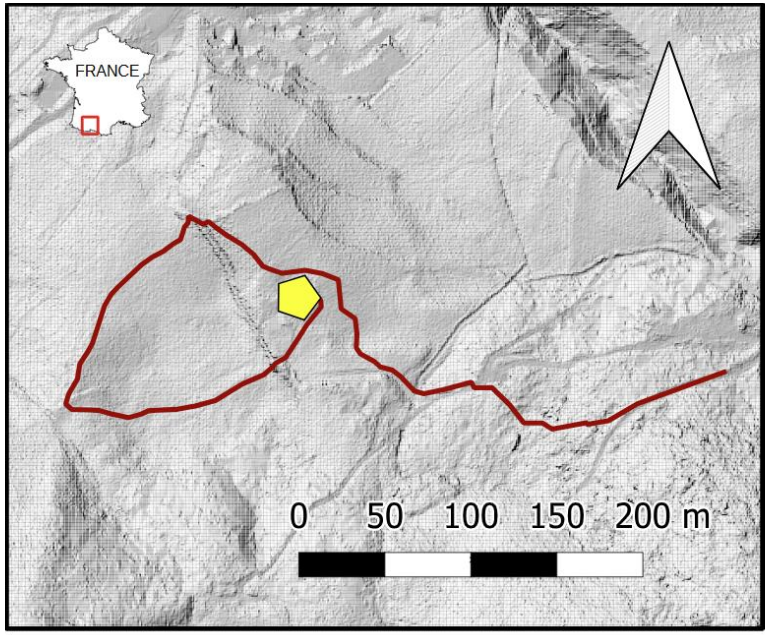
- Fiber length: 91 km.
- Amount of spatial virtual sensors: 19k channels.
- Sampling frequency: 250 Hz.
- Data Volume: 19x 10-min files

Key message: Pyrenees site:
- Regional scale
- Strong environmental and instrumental variability.



DATASET

Data Acquisition Site 2: Viella



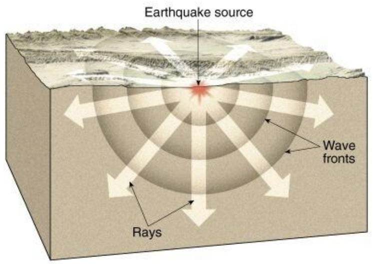
Legends

Fiber Optic (800 m) — Start Fiber

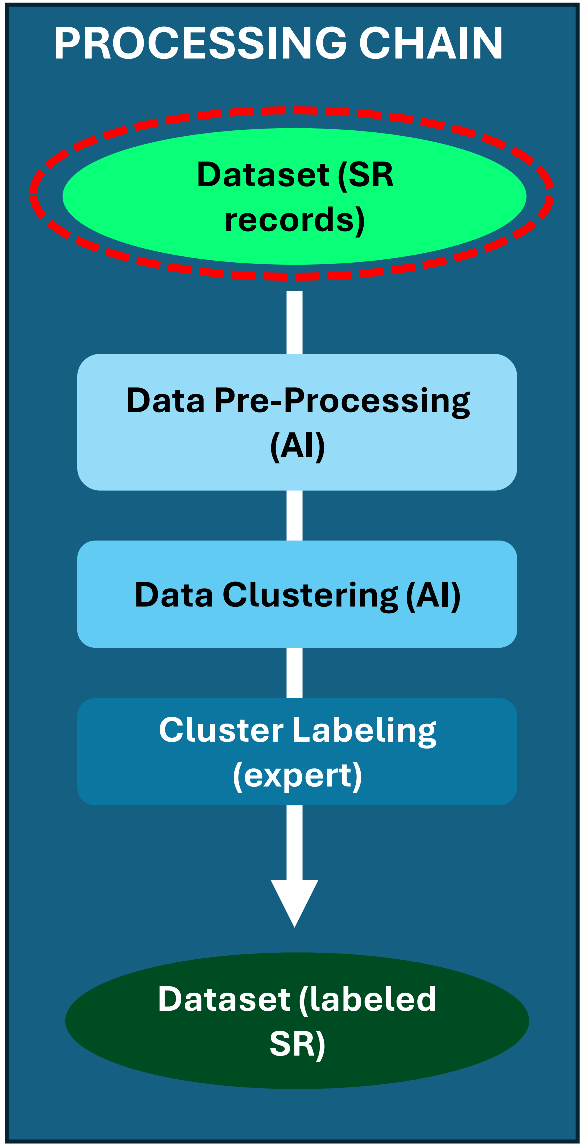
Location: Viella landslide

- Measurement for 44 days.
- Fiber length: 800 m.
- Amount of spatial virtual sensors: 330 channels.
- Sampling frequency: 400 Hz.
- Data Volume: 3 TB.

Events

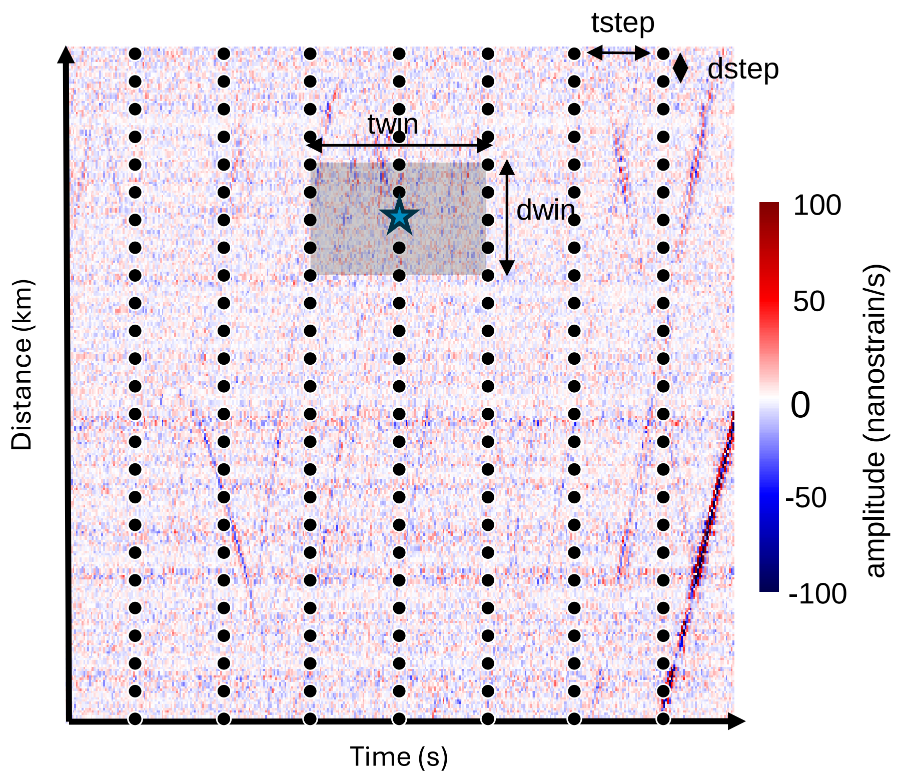


Key message: Viella site:
- Continuous acquisition
- Possible observation of weekly/seasonal variation.



PROCESSING

Data stream processing: Sampling grid



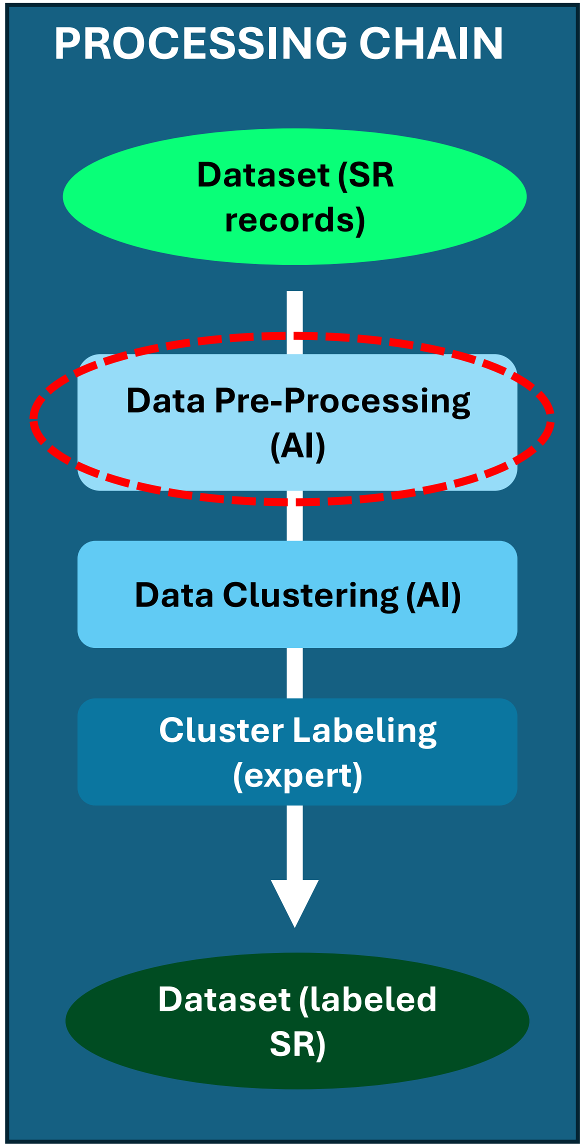
Data is processed in **stream.**

→ Large-scale scalability

Window size:

- Pyrenees: 1000 m, 1 min
- Viella: 100 m, 1 min

Key message: Data stream processing
- Compatible for DAS real-time monitoring in large scale.

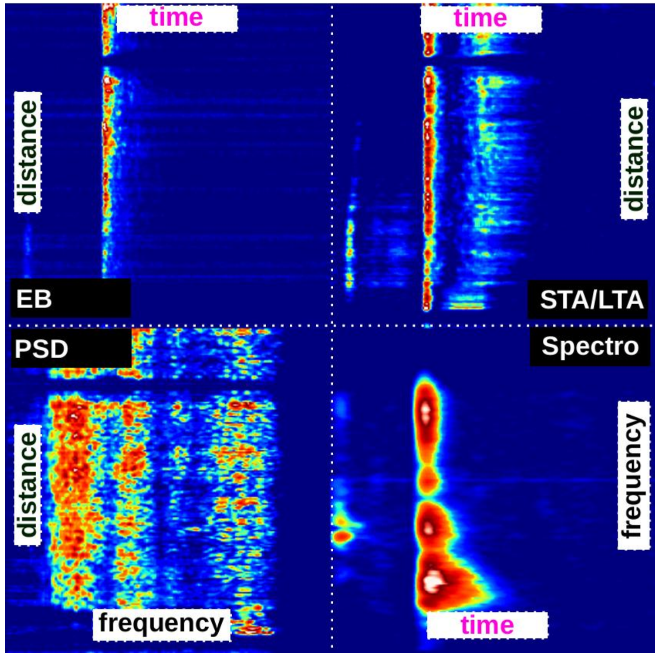


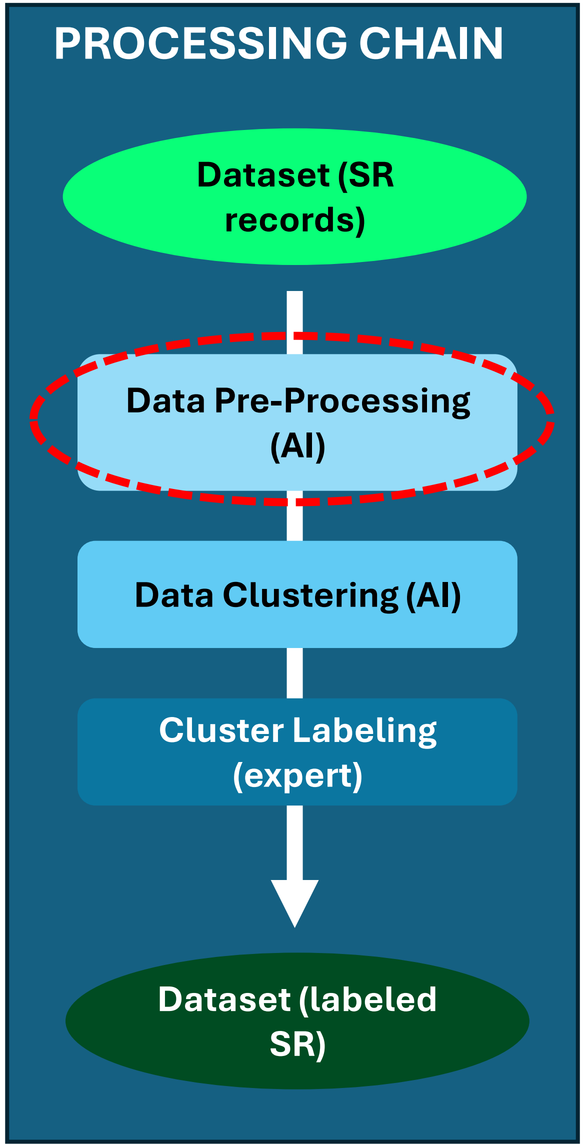
PROCESSING

Input image build for SSL

A geophysical approach based on several commonly used data plots

- Energy Band : Spectrum integration in the [1/60, 30] Hz band.
- STA/LTA : Ratio between the average of the signal taken over 1 s and over 10 s.
- PSD : Spectrum in the [1/60, 30] Hz band for all channels.
- Spectrogram : Mean over channels of spectrum computed using sliding temporal window in the [1/60, 30] Hz band.

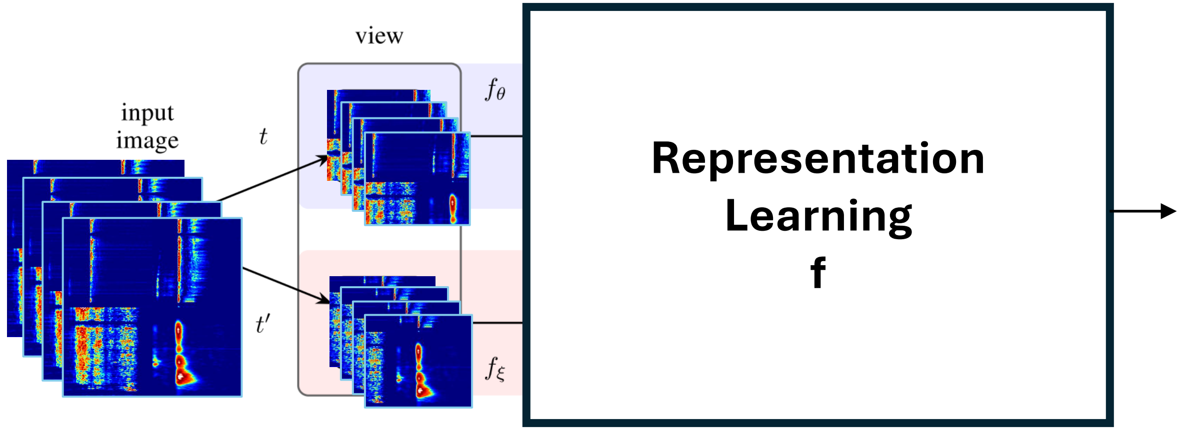


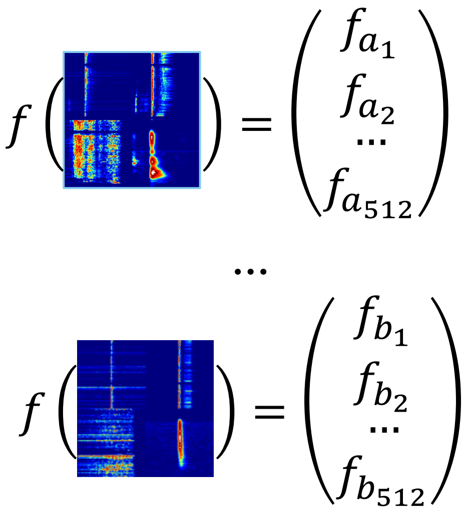


Key message: We build data image input based on the physic of the vibrating events.

PROCESSING

SSL (BYOL): learn data features without labels



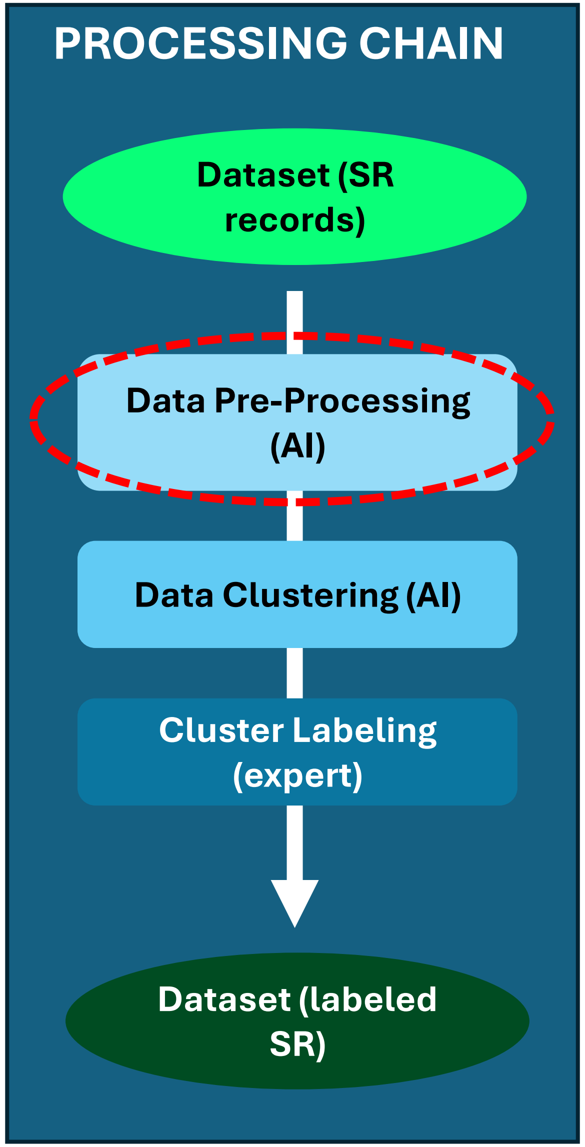


BYOL algorithm representation. *Adapted from Grill et al. 2020*

Goal: Find the best representation transform *f* that reduces the prediction difference between 2 different views of a same input image.

Output: High dimension representation

Key message: SSL-BYOL usage:
- Build features insensitive to different views build from same output image.

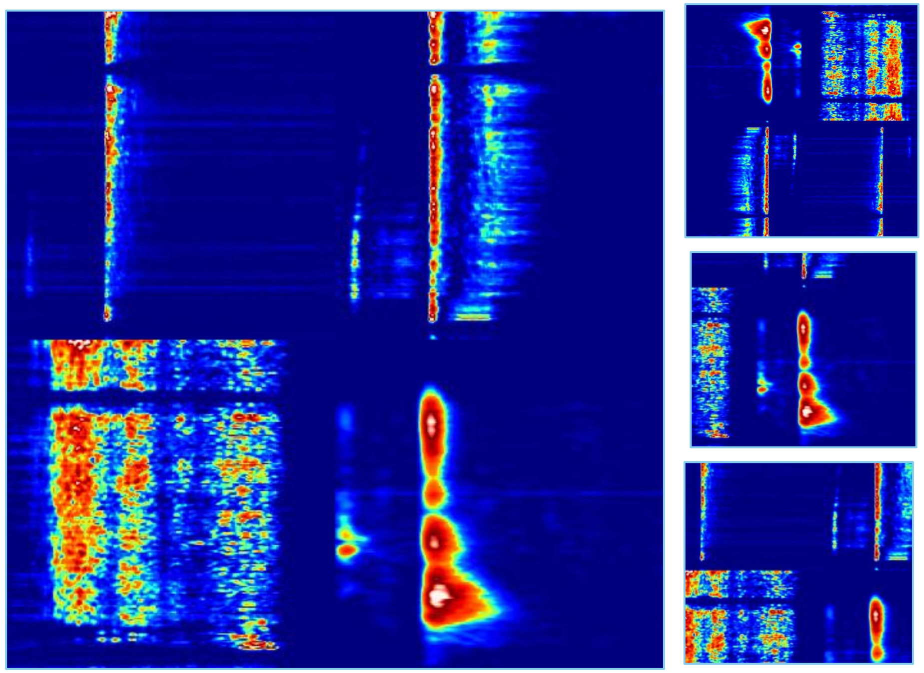


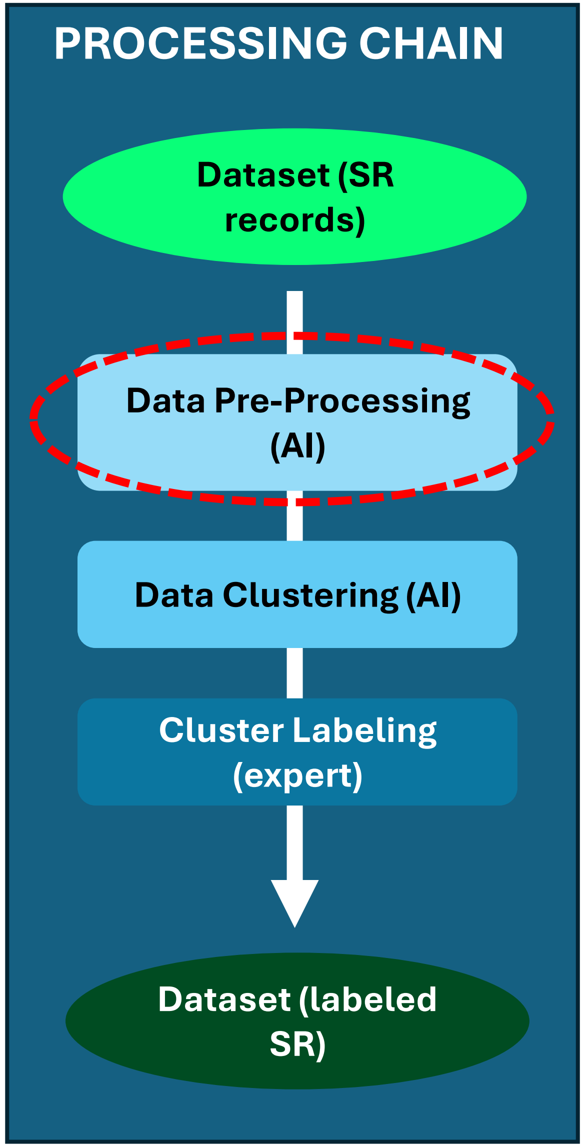
PROCESSING

Data image augmentation

Used augmentation:

- Random flipping
- Random rotation
- Random cropping
- Random noise addition



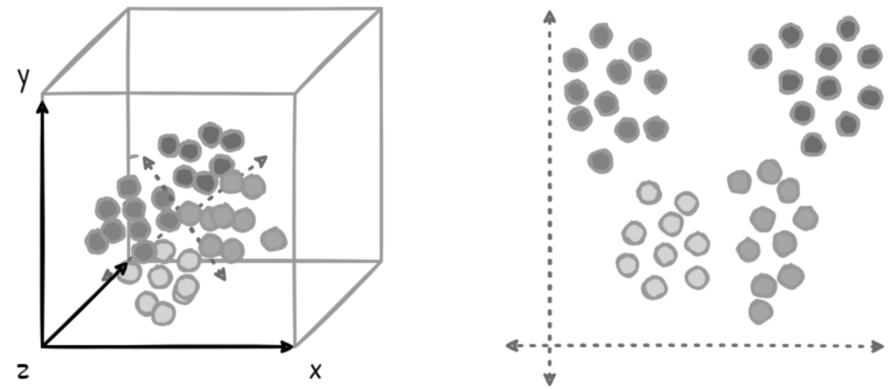


Key message: We extract different views from data image input to feed SSL algorithm during training.

SSL RESULTS

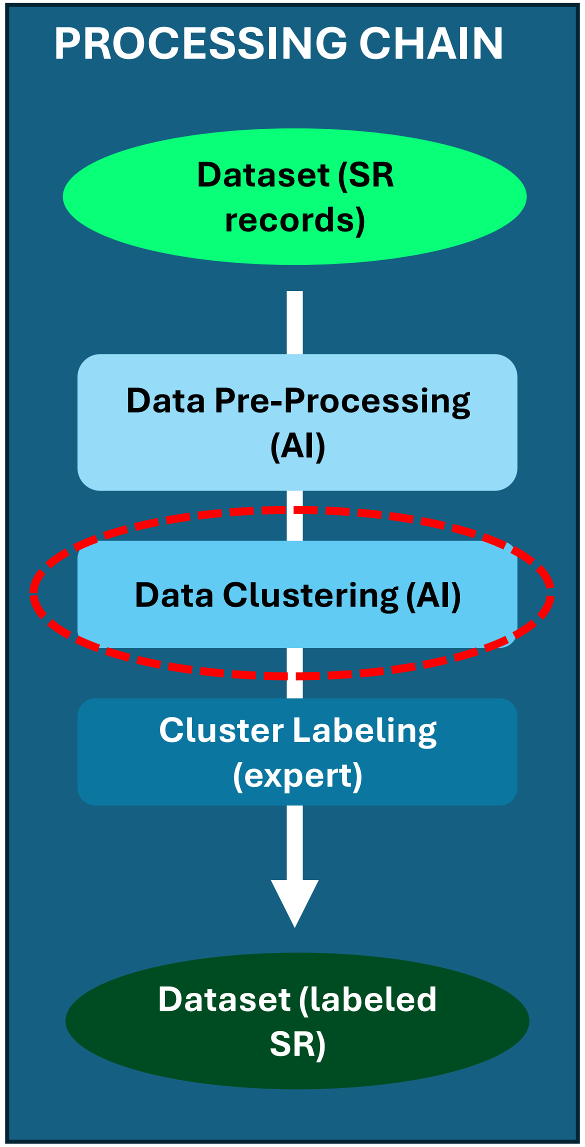
Dimension reduction algorithm for visualization (T-SNE)

Goal: Provide a human-readable representation of the produced representation
!!! Only for visualisation purpose !!!



From 3d representation to 2d representation of the data. Adapted from *mlguru.ai*.

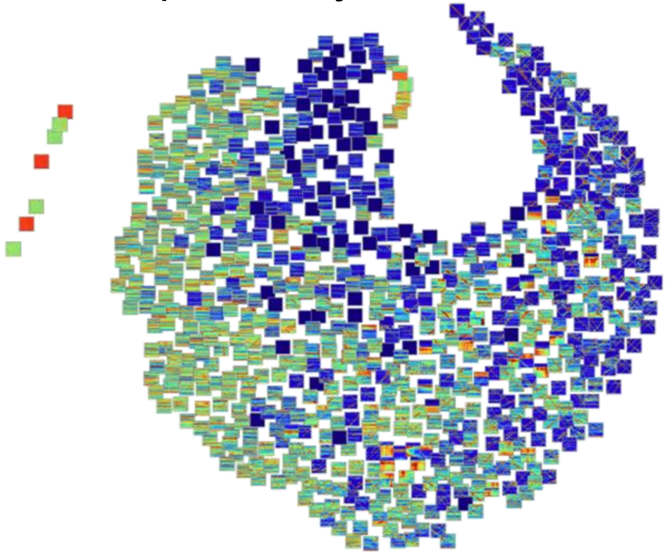
Key message: T-SNE representation:
- A 2-D map
- Signals exhibiting similar behavior (in terms of the representation produced) are close in the map.



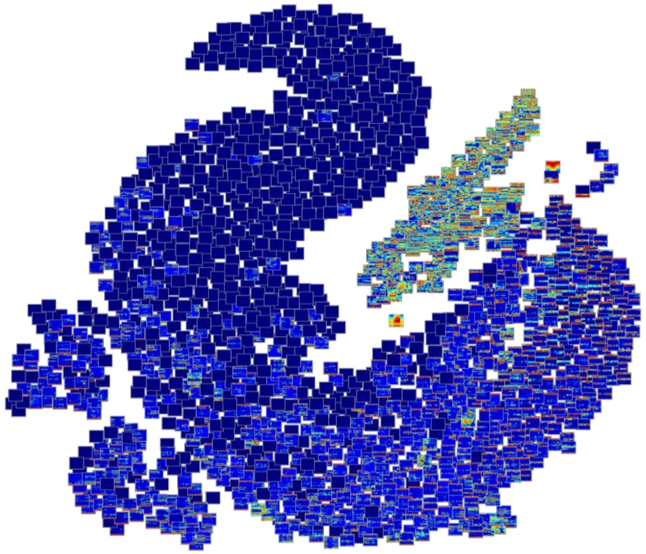
SSL RESULTS

2-D maps, Pyrenees vs Viella

Latent space 1: Pyrenees dataset

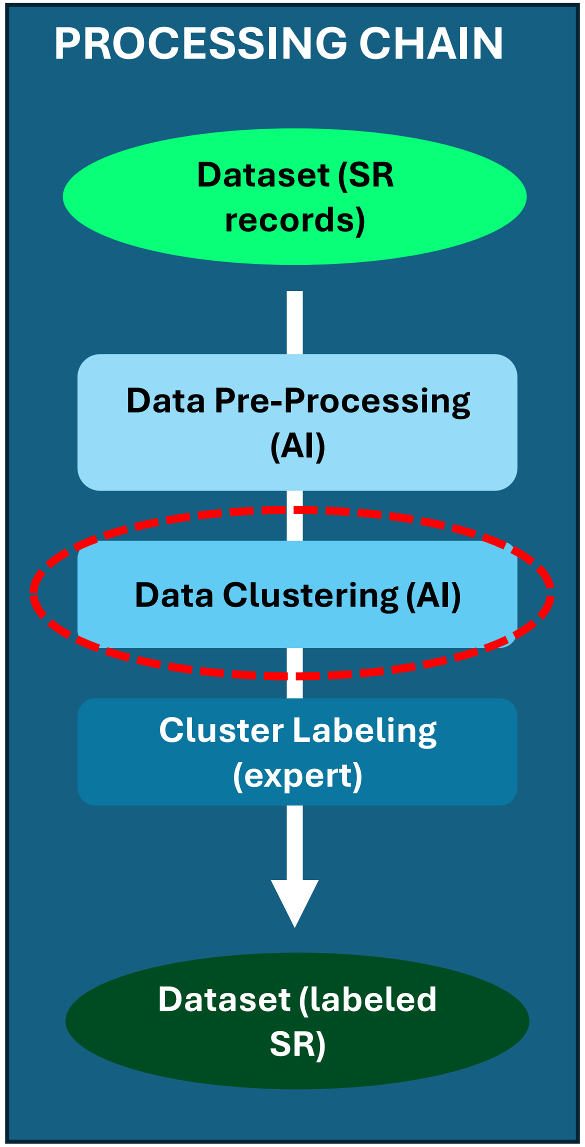


Latent space 2: Viella dataset



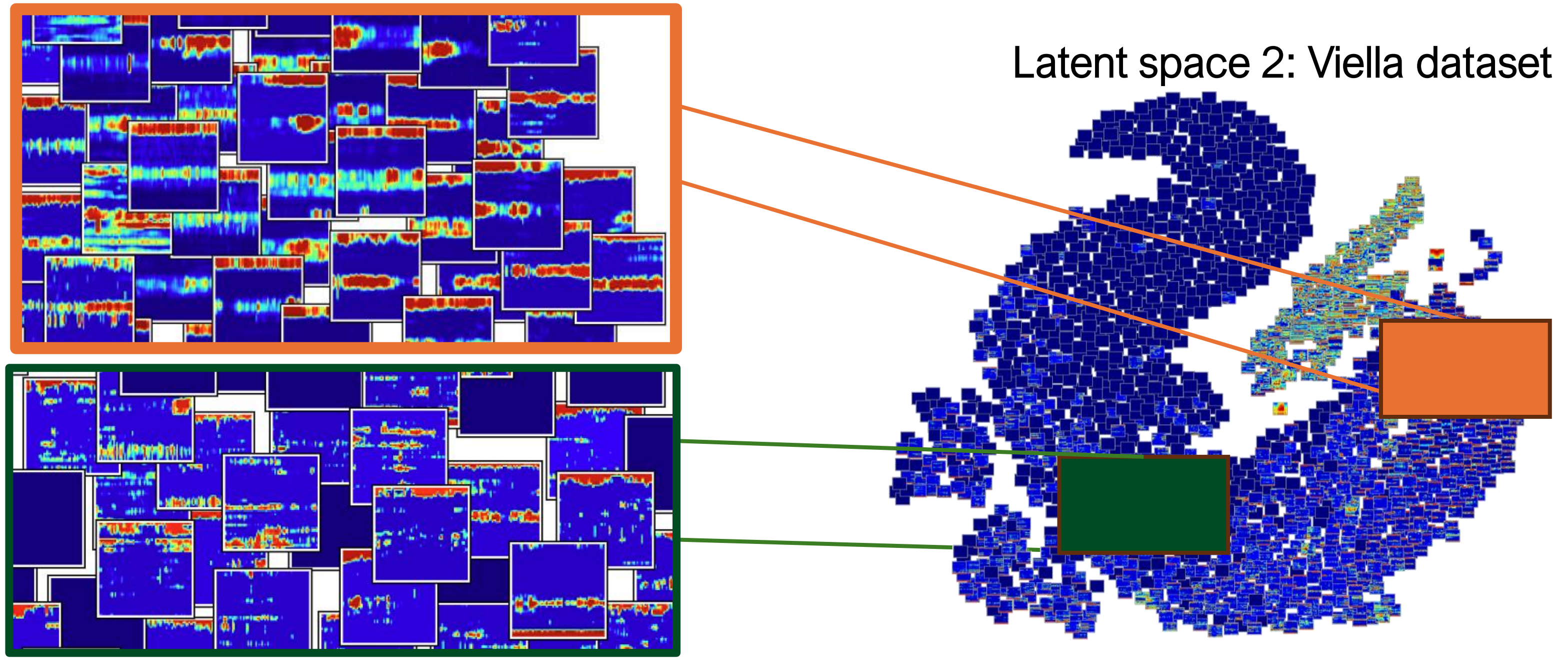
Energy band is represented in each thumbnail.

Key message: Groups of similar data appear.



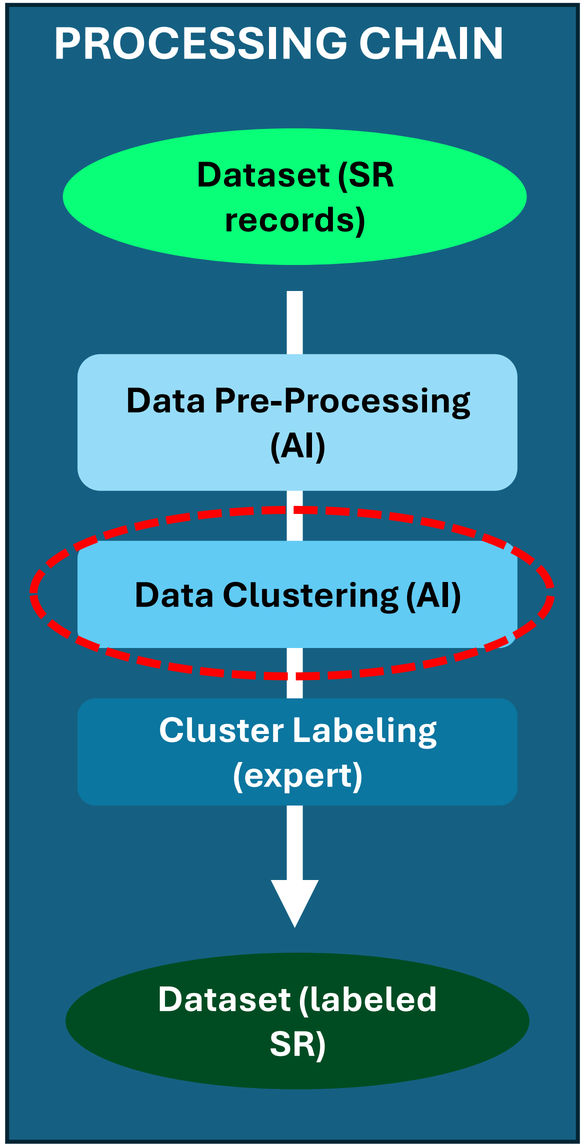
SSL RESULTS

2-D maps, Pyrenees vs Viella



Energy band is represented in each thumbnail.

Key message: Groups of similar data appear.

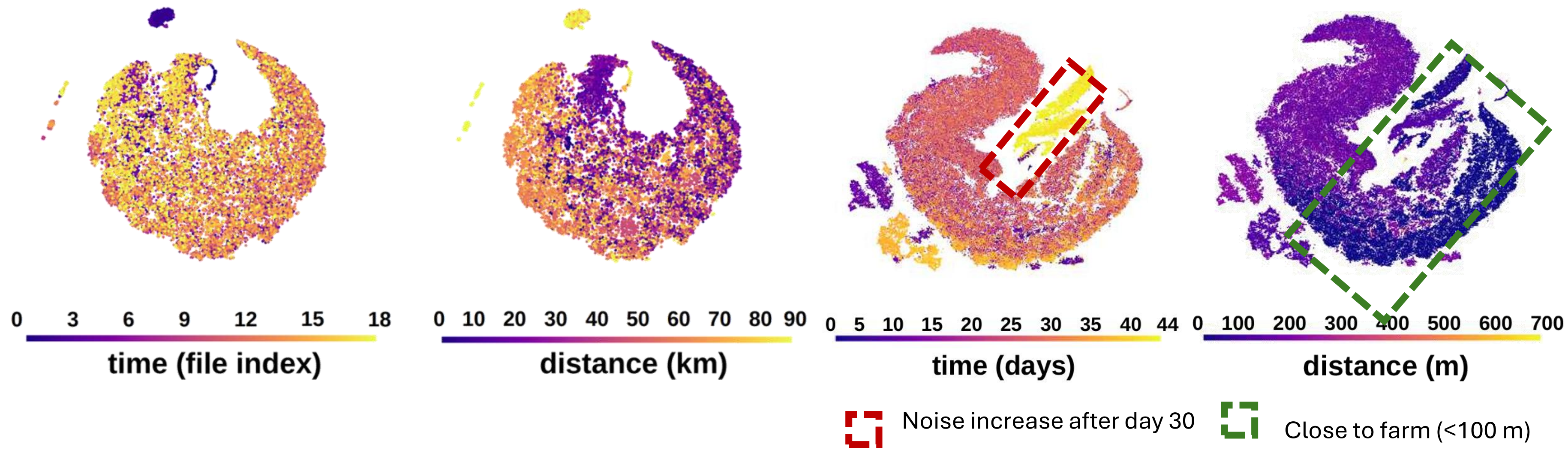


SSL RESULTS

Influence of instrumental and environmental context on DAS data

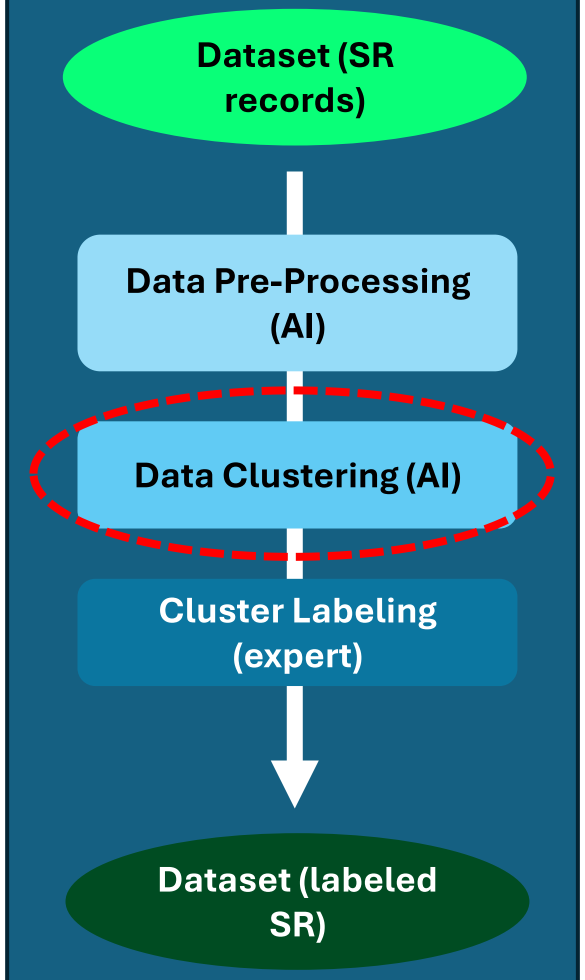
Latent space 1 Pyrenees dataset

Latent space 2 Viella dataset



Key message: The produced representation contains information of the instrumental and environmental context.

PROCESSING CHAIN



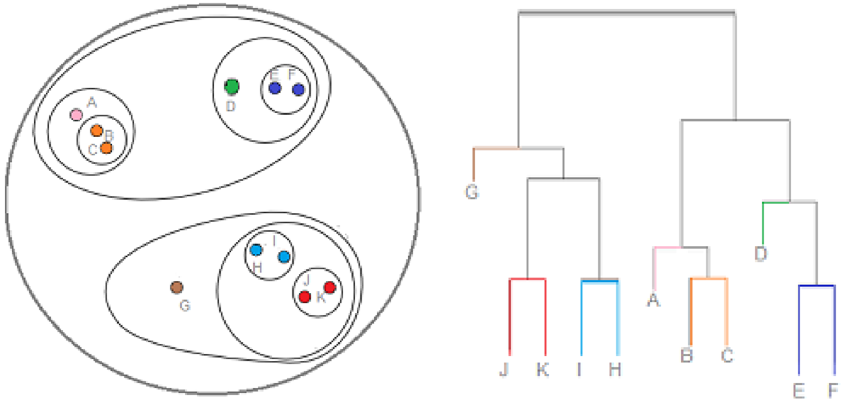
SSL-SPACE INTERPRETATION

Hierarchical clustering: principle and usage

- **Goal**: Propose multi-level data clustering based on their feature representation.

  → Useful for data exploration with no label and to choose the adapted level of details (coarse → fine).

  → The tree can be truncated, each resulting leaves can be independently labelized.



Equivalent representation of the hierarchical clustering result: nested clusters (left), dendrogram (right). From *statisticshowto.com/hierarchical-clustering/*

Key message: Hierarchical clustering helps to manually choose the desired granularity (coarse → fine).

PROCESSING CHAIN

Dataset (SR records)

Data Pre-Processing (AI)

Data Clustering (AI)

Cluster Labeling (expert)

Dataset (labeled SR)

SSL-SPACE INTERPRETATION

TSNE representation after cluster labelization

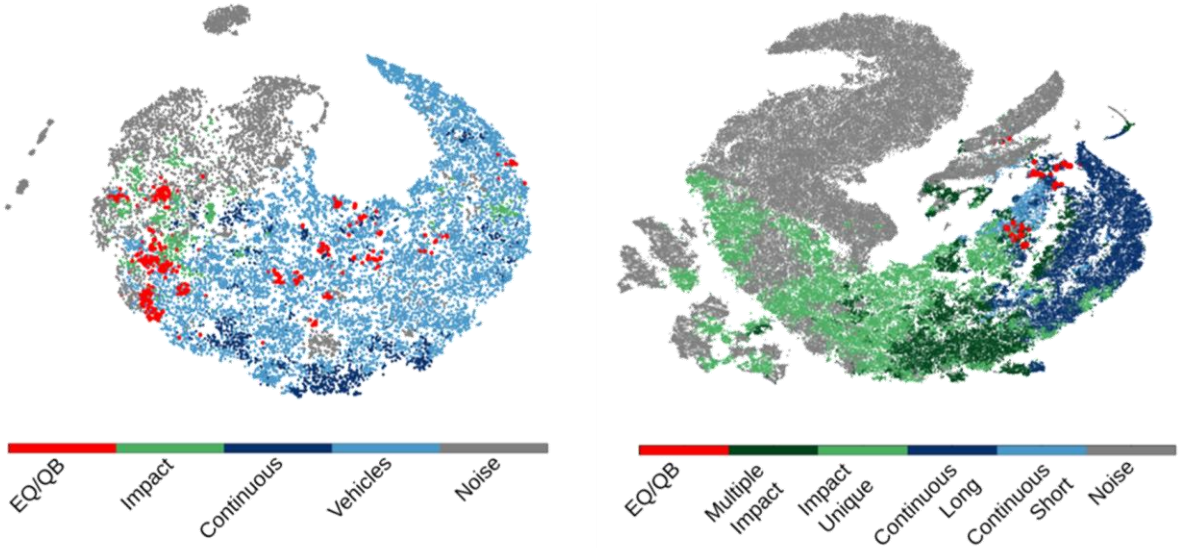
Amount of clusters to labelize:

- Pyrenees: 570.
- Viella: 640.

Amount of class:

- Pyrenees: 5.
- Viella: 6.

Latent space 1: Pyrenees dataset | **Latent space 2: Viella dataset**



Each color represents a label, built by manually aggregating cluster with similar content.

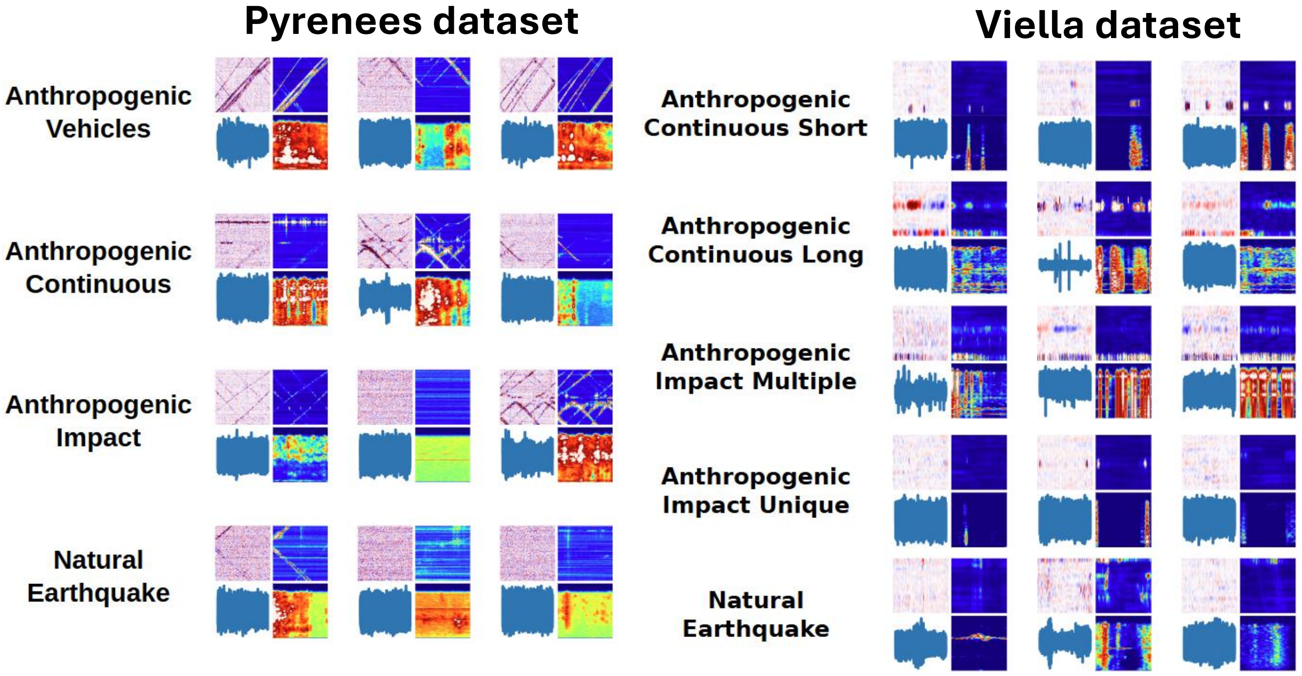
Key message: Hierarchical clustering advantage:
- group-wise label assignment
  → required time for labeling: independent of the amount of data.

PROCESSING CHAIN

Dataset (SR records)
↓
Data Pre-Processing (AI)
↓
Data Clustering (AI)
↓
Cluster Labeling (expert)
↓
Dataset (labeled SR)

RESULTS DISCUSSION

Several examples of clusters



How to read the subplots ? (SR: strain rate, EB: energy band)

SR	EB
Trace	Spectrogram

Key message: Consistent signatures for events of similar class.

PROCESSING CHAIN

Dataset (SR records)

Data Pre-Processing (AI)

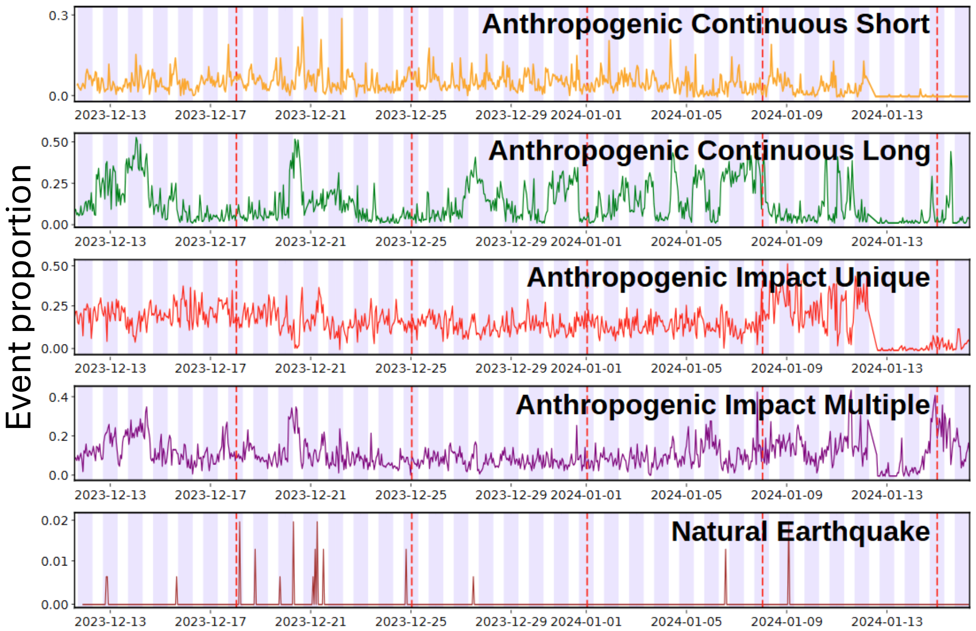
Data Clustering (AI)

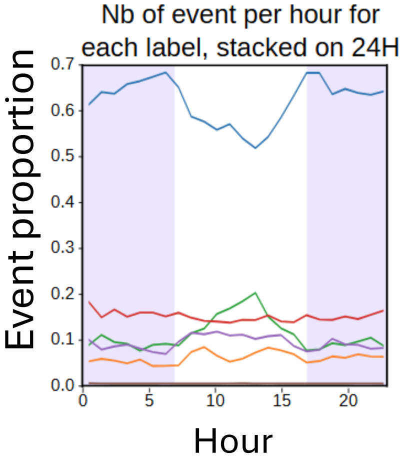
Cluster Labeling (expert)

Dataset (labeled SR)

RESULTS DISCUSSION

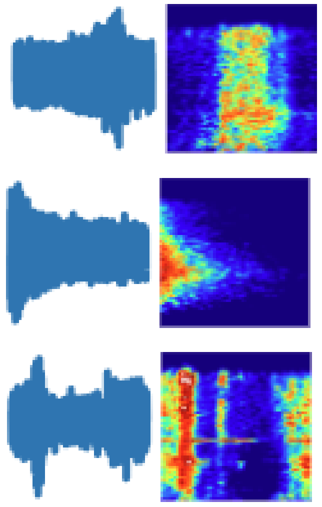
Temporal periodicity – Viella dataset



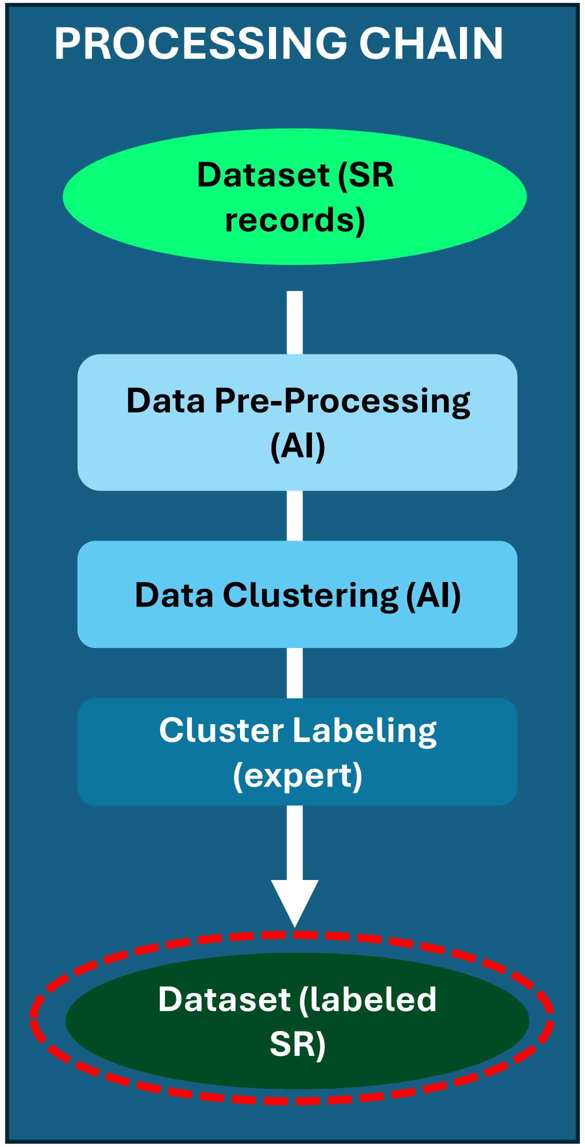


Each color in this plot corresponds to the same event class as one the left plot.

Examples of labelized earthquakes *(Trace & Spectrogram)*



Key message: Clusters reveals routines related to different class of events.



CONCLUSION & PERSPECTIVE

Conclusion:

- Learn robust DAS representations **without labels or prior models**
- Automatically uncover dominant behaviours and rare events.
  → Enable human-in-the-loop interpretation and expert-driven insight.
- Greatly accelerate event labeling at large scale.

Perspective:

- Use labelized clusters to train a supervised model for real-time monitoring.

Take-home message

→ SSL combined with unsupervised learning makes DAS a proactive multi-phenomena monitoring framework.

Contact: camille.huynh@febus-optics.com

Download the scientific paper here!

SCAN ME

Thank You

We Welcome Your Questions!

FEBUS optics

Institut Terre & Environnement de Strasbourg | ITES | UMR 7063
l'Université de Strasbourg cnrs & ENGEES

NORSAR