

Environmental Green AI: the New Zealand TAIAO Project

Albert Bifet

AI Institute, University of Waikato · Director, TAIAO



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato



UC UNIVERSITY OF
CANTERBURY
*Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND*



THE UNIVERSITY OF
AUCKLAND
*Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND*



MetService



Beca



Please note, there is ongoing maintenance and restoration work in progress at Hobbiton Movie Set. [Learn more here.](#)



📍 The Shire's Rest: 501 Buckland Road, Hinuera [Get Directions.](#) [Buy a Gift Voucher here!](#)

[Experiences](#)

[Plan Your Adventure](#)

[Events](#)

[Discover](#)

[Online Shop](#)

[Book Now](#)

OFFICIAL SITE

Step into the enchanting village of Hobbiton Movie Set

[Check Availability →](#)

[Our Experiences →](#)

[FAQs →](#)

Bookings essential



Outline

1

Green AI has two meanings

AI for the environment, and AI that is itself green

2

TAIAO and Aotearoa

The programme, the team, Te Ao Māori, legal persons

3

Research highlights

Continual learning, floods, quantum, forests, weather, solar

4

Open tools

Why streaming is green; MOA, CappyMOA, Weka.ai

5

TuiML and autoresearch

Machine learning for AI agents

6

Open science and beyond

FinAi-Lab

"Green AI" means two things



AI for the environment

Rivers, forests, coasts, species, floods.



AI that is itself green

Small energy, small hardware, small data movement.

TAIAO's position: it has to be both.

AI's footprint is no longer a rounding error

415 TWh

Data centres, 2024

About 1.5% of world electricity

945 TWh

Projected, 2030

Just under 3%: a doubling

30%

Growth per year of AI servers

Half of the net increase to 2030

Source: IEA, Energy and AI (2025)

TAIAO at a glance

Time-Evolving Data Science and AI for Advanced Open Environmental Science

NZ\$13M

MBIE Strategic Science
Investment Fund

7 years

Methods, tools and people

6 partners

Waikato, Canterbury, Auckland,
Victoria, MetService, Beca

The TAI AO team

Meet the team

The TAI AO team brings together a diverse collective of PHD researchers, postdoctoral experts, data scientists, environmental scientists, and software engineers.



Prof. Albert Bifet

Project Leader, AI Institute, University of Waikato



Prof. Karin Bryan

School of Environment, University of Auckland



Prof. Bernhard Pfahringer

AI Institute, University of Waikato



Prof. Eibe Frank

AI Institute, University of Waikato



Assoc. Prof. Te Kaka Keegan

AI Institute, University of Waikato



Prof. Geoff Holmes

Pro-Vice Chancellor - Health, Engineering, Computing and Science,



Prof. Yun Sing Koh

School of Computer Science, The University of Auckland



Dr. Varvara Vetrova

College of Engineering, University of Canterbury



Dr. Heitor Murilo Gomes

School of Engineering and Computer Science, Victoria University of



Dr. Nick Lim

Postdoctoral Fellow, AI Institute, University of Waikato

Four aims, one platform



New data science for time-evolving environmental data



Open tools and open data anyone can use



Grow New Zealand's data science capacity



Advance Vision Mātauranga with iwi and hapū

Te Ao Māori: the environment as family, not resource

Ko au te awa, ko te awa ko au

I am the river, and the river is me.

Kaitiakitanga: guardianship of land and water across generations.
TAIAO works under co-governance, building Māori data science capacity on iwi whenua.



In Aotearoa, a river and a mountain are legal persons

2014

Te Urewera

Former national park, now a legal entity in its own right

2017

Te Awa Tupua

The Whanganui River, with rights, duties and guardians

2025

Te Kāhui Tupua

Taranaki Maunga and its peaks, governed half iwi, half Crown

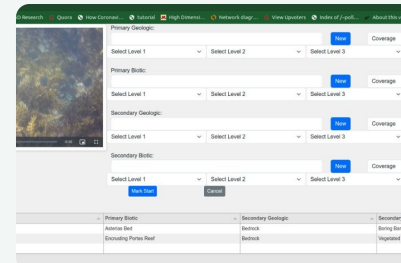
If a river is a person, monitoring it is health care.

AI for the environment: four case studies



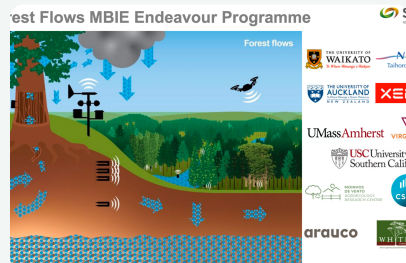
Aotearoa Species Classifier

11,047 endemic species, on a phone, offline



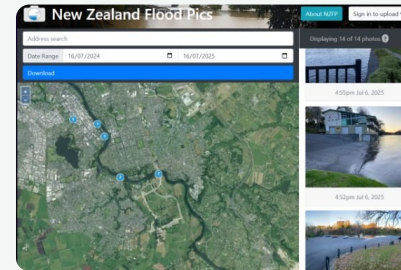
Benthic Habitat Annotator

Reef imagery annotated with DOC



Forest Flows

Water dynamics in planted forests



Flood Prediction

Early warnings for the Coromandel



Continual learning: beyond one memory

PROBLEM

Replay-based continual learning treats memory as one store of past data; the architecture that governs old versus new data is overlooked.

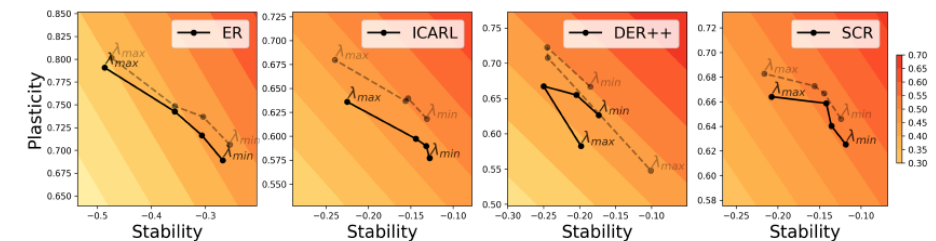
METHOD

A dual-memory system: short-term for quick updates, long-term curated samples for replay; a Memory Capacity Ratio (MCR) links online and offline CL.

RESULT

Smaller MCRs give better stability-plasticity trade-offs and stronger generalisation under equal budgets.

Yaqian Zhang, Bernhard Pfahringer, Eibe Frank, Albert Bifet



Stability vs plasticity across four replay methods

Anomaly detection for ship paths

PROBLEM

Static maritime anomaly detectors go stale as routes, traffic and regulations change, missing illegal fishing, deviations and hazardous operations.

METHOD

CLAISA: continual learning for AIS anomalies, predicting trajectories as routes evolve and adapting to new behaviour without full retraining.

RESULT

Adaptive knowledge bases that keep the detector current as the world changes.



CLAISA predictions on trajectories in the Kattegat

Jack Julian, Yun Sing Koh, Albert Bifet

Flood and hazard platform

PROBLEM

Flooding is New Zealand's most frequent and second-costliest natural disaster; existing systems lack accuracy and lead time.

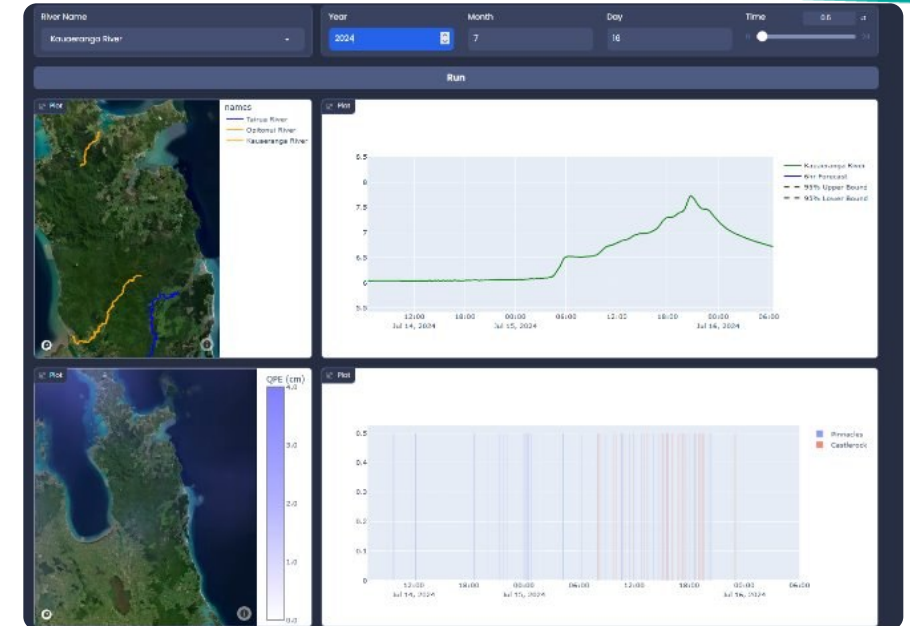
METHOD

visualizer.nzflood.services: real-time forecasting from radar and river data with neural models for 6 to 9 hour predictions; automated flood-image captioning with NZ Flood Pics.

RESULT

A live platform with Waikato Regional Council and NIWA-MetService, and plans for national expansion.

Nick Lim, Phil Mourot • with WRC, NIWA, MetService, Christo Rautenbach, Andrew Lorrey



Real-time forecasts with Waikato Regional Council and NIWA-MetService

Quantum tools for climate resilience

PROBLEM

Insurers and communities need flood risk anticipated precisely, with interpretable and fair parametric insurance decisions.

METHOD

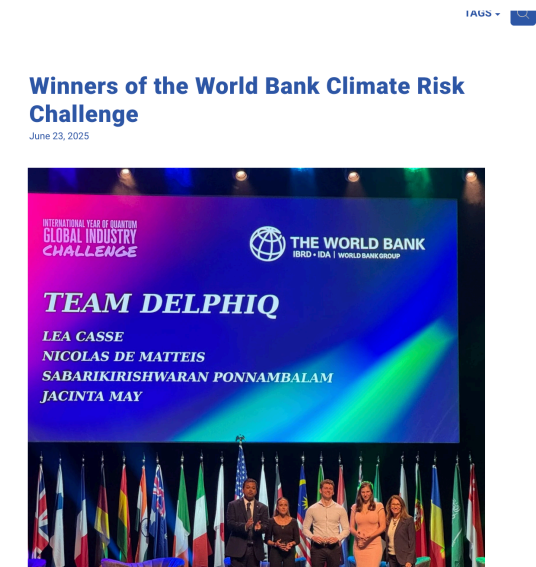
A Quantum Re-Uploading Unit (QRU) forecasts 24 to 48 h flood probability from TAIAO data and MetService rainfall grids; a QAOA-CVaR optimiser allocates insurance cover.

RESULT

Won the World Bank track of the Global Industry Challenge; presented at Quantum World Congress, Washington DC.

Léa Cassé, Sabarikirishwaran Ponnambalam, Jacinta May, Nicolas De Matteis

Source: University of Waikato news; PQIC Global Industry Challenge 2025



Team DelphiQ: winners of the World Bank Climate Risk Challenge, 2025

Explainable tree growth in plantation forests

PROBLEM

Tree-growth forecasts need soil and climate interactions; accurate AI models lack transparency, limiting trust.

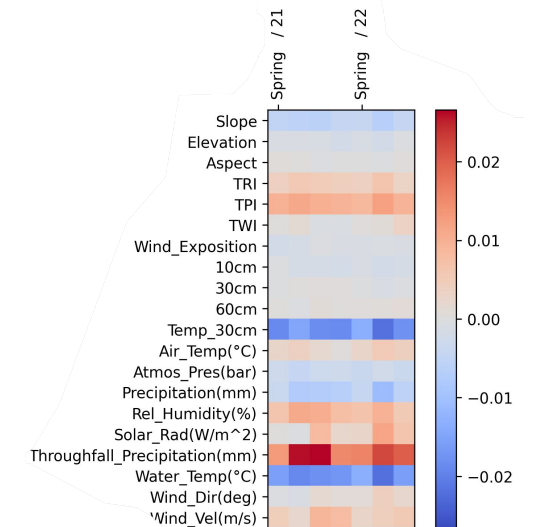
METHOD

TRE-EX: a model-agnostic occlusion framework over dendrometer, soil, climate and topographic data, with temporal sensitivity analysis by systematic perturbation.

RESULT

Biologically plausible drivers of daily radial growth of *Pinus radiata*, alongside accurate forecasts.

Guilherme Cassales, Serajis Salekin, Dean Meason, Nick Lim • with Scion



Feature attribution across spring 2021 and 2022

Global weather forecasting in a single epoch

PROBLEM

Weather is governed by processes at many spatio-temporal scales; current graph forecasters forget global trends and are expensive to train, with large CO2 emissions.

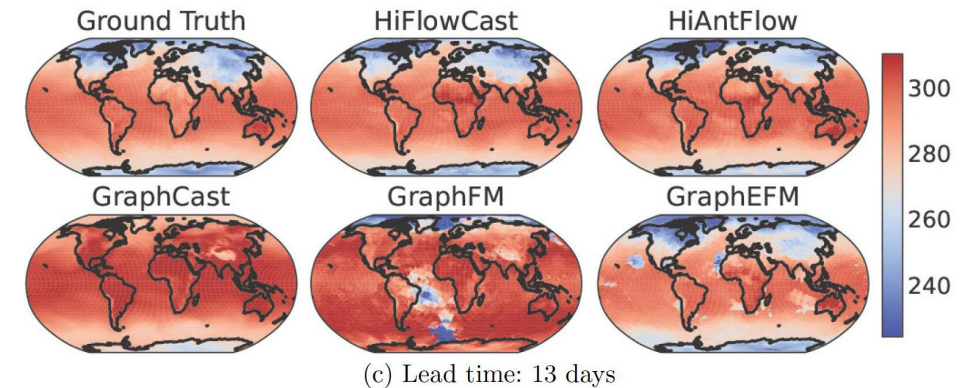
METHOD

HiFlowCast: a hierarchical graph neural network that predicts fields at several resolutions and embeds solar-radiation forcing across them.

RESULT

With pretrained weights the model trains in a single epoch: a green forecaster by construction.

Thomas Bailie, Yun Sing Koh, S. Karthik Mukkavilli, Varvara Vetrova



Ground truth vs HiFlowCast, GraphCast, GraphFM at 13 days lead

Real-time solar forecasting at the edge

PROBLEM

Solar power and price forecasts need low latency; cloud pipelines add delay, and edge devices have little memory.

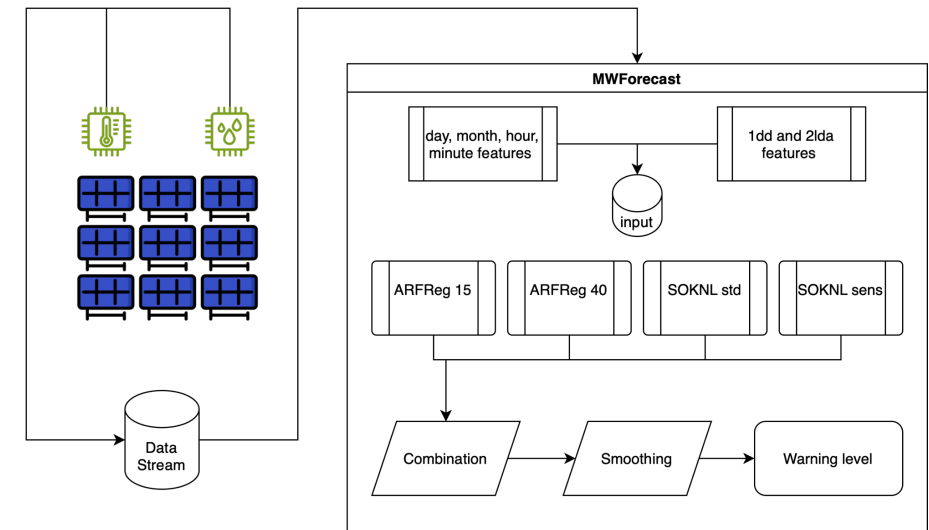
METHOD

Stream algorithms (ARFReg, SOKNL) benchmarked against baselines; MWForecast, a lightweight edge architecture combining models for 24 h forecasts and warnings under single-pass, low-memory constraints.

RESULT

Adaptive predictions on resource-constrained devices, critical for grid stability and renewables.

Guilherme Cassales, Ioan Petri, Heitor Gomes, Omer Rana, Albert Bifet



MWForecast: data stream, models, combination, smoothing, warning level

Why stream learning is green by construction



BATCH RETRAINING

- Store all n examples
- Retrain: $O(n)$ per update
- Memory grows with n



STREAM LEARNING

- See once, then discard
- Update: $O(1)$ per example
- Memory bounded, drift-aware

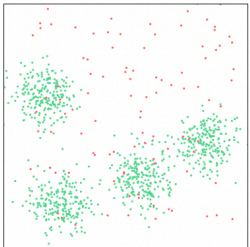
Energy $\propto$ data rate, not archive size. The sensor sets the budget.

Anomaly detection for streams

Online Isolation Forest

Leveni et al. • ICML 2024

Multi-resolution histograms learn new points and forget old ones. Space $O(\tau\sqrt{(\omega/\eta) + \omega})$: independent of stream length.

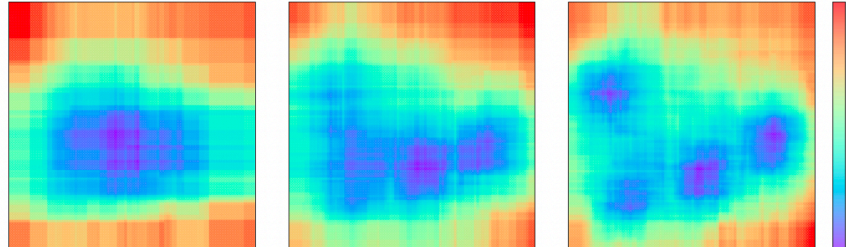


(a) Data stream $\mathbf{x}_1, \dots, \mathbf{x}_t \in \mathbb{R}^d$.

Streaming Isolation Forest

Liu et al. • PAKDD 2025

Trees updated incrementally; sliding window for drift; no history stored.



(b) Anomaly scores s at different time instants t , from left to right.

Adaptive Isolation Forest

Liu et al. • DS 2025

MinTreeMaxMass scores each tree for replacement: stability vs plasticity, linear time.

Anomaly scores refine as the stream evolves, in bounded memory.

Open-source tools

Tool	What it is	Why it is green
MOA	Data stream mining framework (Java)	Bounded memory
CapyMOA	Stream learning in Python	Efficient by design
River	Online ML in Python	Lightweight models
Weka / Weka.ai	ML workbench, now on the web with GPU support	Speed, simply
TuiML	Machine learning for AI agents	Fewer wasted runs



CapyMOA



Machine learning library tailored for data streams. Featuring a Python API tightly integrated with MOA ([Stream Learners](#)), PyTorch ([Neural Networks](#)), and scikit-learn ([Machine Learning](#)). CapyMOA provides a **fast** python interface to leverage the state-of-the-art algorithms in the field of data streams.

CapyMOA: stream learning in Python

See each example once. Update in constant time. Never store the past.

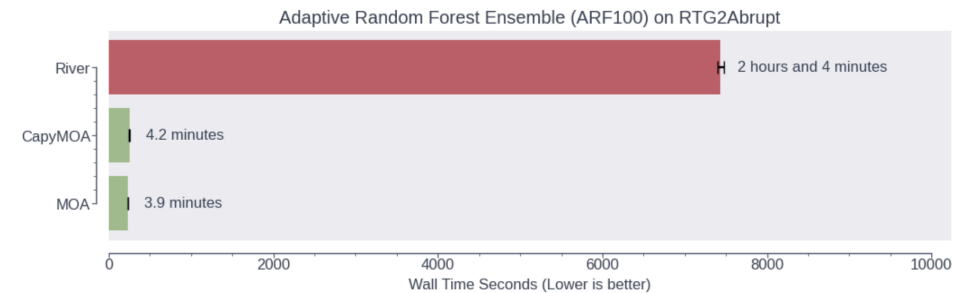
```
from capymoa.classifier import HoeffdingTree
ht = HoeffdingTree(schema=stream.get_schema())
while stream.has_more_instances():
    x = stream.next_instance()
    y_hat = ht.predict(x) # test
    ht.train(x) # then train
```

capymoa.org · pip install capymoa

CapyMOA

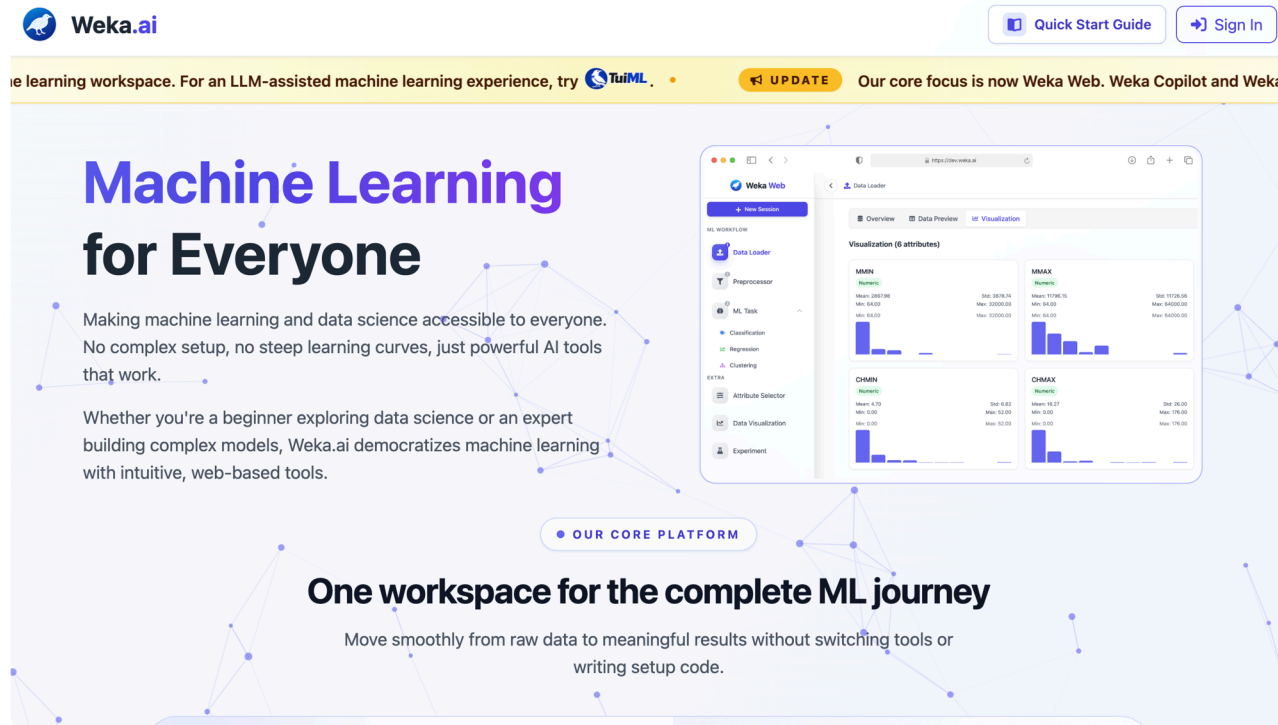


Machine learning library tailored for data streams. Featuring a Python API tightly integrated with MOA (**Stream Learners**), PyTorch (**Neural Networks**), and scikit-learn (**Machine Learning**). CapyMOA provides a **fast** python interface to leverage



ARF ensemble of 100 trees, same task: River 2 h 4 min · CapyMOA 4.2 min · MOA 3.9 min

Weka.ai: machine learning for everyone



The screenshot shows the Weka.ai website with a teal header. The main content area has a light blue background with a network diagram. On the left, the text reads: 'Machine Learning for Everyone', 'Making machine learning and data science accessible to everyone. No complex setup, no steep learning curves, just powerful AI tools that work.', and 'Whether you're a beginner exploring data science or an expert building complex models, Weka.ai democratizes machine learning with intuitive, web-based tools.' In the center, there's a screenshot of the Weka Web interface showing a 'Data Loader' sidebar and a 'Visualization' panel with four histograms labeled MIN, MAX, CHMIN, and CHMAX. Below this, a button says 'OUR CORE PLATFORM'. At the bottom, the text reads: 'One workspace for the complete ML journey' and 'Move smoothly from raw data to meaningful results without switching tools or writing setup code.' The top navigation bar includes 'Weka.ai', 'Quick Start Guide', 'Sign In', and a banner about 'TuiML' and 'Weka Copilot'.

Machine Learning for Everyone

Making machine learning and data science accessible to everyone. No complex setup, no steep learning curves, just powerful AI tools that work.

Whether you're a beginner exploring data science or an expert building complex models, Weka.ai democratizes machine learning with intuitive, web-based tools.

OUR CORE PLATFORM

One workspace for the complete ML journey

Move smoothly from raw data to meaningful results without switching tools or writing setup code.

WEKA WEB

The full Weka toolkit in the browser, no setup

WEKA COPILOT

An LLM assistant for code, analysis and interpretation

WEKA TERMINAL

A Python environment for experiments

Weka: 10 million downloads, 18,000 citations. Now one workspace for the whole ML journey.

TuiML in practice

 ALPHA

ContributeGetting StartedAPI ReferenceTutorialsBenchmarksChangelog

Machine learning for AI agents

Ask your agent to train a model, tune it, compare it to the last run, or find an algorithm that fits your data. It just does it. No code. No guesswork. No forgotten context.

```
~ agent · tuiml connected [🔥 claude]

> Find the best classifier for iris, tune it, and show me a comparison.
→ tuiml_list { category: "algorithm", type: "classifier" }
[ok] 99 classifiers · picked 3 candidates for a small numeric dataset
→ tuiml_benchmark { algorithms: ["RandomForestClassifier", "XGBoostClassifier", "weka.J48"], data: "iris", ...
RandomForestClassifier ██████████ 0.967
XGBoostClassifier ██████████ 0.953
weka.J48 ██████████ 0.947
→ tuiml_tune { algorithm: "RandomForestClassifier", param_grid: { n_estimators: [100, 300], max_depth: [8, ...
[ok] best: n_estimators=300, max_depth=8 · accuracy 0.973 (+0.6pp)
[tuiml] RandomForest, tuned to 0.973 — clear winner. Want it served?
> Yes — serve it and verify with one prediction.
→ tuiml_serve_model { model_id: "7a3c91f2", port: 8080 }
[ok] live at http://127.0.0.1:8080/predict · server_id srv_7f2e
```

ASK

Find the best classifier for this data, tune it, compare to the last run

THE AGENT CALLS

`tuiml_list`, `tuiml_benchmark`, `tuiml_tune`,
`tuiml_serve_model`

YOU GET

A tuned model served locally, with every step traced

TuiML: one typed contract for every algorithm

AGENT CALLS A TOOL

```
tuiml_train(  
  algorithm="random_forest",  
  params={"n_estimators": 200},  
  data="river_gauges.csv",  
  target="flood", seed=0)
```

RUNTIME ANSWERS WITH STRUCTURE

```
{ "model_id": "m_2f9a",  
  "metrics": {"f1": 0.81, "auc": 0.93},  
  "seed": 0, "trace_id": "t_118" }
```

REGISTRY

Every algorithm, transform and metric has a name and a JSON Schema for its parameters

VALIDATED BEFORE ANY FIT

Unknown name, misplaced parameter or unsupported estimator: a structured error with a suggested fix, zero compute spent

SAME CONTRACT EVERYWHERE

MCP, Python API, CLI and local serving all drive the same specification layer

TuiML: reproducible and extensible by construction

Reproducible

- Every call records its random seed
- Lineage: which data and transforms produced each model
- JSON Lines trace: arguments, duration, status
- Session exports as a runnable notebook

Extensible by agents

- `tuiml_algorithm_skeleton`: a typed template
- `tuiml_create_algorithm`: register a new learner
- New components are searchable and callable at once
- `tuiml_self_update`: the runtime updates itself

Green consequence: no malformed run ever executes, and no result is ever recomputed.

Autoresearch: an agent runs the research loop in TuiML



```
baseline = experiment(algo, data, seed=0)
for step in range(N):
    cand = agent.propose(algo, research_log)
    score = experiment(cand, data, seed=0)
    if score > baseline: algo, baseline = cand, score
```

Same seed, same data, one metric: only real improvements survive. Every run traced, exported as a notebook.

Open data, open notebooks, open education



Datasets

Open licences, searchable by tag



Notebooks

Run once, reuse everywhere



AI Education NZ

Free courses with the AI Forum NZ

Every duplicated experiment is energy spent twice.

Beyond the environment: the same ideas in finance



FinAi-Lab: Télécom Paris, Institut Polytechnique de Paris, BNP Paribas

FRAUD DETECTION

Transaction streams monitored for anomalies in real time

TRADING

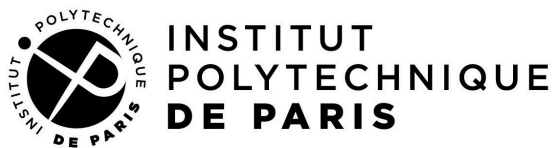
Algorithmic decisions on live market data

CYBER-ATTACK DETECTION

Anomalies in network traffic, stopped as they happen

TELECOM AND IOT

Network optimisation to prevent congestion and predict outages



BNP PARIBAS



NGĀ MIHI • THANK YOU

Environmental Green AI: the New Zealand TAI AO Project

Albert Bifet

AI Institute, University of Waikato • Director, TAI AO



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato



THE UNIVERSITY OF
AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND



MetService

