GeoAI Challenge: everything happens somewhere – applying machine learning to geospatial analysis

Maria Antonia Brovelli

maria.brovelli@polimi.it

POLITECNICO
MILANO 1863

From Maps to the Digital Twin Earth and the Metaverse

[Mappedimappe: Tabula Peutingeriana (luciodp.altervista.org)]

From Maps to the Digital Twin Earth and the Metaverse

Credits: ESA - Digital Twin Earth, quantum computing and AI take centre stage at ESA's Φ-week

Everything happens somewhere

Sky

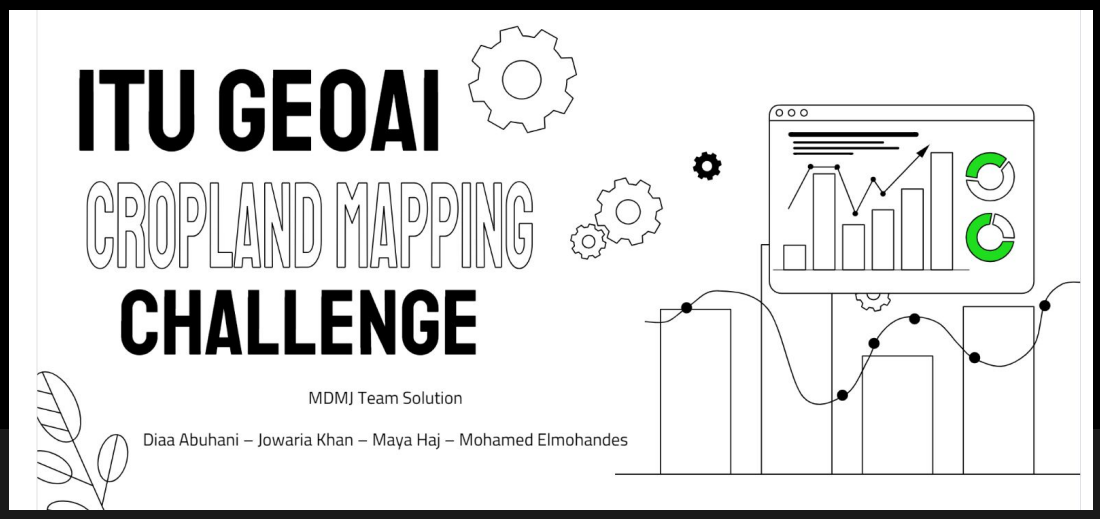
Optical (images & videos), LIDAR, SAR, Multi/Hyper Spectral and GNSS

Ground

User Generated Content

Water + below surface level

Optical (images & videos), SONAR, LIDAR and Bathymetric



Jowaria Khan

Maria Brovelli

Diaa Abuhani

Mohamed ElMohandes

Maya Haj Hussain

Andrea Manara

KYOUNG-SOO EOM

Zhongxin Chen

Pengyu Hao

Julien Smith

ITU

GEOAI Challenge 2023

Compute platform

ITU provides a state-of-the-art, free-of-charge compute platform to participants of the Challenge who do not have adequate access to compute in their respective institutions. The compute platform will provide participants with access to:

- Free GPUs and CPUs
- Hosted Jupyter notebook server
- Python kernel
- Pre-installed machine learning packages, e.g. PyTorch and Tensorflow

GeoAI Challenge Timeline

7 July 2023
Start

31 October 2023
Deadline Project

30 November
Evaluation

December
Challenge Finale

Challenge 1

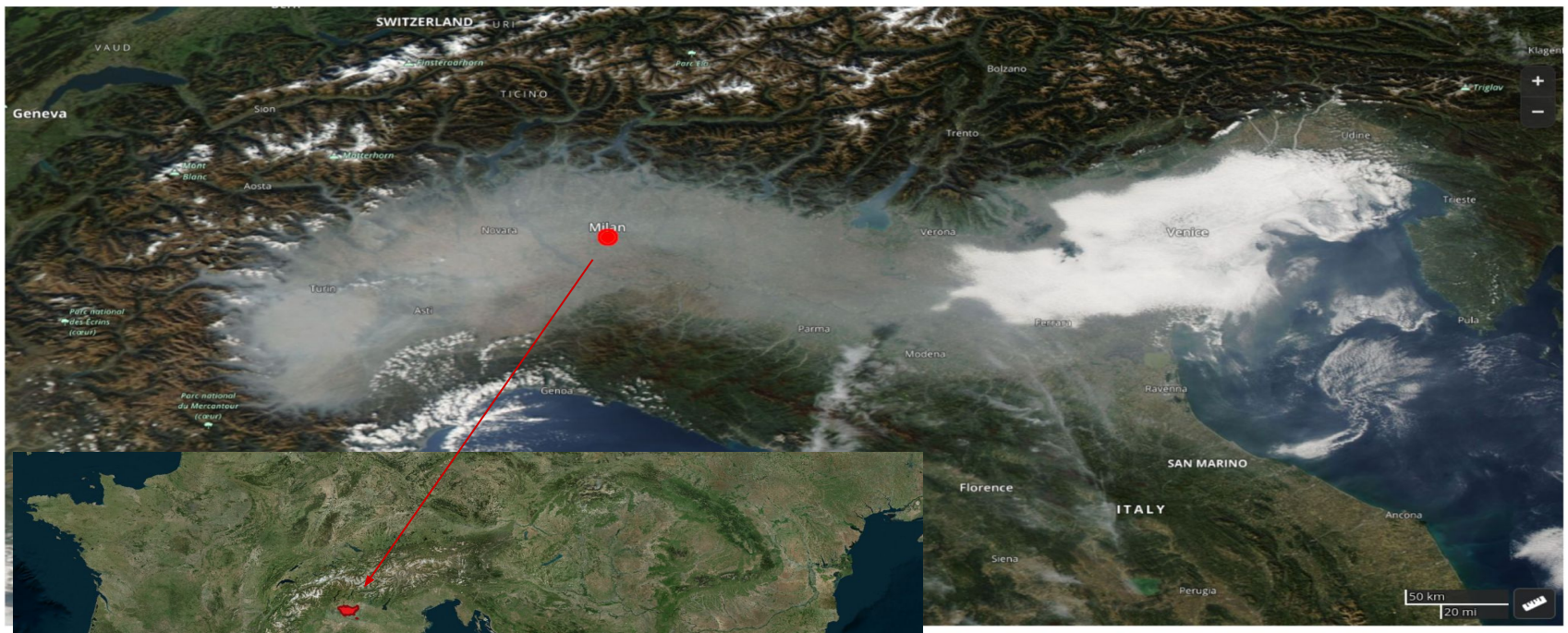
Air pollution susceptibility mapping in Milan, Italy

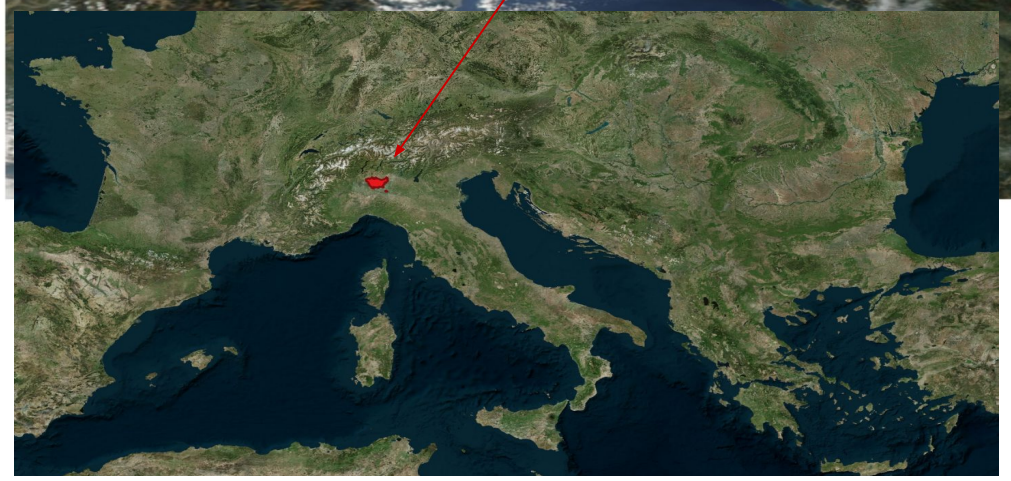
POLITECNICO MILANO 1863

Angelly Pugliese-Viloria

PhD student in Environmental Engineering at Polimi and consultant at the World Bank.

GEOLaB Geomatics and Earth Observation lab





11 SUSTAINABLE CITIES AND COMMUNITIES

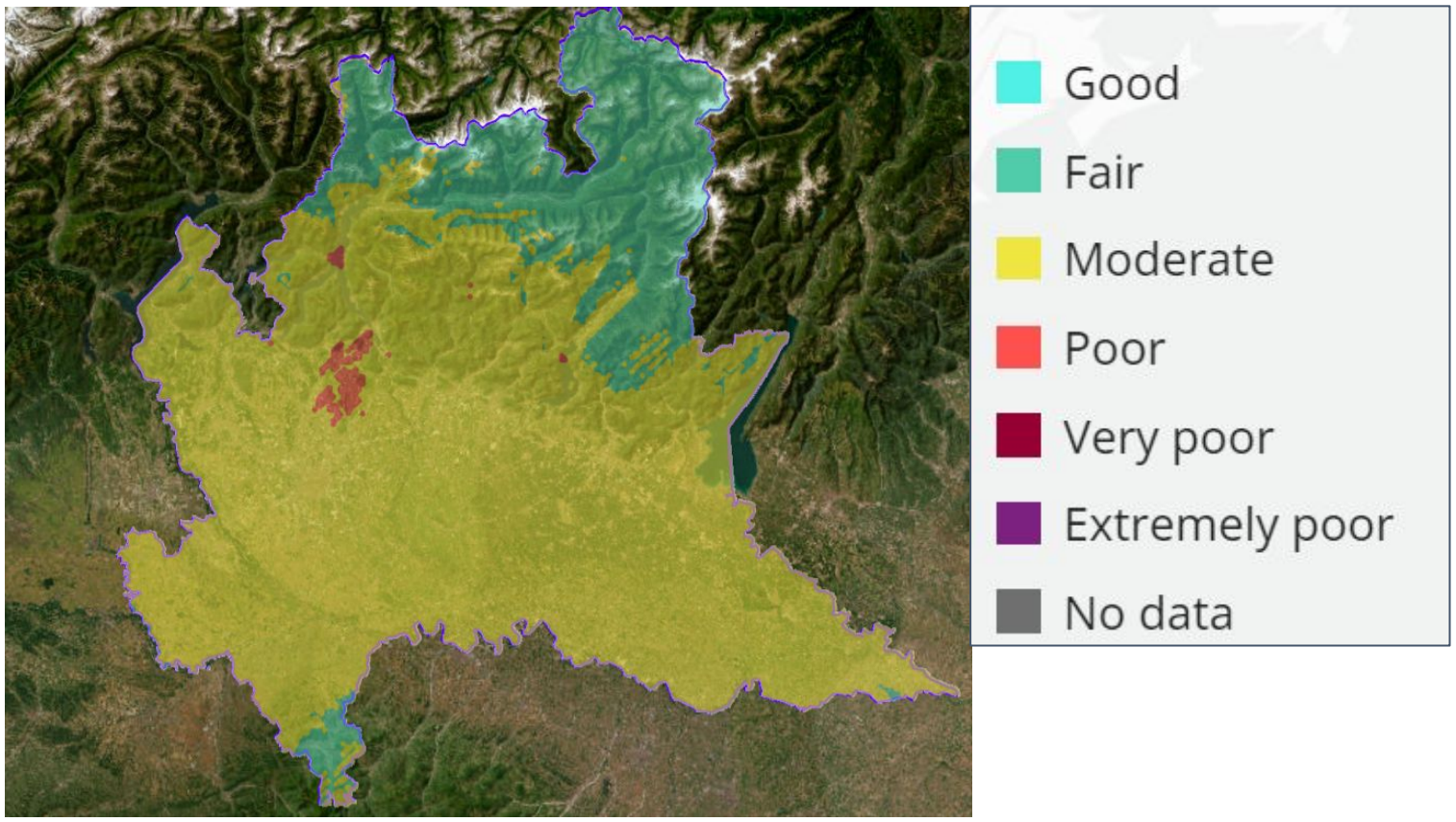
Make cities and human settlements inclusive, safe, resilient and sustainable

13 CLIMATE ACTION

Take urgent action to combat climate change and its impacts

Challenge 1

The objective is to produce seasonal air pollution susceptibility maps in the city of Milan using machine learning which will support decision-making to improve the resilience of the city.



Air Quality Index (AQI) classification based on the European Environment Agency (EEA) guidelines.

To keep in mind

- All the data provided - and additionally used - must be open.
- The proposed approach must be reproducible, publicly available, and is limited to Python.

Time series data
- Pollutants (NO2, SO2, O3, PM2.5, PM10), **individual AQI**, and general AQI.
- Meteorological variables (temperature, precipitation, relative humidity, solar radiation, wind speed and direction).

‘Static’ data
- Digital terrain model.
-Land cover.
-Population density.
-Road, railways, and underground network, etc

The data (points) is provided by ARPA Lombardia (local environment agency) and Geoportale Lombardia. Participants are encouraged to incorporate additional relevant open data (e.g., ERA5 or aerosol data).

Challenge 2

Landslide susceptibility mapping in Northern Italy

POLITECNICO MILANO 1863

Vasil Yordanov, PhD

Research Fellow at the Department of Civil and Environmental Engineering at Politecnico di Milano..

GEOLaB Geomatics and Earth Observation lab

11 SUSTAINABLE CITIES AND COMMUNITIES

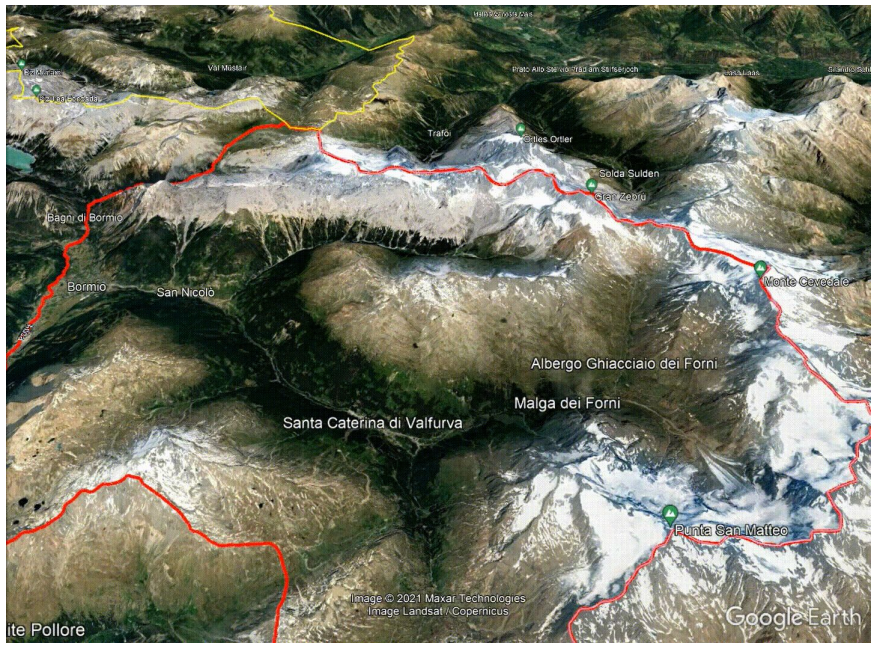
Make cities and human settlements inclusive, safe, resilient and sustainable

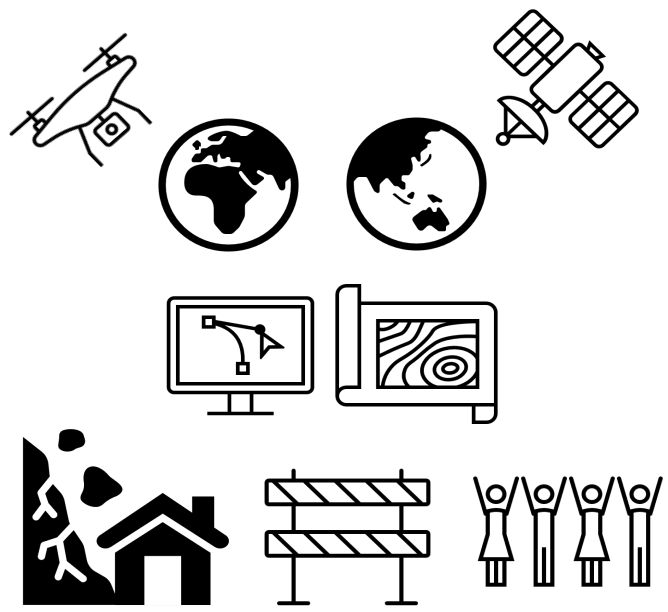
13 CLIMATE ACTION

Take urgent action to combat climate change and its impacts

Challenge 2

Developing machine learning algorithms that can analyze large datasets and identify patterns indicating high probability of landslide occurrence in the Valtellina Valley, Northern Italy.





To keep in mind

- All the data provided - and additionally used - must be open.
- The proposed approach must be reproducible, publicly available, and is limited to Python and GEE JavaScript.

Provided datasets:

The base training data is provided by ISPRA and filtered for the needs. Basic zero-case training data is produced by GISGeoloab@PoliMI.

Participants are **encouraged** to incorporate additional relevant open data

Additional datasets are made available:

- a Digital Terrain Model (DTM) in a raster format at a 5 m/pixel scale (*source: Lombardy Region*),
- a road network at a 1:10,000 scale in vector format (source: Lombardy Region)
- a river network at a 1:10,000 scale in vector format (source: Lombardy Region),
- geological fault zones map at a 1:10,000 scale in vector format (source: Lombardy Region),
- a land use/land cover map at a 1:10,000 scale in vector format (source: Lombardy Region),
- Interpolated yearly averaged hour precipitation for the year of 2020 (source: ARPA Lombardia),
- 90th percentile of the hour precipitation for the year of 2020 (source: ARPA Lombardia).

Challenge 3

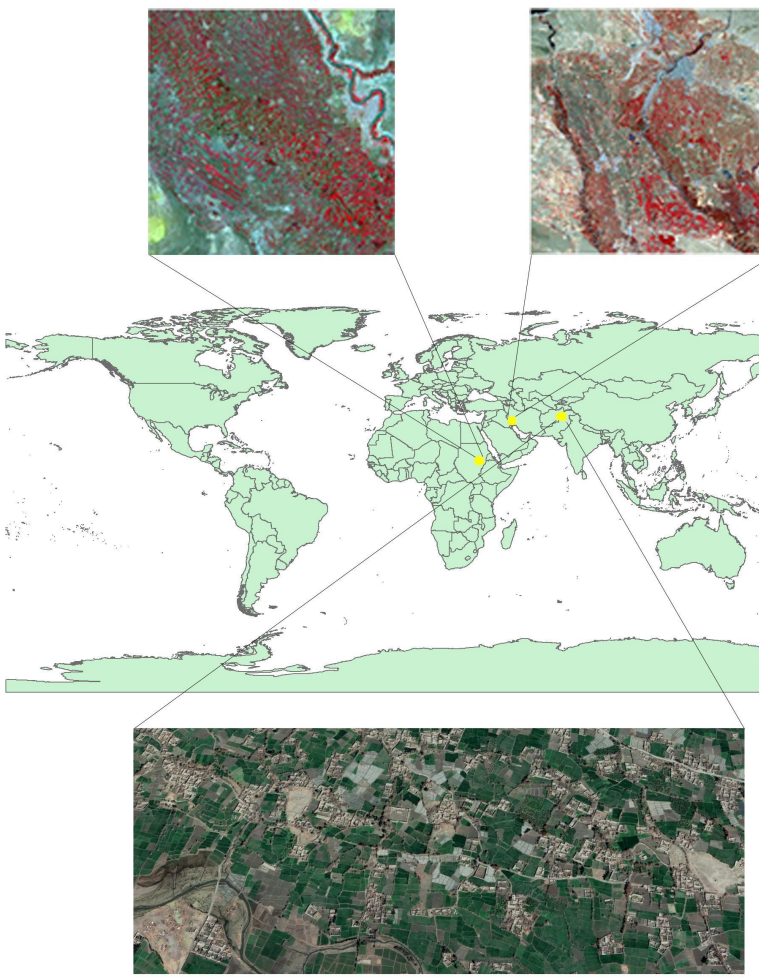
Cropland mapping with satellite imagery

Dr. Pengyu Hao
Information Technology Officer
Digitalization and Informatics Division, FAO

Lorenzo Vita
Geospatial Information Officer
Research and Trend Analysis Branch, UNODC

Food and Agriculture Organization of the United Nations

UNODC
United Nations Office on Drugs and Crime



- Participants will be challenged to develop algorithms for cropland mapping in three test regions.
- The Challenge are cataloged into two sections:

(1) Yearly cropland extent mapping in Iran and Sudan

(2) Temporal cropland extent mapping in AFG

Section 1
Yearly cropland extent mapping in Iran and Sudan

- Objective:

Using time series satellite to identify the yearly cropland extent of the test region

- We provide:

(1) 15-day composited satellite data collected between July 2019 and June 2020,

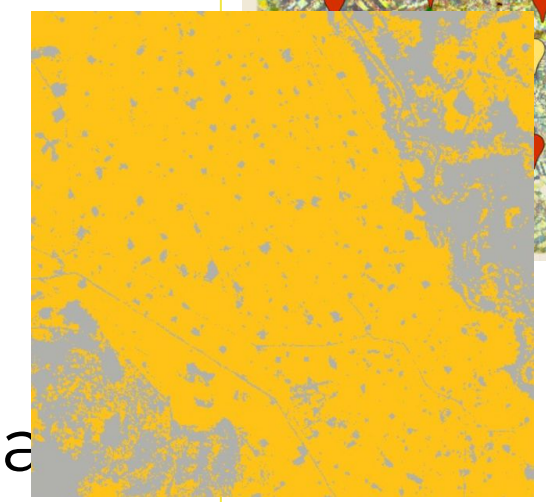
(2) A part of the training data

(3) Example script

- Participants' tasks:

(1) Cropland mapping of the ta time period.

(2) Sharing script and document of proposed workflow.



Section 2
Temporal cropland extent mapping in AFG

Objective:

To develop a model for classifying cropland extent in Nangarhar province for a specific season of the year.

•We provide:

(1) A part of the training data,

(2) An example of python script for processing and analyzing Sentinel-2 images on the Earth Engine platform in the Google cloud.

Participants' tasks:

(1) Delivering cropland masks for the periods April-May 2022 and 2023.

(2) To share full dataset scripts used to develop the model.

Challenge 4

The HYPERVIEW Challenge: Estimating Soil Parameters from Hyperspectral Images

esa KP LABS Mission Complete QZSolutions

This challenge contributes to the SDGs 2nd 3rd and 15th.

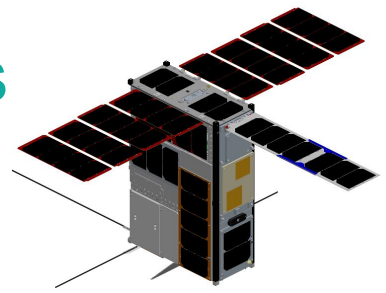
Revolutionizing Future Farming with In-Orbit Processing

- KP Labs, in collaboration with ESA and partner QZ Solutions, presents an extraordinary challenge to revolutionize farming: The goal is to enhance agricultural management practices through Earth observation and artificial intelligence.
- This innovation aims to ensure farm sustainability and enable affordable, planet-friendly agriculture.
- The challenge focuses on advancing soil parameter retrieval from hyperspectral data for the Intuition-1 mission: Intuition-1 is a 6U-class satellite designed by KP Labs, featuring a hyperspectral instrument and onboard AI processing capabilities.
- It will be the world's first satellite with advanced processing power for hyperspectral images in orbit.

This challenge contributes to the SDGs 2nd 3rd and 15th.

HYPERVIEW – data & challenge

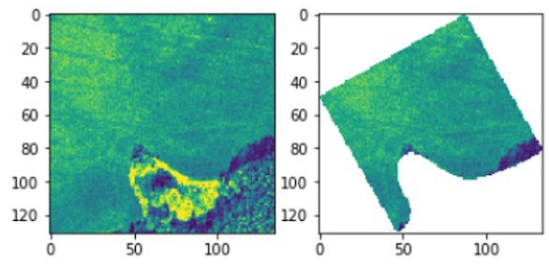
Toward scalable soil analysis



The data

- **Hyperspectral images** acquired on March 3, 2021, using the imagers mounted on the Piper PA-31 Navajo aircraft.
  SWIR-384: 288 bands, 930–2500 nm (5.45 nm resolution),
  VNIR-1800: 186 bands, 400–1000 nm (3.26 nm resolution).
  - **Pre-processing:** radiometric, geometric and atmospheric correction, mosaicing, patch extraction (the patches of various sizes, 60×60 pixels on average) to capture 0.5–4 ha fields of interest (Fig. 1).
  - **Image data was resampled to match the images acquired on board Intuition-1:** 150 bands, 462–942 nm (3.2 nm).
- **In-situ measurements** acquired in Summer/Autumn 2020 (August 2020–November 2020) using the Mehlich-3 methodology: 12 soil samples mixed together and analyzed in the laboratory setting to obtain one tuple of four ground-truth values per area (K, P, Mg, and pH).

- **Challenge objective: To develop a solution for automated estimation of four soil parameters (K, P, Mg, and pH) for each field of interest based on the provided ground-truth data (hyperspectral images and *in-situ* measurements).**
  - **Training data:** 1732 patches with ground truth.
  - **Test data (ψ):** 1154 patches (the ground truth for the test set is not available to the participants).



J. Nalepa, B. Le Saux, N. Longépé, L. Tulczyjew, M. Myller, M. Kawulok, K. Smykala; Michal Gumiela, "The Hyperview Challenge: Estimating Soil Parameters from Hyperspectral Images," 2022 IEEE International Conference on Image Processing (ICIP), 2022, pp. 4268-4272, doi: 10.1109/ICIP46576.2022.9897443.
(https://ieeexplore.ieee.org/document/9897443)

Challenge 5

Location Mention Recognition (LMR) from Social Media Crisis-related Text, Qatar

Hamad Bin Khalifa University (HBKU), a member of Qatar Foundation for Education, Science, and Community Development (QF), was founded in 2010 to continue fulfilling QF's vision of unlocking human potential.

Reem Suwaileh
Research Associate @ HBKU

Qatar Computing Research Institute (QCRI) is a national research institute, established in 2010 by QF. QCRI operates under the umbrella of Hamad bin Khalifa University (HBKU).

Muhammad Imran
Senior Scientist @ Qatar Computing Research Institute (QCRI-HBKU)

Qen Labs © Inc. is a Geospatial AI (GeoAI) company focused on enabling and measuring growth towards Sustainable Development Goals (SDG) through the fusion of geospatial data with other public and private datasets. They develop AI algorithms that leverage the spatiotemporal attributes of both structured and unstructured data to provide actionable insights.

Lokendra Chauhan
Founder & CEO, Qen Labs Inc.

This challenge contributes to the SDGs 10 and 11.

LMR from Social Media Crisis-related Text

Overview

At disaster onset, tweets with critical information such as medical assistance, food or shelter needs, severely damaged infrastructure, are useful for response authorities.

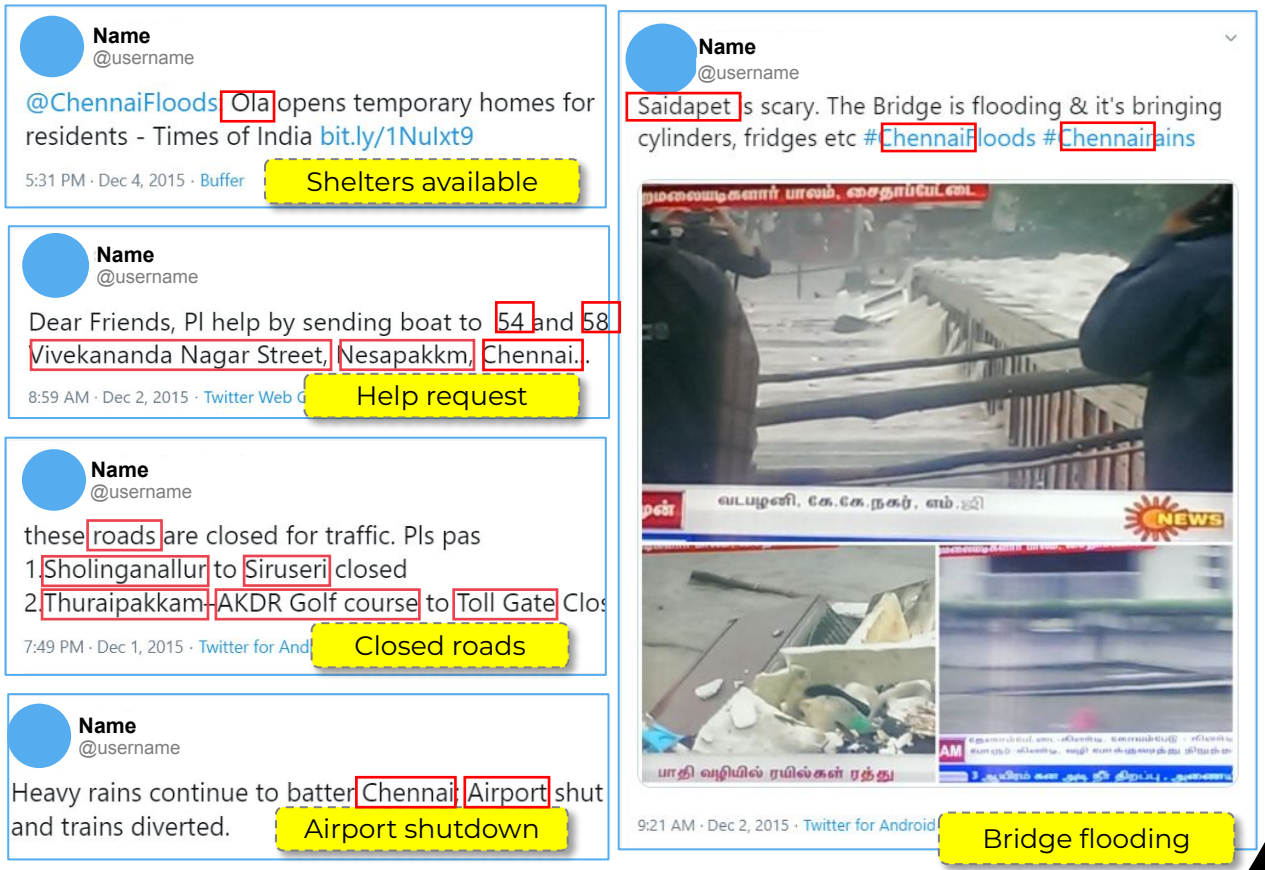
An essential dimension that makes Twitter data invaluable and actionable is the geolocation information.

Nevertheless, Twitter had removed the geotagging feature from tweets since June 2019, which necessitates the development of geolocation extraction tools.

We focus on the ***Location Mention Recognition*** task that has two setups:

- ***Type-less***: detecting LMs regardless of their location types.
- ***Type-based***: detecting LMs while distinguishing between their types.

Example Tweets with Location Mentions (LMs)



10 REDUCED INEQUALITIES

11 SUSTAINABLE CITIES AND COMMUNITIES

This challenge contributes to the SDGs 10 and 11.

LMR from Social Media Crisis-related Text

Objectives

1. Directing the attention of researchers, students, and technicians to the LMR task over Twitter for disaster management.

2. Forming a community that brings together researchers, students, and technicians as well as potential stakeholders.

3. Encouraging the design of new machine learning-based LMR models in disaster-related tweets that tackle both effectiveness and efficiency aspects.

4. Understanding the limitations of the current systems and inspiring the communities.

Evaluation Dataset

Large in size: 20,514 English tweets collected during different disaster events happened 2016-2019

Contains 19 disaster events: 6 hurricanes, 5 earthquakes, 4 floods, 3 wildfires, and 1 cyclone events.

Information types: caution and advice, displacement, damage, casualties, missing people, requests, etc.

Download

github.com/rsuwaileh/IDRISI

Happened in 14+ countries: United States, Italy, Canada, Ecuador, Mexico, New Zealand, Sri Lanka, India, Mozambique, Malawi, …

21,879 location mentions: Country, State, Province, District, County, City, Neighborhood, street, POI, Island & Others

Standard splits:
70% training
10% development
20% test
Per disaster event

Evaluation measures

Precision (P), Recall (R), and F1 score, per entity.

10 REDUCED INEQUALITIES

11 SUSTAINABLE CITIES AND COMMUNITIES

This challenge contributes to the SDGs 10 and 11.

Where to find the challenges?

https://aiforgood.itu.int/about-ai-for-good/geoai-challenge/

Kudos to the great ITU Team:

Reinhard Scholl

Andrea Manara

Thomas Basikolo

Spread the voice
(submission deadline 31 October 2023)
and good luck to everyone!!

Thanks also to:

UN Open GIS Initiative

UN GEOSPATIAL NETWORK
UNITED NATIONS COMMITTEE OF EXPERTS ON
GLOBAL GEOSPATIAL INFORMATION MANAGEMENT

UN-GGIM
ACADEMIC NETWORK