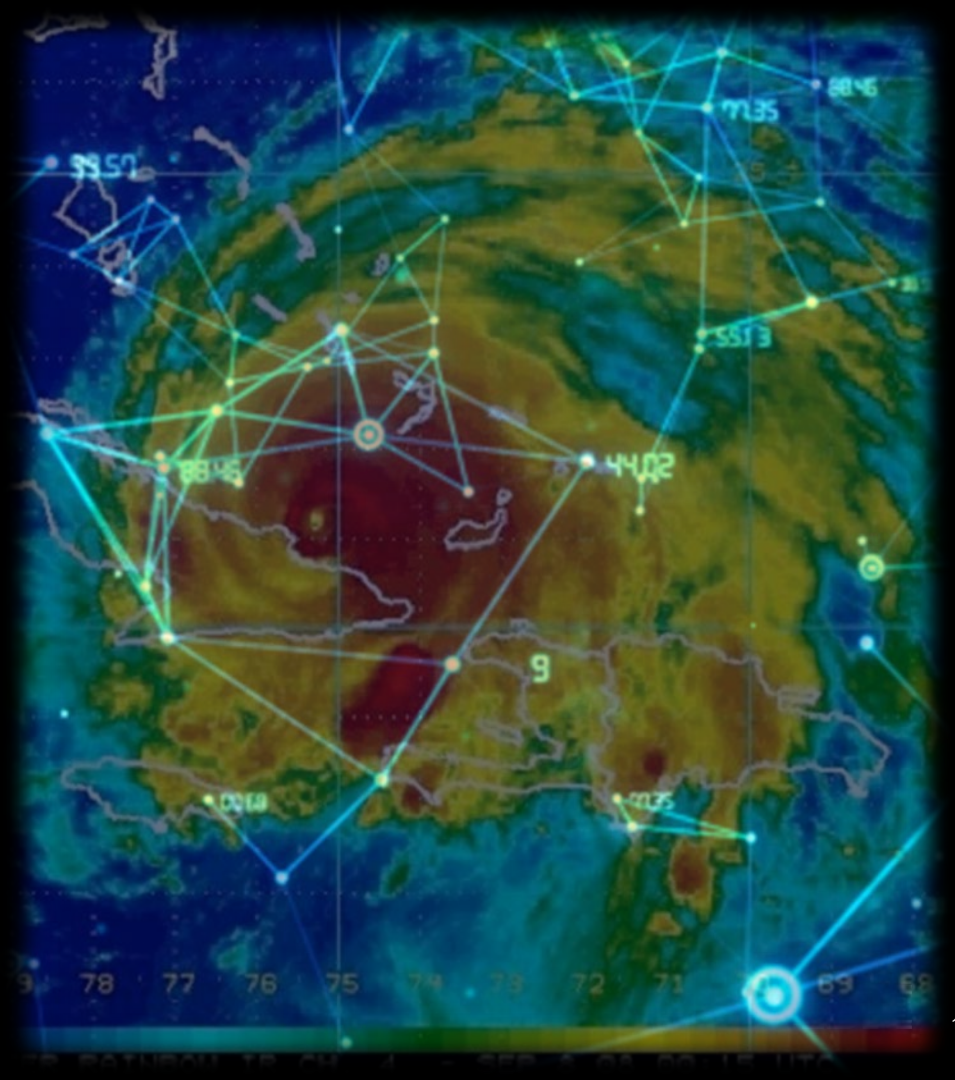# IBM Climate Network for Early Warning Systems

WMO Meeting – Nov. 2/3/4, 22, Geneva

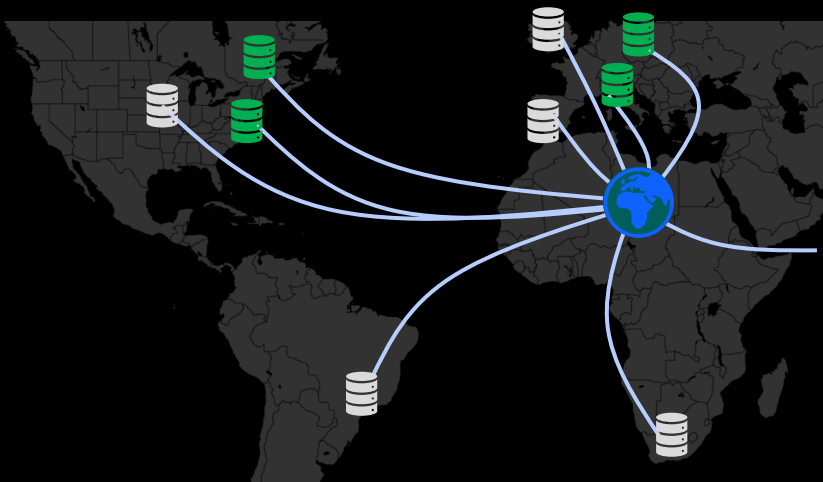Jonas Weiss

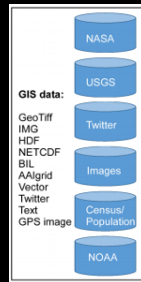Physics and AI for Climate Impact Group

jwe@zurich.ibm.com

The Climate Network

- Volume of geospatial data exceeds capabilities of single stakeholder/data center (storing, indexing, exploiting)
- Complex insights require many modalities, i.e. access to distributed data
- Data transport is expensive and time consuming
- Collaborating partners may have security, privacy and governance constraints -> maintain partial control

⇒ Need to run complex & distributed workflows with data and model federation across instances on a global scale!
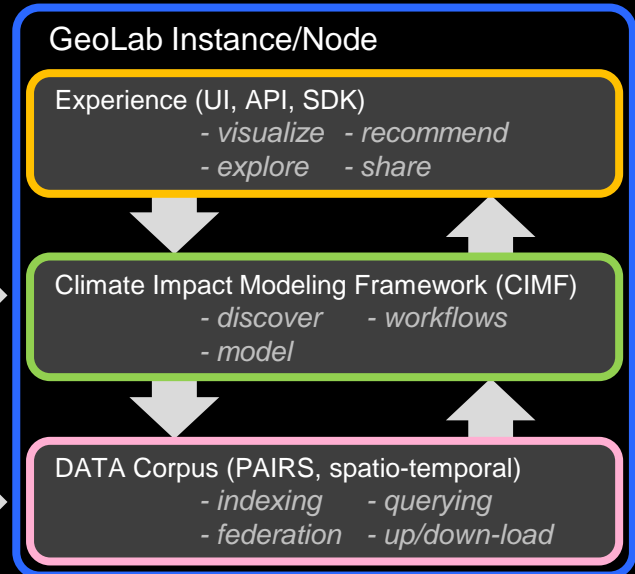

Public & Private Data Sources incl. TWC

GIS data:
GeoTiff
IMG
HDF
NETCDF
BIL
AAIgrid
Vector
Twitter
Text
GPS image

NASA
USGS
Twitter
Images
Census/Population
NOAA

Ingest

Curate

GeoLab Instance/Node

Experience (UI, API, SDK)
- *visualize*   - *recommend*
- *explore*   - *share*

Climate Impact Modeling Framework (CIMF)
- *discover*   - *workflows*
- *model*

DATA Corpus (PAIRS, spatio-temporal)
- *indexing*   - *querying*
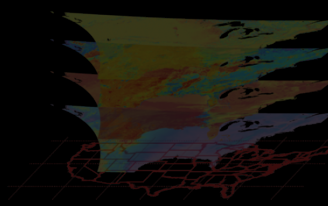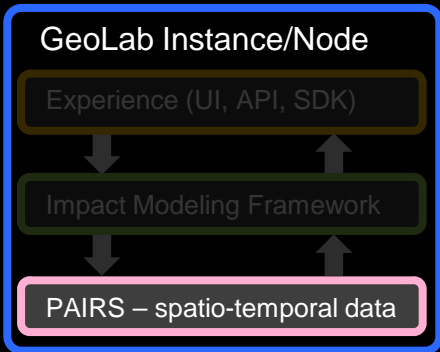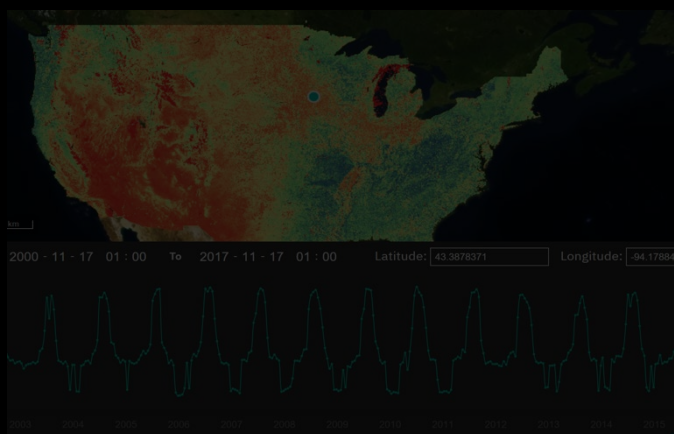- *federation* - *up/down-load*

Planner Coord.

Risk Analyst, Data Scientist

The Climate Network – Geo Time-Series & Weather Data

PAIRS – Spatio-Temporal Data Storage

Distributed Cluster
- > 750 layers
- + 10 Tbytes/day
- > 6 Petabyte total
- > 100 supported file formats
- Exposed through UI & API
- Technology to be open-sourced
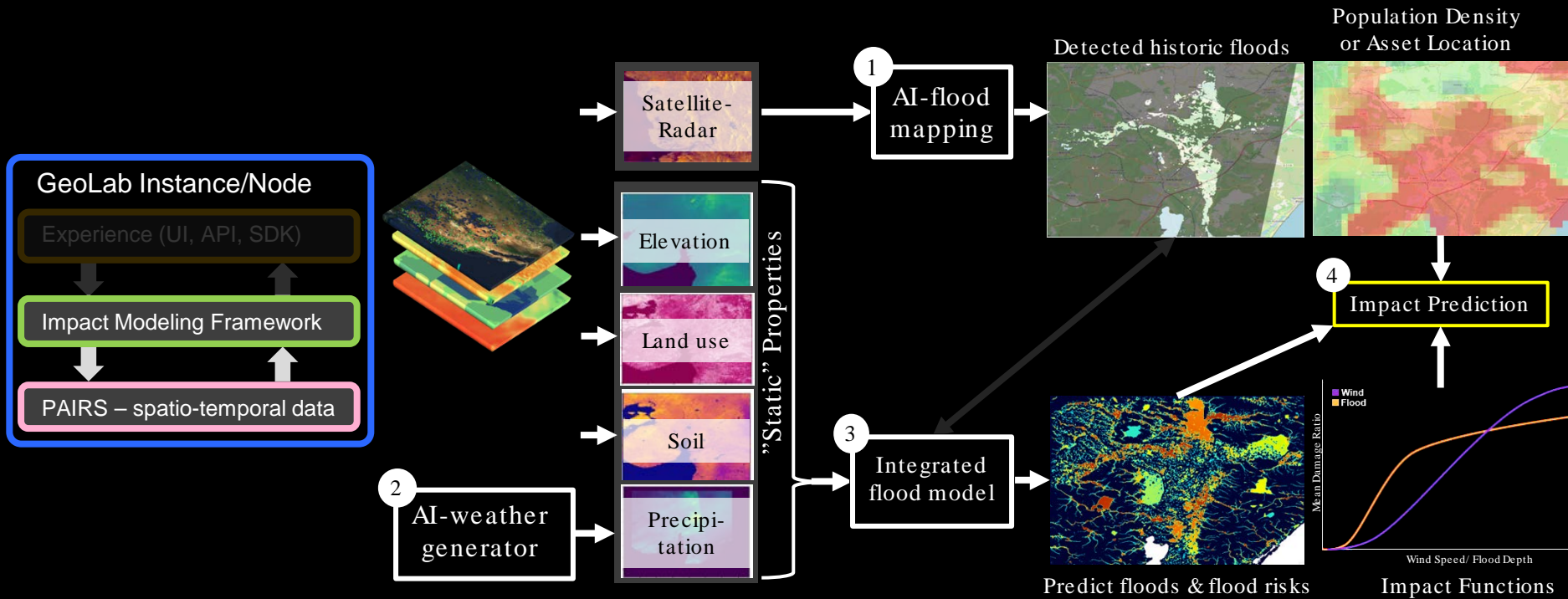- Blocks of time-series instead of raster-data (linked in space & time)

GeoLab Instance/Node
- Experience (UI, API, SDK)
- Impact Modeling Framework
- PAIRS – spatio-temporal data

The Weather Company
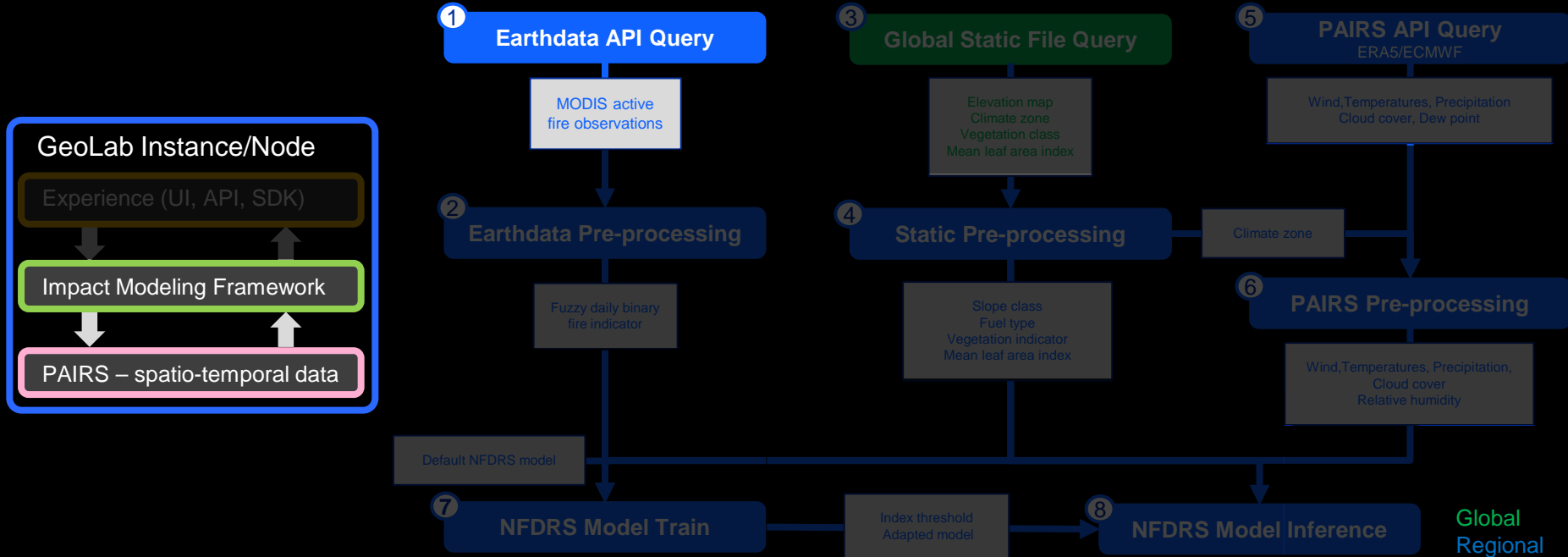**An IBM Business**

Every 60 *seconds* Condition Updates

Every 15 *minutes* a new Forecast

Data Sources

The Weather Company Proprietary and Sourced Data
- The Weather Company Weather Models
- 156K Global Stations
- 40M+ Mobile Phones
- 50K Flights a Day
- Global Lightning
- Air Quality and Pollen
- Traffic / Incident Data

Open & Govt Data
- National Weather Service Weather Stations
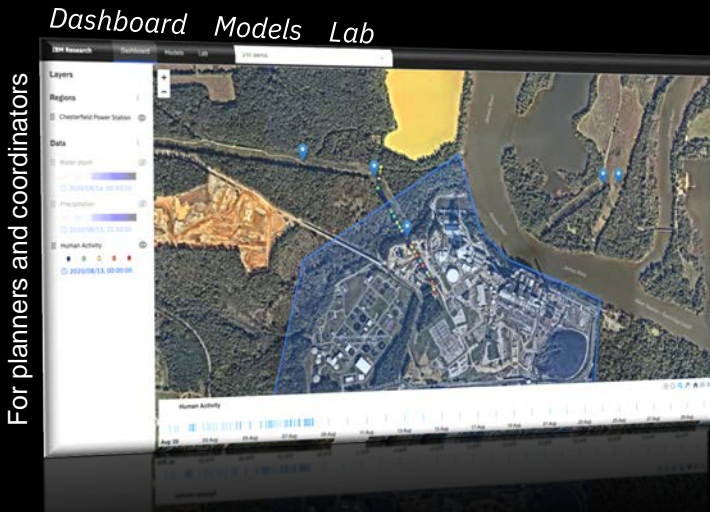- High Resolution Radar
- Oceanographic Data

Types
- Weather
- Atmosphere
- Historical
- Current
- Predictive
- Global
- Ultra-local

The Climate Network – Impact Modeling Framework

GeoLab Instance/Node

Experience (UI, API, SDK)

Impact Modeling Framework

PAIRS – spatio-temporal data

Satellite-Radar

Elevation

Land use

Soil

Precipi-tation

"Static" Properties

1 AI-flood mapping

2 AI-weather generator

3 Integrated flood model

4 Impact Prediction

Detected historic floods

Population Density or Asset Location

Predict floods & flood risks

Impact Functions

Mean Damage Ratio

Wind Speed/ Flood Depth

Wind
Flood

The Climate Network – Containers & Workflows → Portability

GeoLab Instance/Node
- Experience (UI, API, SDK)
- Impact Modeling Framework
- PAIRS – spatio-temporal data

① **Earthdata API Query**

MODIS active fire observations

③ **Global Static File Query**

Elevation map
Climate zone
Vegetation class
Mean leaf area index

⑤ **PAIRS API Query**
ERA5/ECMWF

Wind,Temperatures, Precipitation
Cloud cover, Dew point

② **Earthdata Pre-processing**

Fuzzy daily binary fire indicator

④ **Static Pre-processing**

Climate zone

Slope class
Fuel type
Vegetation indicator
Mean leaf area index

⑥ **PAIRS Pre-processing**

Wind,Temperatures, Precipitation,
Cloud cover
Relative humidity

Default NFDRS model

⑦ **NFDRS Model Train**

Index threshold
Adapted model

⑧ **NFDRS Model Inference**

Global
Regional

The Climate Network – GeoLab User Experience

Dashboard   Models   Lab
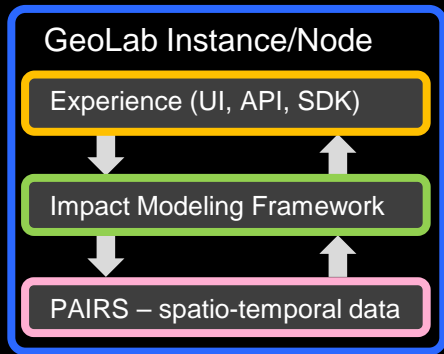
Application Programming Interface

For planners and coordinators

For analysts

For data scientists

GeoLab Instance/Node

Experience (UI, API, SDK)

Impact Modeling Framework

PAIRS – spatio-temporal data

Software Development Kit

```
In [ ]: #!pip install ibm-geolab

geolab.explore("2m temperature")
Node = 1, LayerID = 94, LayerName = ERA5 2m Temperature
Node = 1, LayerID = 106, LayerName = CPC 2m Temperature
Node = 1, LayerID = 107, LayerName = CPC 2m Dew Point Temperature

data1 = geolab.load_collection("1, 94")
data1 = data1.filter_temporal(
    start="2010-01-01T00:00:01Z",
    end="2020-01-02T00:00:01Z")

geolab.explore("Precipitation")
Node = 2, LayerID = 56, LayerName = ERA5 Precipitation
Node = 2, LayerID = 67, LayerName = CHIRPS Precipitation

data2 = geolab.load_collection("2, 67")
data2 = data2.filter_temporal(
    start="2010-01-01T00:00:01Z",
    end="2020-01-01T00:00:01Z")

pearsonr = scipy.stats.pearsonr(data1, data2)

#!import torchgeo
```
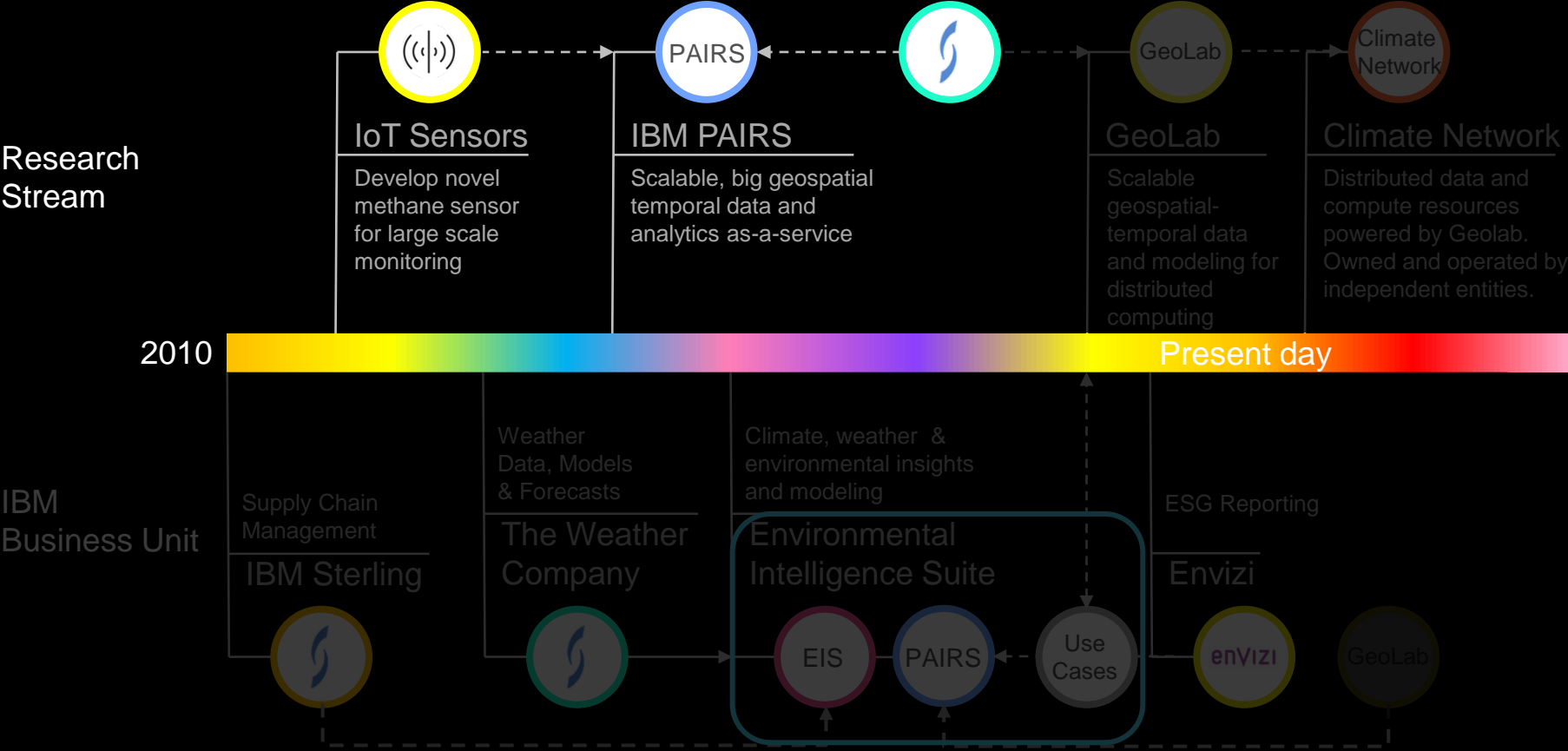
Future contribution to pytorch-geospatial comm.

How everything fits together

Research Stream

IoT Sensors
Develop novel methane sensor for large scale monitoring

IBM PAIRS
Scalable, big geospatial temporal data and analytics as-a-service

GeoLab
Scalable geospatial-temporal data and modeling for distributed computing

Climate Network
Distributed data and compute resources powered by Geolab. Owned and operated by independent entities.

2010 — Present day

IBM Business Unit

Supply Chain Management
IBM Sterling

Weather Data, Models & Forecasts
The Weather Company

Climate, weather & environmental insights and modeling
Environmental Intelligence Suite

EIS — PAIRS — Use Cases

ESG Reporting
Envizi

GeoLab

IBM Environmental Intelligence Suite – (Climate) Use Cases

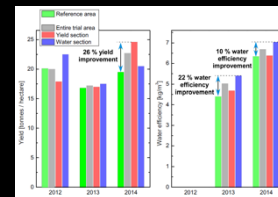Renewable energy
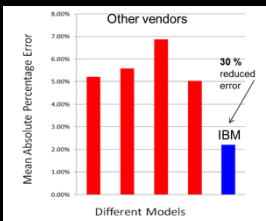
Long-term forecasting
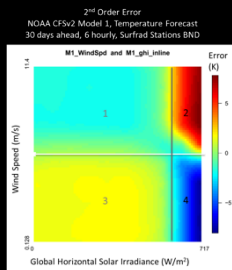
AQ modeling
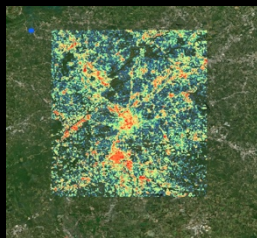
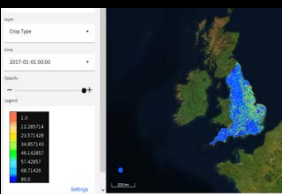Disaster management: Wildfire, Floods etc.

Irrigation management

GHG monitoring

Optimal land identification
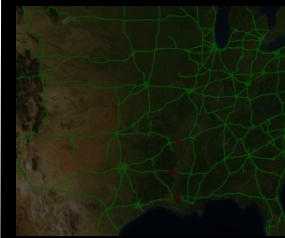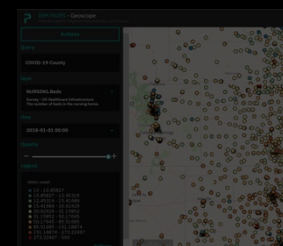
Vegetation management

Crop protection

Crop forecasting
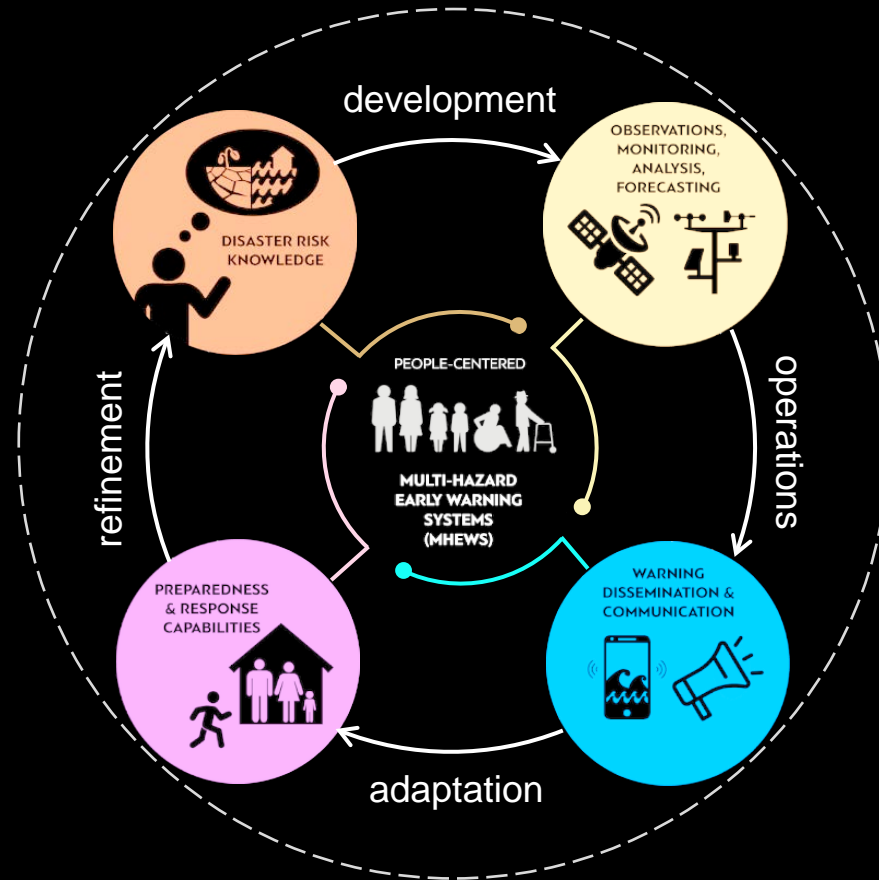
Logistics planning

Geospatial Asset management

Multi Hazard Early Warning Systems

Multi Hazard Early Warning Systems – IBM Ecosystem

DISASTER RISK KNOWLEDGE

OBSERVATIONS, MONITORING, ANALYSIS, FORECASTING

COMMUNICATION

PREPAREDNESS & RESPONSE CAPABILITIES

IBM PAIRS – Part of IBM EIS

Distributed Cluster
- > 750 layers
- + 10 Tbytes/day
- > 6 Petabyte total

IBM GeoLab & Climate Network

- Experience
  UI & SDK
- Modeling
  CIMF
- Data
  PAIRS

The Weather Company

Global forecasts, every 5 mins., reaching Millions today

The Weather Company
An IBM Business

Operational Risk Insights

Multiple risk factors, configurable alerst Operational (in use)

IBM Deep Search – NLP

Identify Impact function Search all kind of scientific and other (NLP) data

TWC: Decision Support for Billions to Prepare for Weather

2.2 Billion weather forecast locations

Data from **40 Million** mobile devices

Atmospheric Data from **50'000** flights per day

200'000+ worldwide weather stations

Existing Channels

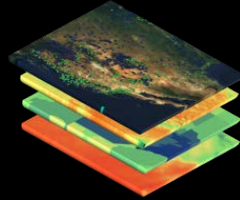Trusted by many Industries

Global Scale Operational Solution

Decision Support for Billions to Prepare for Incoming Weather

26 billion forecasts a day at your fingertip

Multi Hazard Early Warning Systems – IBM Ecosystem

DISASTER RISK KNOWLEDGE

OBSERVATIONS, MONITORING, ANALYSIS, FORECASTING

WARNING DISSEMINATION & COMMUNICATION

PREPAREDNESS & RESPONSE CAPABILITIES

IBM PAIRS – Part of IBM EIS

Distributed Cluster
- > 750 layers
- + 10 Tbytes/day
- > 6 Petabyte total

IBM GeoLab & Climate Network

Experience
UI & SDK

Modeling
CIMF

Data
PAIRS

The Weather Company

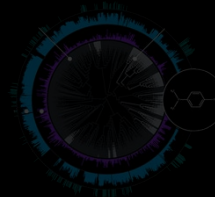Global forecasts, every 5 mins., reaching Millions today

The Weather Company
An IBM Business

Operational Risk Insights (ORI)

Multiple risk factors, configurable alerst Operational (in use)
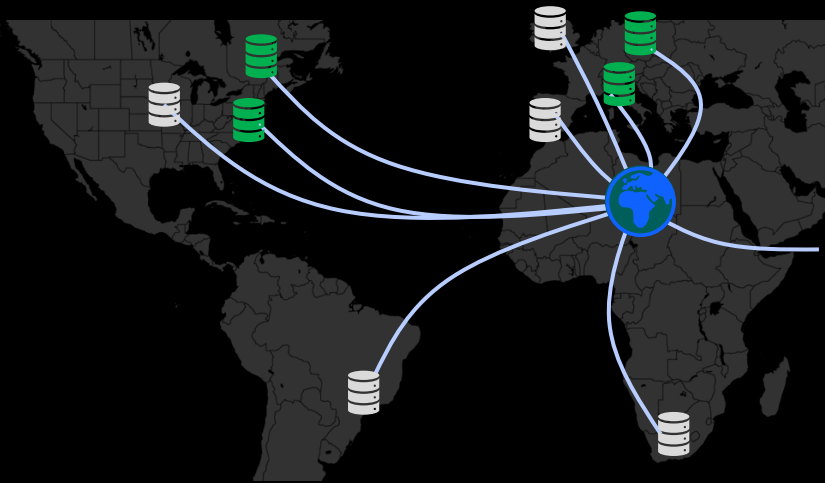
IBM Deep Search – NLP KGs

Identify Impact functions Search & contextualize all kind of scientific and other (NLP) data
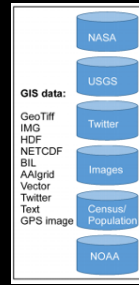
Collaborations and Partnerships

- IBM and IBM Research have a long tradition of collaborating with industry, academia and governments

- With research and development labs/centers on all major continents, cultural and geographic differences and similarities are naturally reflected in our inclusive work

- Our processes and mindset allows for low entry-barrier collaborations

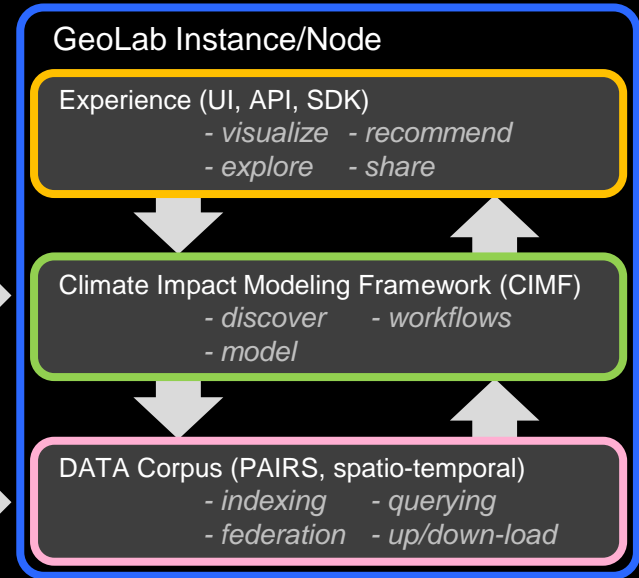# The Climate Network – Data and Model Collaboration Platform

GeoLab Instance/Node

Experience (UI, API, SDK)
- *visualize*   - *recommend*
- *explore*   - *share*

Climate Impact Modeling Framework (CIMF)
- *discover*   - *workflows*
- *model*

DATA Corpus (PAIRS, spatio-temporal)
- *indexing*   - *querying*
- *federation*   - *up/down-load*

Public & Private Data Sources incl. TWC

GIS data:
GeoTiff
IMG
HDF
NETCDF
BIL
AAlgrid
Vector
Twitter
Text
GPS image

NASA
USGS
Twitter
Images
Census/Population
NOAA

Ingest

Curate

Planner Coord.

Risk Analyst, Data Scientist

- Data- and model-federation for global collaboration, data and model sharing across frontiers
- Privacy, security and governance to preserve national and stake-holder interests by "local" ownership
- Public and on-prem cloud enabled
- Kubernetes/Openshift enabled workflows for infrastructure portability and model-sharing flexibility

Climate Network Summit on Jan 19th-20th

→ Marina Rakhlin: mrakhlin@us.ibm.com

Jonas Weiss, *jwe@zurich.ibm.com*