PROGRAMME OF THE EUROPEAN UNION

Copernicus

European Commission

Empowering Disaster Risk Management with eXplainable AI

Guido Fioravanti, Arthur Essenfelder Hrast, Andrea Toreti

European Commission - Joint Research Centre

Bruxelles

06/17/2025

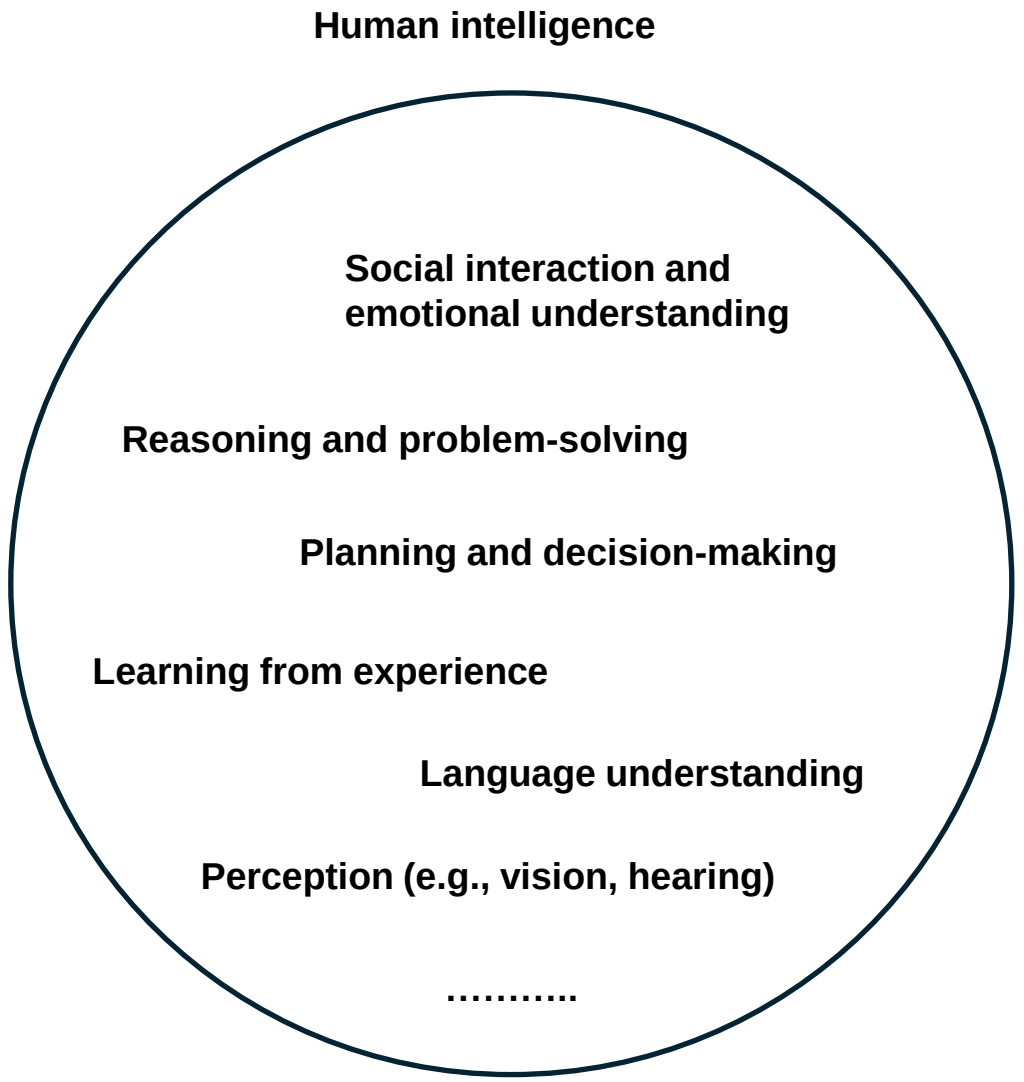
COPERNICUS EMERGENCY MANAGEMENT SERVICE

What is AI?

AI is a set of technologies designed to perform tasks that typically require human intelligence

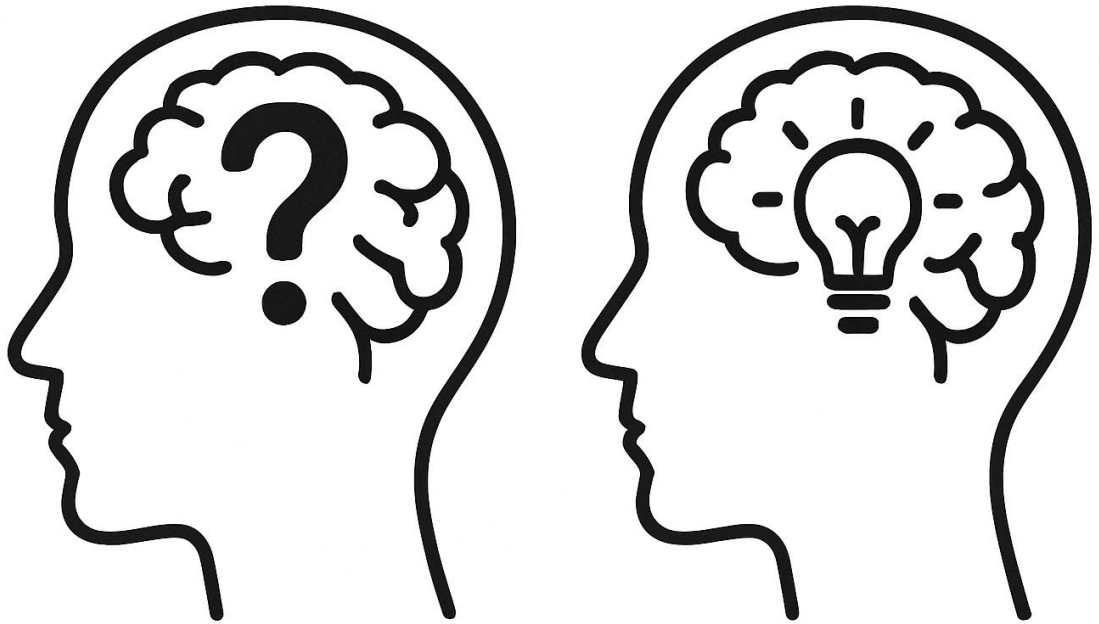
Various approaches to building intelligent systems, such as *knowledge representation*, *reasoning*, *planning*, *robotics*, *natural language processing*, and *computer vision*, many of which do not rely **exclusively on Machine Learning (ML)**.

Russell, S. J., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson Education.



eXplainable AI (XAI)

A set of tools/techniques and practices designed to help humans understand why an AI model makes a certain prediction.



Wthouth explainability:
The model predicts a drought in this area and gives no reason (just outputs a label). Not explainable, even if it's accurate

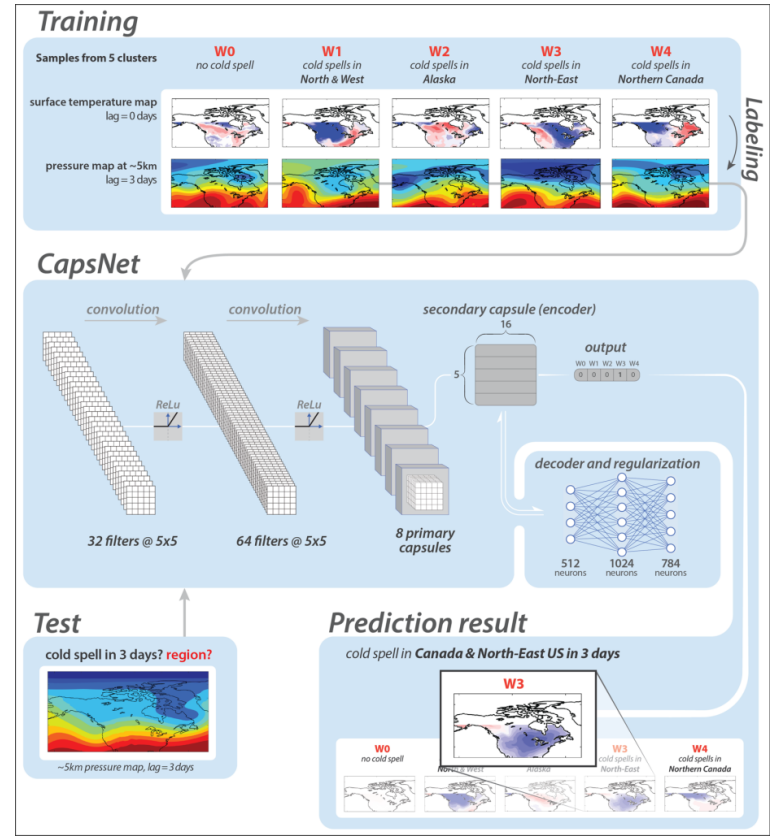
With explainability:
The model predicts a drought because the soil moisture is low, precipitation is below average, and vegetation stress is increasing.

Initiatives and Research in AI/ML for supporting Disaster Risk Management

CLINT
CLIMATE INTELLIGENCE
EXTREME EVENTS
HOTSPOTS
MACHINE LEARNING

XAIDA

Detection and attribution



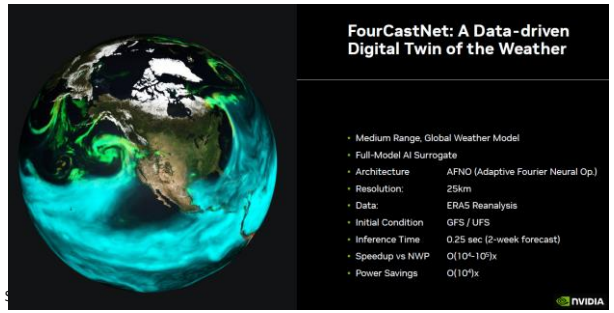
Source: Chattopadhyay et al., 2020, https://doi.org/10.1029/2019MS001958, CC BY 4.0

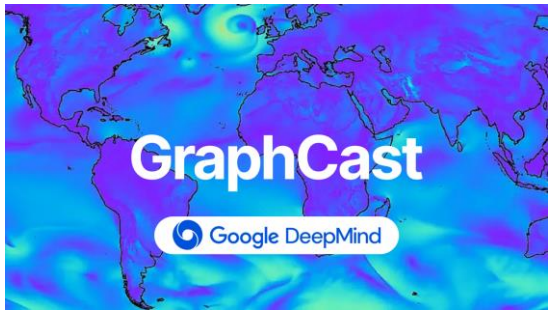
"*To ensure decision makers can have trust and confidence in new built tools and algorithms in DestinE, the use of emerging AI technologies must align with principles of ethics and responsibility.*"
(**Source:** Artificial Intelligence in DestinE – The explainer)

NOAA AI
CENTER FOR ARTIFICIAL INTELLIGENCE

MACHINE LEARNING for DISASTER RISK MANAGEMENT

GFDRR





MedEWSa

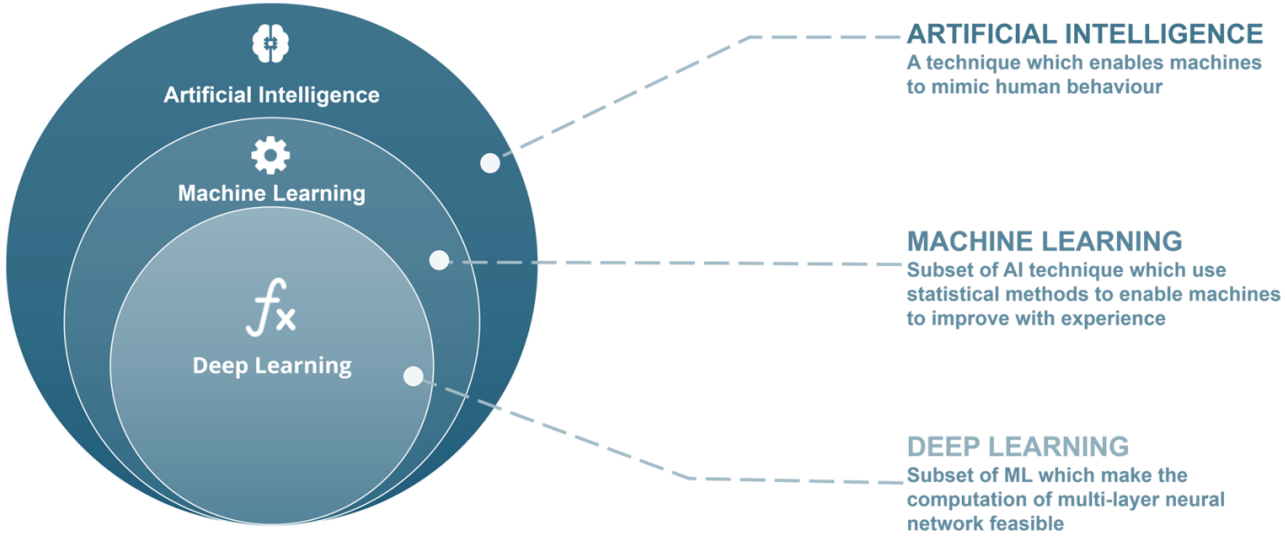
AI weather models

AI: AI for Natural Disaster Management
ITU Focus Group

Machine Learning & Deep Learning

Machine Learning (ML) is the study of programs that can improve their performance on a given task automatically (Wikipedia)

A type of Machine Learning that uses **neural networks (NN)** with many layers to automatically learn patterns from data.

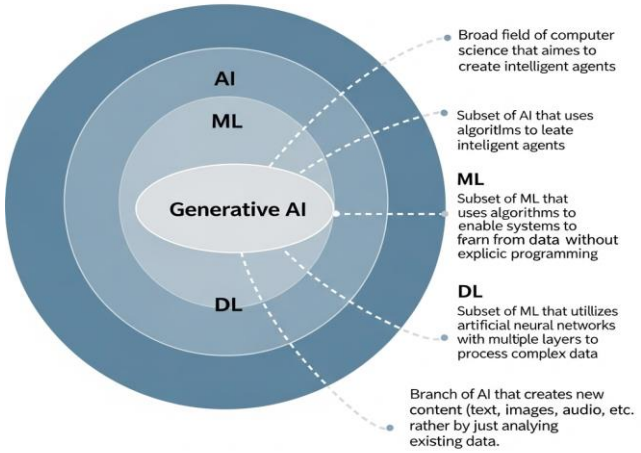


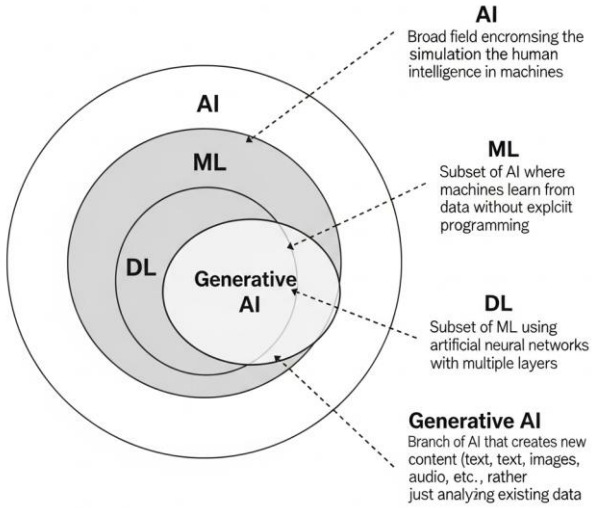
AI ≠ ChatGPT

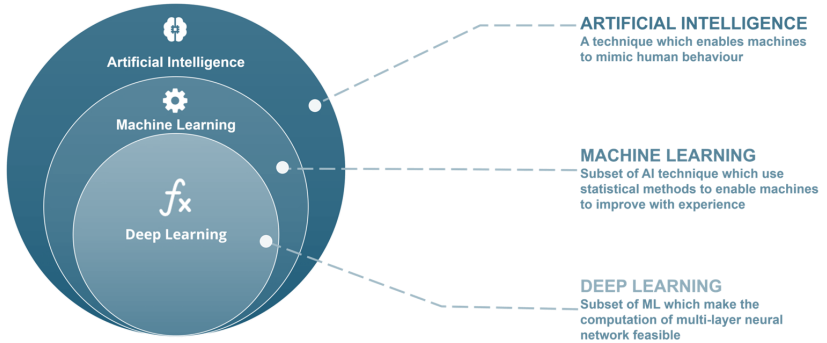
Generative AI (gen AI) is the branch of AI that creates new content, rather than just analyzing existing data

A key part of Generative AI involves **Large Language Models (LLMs)**

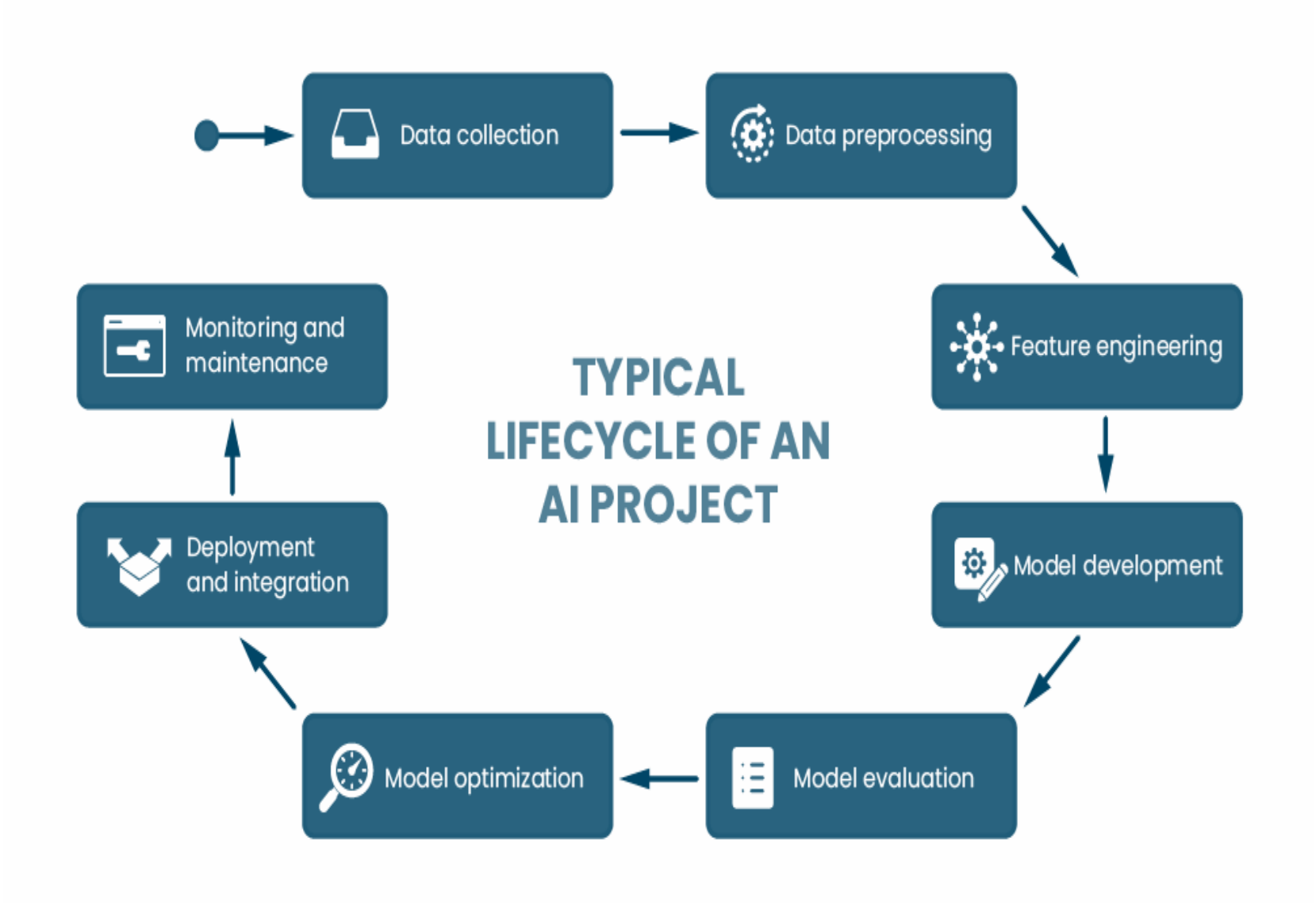
Gen AI poses novel risks (e.g., hallucinations, inaccurate or biased outputs), which threaten to undermine the trust in the technology



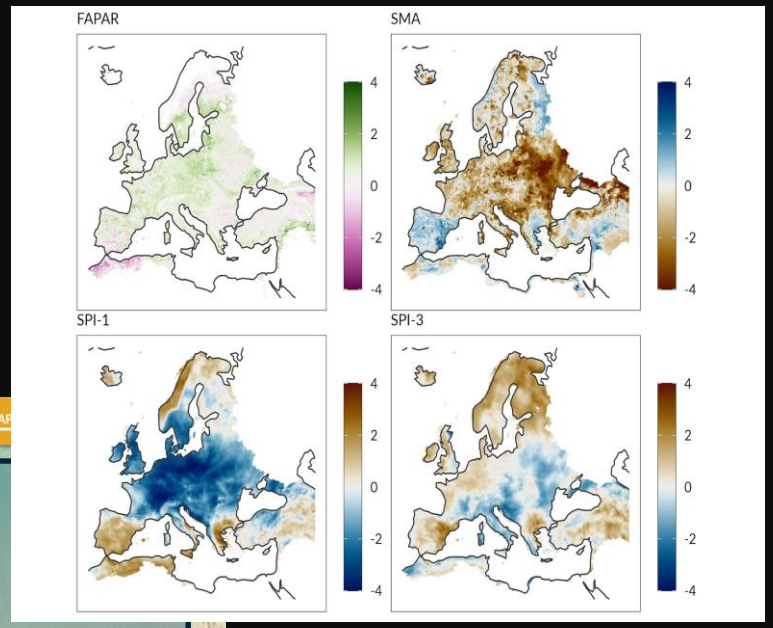


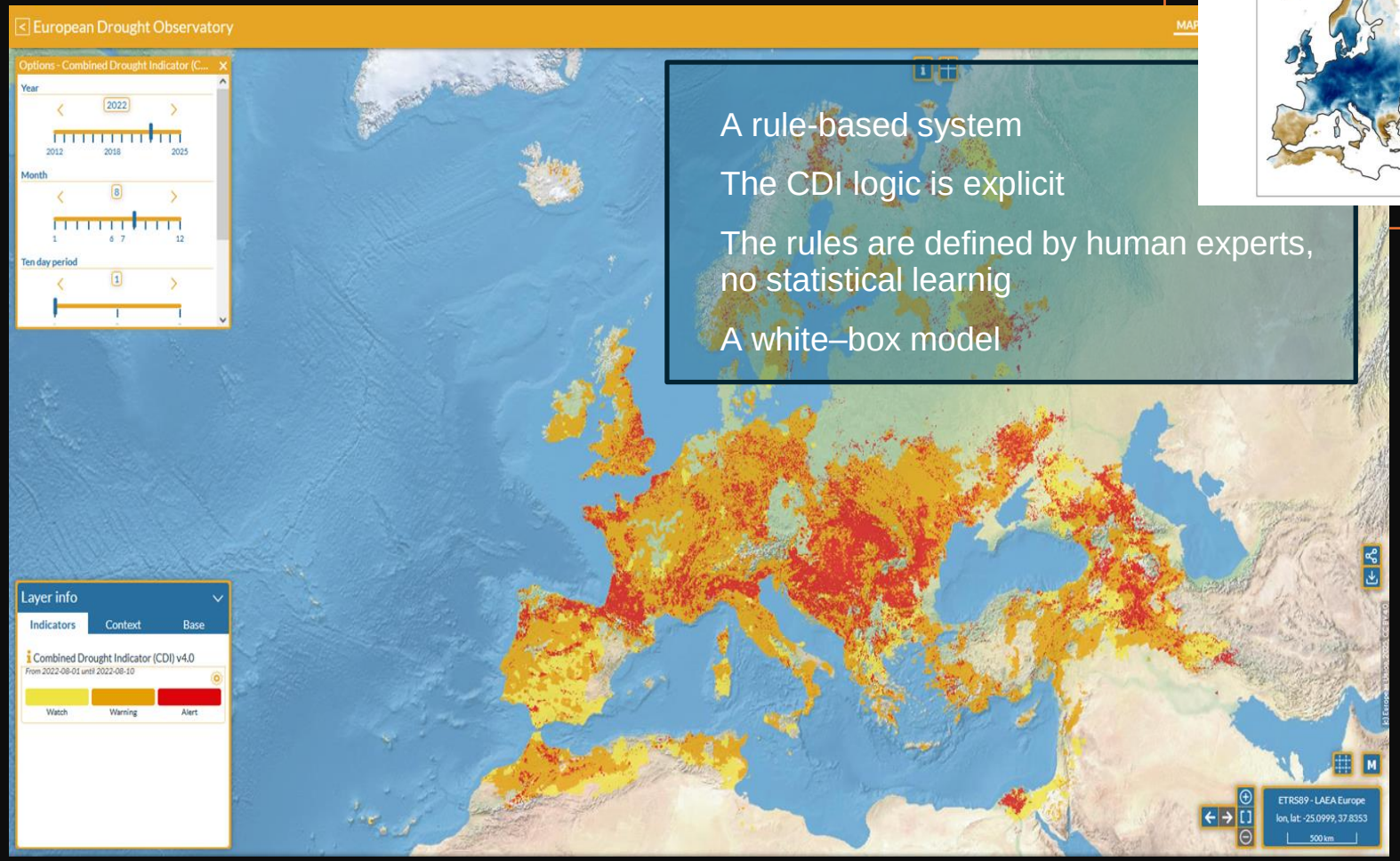


The images on the left have been generated using the generative AI capabilities of ChatGPT, and were designed to mimic the style of the Venn diagram on the top. While the diagrams may **contain inaccuracies**, they serve as an example of the potential of generative AI to create visual content.


Data collection
Data preprocessing
Feature engineering
Model development
Model evaluation
Model optimization
Deployment and integration
Monitoring and maintenance
TYPICAL LIFECYCLE OF AN AI PROJECT

The Combined Drought Indicator





A rule-based system

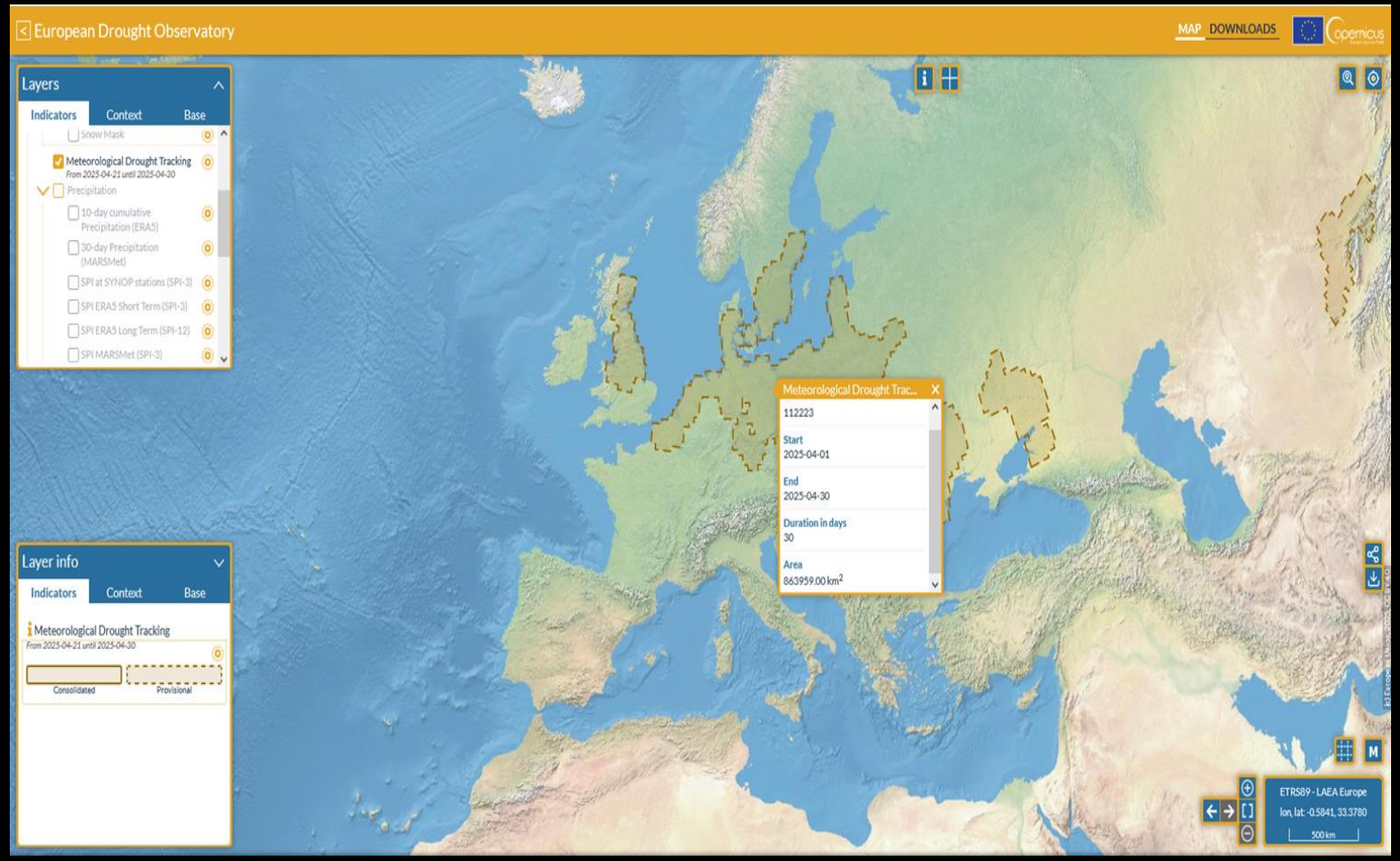
The CDI logic is explicit

The rules are defined by human experts, no statistical learnig

A white–box model

GOFAI (Good Old Fashioned AI)

Meteorological Drought Tracking



Machine Learning Alogirthm (DBSCAN)

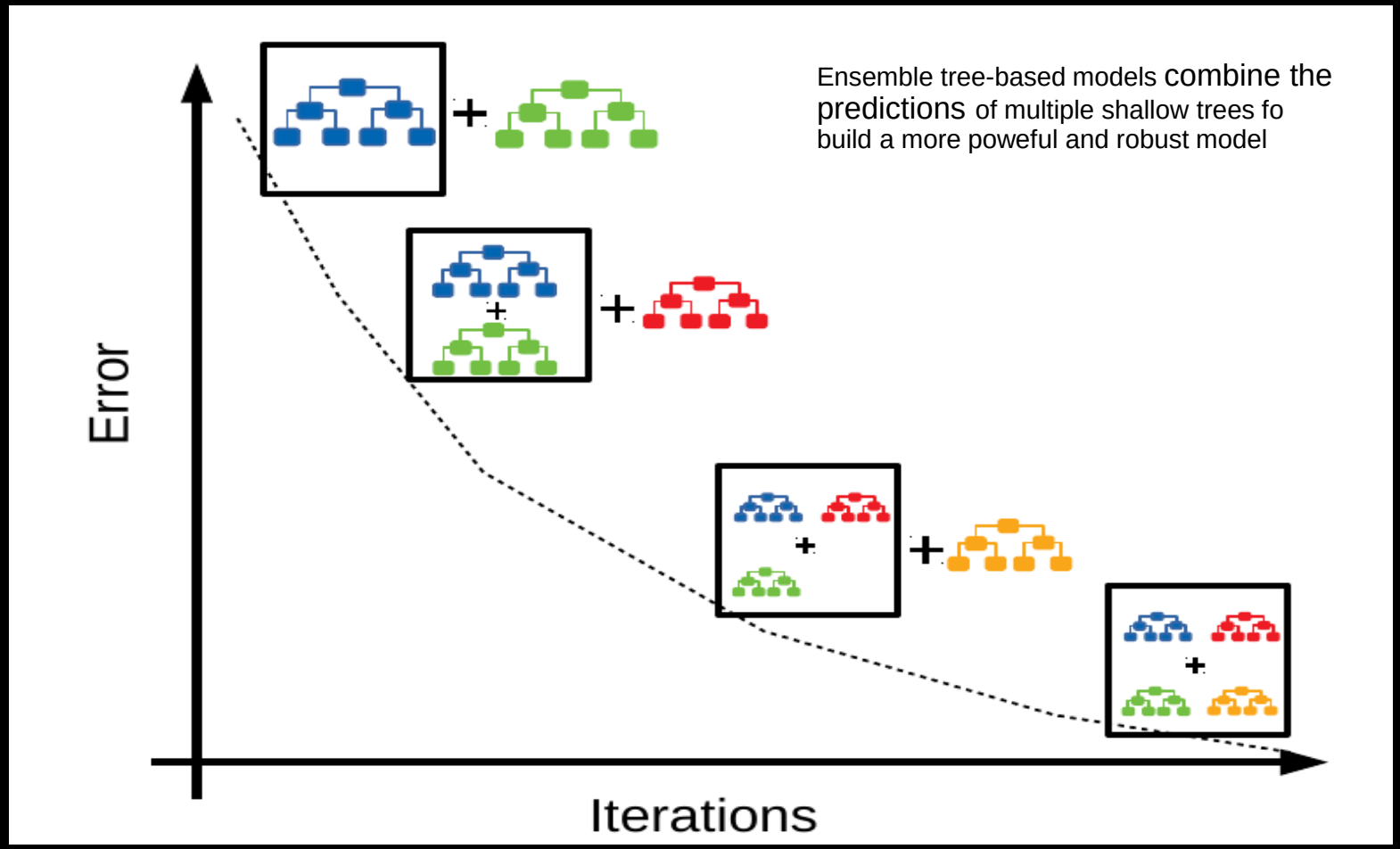
Implicit logic (Learned Patterns)

Data-Driven Rules/Patterns

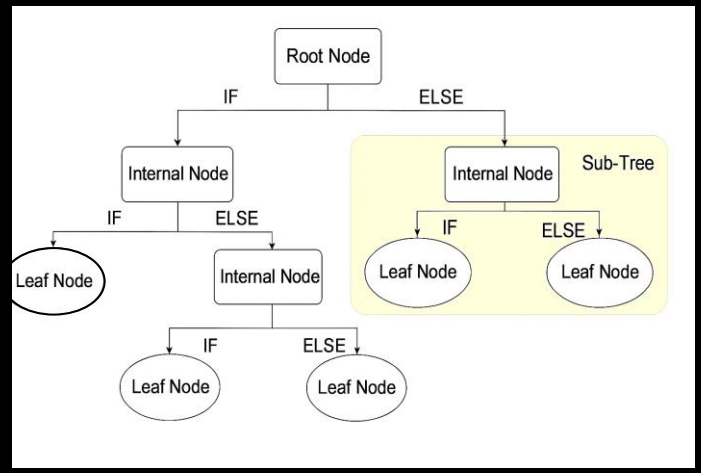
White–box model

Unsupervised Learning

Decision Tree & Ensemble Learning Models



Usually, a single decision tree is not strong enough to be used in Practice (weak learner)

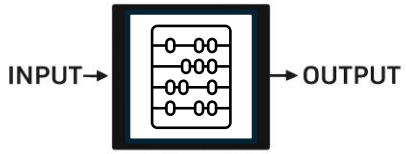


How do we build ensemble models?

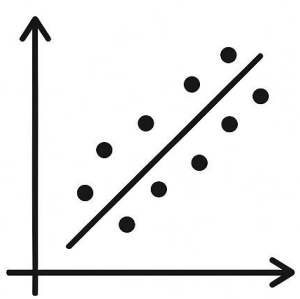
Bagging: builds trees independently using random subsets of the data (Random Forest)

Boosting: builds trees sequentially, where each new tree focuses on correcting the errors of the previous ones (Gradient Boosting Trees and XGBoost)

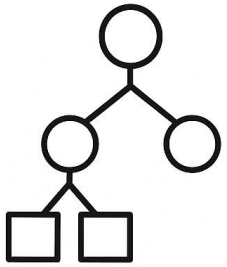
AI ≠ Black-box models



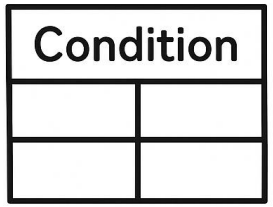
Inherently Interpretable Models



Linear Regression



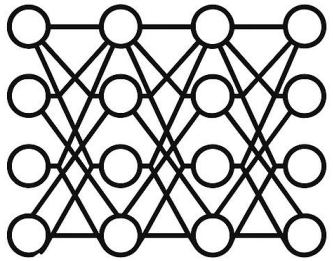
Simple Decision Tree



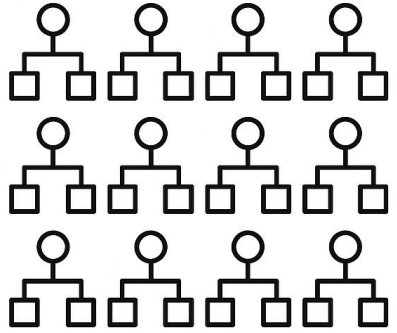
Decision Table

White//black box models describe how **transparent and interpretable** a model is to humans

Black Box Models



Neural Network



Collection of Decision Trees

The black-box issue refers to the lack of transparency and understanding about how an AI system makes its decisions.

Lack of trust
Accountability problems
Ethical and legal risk
Debugging difficulty

References

- Ghaffarian, S. et al., 2023. Explainable artificial intelligence in disaster risk management: Achievements and prospective futures, International Journal of Disaster Risk Reduction, https://doi.org/10.1016/j.ijdrr.2023.104123
- Giovine, C. and Roberts, R., 2024. Building AI trust: The key role of explainability, https://www.mckinsey.com/capabilities/quantumblack/our-insights/building-ai-trust-the-key-role-of-explainability#/
- Maier, H. R., 2024. How much X is in XAI: Responsible use of “Explainable” artificial intelligence in hydrology and water resources, Journal of Hydrology X, https://doi.org/10.1016/j.hydroa.2024.100185
- Melkamu Mersha, M. et al., 2024. Explainable artificial intelligence: A survey of needs, techniques, applications, and future direction, Neurocomputing, https://doi.org/10.1016/j.neucom.2024.128111

STAY CONNECTED

EVENTS, ONLINE, and MAP VIEWERS

@CopernicusEMS

emergency.copernicus.eu

activations.emergency.copernicus.eu

Expert validation of AI-generated graphs on disaster events

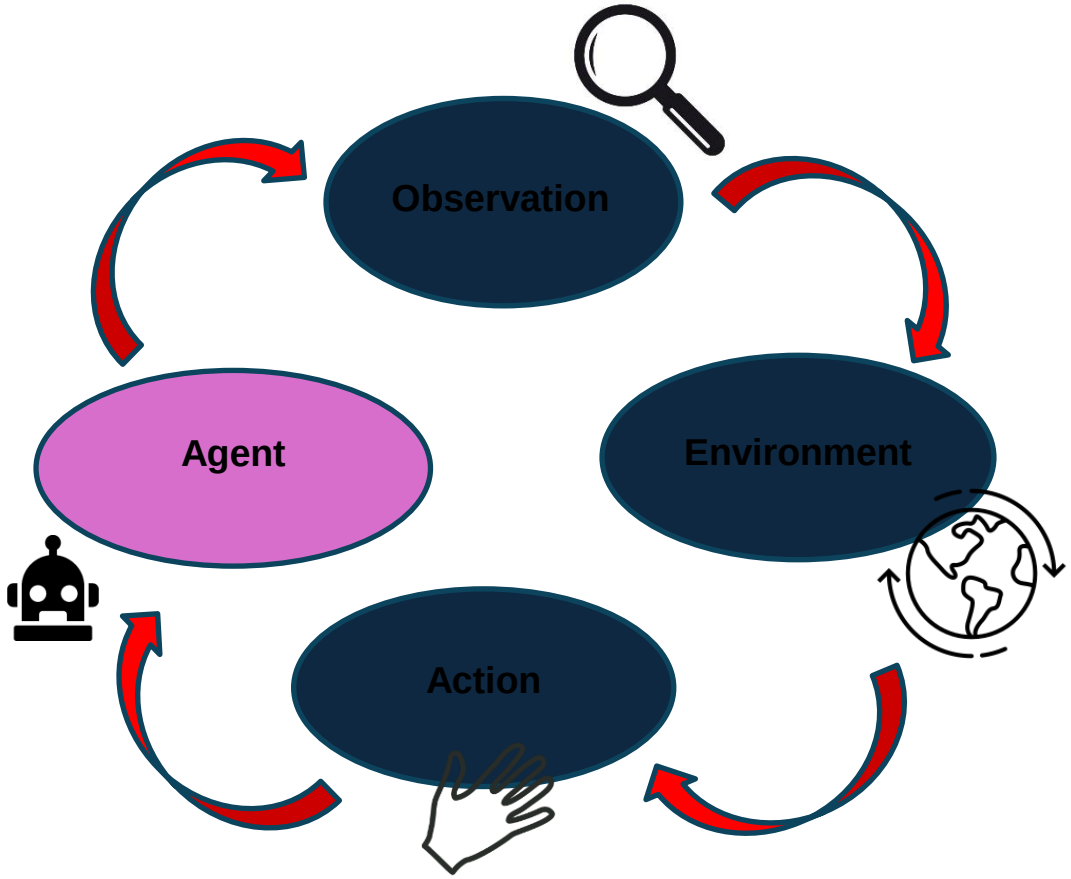
Michele Ronco, AI Specialist, JRC.E1

Sergio Consoli (JRC.F7), Lorenzo Bertolini (JRC.F7), Luca Bandelli (JRC.T5), Alessio Spadaro (JRC.T5), Marco Verile (JRC.T5), Christina Corbane (JRC.E1)

AI For Preparedness: Building capacity for AI-powered Disaster Risk Management,
Brussels, 17/06/2025

AI Agents

LLM-powered Agents are artificial entities that enhance LLMs with essential capabilities, enabling them to sense their environment, make decisions, and take actions



- **Sam Altman:** “GPTs and Assistants are precursors to agents. They will gradually be able to plan and to perform more complex actions on your behalf. These are our first step toward AI Agents.”
- **Bill Gates:** “Agents are not only going to change how everyone interacts with computers. They’re also going to upend the software industry, bringing about the biggest revolution in computing since we went from typing commands to tapping on icons.”

AI DRM Agent

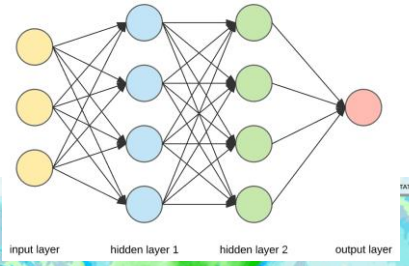
IPC

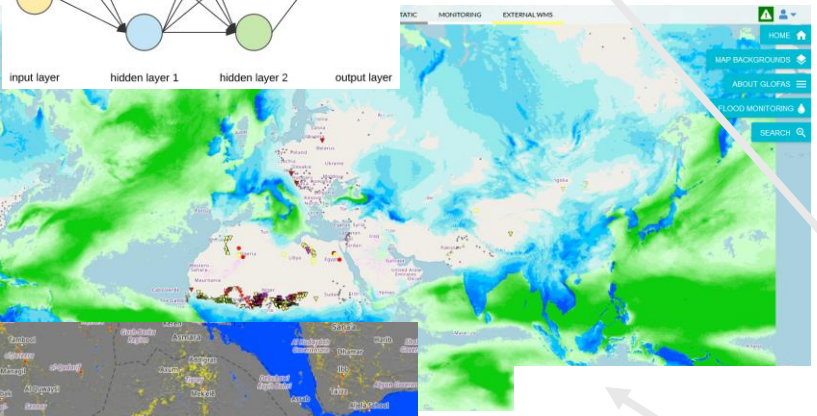
iDMC internal displacement monitoring centre

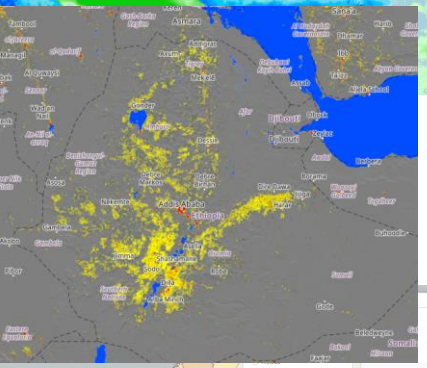
EM-DAT The International Disaster Database Centre for Research on the Epidemiology of Disasters

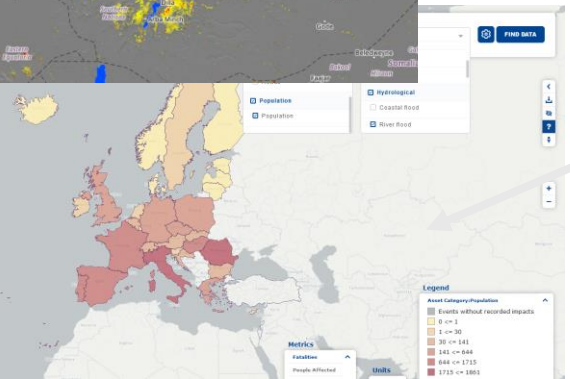
emm

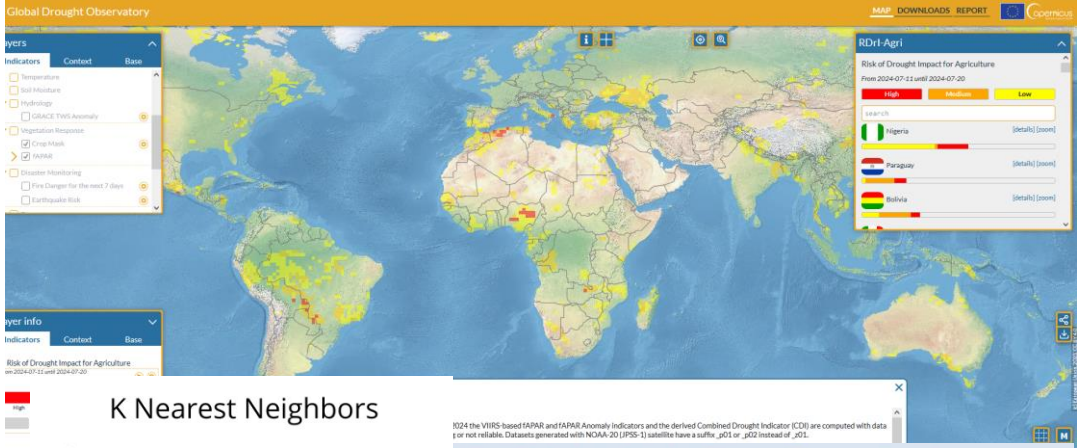
GPT@JRC

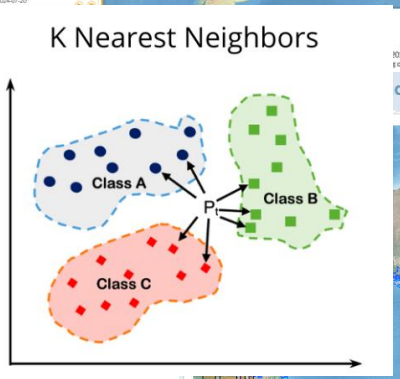


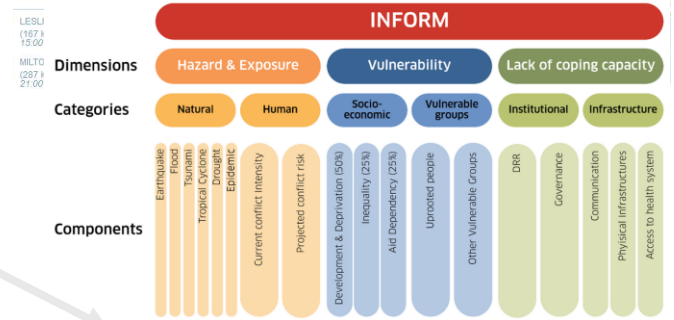












What can I help with?

What humanitarian and economic impacts could a cascading crisis caused by a flood in Ethiopia have within the next month?

Create image | Summarize text | Help me write | Get advice | More

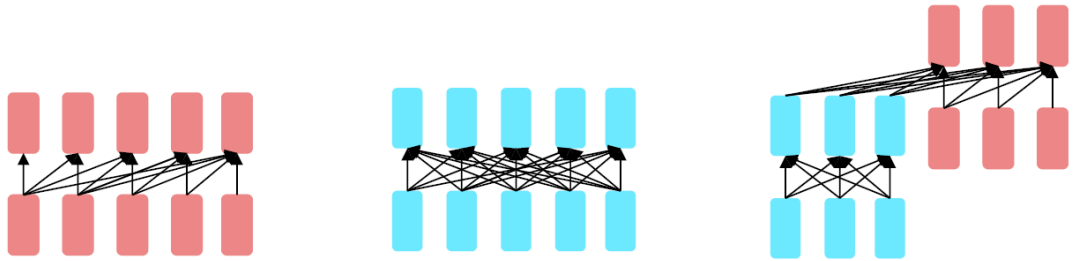
Must have..

- Interactive Capabilities with Digital World
- Advanced Reasoning and Action Planning
- Hierarchical Problem-Solving
- Understanding of Complex Scenarios
- Transparent Decision-Making

Image generated by DALL-E3

What is a Large Language Model (LLM)?

- AI models trained by learning to guess what word comes next
- LLMs can generate text by sampling possible next words
- Learn a lot of useful language knowledge
- Trainable parameters: 1-1000 billions
- Common tasks: machine translation, sentiment analysis, classification, speech recognition



Decoders
GPT, Claude, Llama Mixtral

Encoders
BERT family, HuBERT

Encoder-decoders
Flan-T5, Whisper

$P(w|$`Q: Who wrote the book ‘‘The Origin of Species"? A:`$)$

[1] *Attention is all you need*, Vaswani et al (2017) https://arxiv.org/abs/1706.03762

Retrieval Augmented Generation (RAG)

RAG is like a smart friend who helps the AI find the right info before answering, making responses more accurate!

Towards agentic behavior..

Non-agentic workflow (zero-shot):

Please type out an essay on topic X from start to finish in one go.

Start

Finish

Agentic workflow (decompose task):

Write an essay on topic X

Do you need web search or documents?

Write a first draft

Consider what parts need revision or more research

Revise your draft

….
….

Combining news and AI to address crisis complexity

- Retrieval Augmented Generation on EMM to extract knowledge on to extract and analyze large-scale disaster data across thousands of events
- LLMs to create disaster narratives and capture main cause-and-effect factors forming impact chains
- Semi-qualitative approach to multi-hazard and multi-risk assessment

Data (EMM) + AI (LLMs)

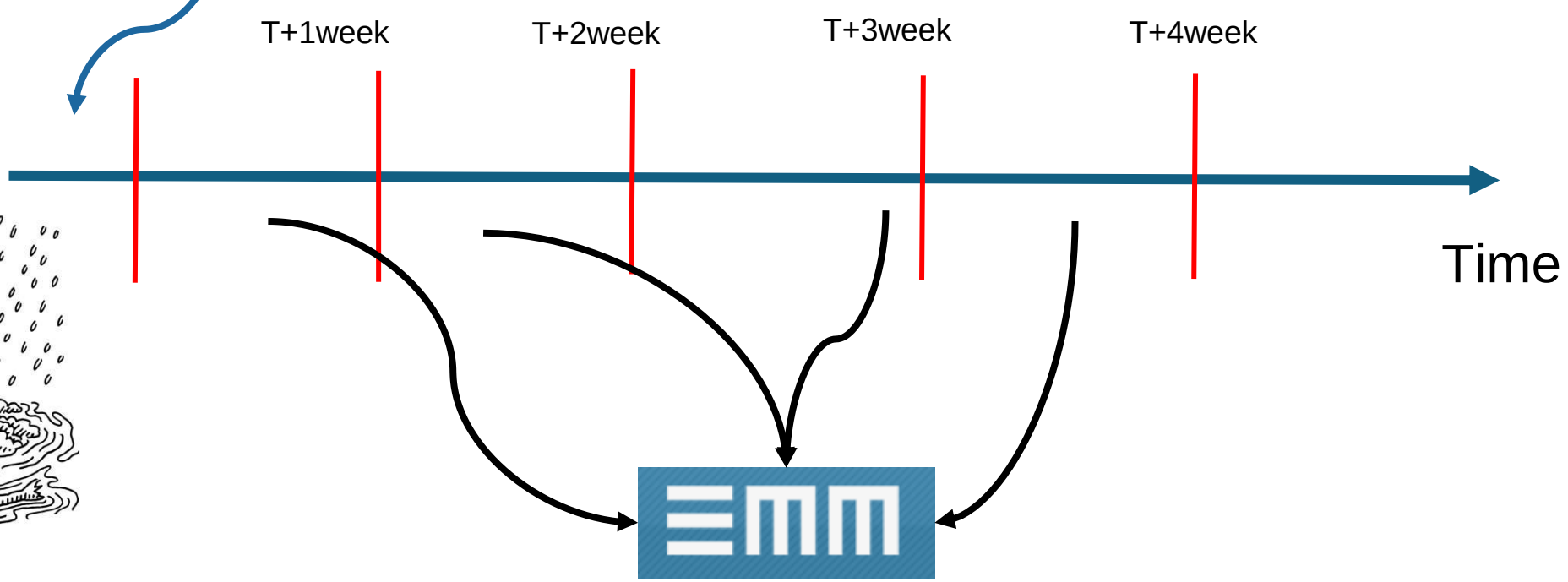
GPT@JRC

Retrieving disaster news

EM-DAT
The International Disaster Database
Centre for Research on the Epidemiology of Disasters

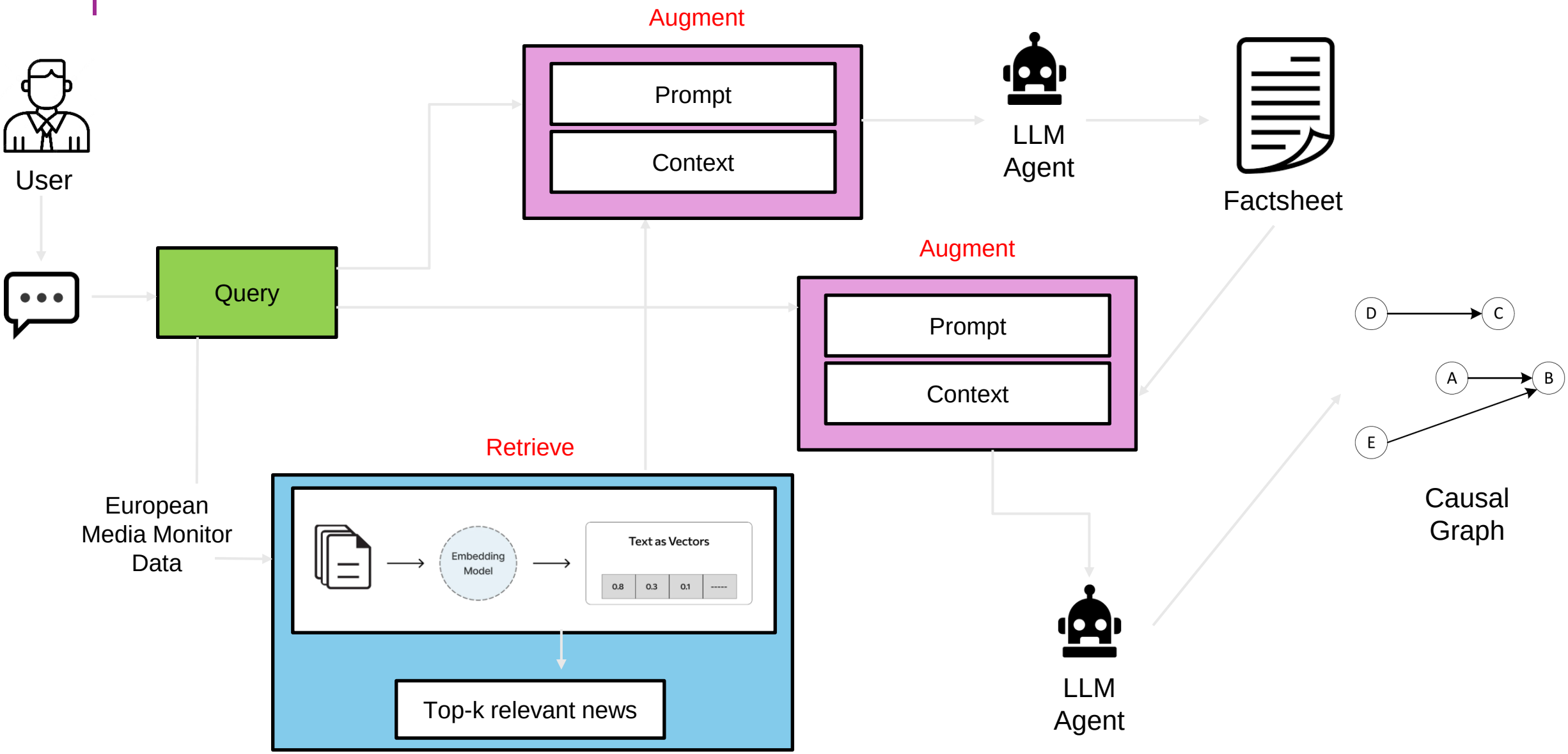
Dataset containing global disaster records (location, date, type)

Given a disaster record we look for news in the following 4 weeks retrieving top-k relevant news for each interval

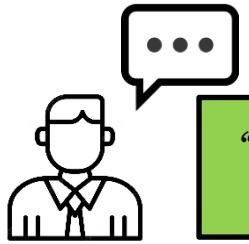


Pipeline overview

Dashboard at: https://huggingface.co/spaces/jrc-ai/crisesStorylinesRAG



Two-step process with RAG & ICL



"What are the latest developments on the wildfire disaster occurred in Greece on October 2024?"

Prompt 1

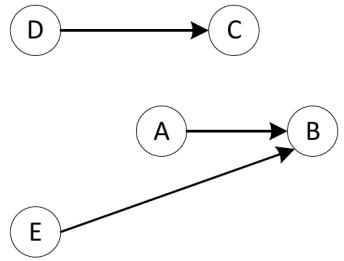
Context (top-k news)

"Complete this factsheet on the wildfire event in Greece: key information, severity, key drivers, main impacts, likelihood of multi-hazard risk, best practices, recommendations"

Prompt 2

Context (storyline)

"Create a knowledge graph that captures the causal relationships presented in the text. You have to use two relationship types: 'causes' or 'prevents'. Example: XXX"

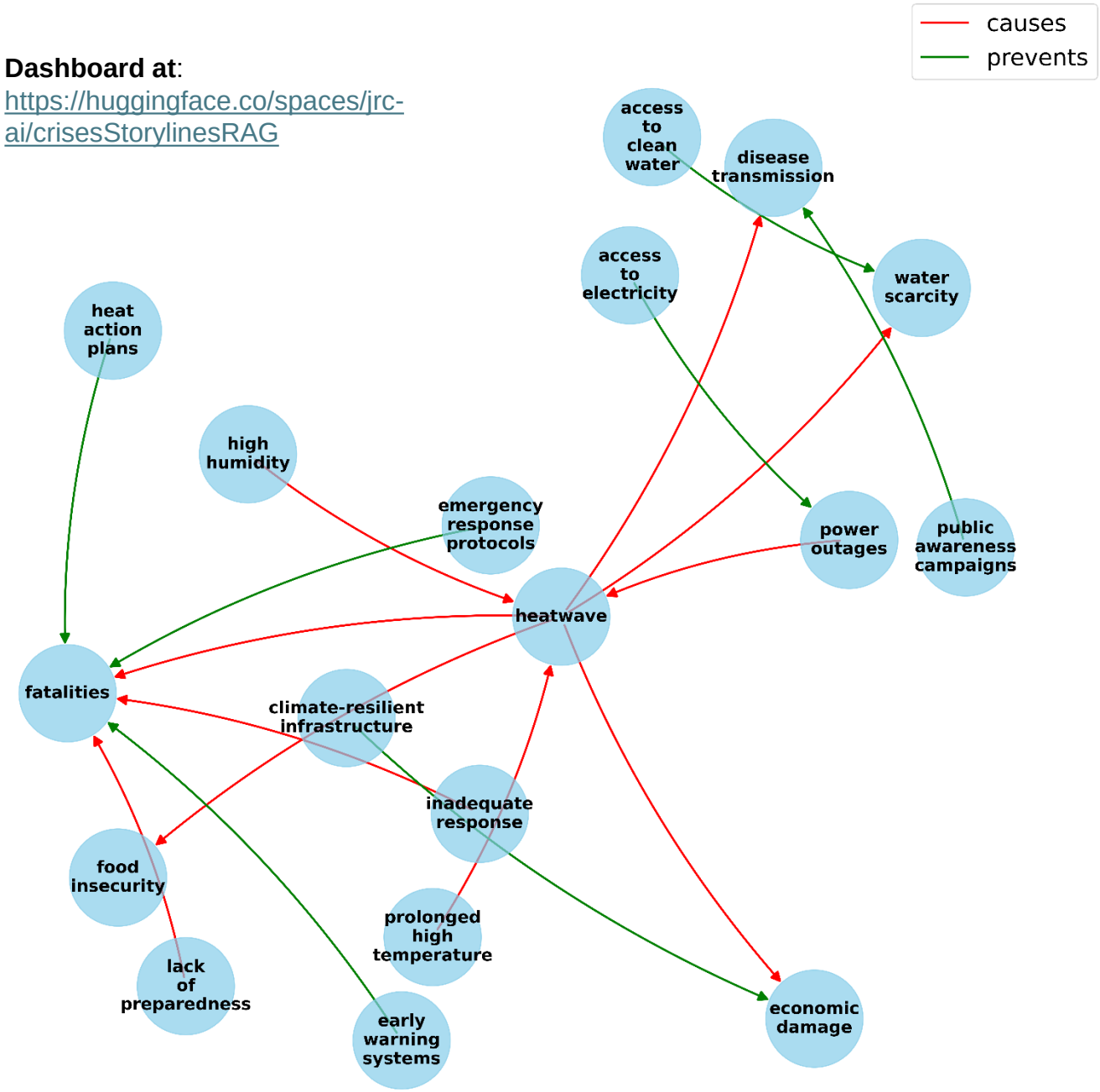


Example outcome

Dashboard at:
https://huggingface.co/spaces/jrc-ai/crisesStorylinesRAG

India-Pakistan heatwave 2015

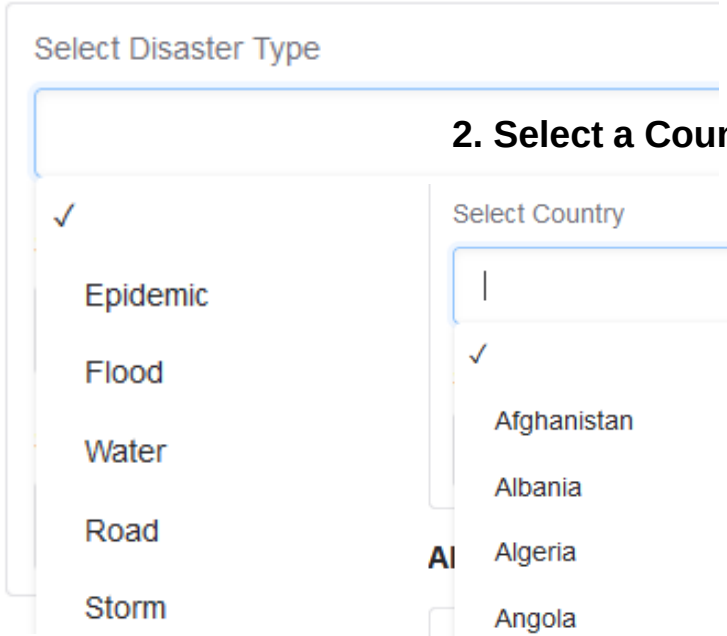
- **Key information:** A severe heatwave struck Pakistan primarily affecting the Sindh and Punjab provinces in June 2015
- **Severity:** High
- **Key drivers:** Prolonged temperatures above 45C 113F for several days, high humidity, power outages which exacerbated the effects of the heat, lack of preparedness and inadequate response by authorities
- **Main impacts, exposure and vulnerability**: Economic damage, people affected, people died and thousands more were affected, fatalities with the majority being elderly or poor. Effects on communities and infrastructure. Hospitals were overwhelmed with heat-related cases. Power outages disrupted daily life and exacerbated the heat. Water scarcity was reported in some areas
- **Likelihood of multi-hazard risk**: Medium to High. The heatwave increased the risk of subsequent hazards such as water scarcity and drought, food insecurity, increased risk of disease transmission e.g. heat-related illnesses waterborne diseases
- **Best management practices**: Establishing early warning systems for heatwaves. Implementing heat action plans including public awareness campaigns and emergency response protocols Improving access to clean water and electricity Enhancing healthcare infrastructure and emergency services Promoting climate-resilient infrastructure and urban planning
- **Recommendations for recovery**: Provide emergency aid including food water and shelter to affected communities Support the establishment of heat action plans and early warning systems Invest in climate-resilient infrastructure including green spaces and climate-resilient buildings Promote public awareness campaigns on heat-related illnesses and prevention measures Support research and development of climate-resilient technologies and practices Provide technical assistance and capacity building programs for local authorities and emergency responders



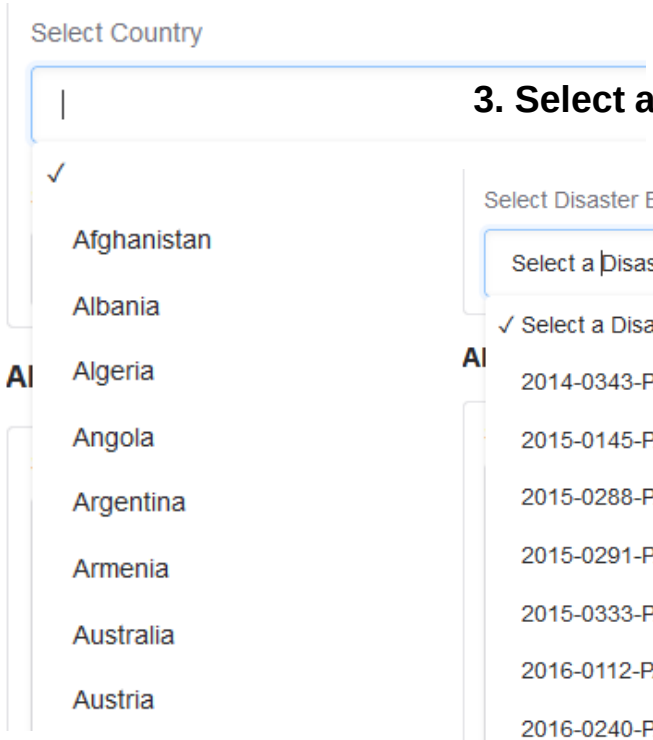
How to use the tool (1)

Go to: https://huggingface.co/spaces/jrc-ai/crisesStorylinesRAG

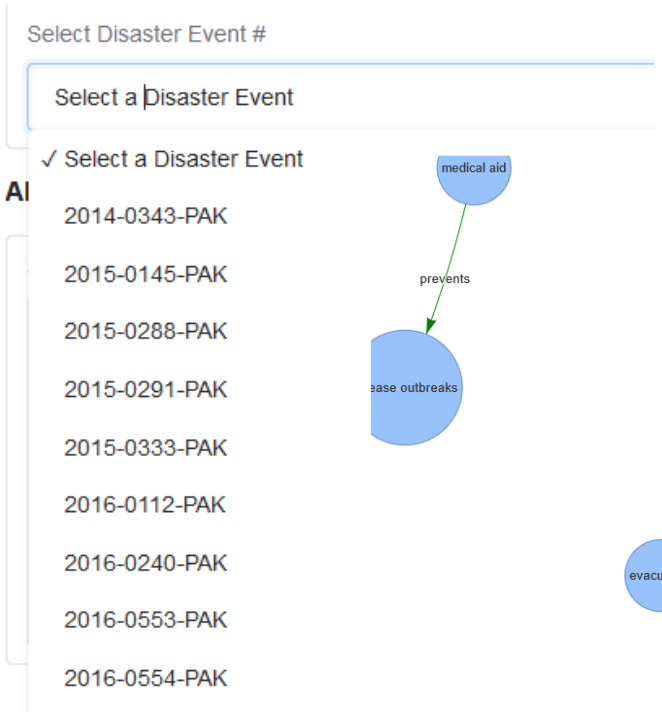
1. Select a Disaster Type



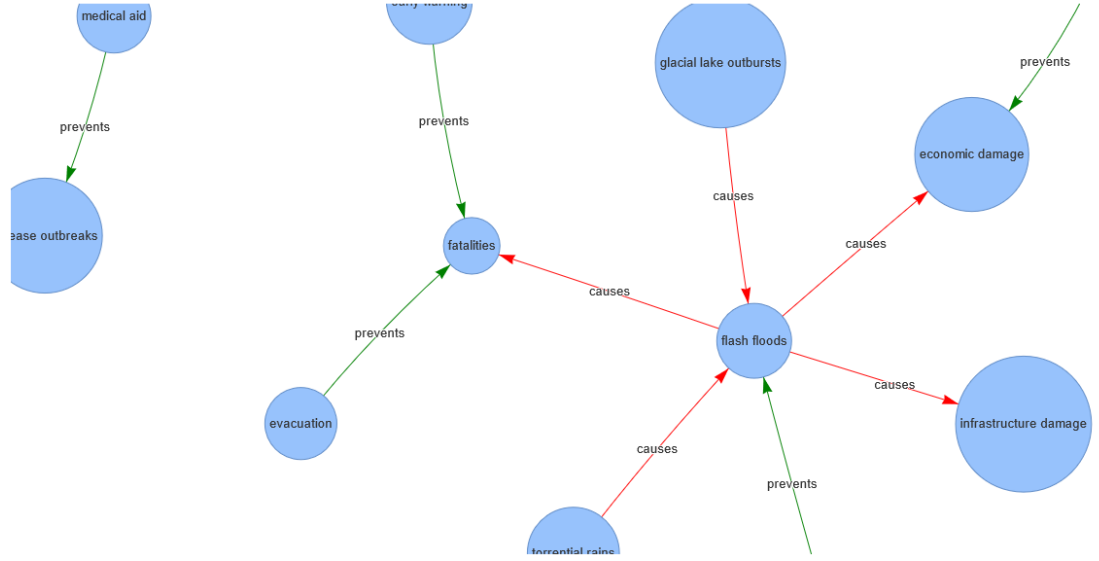
2. Select a Country



3. Select a EM-DAT event



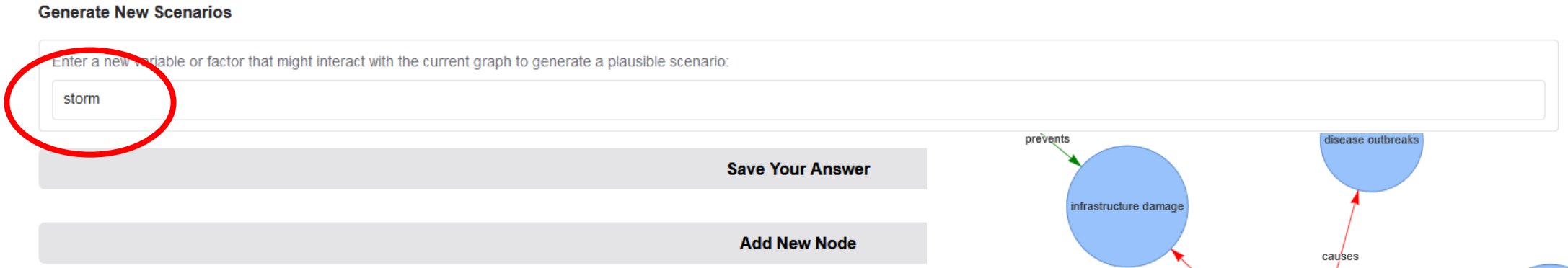
Analyse the AI-generated storylines and impact-chain graphs!



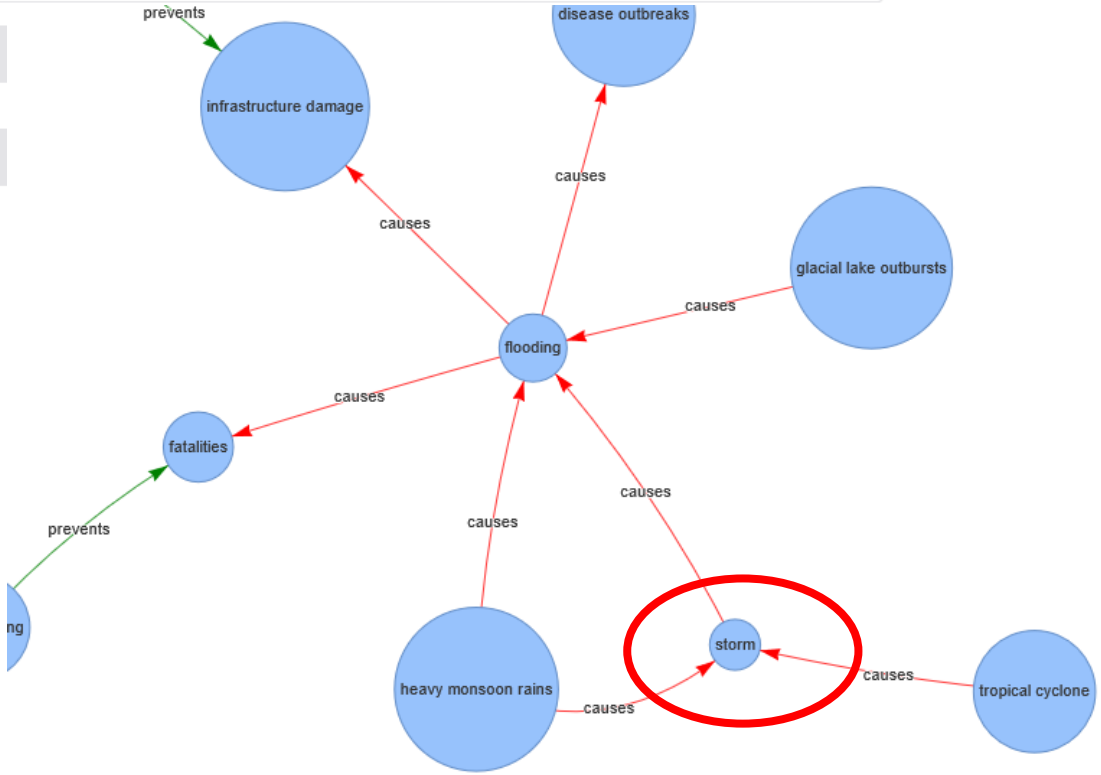
How to use the tool (2)

Go to: https://huggingface.co/spaces/jrc-ai/crisesStorylinesRAG

4. Play with new scenarios

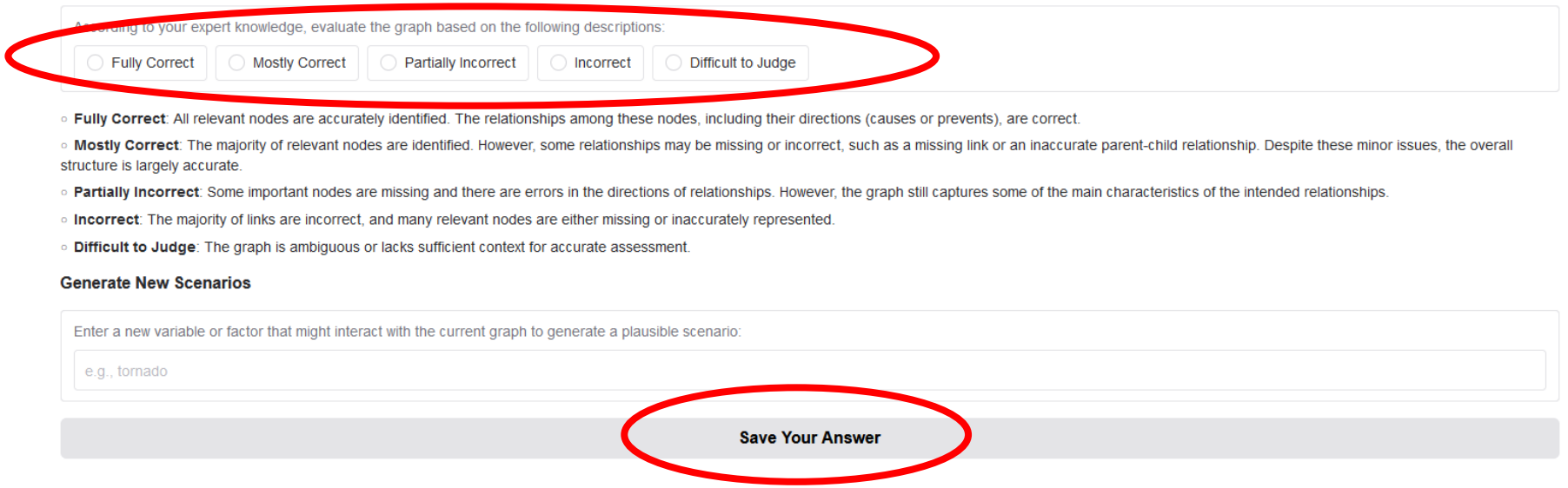


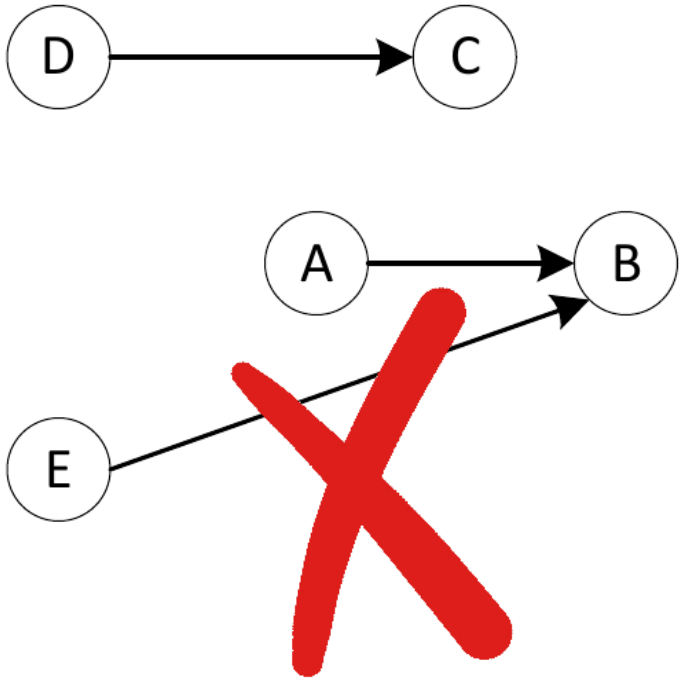
Insert a new relevant factor (driver, impact, hazard, ..anything) and generate a new plausible scenario!

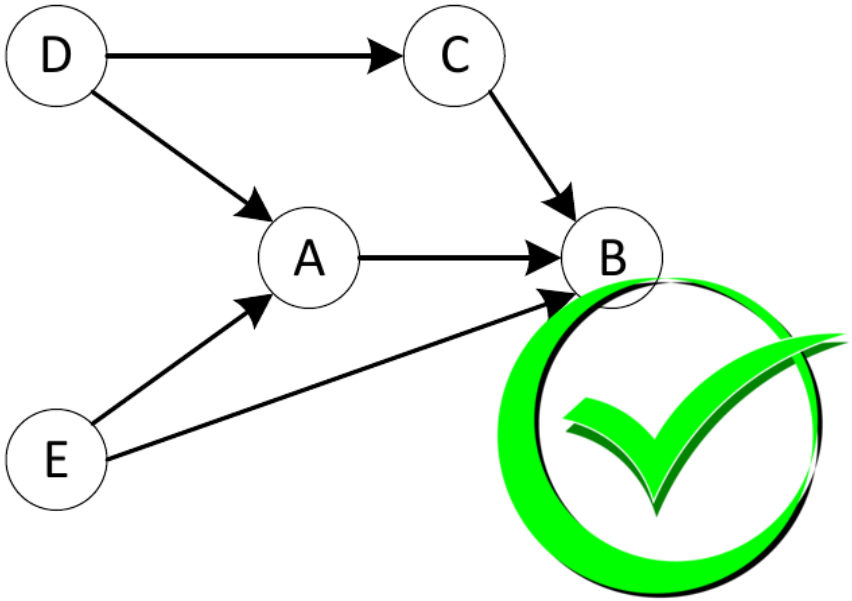


Let’s evaluate it together!

Dashboard at:
https://huggingface.co/spaces/jrc-ai/crisesStorylinesRAG







Thanks for your attention