

Module 5: Advanced technologies for environmental management

The Φ -lab Innovation Hub



• **Part of ESA's Climate Action, Sustainability and Science Department**

Focused on integrating AI, deep learning, and machine learning for real-time environmental monitoring → ESA's mission is to explore and protect the earth

• **Efficient analysis of massive satellite data**

Utilizes data from missions like **Copernicus** with unprecedented speed and precision

Key monitoring areas:

- Climate change
- Deforestation
- Sea level rise
- Air quality
- Rapid urbanization

The Φ -lab Innovation Hub



Advanced capabilities:

- AI detects and analyzes land cover changes to protect forests and biodiversity.
- Edge computing & AI onboard satellites reduce data processing time, enabling faster crisis response (wildfires, floods, disasters).

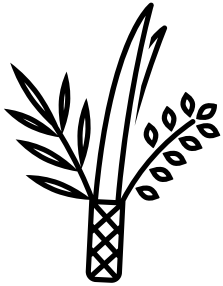
Deep learning applications:

- Improve land cover classification and global forest mapping.
- Support continuous deforestation monitoring and ecosystem impact analysis.

Beyond environment:

- Combines satellite data with epidemiological models to predict vector-borne diseases (e.g., dengue).
- Enables rapid, targeted public health interventions to reduce risks and boost community resilience.

Precision agriculture & air quality monitoring



Precision agriculture:

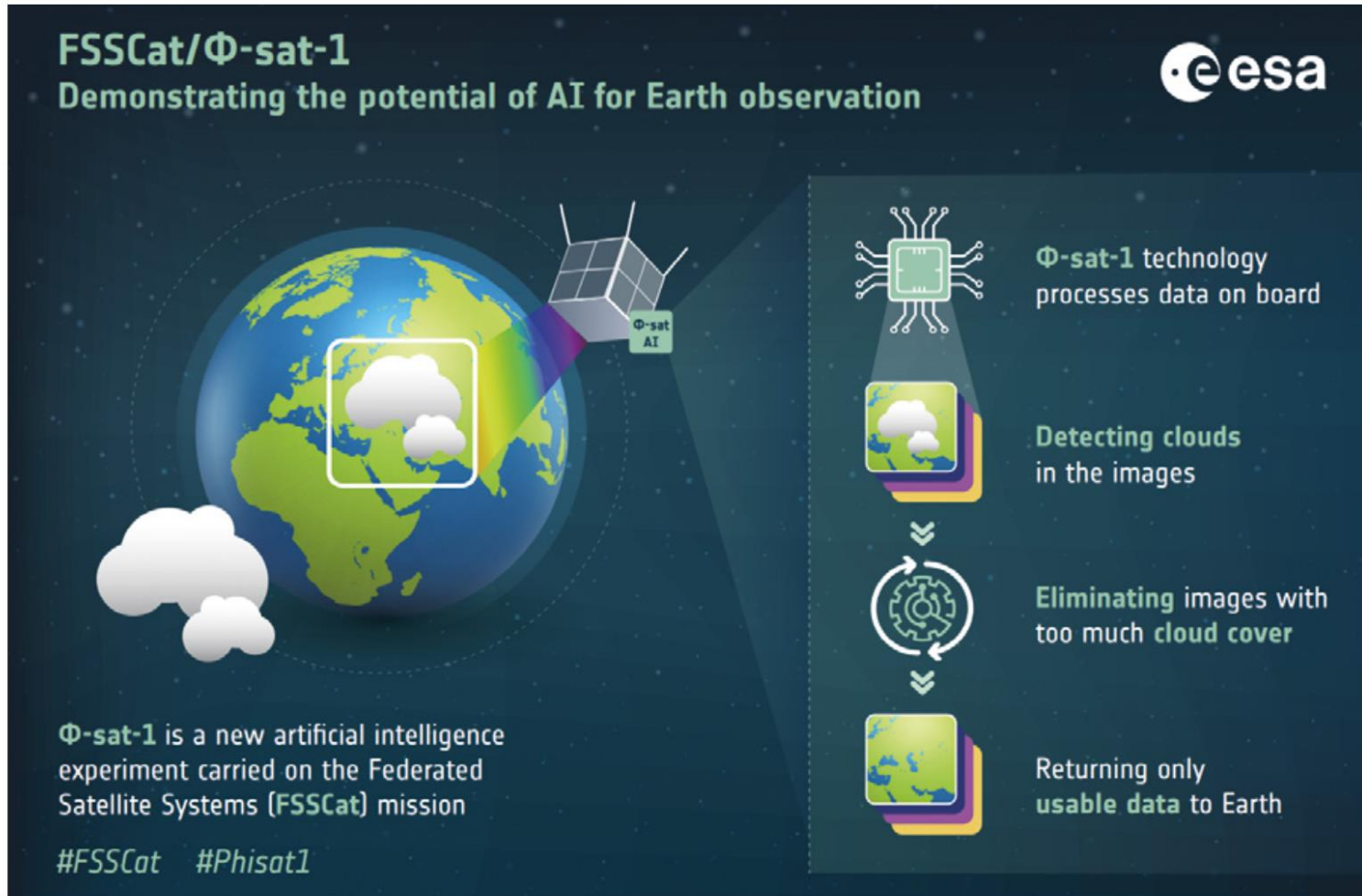
- AI algorithms analyze satellite images to monitor crop health, predict yields, and optimize natural resource use (water, fertilizers).
- Enhances production efficiency and reduces waste.
- Supports sustainable, profitable farming and addresses global food challenges amid climate change.



Air quality management:

- Φ-lab leverages satellite data (e.g., Sentinel-5P) to monitor pollutants like nitrogen dioxide (NO₂) and carbon monoxide (CO).
- Provides essential data for air quality policies.
- Supports new economic models to reduce emissions and improve public health.

Open platforms and democratization



FSSCat/ Φ -sat-1
Demonstrating the potential of AI for Earth observation

esa

Φ -sat-1 is a new artificial intelligence experiment carried on the Federated Satellite Systems (FSSCat) mission

#FSSCat #Phisat1

Φ -sat-1 technology processes data on board

Detecting clouds in the images

Eliminating images with too much cloud cover

Returning only usable data to Earth

COENIT-ESA/70774

Example: *Sentinel Hub*

Democratizes access to satellite data

Enables:

- Development of new applications
- Commercial solutions
- Competition between startups and large companies
- Dynamic ecosystem for innovation
- Maximizing technology potential for environmental sustainability

Environmental data more accessible to everyone

CONVERSATIONAL ARTIFICIAL INTELLIGENCE HAS ENORMOUS POTENTIAL TO MAKE ENVIRONMENTAL DATA MORE ACCESSIBLE BY ALLOWING USERS TO ASK QUESTIONS IN NATURAL LANGUAGE. WHAT TOOLS ARE AVAILABLE AND NEED TO BE DEVELOPED? WHAT ARE THE CHALLENGES TO ACHIEVE EFFECTIVE SOLUTIONS AND RELIABLE ANSWERS?

Enviromental data more accessible to everyone

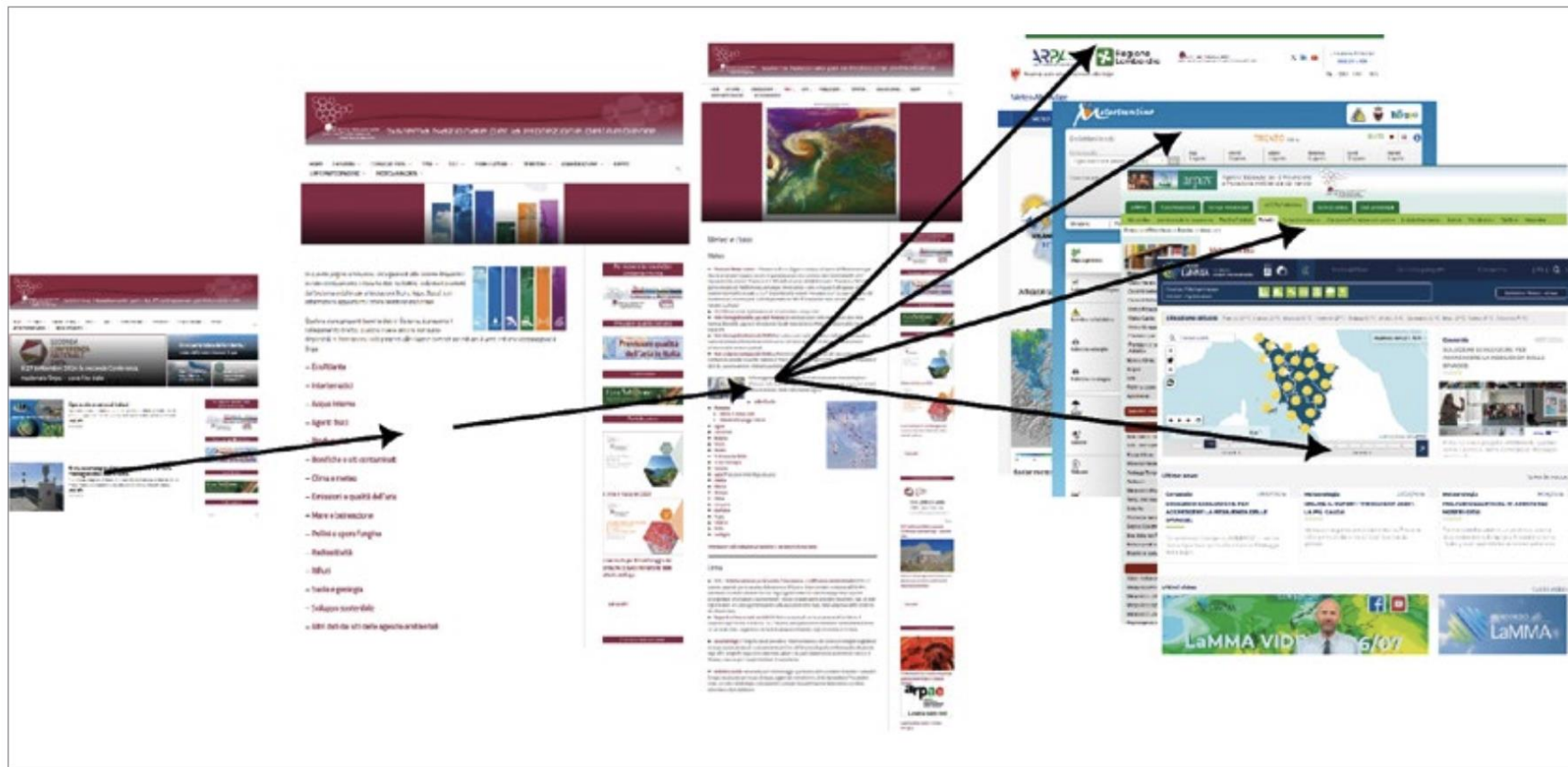


IMMAGINE: G. SCATEVA - S. CIATTONI - CC BY-NC-SA 4.0

FIG. 1 RICERCA DATI CLIMATICI

Rappresentazione del percorso di un utente alla ricerca di dati climatici per un confronto tra dati di regioni diverse.

AI tools enabling natural language access to data

- LLMs (Large Language Models) and RAG (Retrieval-Augmented Generation)



Limitations of LLMs

- May produce:
 - Inaccurate or outdated information
 - Confusion between similar terms in different contexts
- Challenges in specialized or rapidly evolving domains



Introducing the RAG Framework

- Combines LLMs with external knowledge sources
- Two main components:
 - **Retriever**: finds relevant info from databases/docs
 - **Generator**: produces natural language responses using retrieved info



Benefits of LLM + RAG

- More **accurate** and **context-aware** answers
- **Always grounded** in updated, authoritative sources
- No need to retrain the entire model
- Ideal for complex, dynamic domains like environmental data

- Transform how we interact with complex environmental data

Challenges of applying LLMs and RAG to environmental data

- Data often **fragmented** and from **heterogeneous sources**
- **Non-standard formats and schemas** complicate integration
- **Complex, context-specific causal relationships**
- **Technical language** difficult for general-purpose models
- **Unstructured and inconsistent** data further complicate processing

What's needed for effective AI solutions?

> **Domain-adapted retrievers**, using:

• **Embedding-based approaches:**

- Understand unstructured textual data

• **Graph-based approaches:**

- Represent spatial/temporal relationships and complex entity links

Challenges of applying LLMs and RAG to environmental data

Improving data quality and accessibility

- **Unify and centralize** data and documents
- Ensure:
 - **Well-structured, complete datasets**
 - **Comprehensive, gap-free document corpora**
- Result:
 - Enhanced retriever performance
 - **More accurate and faster responses**

Case study



IMMAGINE: G. SCATENA - S. CIATTONI - CC BY-NC-SA 4.0

FIG. 2 QUIVR

Un esempio di creazione di un sistema con Quivr: una volta aggiunta la conoscenza (a) è possibile interrogarla in linguaggio naturale in modo conversazionale (b).

Case study



Ciao, come posso aiutarti ?

Dimmi il dato di ar_id_p3 per la Regione Toscana e quella del Lazio a confronto

Ecco i dati richiesti per la variabile ar_id_p3 (Superficie delle aree a pericolosità idraulica elevata in kmq) per le regioni Toscana e Lazio:

- **Toscana:** 1413.166 kmq
- **Lazio:** 442.498 kmq

Questi dati mostrano che la superficie delle aree a pericolosità idraulica elevata è significativamente maggiore nella Toscana rispetto al Lazio.

cod reg	regione	ar id p3	Source
9	toscana	1413.166	📄
12	lazio	442.498	📄

regioni_pir.xlsx.xlsx

IMMAGINE: G. SCATENA - S. CIATTONI - CC BY-NC-SA 4.0

FIG. 3 RAGFLOW

Esempio di interazione con RagFlow: l'agente, a cui è stata creata una base di conoscenza costituita unicamente da fogli di calcolo, comprende la domanda, seleziona e cita la fonte dati pertinente, restituisce il valore desiderato estraendola da essa, spiegandone il significato ed elaborando una valutazione contestuale.



BREAK