



ARSINOE



**ARSINOE**  
CLIMATE RESILIENT-REGIONS THROUGH SYSTEMIC SOLUTIONS AND INNOVATIONS

## SustainGraph: a Knowledge Graph for tracking the progress and the interlinking among the SDGs

Eleni Fotopoulou, Ioanna Mandilara, Christina-Maria Androna,  
Anastasios Zafeiropoulos, Symeon Papavassiliou

Institute of Communications and Computer Systems (ICCS),  
National Technical University of Athens (NTUA)



This project has received funding from the European Union's Horizon H2020 innovation action programme under grant agreement 101037424.



# The SustainGraph Team



**Eleni Fotopoulou**

[efotopoulou@netmode.ntua.gr](mailto:efotopoulou@netmode.ntua.gr)



**Ioanna Mandilara**

[ioannamand@netmode.ntua.gr](mailto:ioannamand@netmode.ntua.gr)



**Christina-Maria Androna**

[andronaxm@netmode.ntua.gr](mailto:andronaxm@netmode.ntua.gr)



**Anastasios Zafeiropoulos**

[tzafeir@cn.ntua.gr](mailto:tzafeir@cn.ntua.gr)



**Prof. Symeon Papavassiliou**

[papavass@mail.ntua.gr](mailto:papavass@mail.ntua.gr)



**NETMODE**  
NETWORK MANAGEMENT & OPTIMAL DESIGN LAB

<https://www.netmode.ntua.gr/>

# Network Management and Optimal Design Laboratory (NETMODE)

- The Institute of Communication and Computer Systems (ICCS) is a research institute of the School of Electrical and Computer Engineering (ECE) of the National Technical University of Athens (NTUA)
- Network Management and Optimal Design Laboratory (NETMODE)
- Consists of: 4 faculty members, 10 postdoctoral researchers, 20+ PhD students and research engineers
- NETMODE research concentrates on the design, optimization and management of heterogeneous networks and distributed dynamic systems, emphasizing on: optimization and orchestration in 5G/6G Networks, knowledge and data analysis, knowledge graphs, machine learning techniques, complex systems analysis, smart cities, and performance evaluation of stochastic systems.
- NETMODE members have extensive experience in Future Internet Research Experimentation and have deployed a Future Internet Testbed. NETMODE is actively involved in several RTD programs sponsored by National and European organizations.

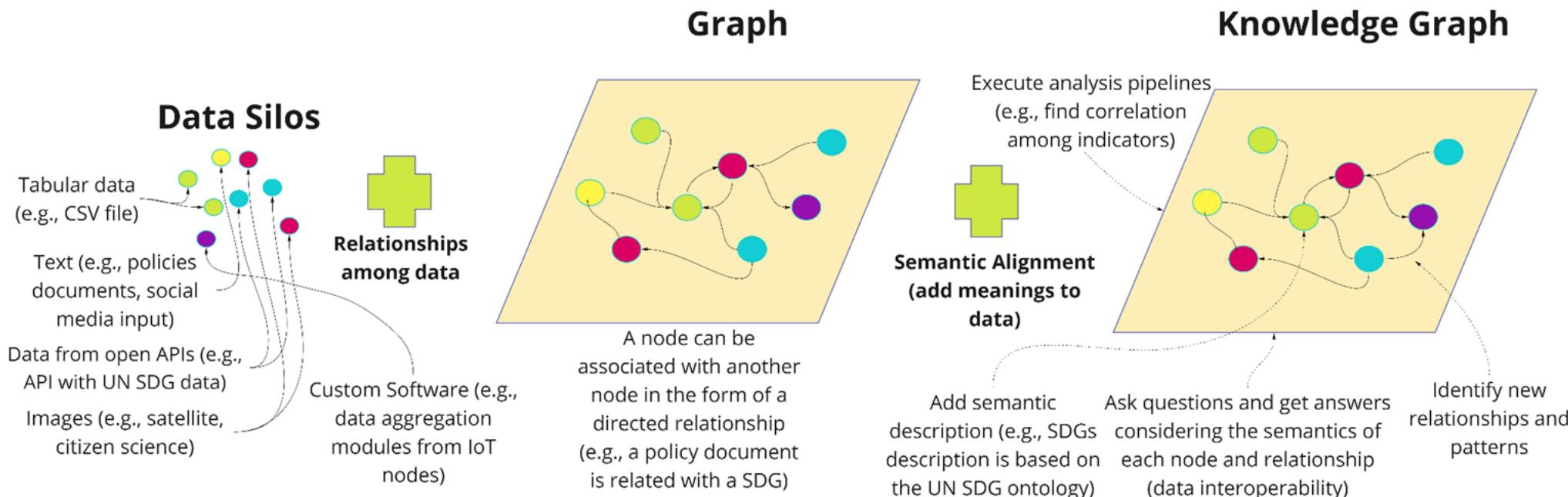
<https://www.netmode.ntua.gr/>

# Why we need SustainGraph? (1/2)

- **Existence of many data silos around the SDGs**
  - SDG datasets in different formats and represented based on different semantics
  - Definition of indicators may differ per geographical area
  - Multiple policies documents and directives around the climate change
  - Customized software and Application Programming Interfaces (APIs) to do one thing
    - Data semantics are usually hidden from end users
  - Need for data interoperability and re-usability to develop resilient adaptation and mitigation solutions for the climate change
- **Development of SustainGraph**
  - a knowledge graph to represent data around the SDGs that can be interconnected and enriched with meaning to explicitly represent knowledge
  - Interlinking of the represented concepts with well-defined semantics
  - alignment of terminologies of the same concepts under different data schemas to facilitate interdisciplinary studies
  - data population mechanisms for time-series data, documents, files in tabular format, considering the time and spatial scale

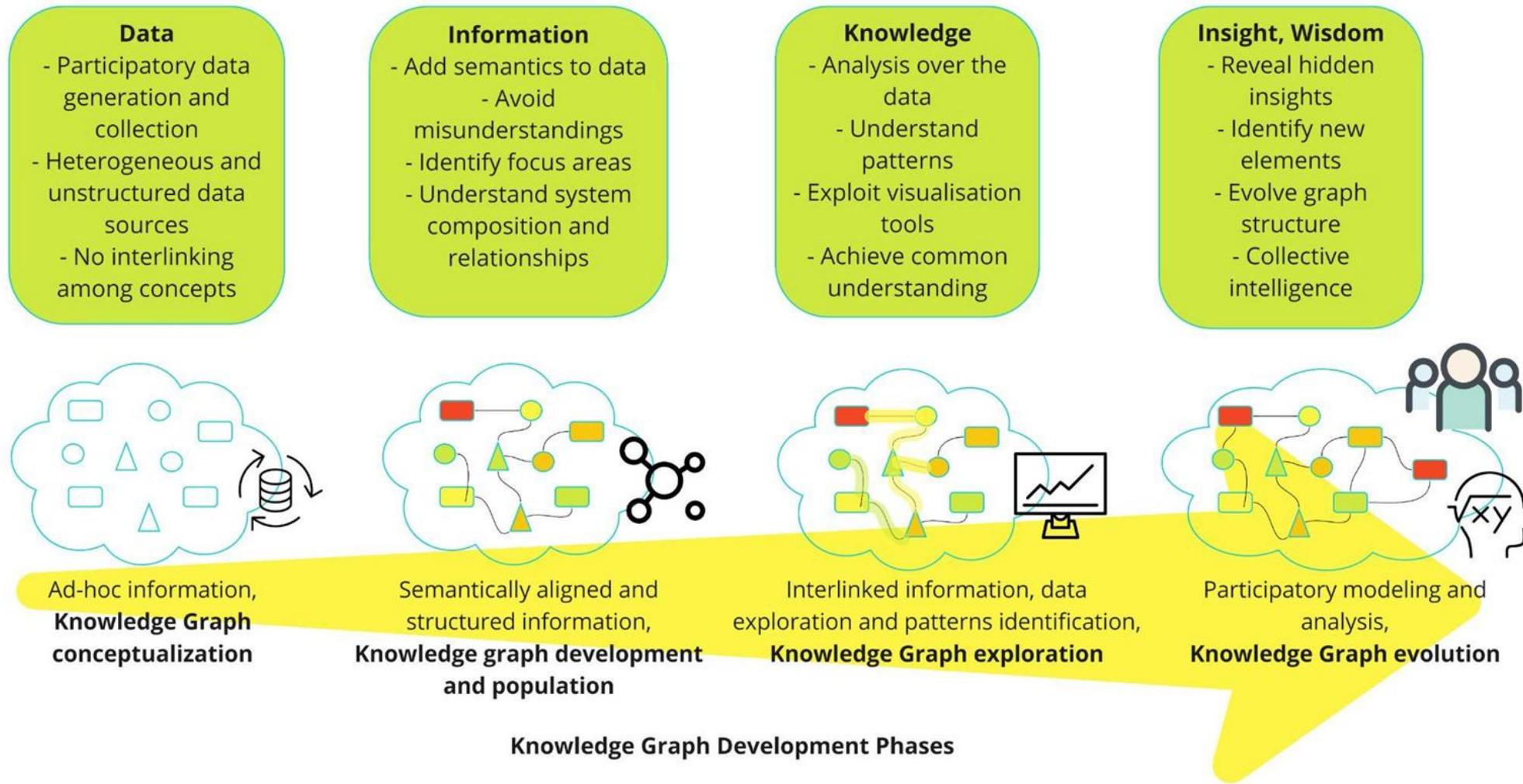
# Why we need SustainGraph? (2/2)

- **Management of data volatility and assurance of data quality**
  - relationships among nodes can be dynamic
  - representation of complex and dynamic socio-ecological systems
  - quality management processes
- **Reasoning and analysis over the available data**
  - identification and prediction of new relationships among entities
  - participatory socio-environmental systems analysis
  - identify transformative patterns, extract new knowledge and insights and assess the impact of climate change scenarios



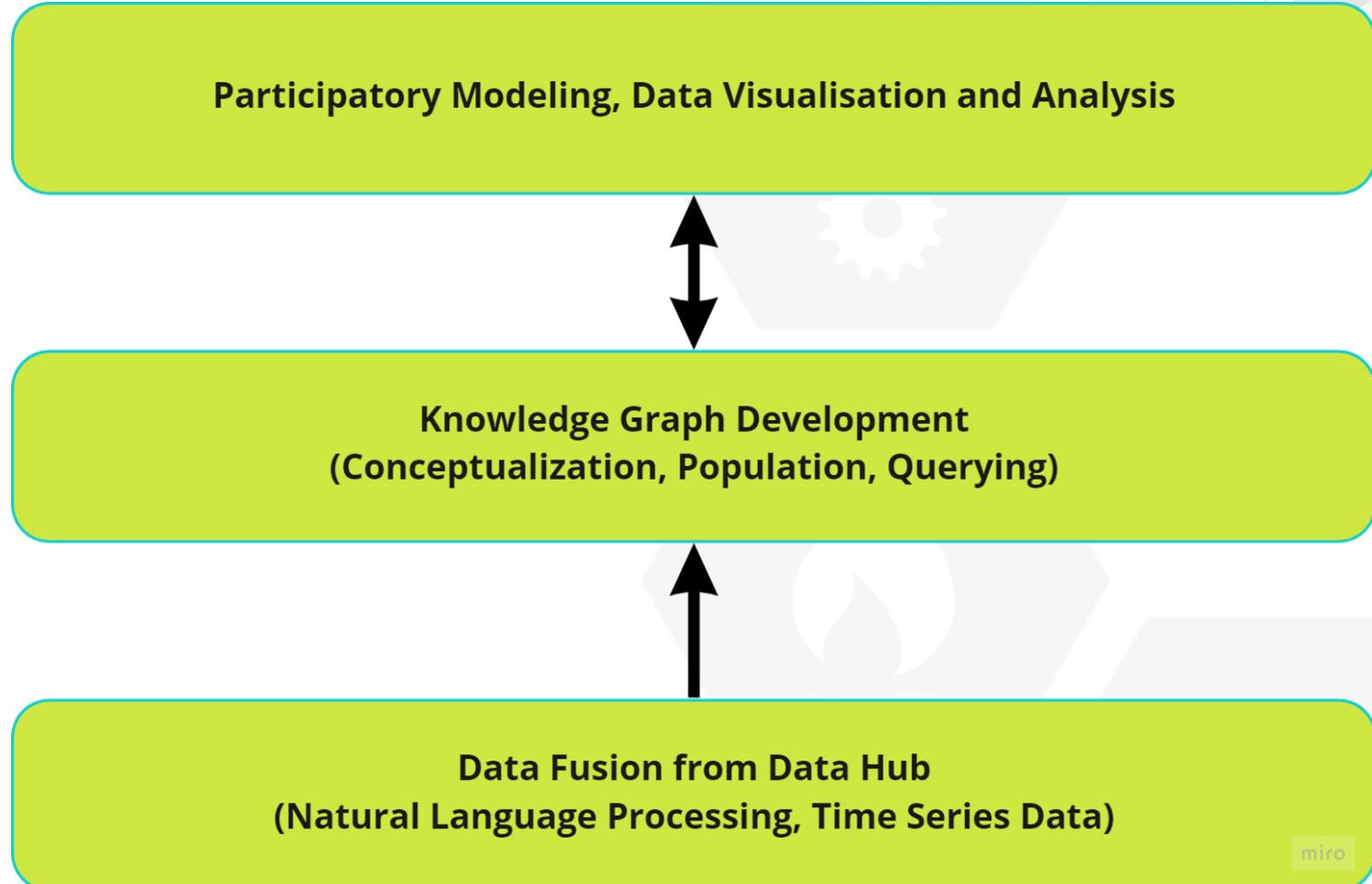
# Interplay between a Systems Innovation Approach....

## .... and a Knowledge Graph development

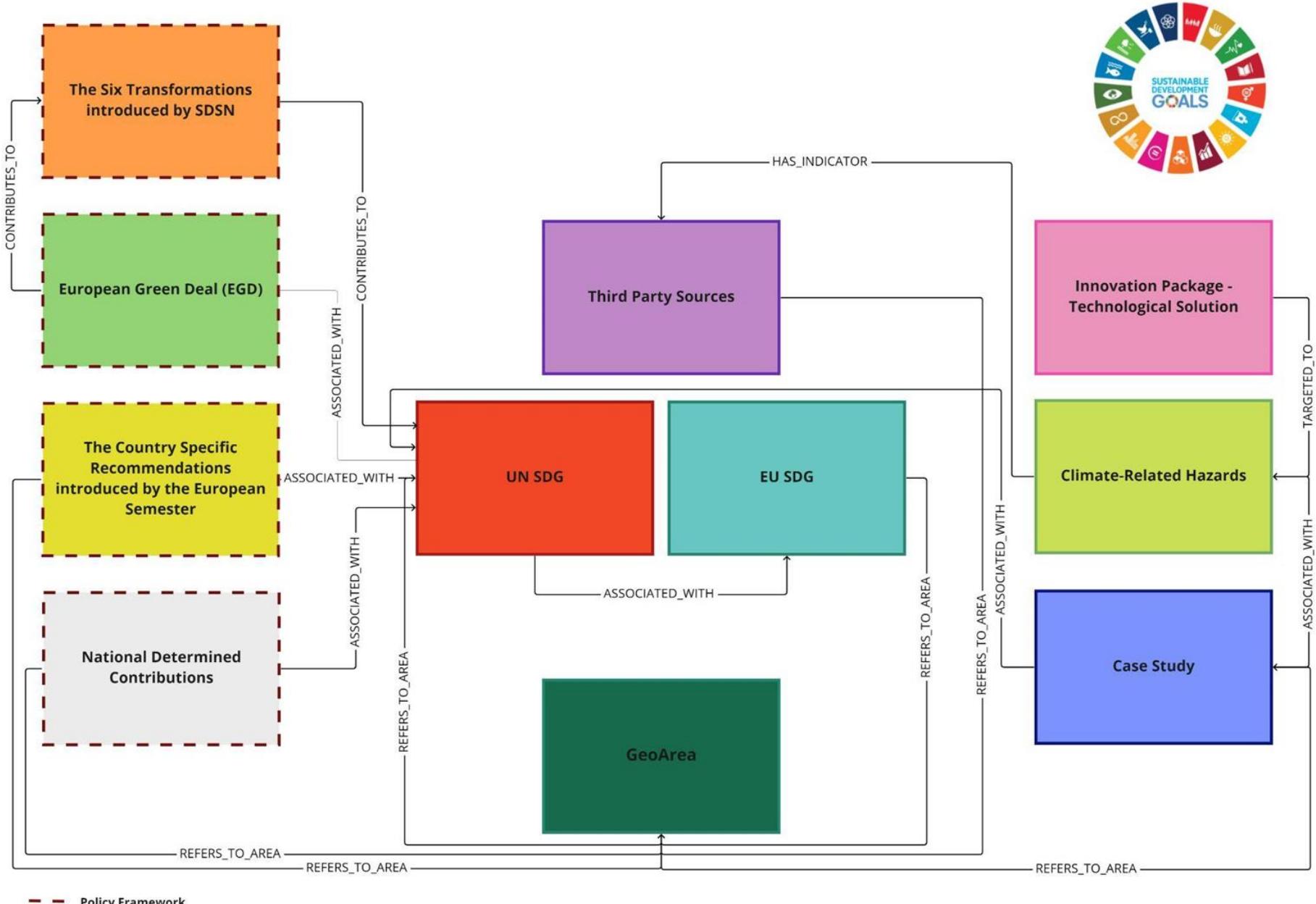


# Knowledge Graph Ecosystem in ARSINOE

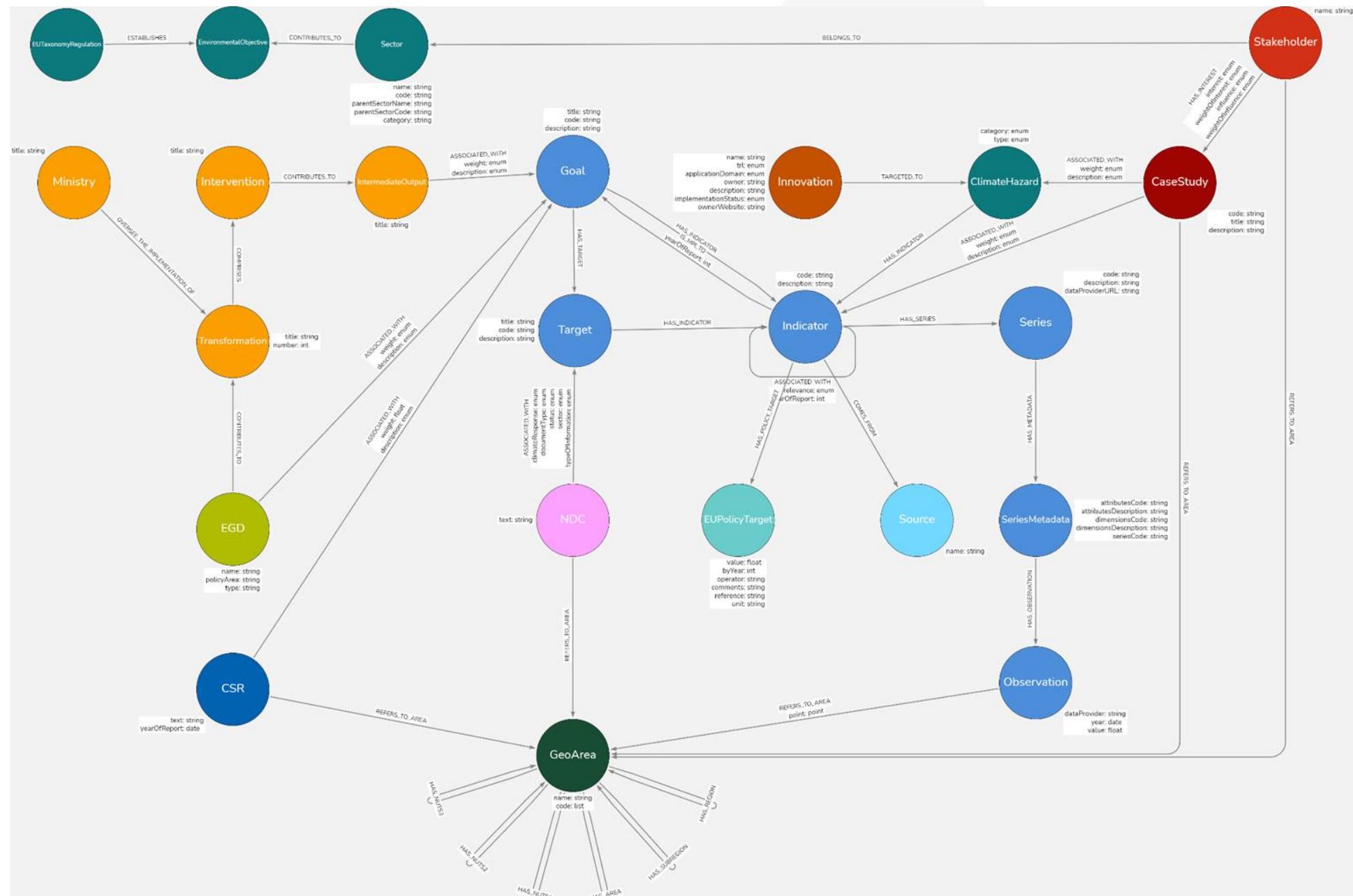
## Layered Approach



# High level view of SustainGraph



# Detailed view of SustainGraph



# SustainGraph Key Characteristics

- Consider **various types** of data:
  - time series data
  - text/documents
  - tabular data
- Where applicable, include **geolocation** characteristics (e.g., geometry points from data coming from Copernicus service)
- Set of **data population pipelines** taking advantage of **ML techniques**
- Multiple **visualization tools** to cover needs of different stakeholders
- Alignment of terms with the **SustainGraph Ontology** (keep as much as possible semantic consistency and alignment)
  - Described using W3C RDF Schema and the Web Ontology Language.

<https://gitlab.com/netmode/sustaingraph>

<https://netmode.gitlab.io/sustaingraph-ontology/>

# Querying and Visualization Tools (1/2)

1 MATCH (g:goal)-[r1:hasTarget]→(t:target)-[r2:hasIndicator]→(i:indicator)-[r3:hasSeries]→(s:serie)-[r4:hasObservation]→(o)  
2 where g.code = "1" and t.code="1.5" and i.code = "1.5.1"  
3 RETURN g,r1,t,r2,i,r3,s

Neo4j Browser

NeoDash

Arsinoe

Main Page

Number of targets per SDGs

Category	Value
Goal	NumberofTargets

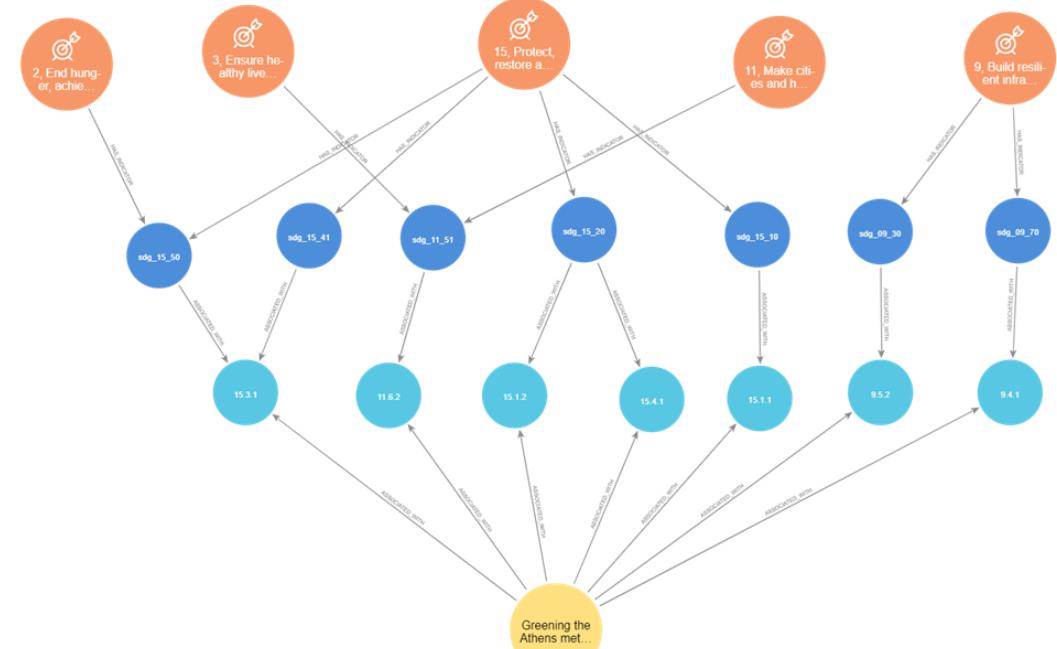
Portfolio investment (Balance of Payments, current United States)

X-value	Y-value
Year	Value

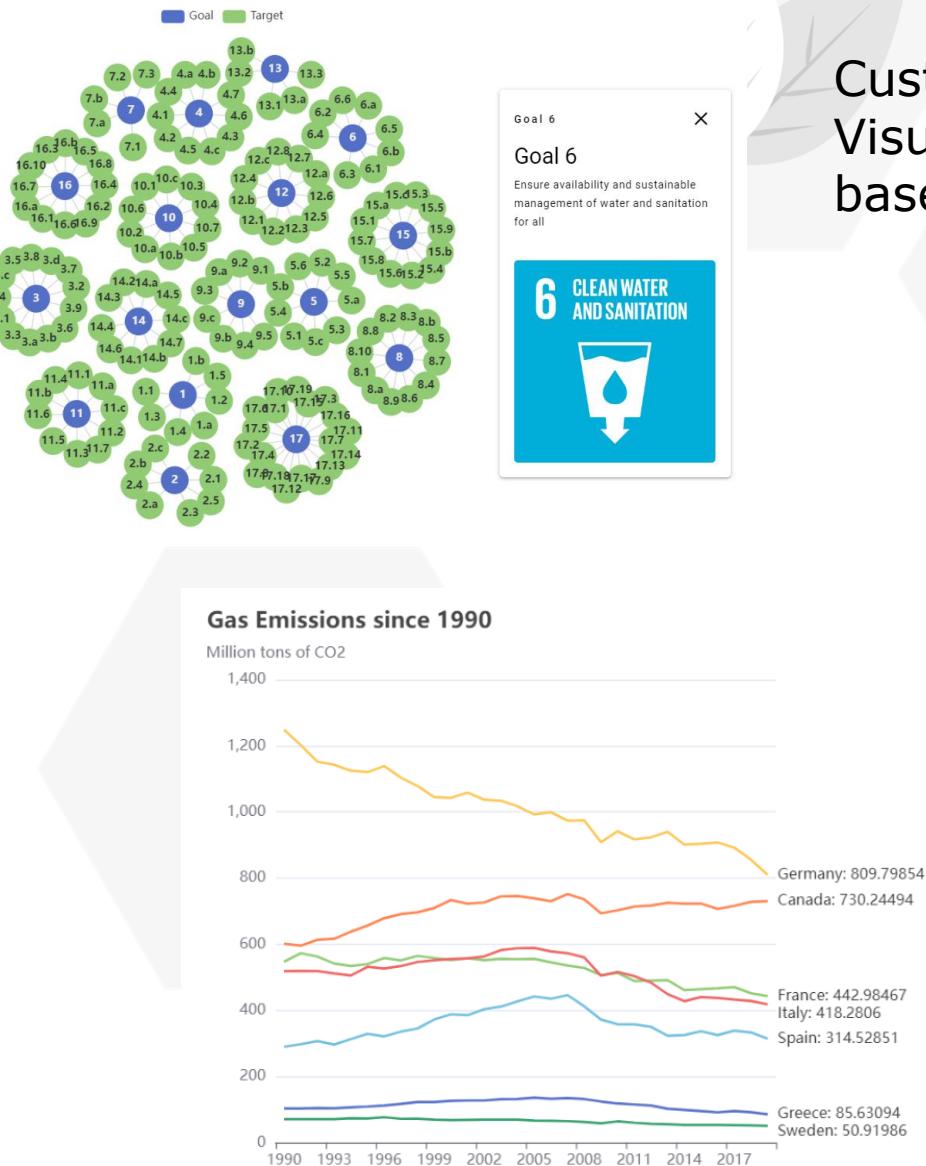
ARSINOE

# Querying and Visualization Tools (2/2)

## Neo4j Bloom

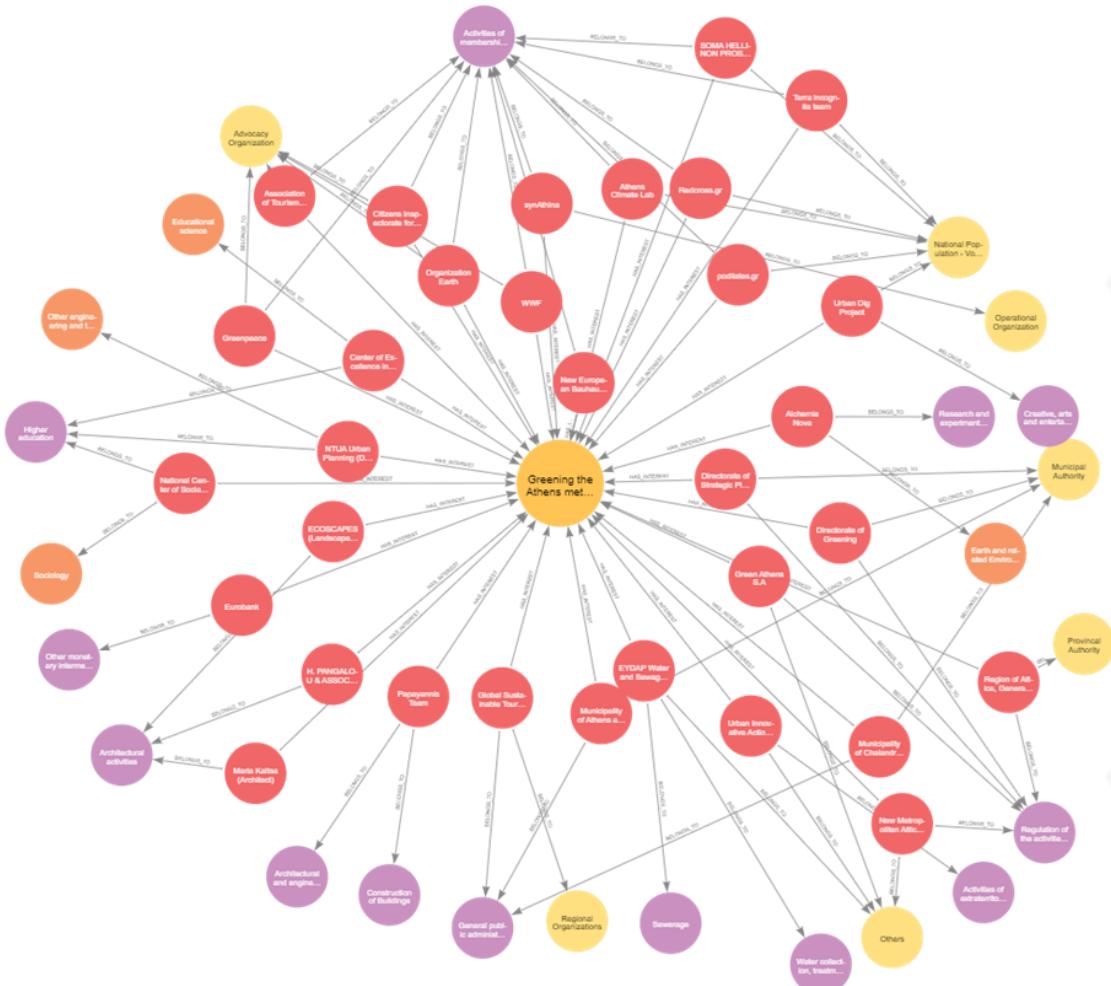


Relationship between the UN  
and the EU SDG indicators



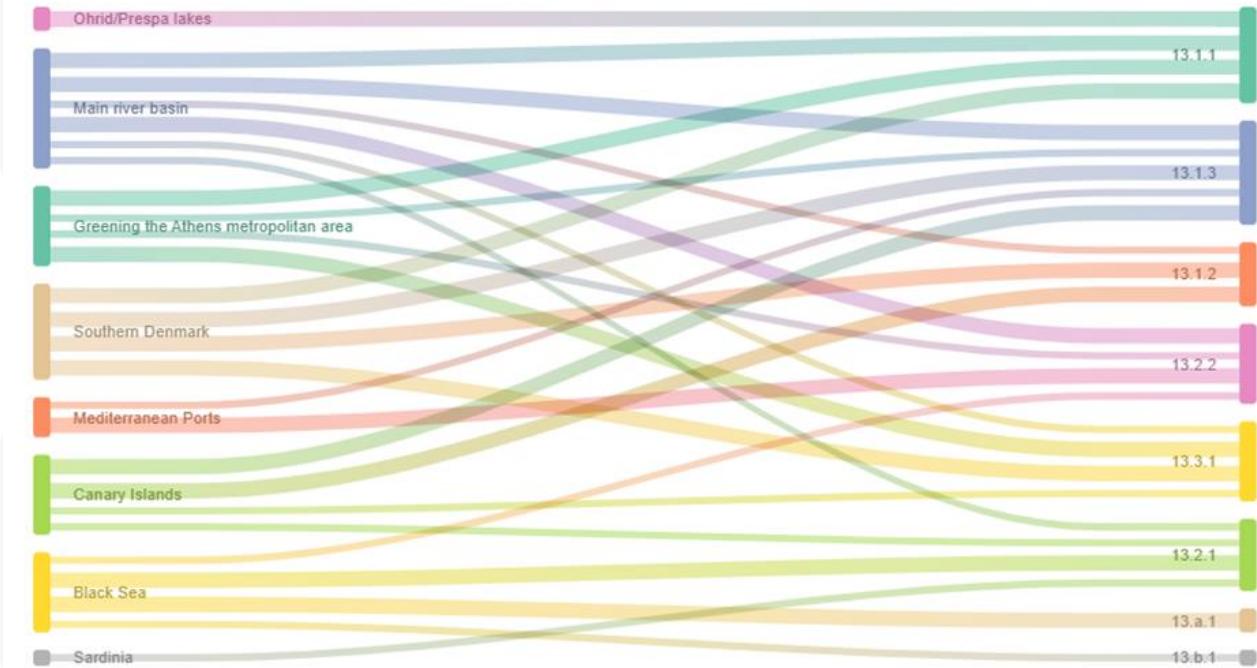
Custom  
Visualization kit  
based on Vue.js

# Indicative Visualizations (1/2)



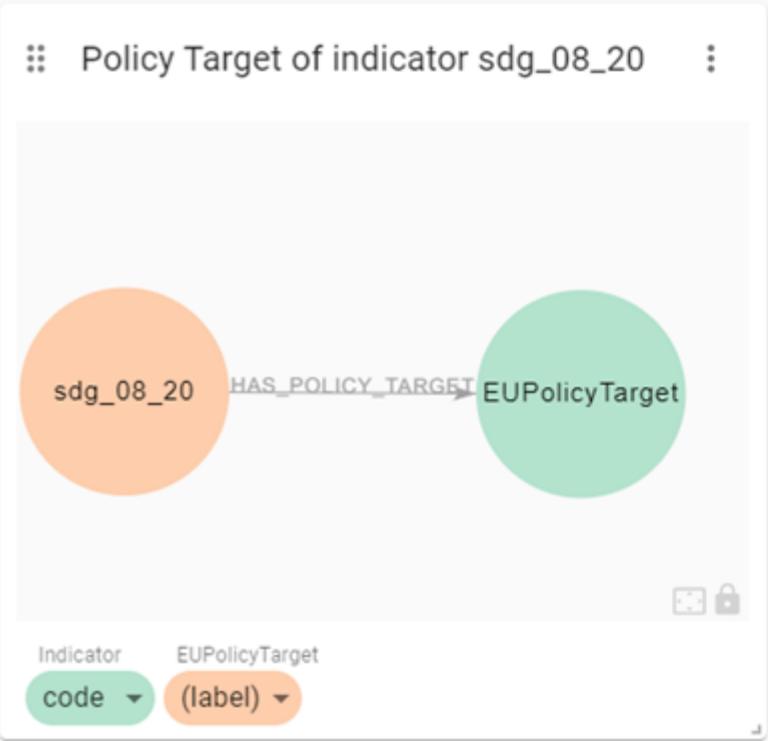
Stakeholders mapping in a  
case study in Athens  
**ARSINOE**

The association between the Case Studies and the UN Indicators of the SDGoal 13

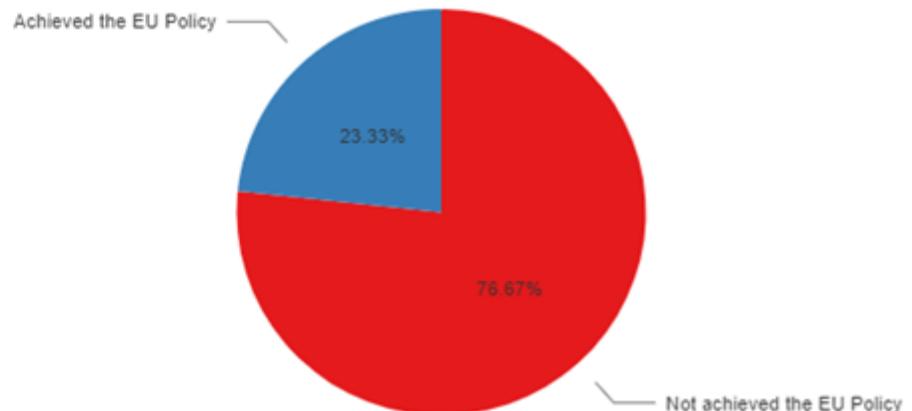


## Indicative Visualizations (2/2)

Policy Target of indicator sdg\_08\_20



Pie Chart of the indicator sdg\_08\_20

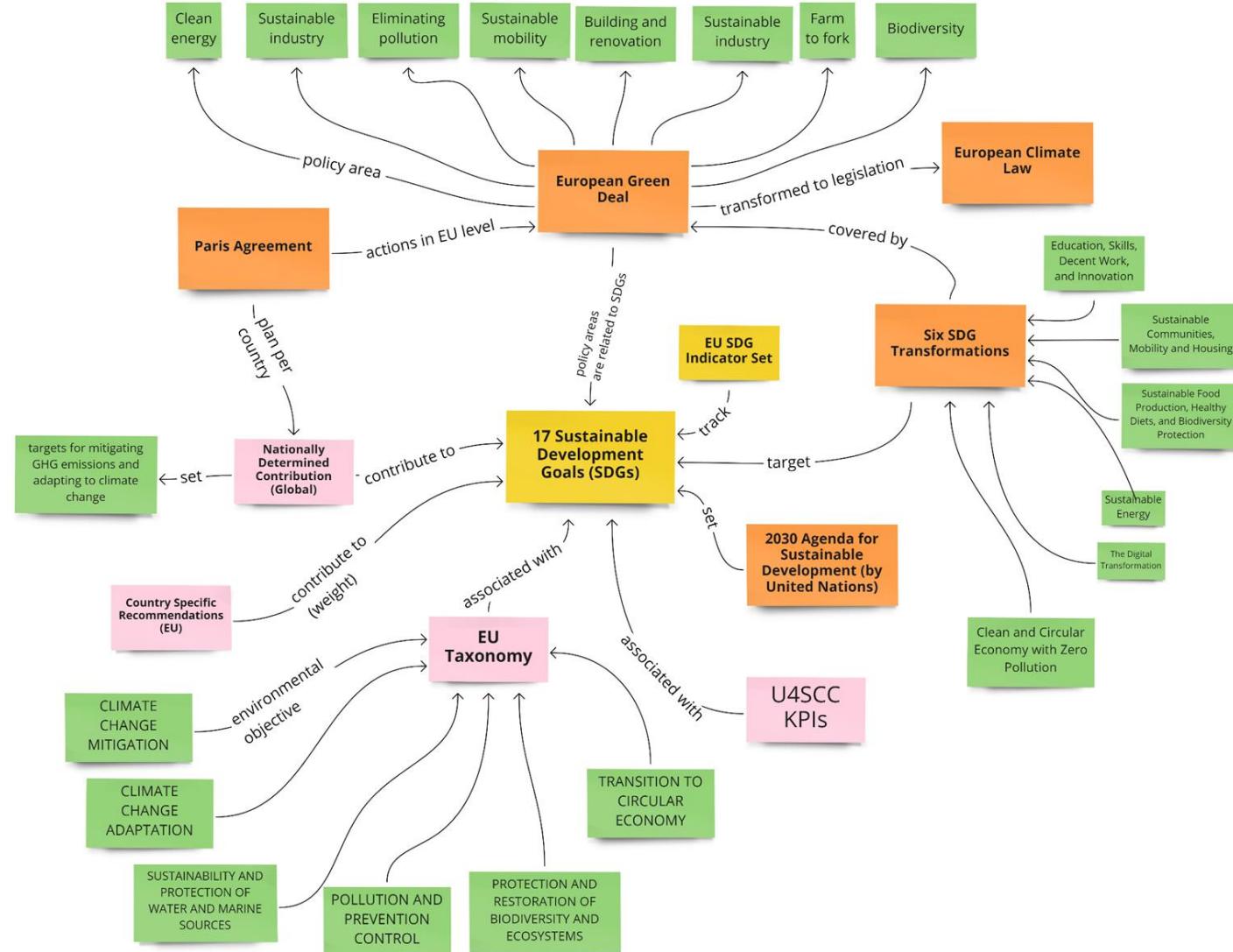


The European countries that have achieved or not the target

GeoArea	Status	Value
Netherlands	Achieved the EU Policy	5.5
Sweden	Achieved the EU Policy	6
Iceland	Achieved the EU Policy	7.3
Slovenia	Achieved the EU Policy	7.3
Norway	Achieved the EU Policy	7.4

Assessment for the achievement of a target for an EU SDG indicator

# Mapping of policies documents to SDGs...not a straightforward task



# Natural Language Processing for SDGs

## Introduce Texts to the SustainGraph

### Country Specific Recommendations:

- Documents prepared by the European Commission for each country analysing its economic situation and providing recommendations on measures it should adopt over the coming 12 months.

### European Green Deal Strategies

- Set of policy initiatives by the European Commission with the overarching aim of making the European Union (EU) climate neutral in 2050.

### NLP Techniques



# NLP Techniques

## Multi-label Classification using pretrained transformer-based models

*Training dataset: OSDG Community dataset - <https://github.com/osdg-ai/osdg-data>*

- Fine-tuning techniques on BERT, XLNet, GPT2 etc models
  - Train the entire architecture
  - Train some layers while freezing others
  - Freeze the entire architecture (*Transfer learning*)

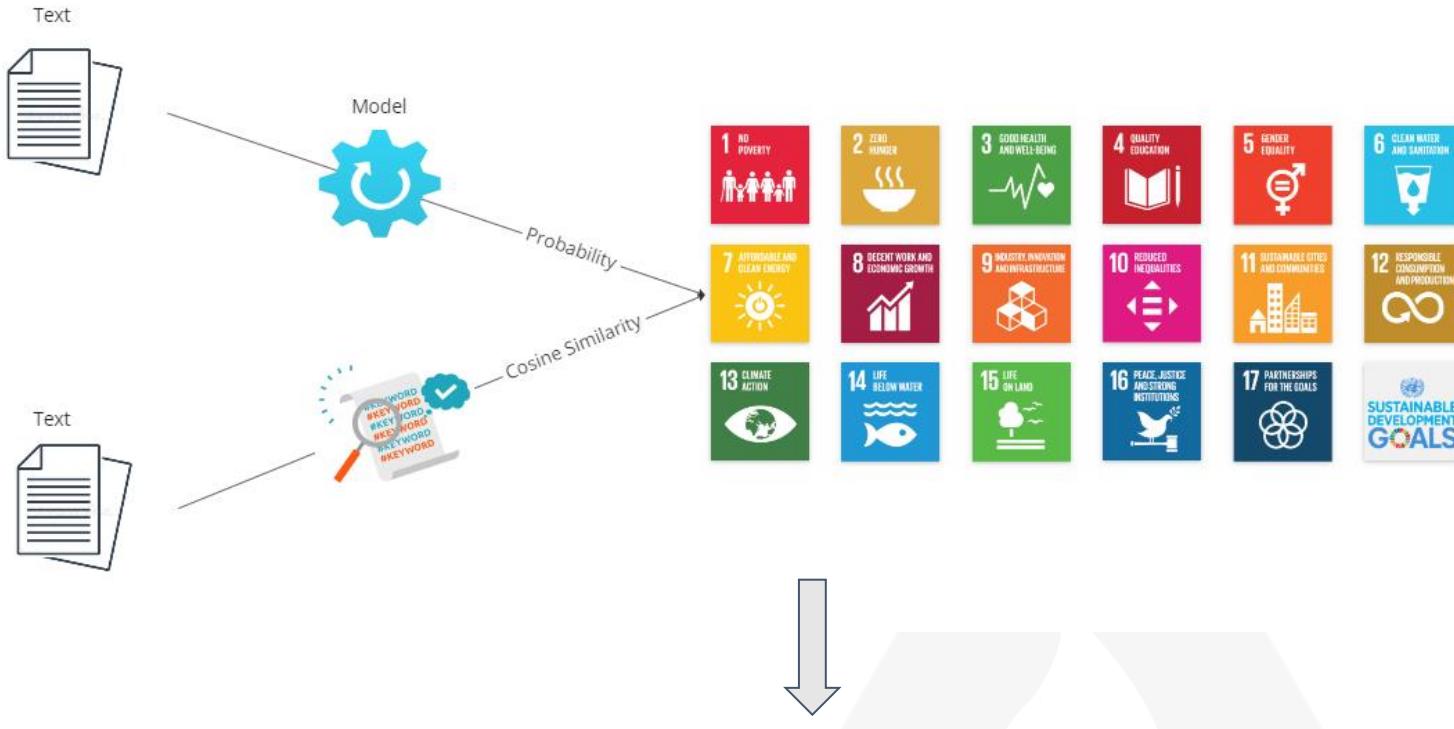
Output: Probability scores with each SDG

## Cosine similarity between the top n keywords of texts with the keywords of the SDGs

- Find the candidate keywords/key phrases of the document (Tokenize and count the word occurrences)
- Convert the keywords/key phrases and the document to numerical data (Embeddings)
- Find the top n keywords of text based on the cosine similarity between the embeddings of text's keywords and document
- Compute the cosine similarity matrix between the top n keywords of text and the keywords of SDGs (<https://sustainability.utoronto.ca/inventories/sustainable-development-goals-sdgs-keywords/> + <https://ap-unssdn.org/regional-initiatives/universities-sdgs/>).

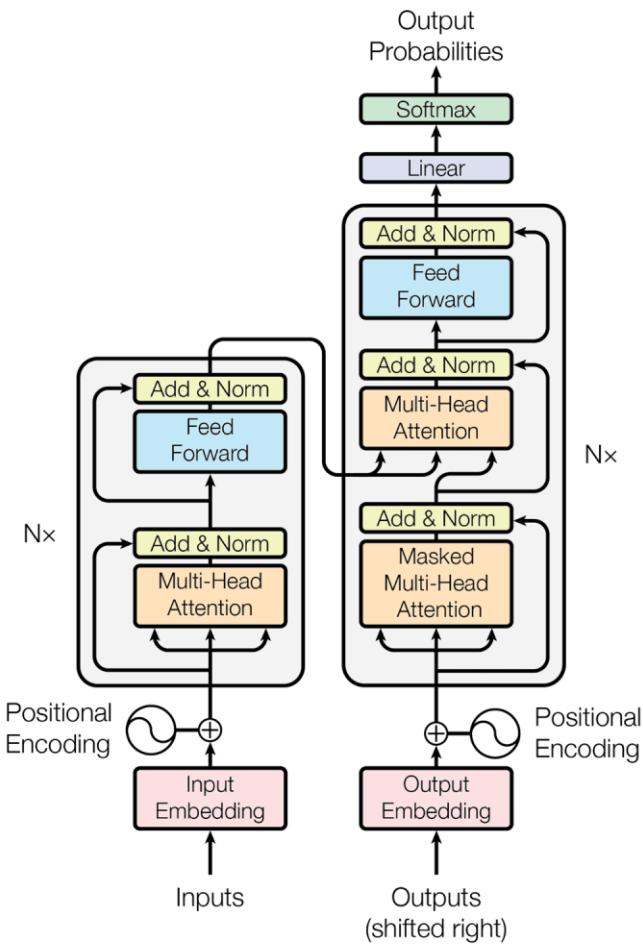
Output: Average Cosine Similarity Score with each SDG

# Combination of the two methods



$$r_{SDG} = \begin{cases} 0.7 * probability + 0.3 * avg\_cosine\_similarity & \text{for SDGs 1-16} \\ 0.5 * avg\_cosine\_similarity & \text{for SDG 17} \end{cases}$$

# Python Library SDGDetector



**Gitlab Repository :** <https://gitlab.com/netmode/sdg-detector/>

# SDGDetector Outcomes

How are the EGD Strategies associated with the SDGs?

A Farm to Fork Strategy

A hydrogen strategy for a climate-neutral Europe

EU Soil Strategy for 2030

EU Solar Energy Strategy

A Renovation Wave for Europe - greening our buildings, creating jobs, improving lives

Powering a climate-neutral economy: An EU Strategy for Energy System Integration

Sustainable and Smart Mobility Strategy - putting European transport on track for the future

A New Industrial Strategy for Europe

EU Strategy to reduce methane emissions

Chemicals Strategy for Sustainability

EU Strategy for Sustainable and Circular Textiles

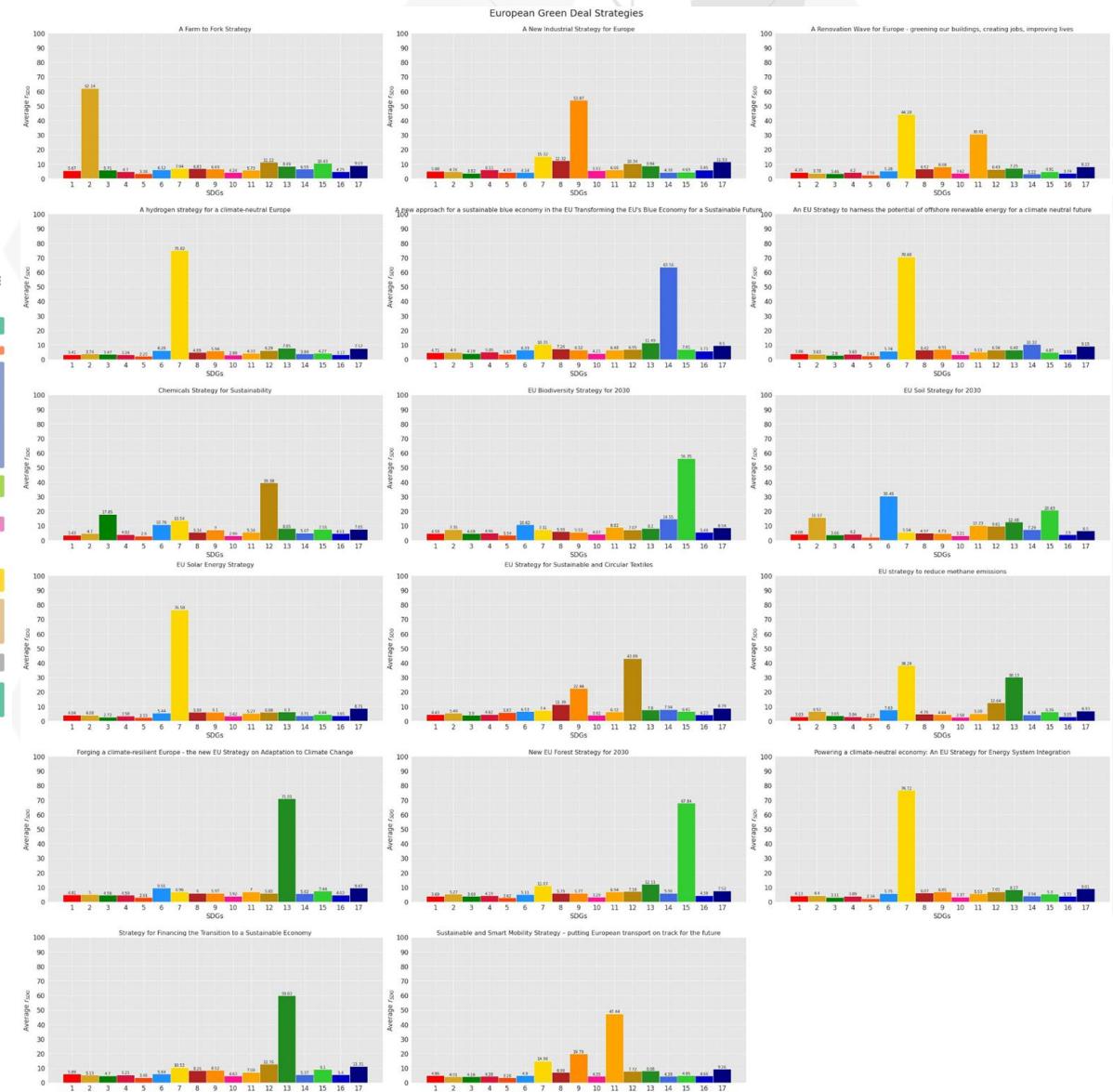
Forging a climate-resilient Europe - the new EU Strategy on Adaptation to Climate Change

Strategy for Financing the Transition to a Sustainable Economy

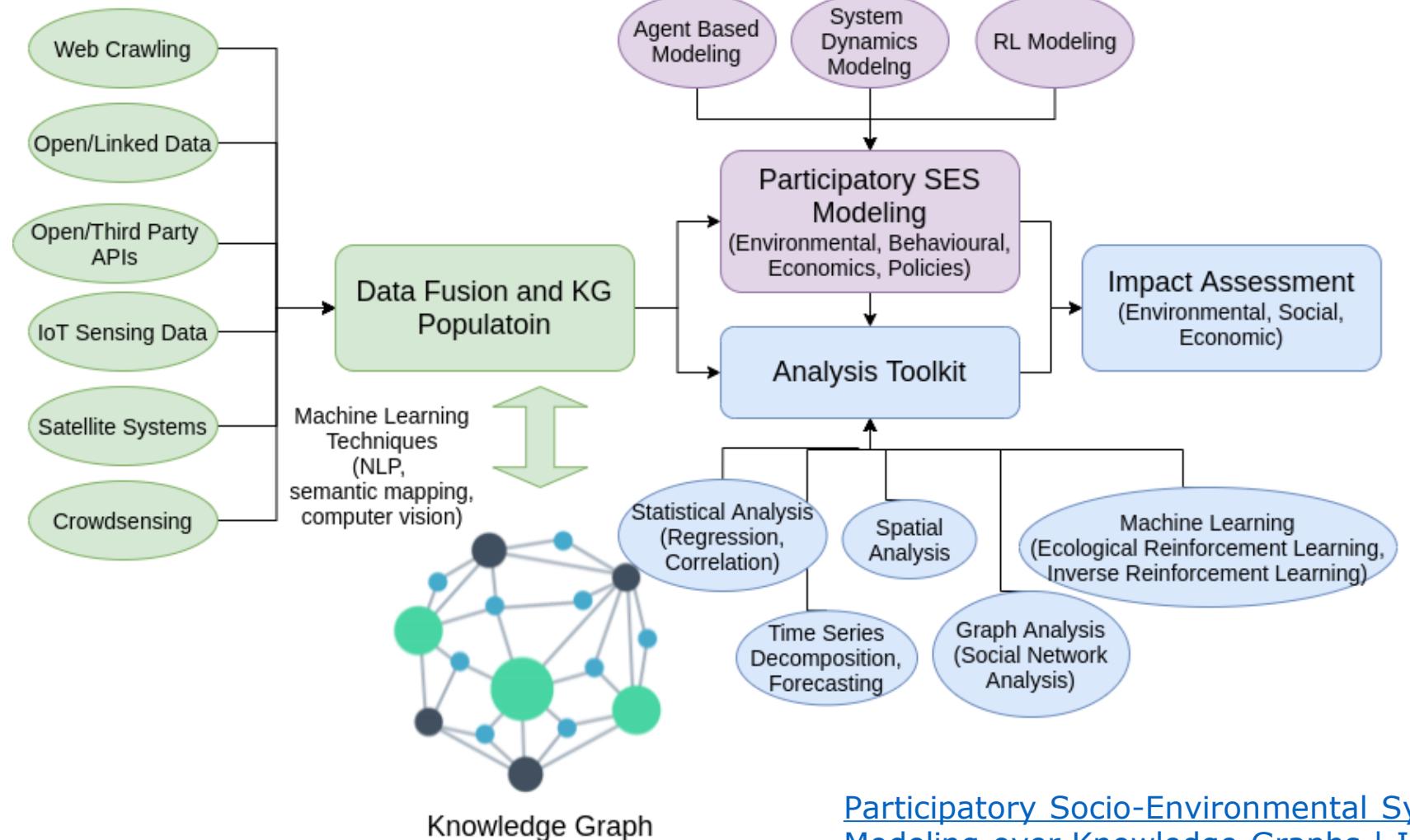
A new approach for a sustainable blue economy in the EU Transforming the EU's Blue Economy for a Sustainable Future

EU Biodiversity Strategy for 2030

New EU Forest Strategy for 2030

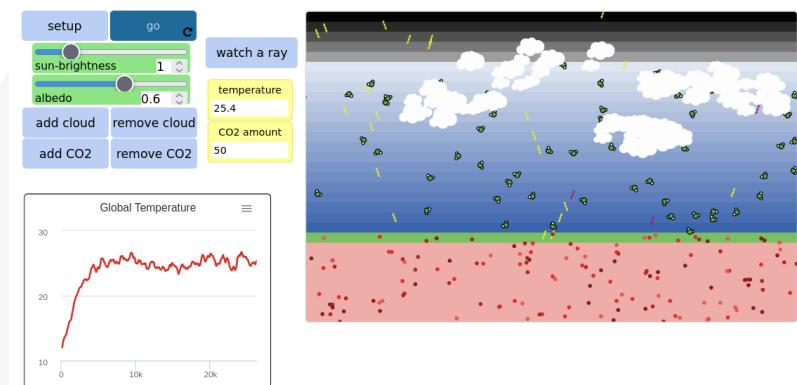
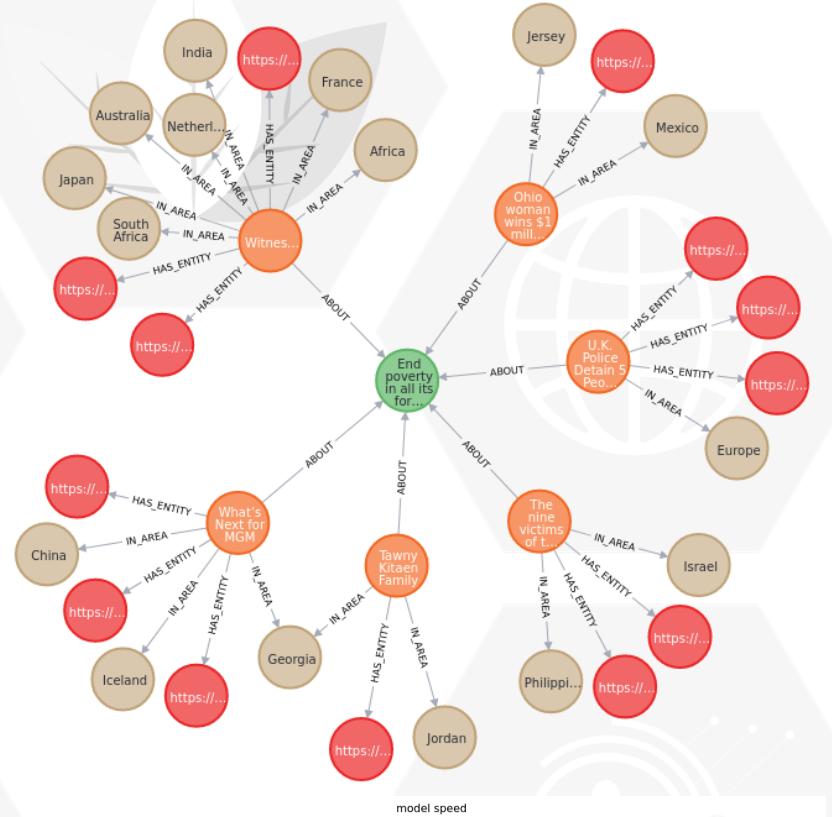
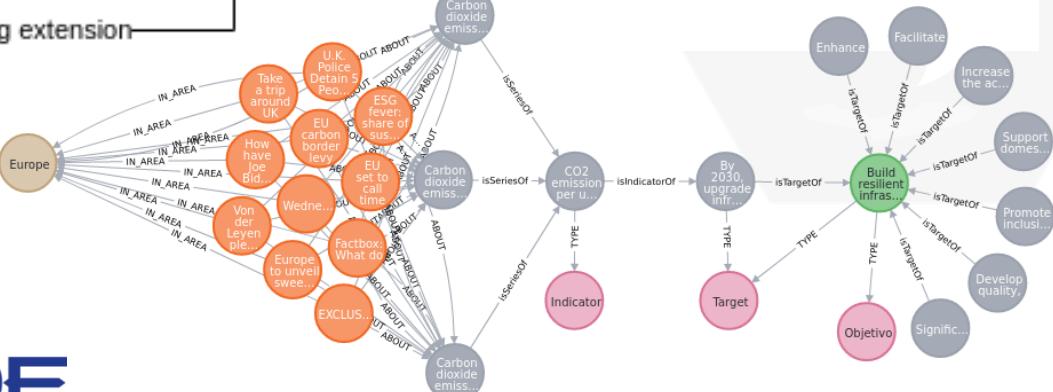
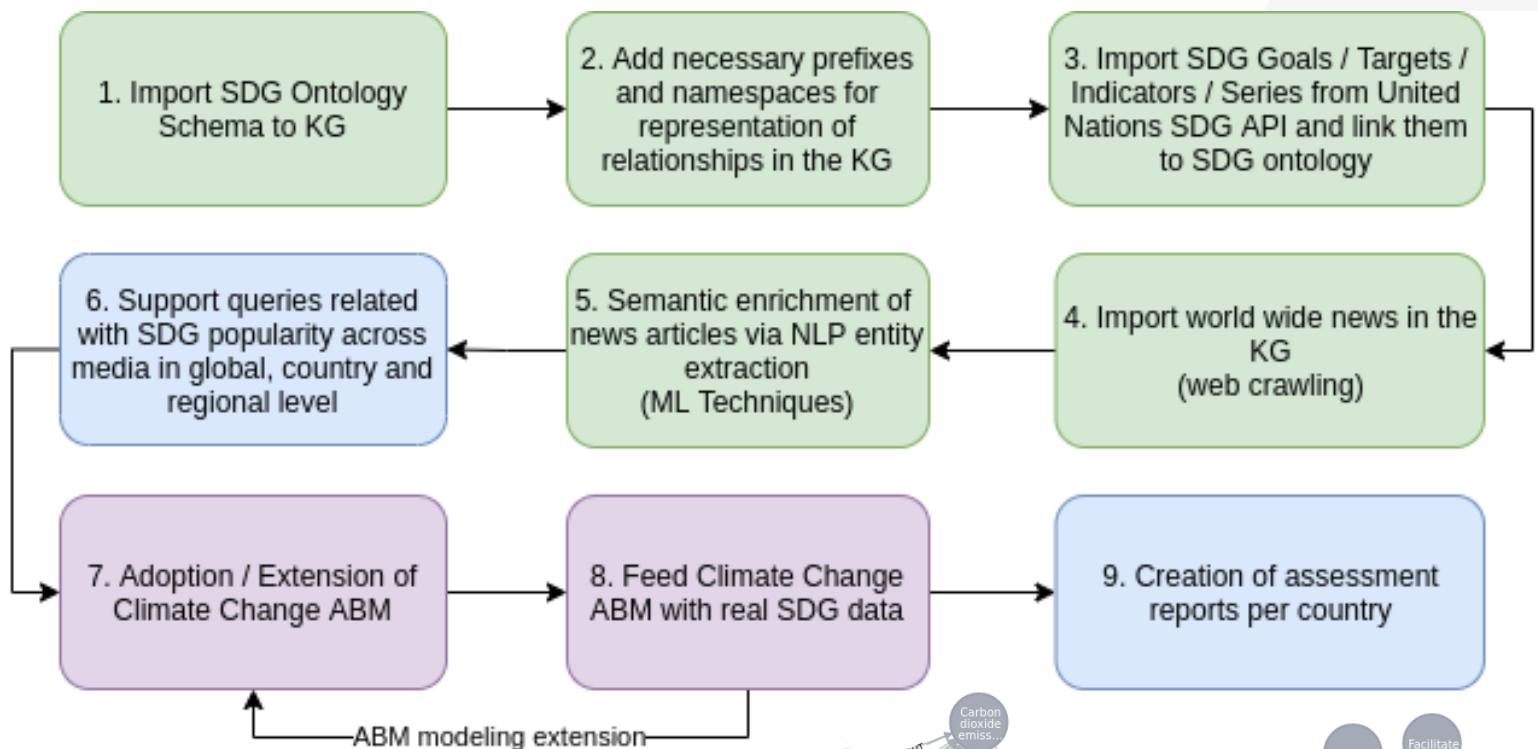


# Participatory socio-environmental systems modeling approach



[Participatory Socio-Environmental Systems  
Modeling over Knowledge Graphs | IEEE  
Conference Publication](#)

# Indicative Usage Scenario



For a live demo...join the Neo4j live session on Wednesday, 8th of February at 17:00 CET

<https://neo4j.com/event/neo4j-live-sustainingraph/>



« All Events

Note: Timings for all events are listed in the local timezone detected from your browser - Eastern European Standard Time

## Neo4j Live: SustainGraph – A Knowledge Graph for Sustainable Development

Wednesday, Feb 08 @ 06:00 pm GMT+2

Presenters: Eleni Fotopoulou, Ioanna Mandilara

Summary: In this talk, we are going to navigate through SustainGraph. SustainGraph is a Knowledge Graph that is developed in the framework of the ARSINOE H2020 project to track information related to the progress towards the achievement of targets defined in the United Nations Sustainable Development Goals (SDGs) at national and regional levels. The SustainGraph aims to act as a unified source of knowledge around information related to the SDGs, by taking advantage of the power provided by the development of graph databases and the exploitation of Machine Learning (ML) techniques for data population, knowledge production and analysis. Upon providing an overview of the main entities represented in SustainGraph, a demonstration session is going to take place. We are going to navigate in the graph based on a set of queries, produce visualizations and examine ML-driven data population pipelines.

Date:

Feb 08

Time:

06:00 pm - 07:00 pm

Event Category:

Live Stream

Language

English

Is Virtual?

Yes

Watch

# THANK YOU



Anastasios Zafeiropoulos



[tzafeir@cn.ntua.gr](mailto:tzafeir@cn.ntua.gr)



[info@arsinoe-project.eu](mailto:info@arsinoe-project.eu)



[www.arsinoe-project.eu](http://www.arsinoe-project.eu)



ARSINOE\_EU



This project has received funding from the European Union's Horizon H2020 innovation action programme under grant agreement 101037424.