

# MSCA Staff Exchanges : Starwars Project

Salem Benferhat

CRIL, CNRS UMR 8188, Université d'Artois,  
benferhat@cril.fr



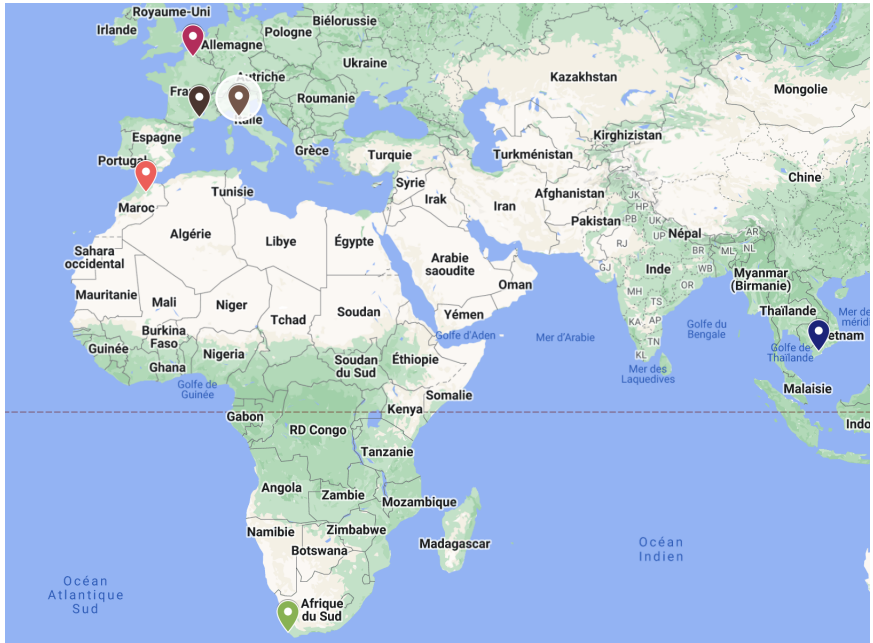
Funded by



# Starwars project

- A few words about the consortium and the project
- Feedback:
  - ▶ Project setup.
  - ▶ Grant agreement.
  - ▶ Project start-up.
- Brief discussion of the interest and impact of the project

# Consortium Partners



# Consortium Partners

- CNRS (Centre National de la Recherche Scientifique).
  - ▶ Project coordinator
  - ▶ CRIL (Centre de Recherche en Informatique de Lens), UMR8188
- UA (University of Artois, associated partner linked to CNRS)
- CNR (Consiglio Nazionale delle Ricerche)
  - ▶ ISTI (Institute for Information Science and Technology Institute), Pisa, Italy.
- IRD (Institut de recherche pour le développement)
- UM (University of Montpellier)
  - ▶ HSM (Research Lab HydroSciences Montpellier) IRD, CNRS, UMR 5569.

# Consortium Partners

- UCT (University of Cape Town)
  - ▶ CAIR (Centre for Artificial Intelligence Research), Cape Town, South Africa.
- USMBA (University of Sidi Mohamed Ben Abdellah),
  - ▶ LSIA (Intelligent Systems and Applications Lab), Fez, Morocco.
- CTU (Can Tho University),
  - ▶ CICT (College of Information and Communication Technology), Can Tho, Vietnam

# Starwars (STormwAteR and WastewAteR networkS heterogeneous data AI-driven management)

# Starwars (STormwAteR and WastewAteR networkS heterogeneous data AI-driven management)

There are three major questions in the acronym:

# Starwars (STormwAteR and WastewAteR networkS heterogeneous data AI-driven management)

There are three major questions in the acronym:

- What are the types of data to be processed?  
**heterogeneous data** of different nature



# Starwars (STormwAteR and WastewAteR networkS heterogeneous data AI-driven management)

There are three major questions in the acronym:

- What are the types of data to be processed?  
**heterogeneous data** of different nature
- What are the main tasks to be carried out?  
**AI-driven tasks**: representing, completing, merging, querying, etc.

# Starwars (STormwAteR and WastewAteR networkS heterogeneous data AI-driven management)

There are three major questions in the acronym:

- What are the types of data to be processed?  
**heterogeneous data** of different nature
- What are the main tasks to be carried out?  
**AI-driven tasks**: representing, completing, merging, querying, etc.
- What is the framework?  
**Stormwater and wastewater** network data

# Starwars (STormwAteR and WastewAteR networkS heterogeneous data AI-driven management)

There are three major questions in the acronym:

- What are the types of data to be processed?  
**heterogeneous data** of different nature
- What are the main tasks to be carried out?  
**AI-driven tasks**: representing, completing, merging, querying, etc.
- What is the framework?  
**Stormwater and wastewater** network data

# Starwars (STormwAteR and WastewAteR networkS heterogeneous data AI-driven management)

There are three major questions in the acronym:

- What are the types of data to be processed?  
**heterogeneous data** of different nature
- What are the main tasks to be carried out?  
**AI-driven tasks**: representing, completing, merging, querying, etc.
- What is the framework?  
**Stormwater and wastewater** network data

# Distributed heterogeneous data

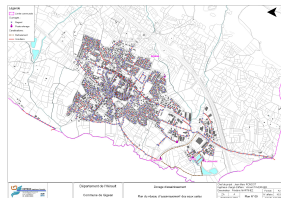


Images

# Distributed heterogeneous data



Images

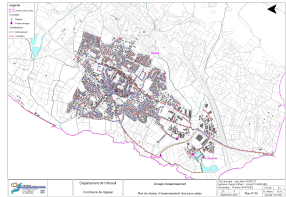


Maps

# Distributed heterogeneous data



Images

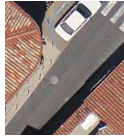


Maps

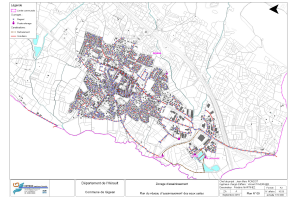
0.16	0.000000	
	0.000000	AC
	0.000000	PVC
	0.000000	PVC
	0.000000	AC
	0.000000	

Missing data

# Distributed heterogeneous data



Images



Maps

0.16	0.000000	
	0.000000	AC
	0.000000	PVC
	0.000000	PVC
	0.000000	AC
	0.000000	

Missing data



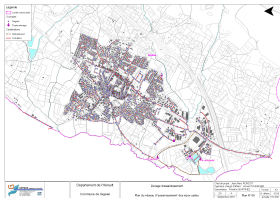
Videos



# Distributed heterogeneous data



Images



Maps

0.16	0.000000	
	0.000000	AC
	0.000000	PVC
	0.000000	PVC
	0.000000	AC
	0.000000	

Missing data



Videos

## Other forms of heterogeneous data/information

Textual data, expert knowledge, ontologies, constraints, etc.

## Imperfect data/information:

Unreliable, imprecise, incomplete, dynamic and conflicting.

Two distinct and important aspects:

Two distinct and important aspects:

- setting up the scientific part project

Two distinct and important aspects:

- setting up the scientific part project
- setting up the part linked to the secondments

# Scientific part project: representation languages



0.16	0.00000	
	0.00000	AC
	0.00000	PVC
	0.00000	PVC
	0.00000	AC
	0.00000	



Other heterogeneous data

# Scientific part project: representation languages



0.16	0.00000	
	0.00000	AC
	0.00000	PVC
	0.00000	PVC
	0.00000	AC
	0.00000	



map to

Common languages

ABoxes/databases

Lightweight ontologies

Uncertainty-based logics

Non-monotonic logics

etc.

Other heterogeneous data

# Scientific part project: representation languages



0.16	0.00000	
	0.00000	AC
	0.00000	PVC
	0.00000	PVC
	0.00000	AC
	0.00000	



Knowledge/  
Ontologies



map to

Common languages

ABoxes/databases

Lightweight ontologies

Uncertainty-based logics

Non-monotonic logics

etc.

Other heterogeneous data

# Scientific part project: representation languages



0.16	0.00000	
	0.00000	AC
	0.00000	PVC
	0.00000	PVC
	0.00000	AC
	0.00000	



Knowledge/  
Ontologies

map to

Common languages

ABoxes/databases

Lightweight ontologies

Uncertainty-based logics

Non-montonic logics

etc.

Query-answering

Other heterogeneous data



# Scientific objectives of the project

- Collecting, completing and representing heterogeneous information/data.
  - ▶ Identifying data sources and collecting different forms of data (including textual data) and knowledge.
  - ▶ Using machine learning, EM algorithms and inference algorithms for extracting, completing and deriving missing data.
  - ▶ Defining ontologies, graph-based and logical-based languages for representing different forms of data and knowledge
- New practical solutions for merging, conditioning and conflict management.

# Scientific objectives of the project

- Development of an intelligent and explainable query processing tool that combines ML and knowledge (ontology).
- Mapping wastewater and stormwater networks for better management and hydraulic modelling.
- Building sustained open access datasets which will benefit both:
  - ▶ the Water Science community to better model hydraulic issues,
  - ▶ the Artificial Intelligence community which lacks examples of real datasets for their research activities: machine learning, reasoning, information fusion, etc.

## Project objectives linked to secondments

- Promoting **knowledge exchange** between partners of this project; in particular between EU and non-EU partners.

## Project objectives linked to secondments

- Promoting **knowledge exchange** between partners of this project; in particular between EU and non-EU partners.
- Producing **new knowledge**: members of this consortium have very **complementary** skills.

## Project objectives linked to secondments

- Promoting **knowledge exchange** between partners of this project; in particular between EU and non-EU partners.
- Producing **new knowledge**: members of this consortium have very **complementary** skills.
- Use the **scheduled secondment plan** to maximize knowledge transfer between partners and for training of young researchers.

# Project objectives linked to secondments

- Promoting **knowledge exchange** between partners of this project; in particular between EU and non-EU partners.
- Producing **new knowledge**: members of this consortium have very **complementary** skills.
- Use the **scheduled secondment plan** to maximize knowledge transfer between partners and for training of young researchers.
  - ▶ Identify the project members and their role in the project

# Project objectives linked to secondments

- Promoting **knowledge exchange** between partners of this project; in particular between EU and non-EU partners.
- Producing **new knowledge**: members of this consortium have very **complementary** skills.
- Use the **scheduled secondment plan** to maximize knowledge transfer between partners and for training of young researchers.
  - ▶ Identify the project members and their role in the project
  - ▶ Identify the skills (or knowledge transfer) that will be acquired following the secondments

# Project objectives linked to secondments

- Promoting **knowledge exchange** between partners of this project; in particular between EU and non-EU partners.
- Producing **new knowledge**: members of this consortium have very **complementary** skills.
- Use the **scheduled secondment plan** to maximize knowledge transfer between partners and for training of young researchers.
  - ▶ Identify the project members and their role in the project
  - ▶ Identify the skills (or knowledge transfer) that will be acquired following the secondments
  - ▶ Outline a staff secondment plan in relation to the scientific objectives of the project



# Project objectives linked to secondments

- Promoting **knowledge exchange** between partners of this project; in particular between EU and non-EU partners.
- Producing **new knowledge**: members of this consortium have very **complementary** skills.
- Use the **scheduled secondment plan** to maximize knowledge transfer between partners and for training of young researchers.
  - ▶ Identify the project members and their role in the project
  - ▶ Identify the skills (or knowledge transfer) that will be acquired following the secondments
  - ▶ Outline a staff secondment plan in relation to the scientific objectives of the project
- Multidisciplinary project.

- Beginning of July

# Grant agreement

- Beginning of July
- Sygma portal.

# Grant agreement

- Beginning of July
- Sygma portal.
- Consortium agreement!

# Grant agreement

- Beginning of July
- Sygma portal.
- Consortium agreement!
- Process reviewer comments

# Grant agreement

- Beginning of July
- Sygma portal.
- Consortium agreement!
- Process reviewer comments
- Deliverables (be sure to include mandatory deliverables)

# Grant agreement

- Beginning of July
- Sygma portal.
- Consortium agreement!
- Process reviewer comments
- Deliverables (be sure to include mandatory deliverables)
- Avoid adding affiliate/associate partners last minute

# Grant agreement

- Beginning of July
- Sygma portal.
- Consortium agreement!
- Process reviewer comments
- Deliverables (be sure to include mandatory deliverables)
- Avoid adding affiliate/associate partners last minute
- Redistribution of the budget is not possible



# Grant agreement

- Beginning of July
- Sygma portal.
- Consortium agreement!
- Process reviewer comments
- Deliverables (be sure to include mandatory deliverables)
- Avoid adding affiliate/associate partners last minute
- Redistribution of the budget is not possible
  - ▶ Try to be as close as possible to reality when setting up the project

# Feedback on project start-up

- From a scientific point of view:
  - ▶ Data risk has been addressed
  - ▶ Several events already organized
  - ▶ Good collaborations
  - ▶ etc
- Secondment agreement template (within consortium agreement)
- Templates for secondment reports
- “A little” administrative for the support of secondments
- Important visa concerns!
  - ▶ “too” long delays for visa processing
  - ▶ need to revise or adapt the secondment initial plan

# Impact of this project

- Impact on CRIL(CNRS)/Artois University:
  - ▶ Improve the visibility of CRIL, financial support, strengthen the policy of international collaboration, training, co-supervision of doctoral students, etc.
  - ▶ Explore new research themes
  - ▶ Strengthen the recent CRIL theme on "Data" with datasets generated by the project
- Datasets/standards are important for research in AI and Water Sciences
- Beyond Water Science applications and technology transfer.
- etc.

# Acknowledgments

- Our research activities received support from the European Union's Horizon 2022 research and innovation programme under the MSCA (Marie Skłodowska-Curie Actions)-SE (Staff Exchanges) grant agreement 101086252; Project title: STARWARS (STormwAteR and WastewAteR networkS heterogeneous data AI-driven management).
- Example : Autumn School in Artificial Intelligence (IA2)
  - ▶ AI and the management of heterogeneous and imperfect information and data  
<https://sites.google.com/view/ecole-ia2-2023/>
- All consortium members and consortium institutes.

**Thank you**