



# A Competitive Intelligence Cloud/PHC Platform for AI -based STI Policy Making.

Jerónimo Arenas -García (Universidad Carlos III de Madrid)  
Lorena Calvo -Bartolomé (Universidad Carlos III de Madrid)

Sep 20, 2022

EGI Conference 2022



# IntelComp's Main Objectives

- Platform for Public Administration
- Tools for evidence-based AI-driven STI policy (all phases)
- Open data
- Innovative analytics services, NLP pipelines and AI workflows
- Deployment in HPC and cloud infrastructure
- Co-creation activities: Artificial Intelligence, Health and Climate Change

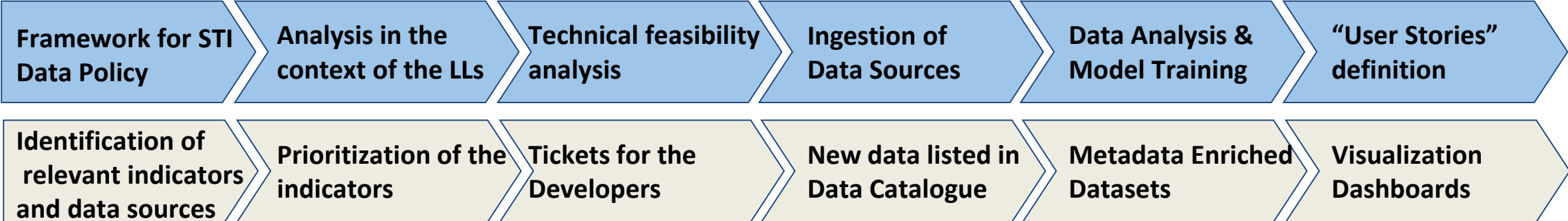
# Statistical classification approach limitations

Our target datasets usually include metadata from well-known taxonomies, but

- Taxonomy updates is (normally) a slow process, so emerging technologies may not be well covered by them
- Different datasets use heterogeneous taxonomies, making it difficult a joint analysis
- Documents are usually labelled in a binary manner. Soft assignments would be preferable to measure semantic similarities across documents.

IntelComp provides complementary information based on the **AI-based analysis of documents** (e.g., abstracts of papers, descriptions of companies, etc)

# IntelComp's Workflow



## NLP and AI services

- Scalable multilingual NLP pipelines
- Last-generation neural-based language models for:
  - Domain classification
  - Topic modelling (static, dynamic, hierarchical)
  - Automatic classification
- Information retrieval based on topic similarity and keyword search
- Scalable document graph generation and analysis (GPU implementations)
- Lead-lag detection (thematic lead/lag between corpora)
- Short- and Long-term Impact analysis

# Implementation and Deployment

- Java 1.9 & Python 3 code
- NLP Pipeline: SparkNLP, Spacy
- Topic library: Mallet, Gensim, Pytorch NN implementation (beta)
- Frontend: Bootstrap JS + D3.js + Banana Lucidworks (AngularJS)
- Search Engine SolR 7.X + Banana Lucidworks
- Apache + Tomcat, Postgres
- Prometheus + Graphana
- CD/CI environment GOCD
- Deployment: Ansible + Kubernetes (K8s) + Dockers containers



**intelcomp**

<https://intelcomp.eu>

**Jerónimo Arenas García (jeronimo.arenas@uc3m.es)**



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101004870. H2020-SC6-GOVERNANCE-2018-2019-2020 / H2020-SC6-GOVERNANCE-2020

