TERRITORI APERTI

DIORAMA: Digital twIn fOR sustAinable territorial MAnagement

A. Bianchi, G. d'Aloisio, A. D'Angelo, **Antinisca Di Marco**,
A. Di Matteo, J. Leone, G. Scoccia, G. Stilo and L. Traini

University of L'Aquila

antinisca.dimarco@univaq.it

UNIVERSITÀ DEGLI STUDI DELL'AQUILA

COMUNE DELL'AQUILA

PARTNER

Social Mining & Big Data Ecosystem SoBigData RESEARCH INFRASTRUCTURE

CNR-ISTI

USRA Ufficio Speciale per la Ricostruzione dell'Aquila

UNIVERSITA' DEGLI STUDI D'ANNUNZIO

CGIL CISL UIL

Outline

- Introduction
- DIORAMA
  - Disaster recovery
  - Fairness and bias
  - Visual Analytics
- Discussion and Future Works

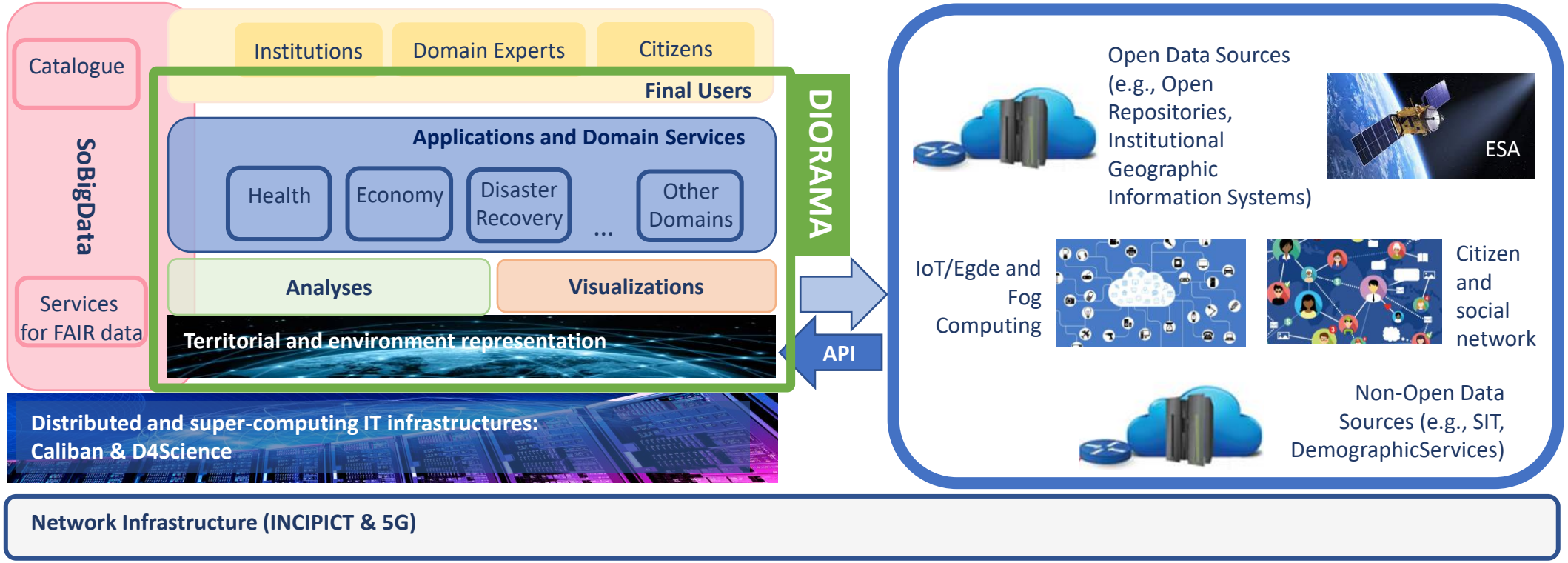
Introduction

- Territorial management is always challenging, given the several domains involved, sometimes in contrast to each other (e.g., health, economy, regulations, urban planning, risk assessment and reduction, etc.).
- Managing a territory after a disaster is even more complex due to the damages suffered by public/private buildings and infrastructures.
- Effective management must be based on the population's wellness.
- Our definition of sustainability: in the territorial management, sustainability is the combination of effective management of resources, quick decisions and wellness of the population.

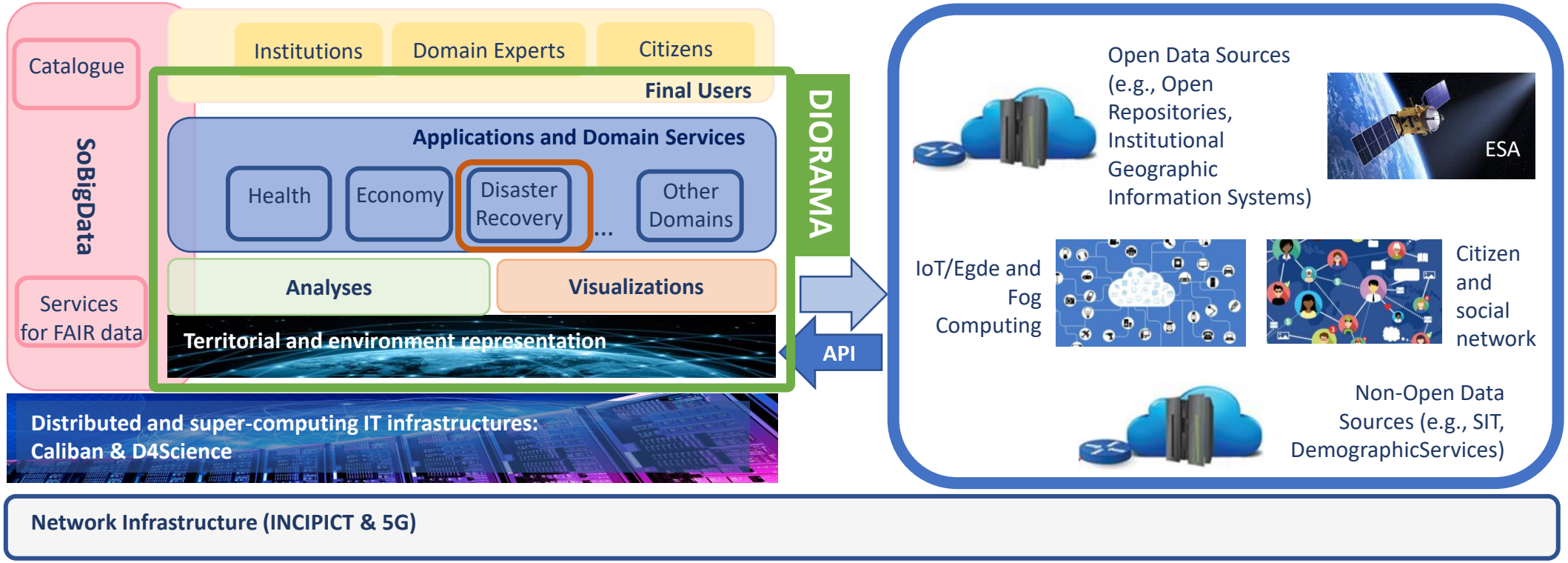
Introduction

- The availability of big data coming from heterogeneous sources (such as open data repositories, geographic information systems, smartphones, and social media) enables the digital representation of the territory (and possibly of its population).
- The digital representation permits to implement a software system(a digital twin) that, leveraging on new technologies (such as, machine learning, big data management, etc), can support decision makers to assess risks and evaluate strategies, to determine a sustainable urban plan, to monitor the territory, to predict and promptly react to events.

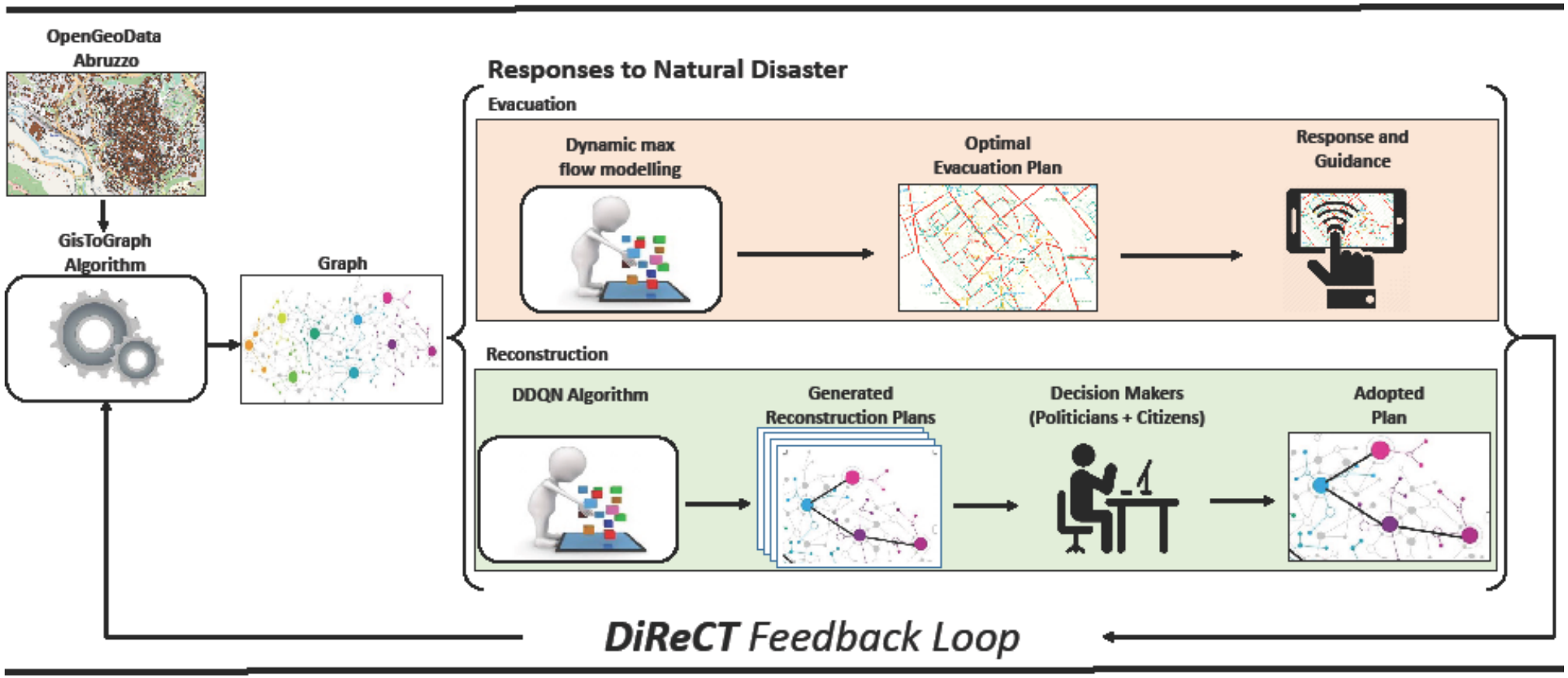
DIORAMA architecture



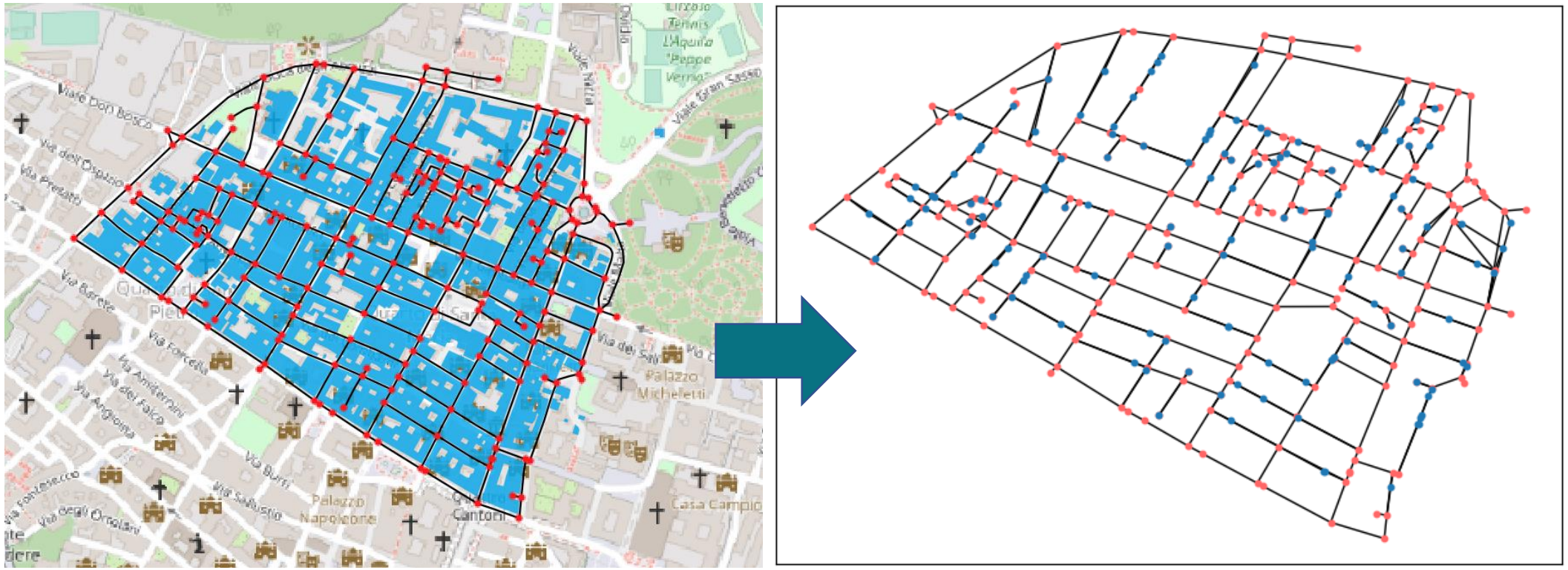
DIORAMA architecture



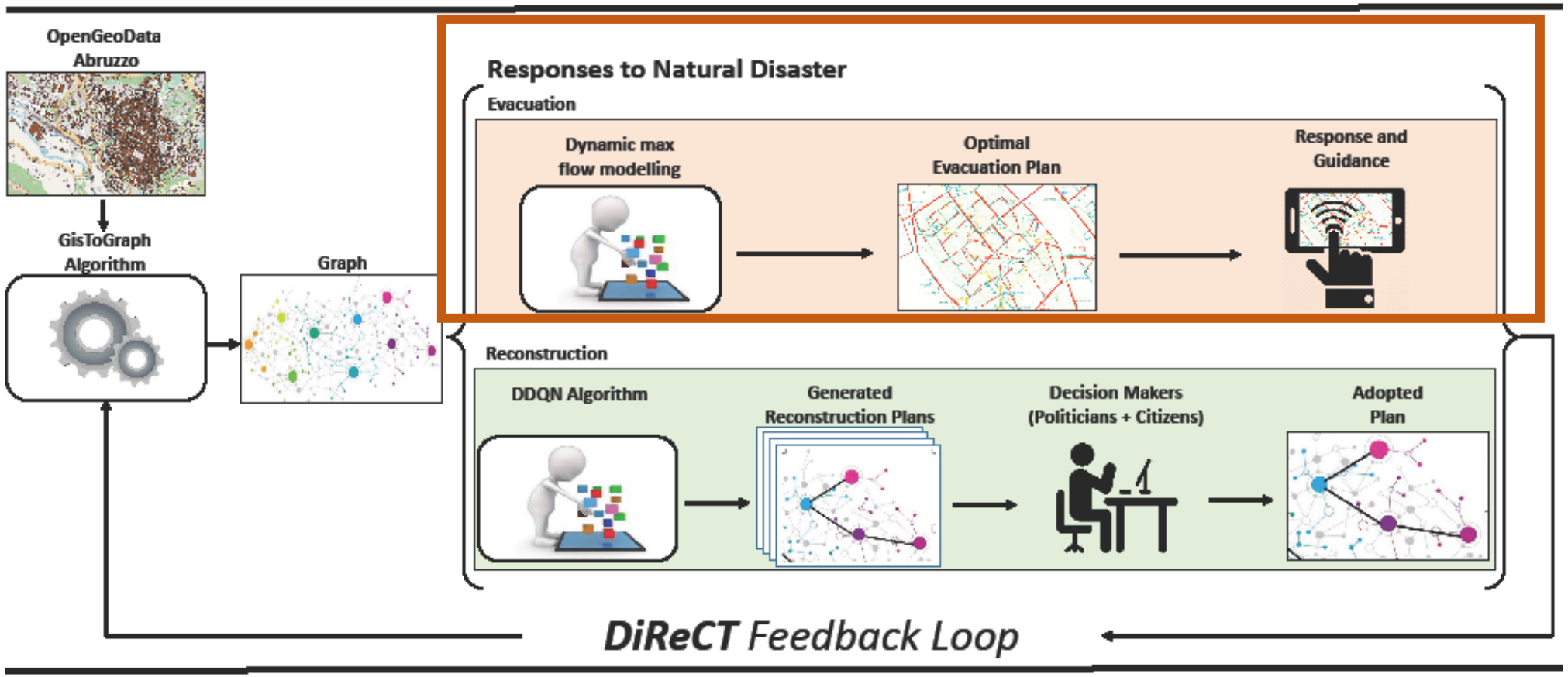
Disaster recovery



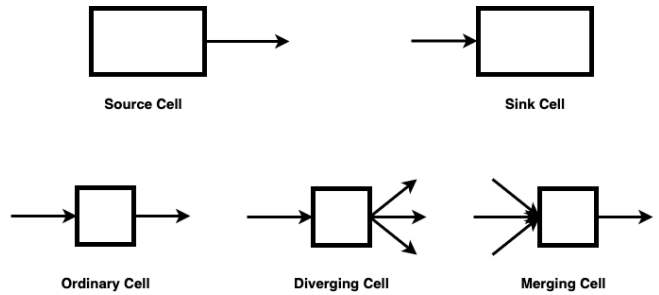
Disaster recovery: GIS2Graph



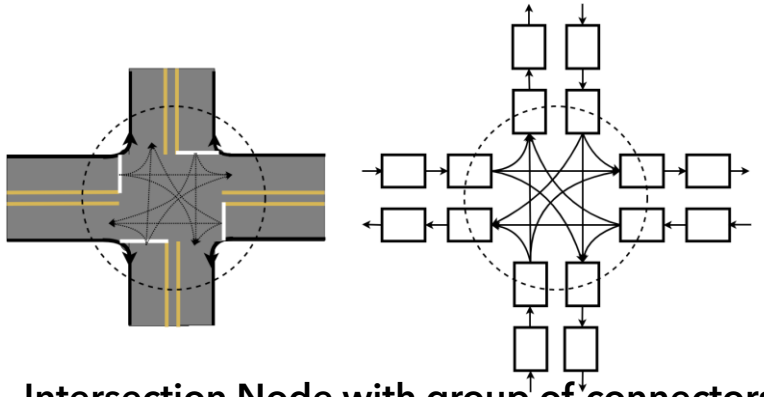
Disaster Recovery: reconstruction plan



Evacuation: Detailing the graph

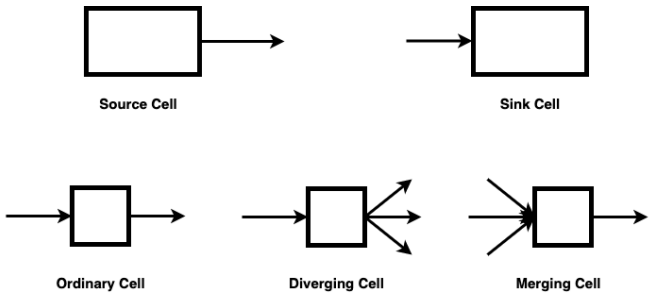


Depiction of the various types of cells used in the model

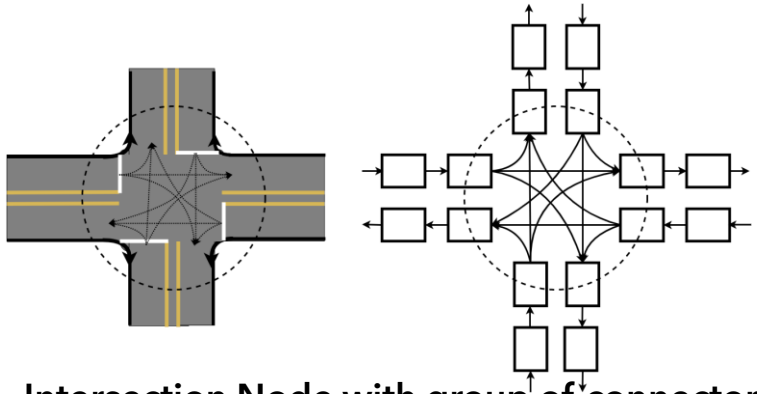


Intersection Node with group of connectors

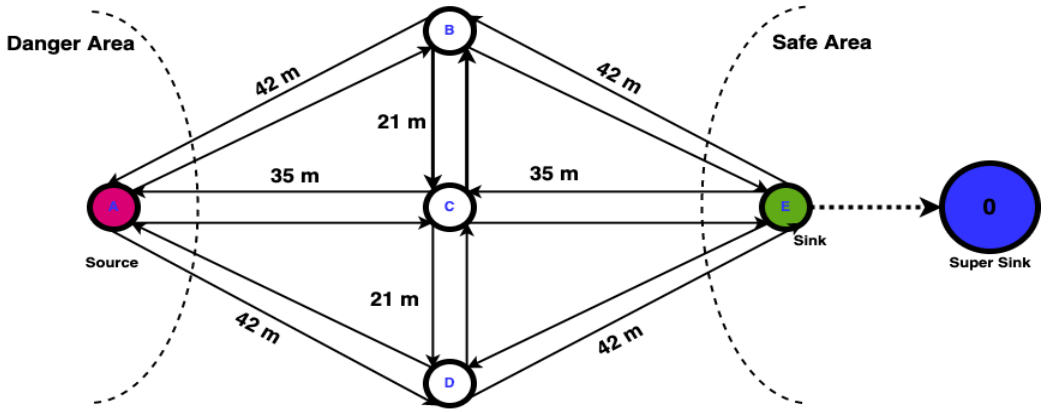
Evacuation: Detailing the graph

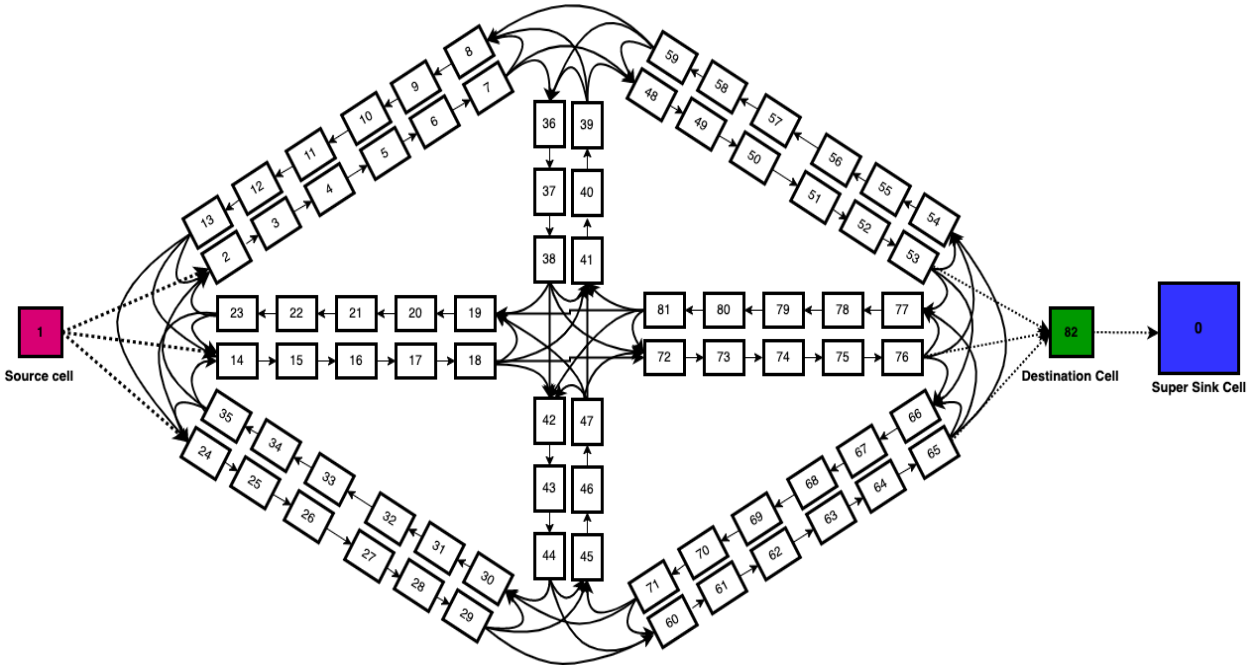


Depiction of the various types of cells used in the model



Intersection Node with group of connectors





Flow model for evacuation plan

Dynamic Cell Transmission Evacuation Planning (DyCTEP) model

$$\text{Problem: SO-DTA (CTM)} : \min \quad \sum_{t\in\mathfrak{T}} \sum_{i\in V\setminus 0} y_i^t \tag{1.10a}$$

$$y_i^t - y_i^{t-1} - \sum_{j:ji\in A} x_{ji}^{t-1} + \sum_{j:ij\in A} x_{ij}^{t-1} = 0, \quad \forall i \in V \setminus \{S \cup 0\}, t \in T, t > 0 \tag{1.10b}$$

$$y_0^t - y_0^{t-1} - \sum_{j:j0\in A} x_{j0}^{t-1} = 0, \quad t \in T, t > 0 \tag{1.10c}$$

$$y_i^t - y_i^{t-1} + \sum_{j:ij\in A} x_{ij}^{t-1} = \begin{cases} q_i, & \text{for } t = 1 \\ 0, & \text{for } \forall t > 1 \end{cases}, \quad \forall i \in S \tag{1.10d}$$

$$\sum_{j:ji\in A} x_{ji}^t \le Q_i, \quad \forall i \in V \setminus \{S\}, t \in T \tag{1.10e}$$

$$\sum_{j:ji\in A} x_{ji}^t \le \delta_i (n_i - y_i^t), \quad \forall i \in V \setminus \{S\}, t \in T \tag{1.10f}$$

$$\sum_{j:ij\in A} x_{ij}^t \le Q_i, \quad \forall i \in V \setminus \{0\}, t \in T \tag{1.10g}$$

$$\sum_{j:ij\in A} x_{ij}^t - y_i^t \le 0, \quad \forall i \in V \setminus \{0\}, t \in T \tag{1.10h}$$

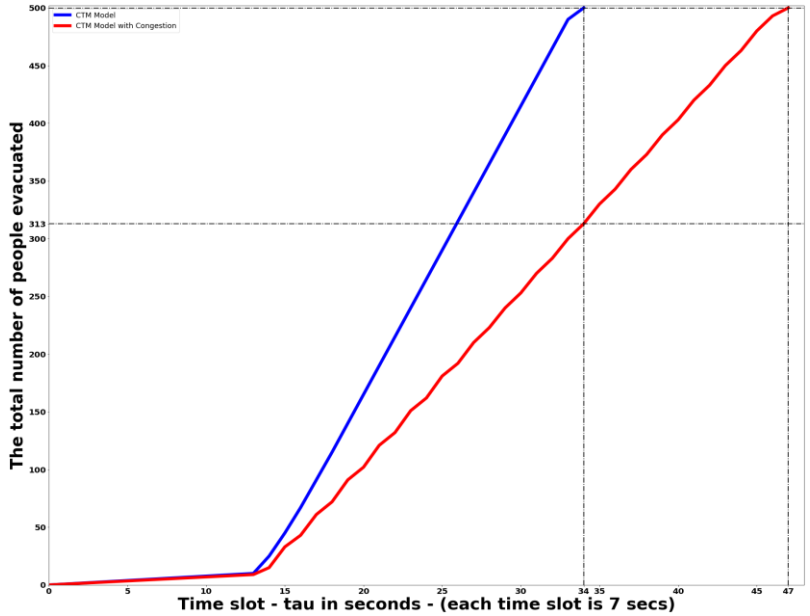
$$0 \le x_{ij}^t + x_{ji}^t \le c_{ij}, \quad \forall (ij) \in A, t \in T \tag{1.10i}$$

$$y_i^0 = 0, \quad \forall i \in V \tag{1.10j}$$

$$x_{ij}^0 = 0, \quad \forall (i,j) \in A \tag{1.10k}$$

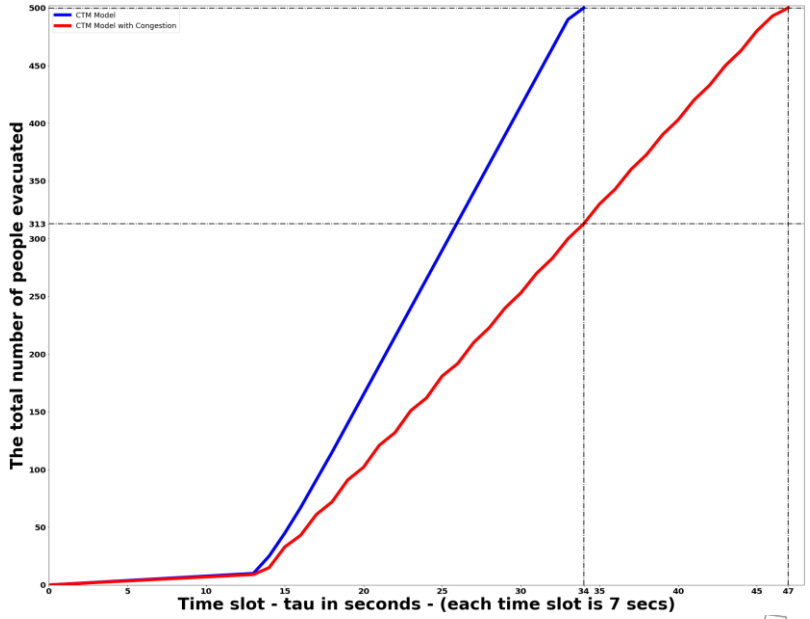
Evacuation plan goals

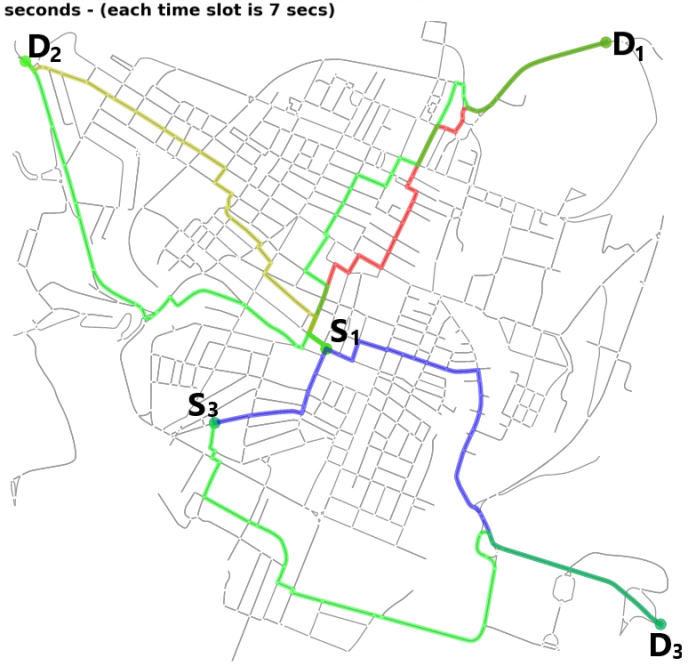
- To determine the time needed to evacuate the whole population
- To determine if the identified safe points are enough to guarantee an easy and quick evacuation



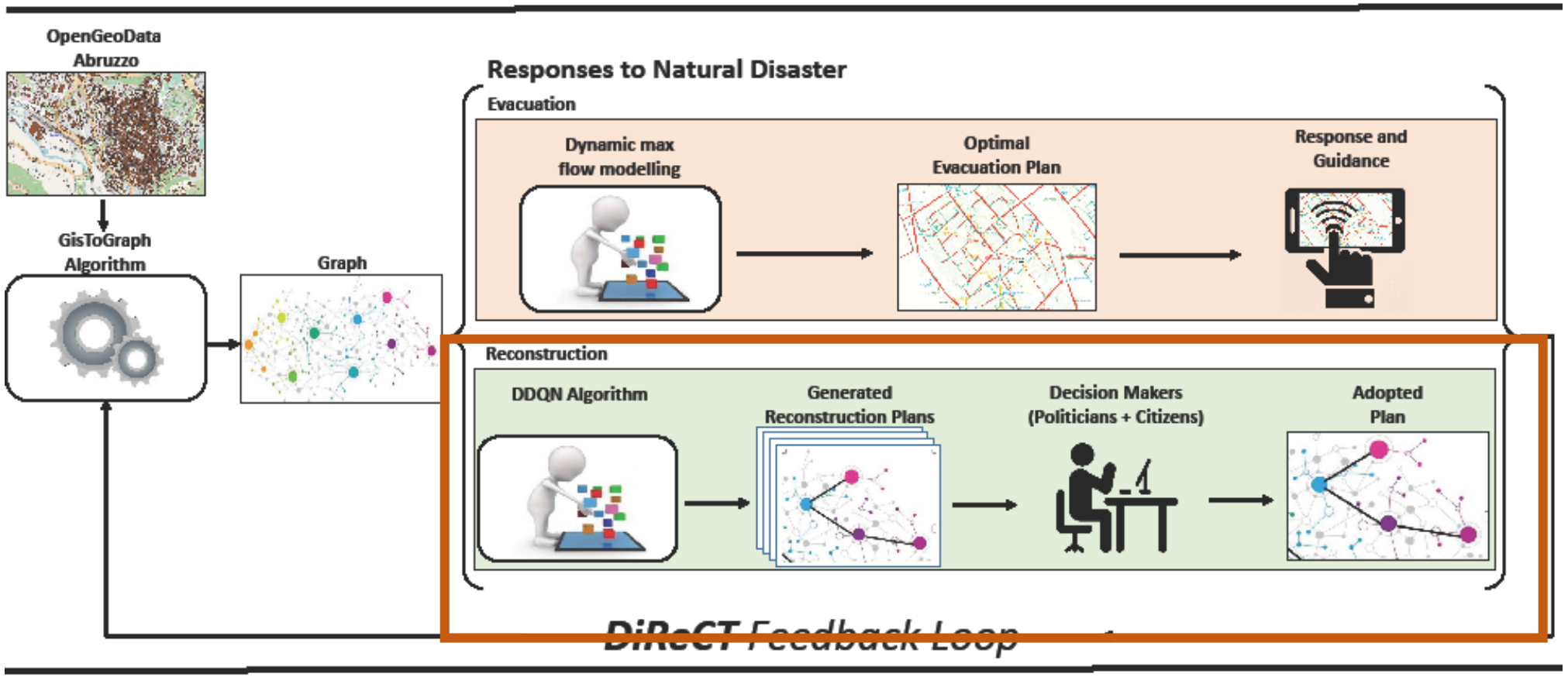
Evacuation plan goals

- To determine the time needed to evacuate the whole population
- To determine if the identified safe points are enough to guarantee an easy and quick evacuation
- To determine safe paths to communicate to citizens
- To evaluate different evacuation plans and evacuation strategies





Disaster Recovery: reconstruction plan



Disaster Recovery: reconstruction plan

Maximize **Social benefits** : it concerns the number of people who will use any unit/building, describe how much the plan is beneficial for the affected community

$$S_P = \sum_{v \in P} S(v).(T_e - T_v)$$

$$S(v) = \left[\alpha.b_v + \beta \left(\sum_{u \in V | s_u = 1} \frac{S(u)}{d(u,v)} \right) \right]$$

$$\alpha, \beta \in [0,1], \ \alpha + \beta = 1$$

Disaster Recovery: reconstruction plan

Maximize **Social benefits** : it concerns the number of people who will use any unit/building, describe how much the plan is beneficial for the affected community

$$S_P = \sum_{v \in P} S(v).(T_e - T_v)$$

$$S(v) = \left[\alpha.b_v + \beta \left(\sum_{u \in V | s_u = 1} \frac{S(u)}{d(u,v)} \right) \right]$$

$$\alpha, \beta \in [0,1], \ \alpha + \beta = 1$$

Time : it concerns the time required to construct any damage unit/building.

$$\sum_{v \in P} T_v \leq T_e$$

Cost: it concerns the cost required to construct any damage unit/building.

$$\sum_{v \in P} C_v \leq Budget$$

Political Priority : it imposes a threshold on the plan in order to guarantee that the building plan respects the set political strategies

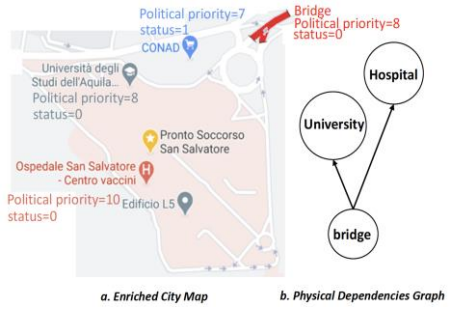
$$\frac{\sum_{v \in P} P_v}{|P|} \geq Th_p$$

Physical dependencies: (directed graph) among reconstruction units (like bridge/flyover) that impose ordering in the building reconstruction

$$\exists \ v \in P \ that \ is$$

$$e = (v, \overline{v})$$

$$\overline{v} \notin p \text{ and } s_v = 1$$



a. Enriched City Map *b. Physical Dependencies Graph*

Disaster Recovery: reconstruction plan

- Double Deep Q-Learnig Network (reinforcement learning algorithm)

Fixed Parameters	**Value**
Optimizer	Adam optimizer, learning rate = 0.001
Loss function	Mean squared error
Q-Learning function	$Q(s,a;\theta) = S_r(v) + \gamma \max_{a' \in A_v} Q'(s', a'; \theta_i^-)$
Batch size	32
Steps before training	15000
Maximum memory size	2000
Political Priority	Minimum=1 , Maximum =10
Exploration strategy	Epsilon greedy policy (Epsilon $\in 10^{-7}, 1$ and self.epsilon_ decay=0.0003)
Reward discount factor	self.discount_factor = 0.95
Input Parameters	'Budget' and 'Time' (T_e)

Reconstruction plan goals

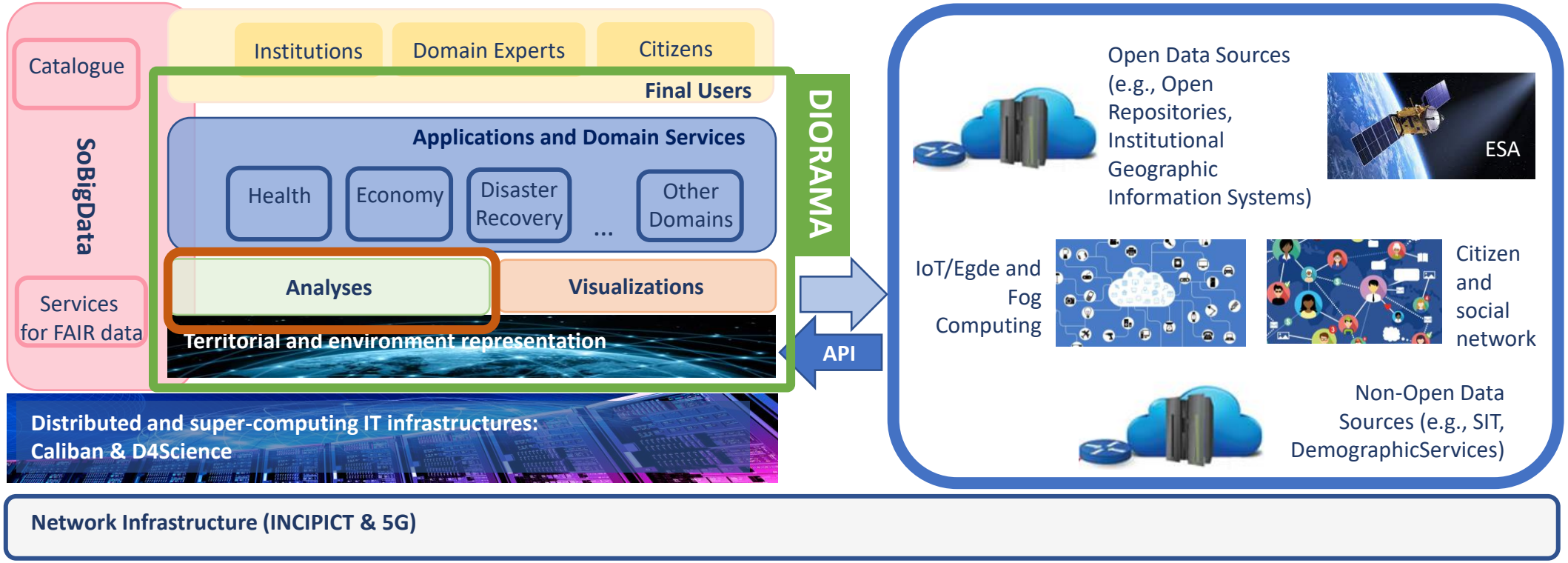
- To determine the reconstruction plans of an area that maximize the social benefits still respecting all constraints.
- To monitor the progress of reconstruction in order to promptly react in case of delay
- To evaluate (by simulation) different reconstruction strategies and policies before to start the reconstruction

Sr.No	Parallel Units
1	[35,690-783]
2	[732,434,1166,432,911,1213-681]
3	[582,85-82]
4	[59,765-116]
5	[116,1014-831]
6	[131-227]
7	[644-600]
8	[600-604]
9	[604,633-472]
10	[241,327-203]
-	-
52	[1131,1072,131]

Sr.No	Parallel Units
1	[65,516-1071]
2	[906-912]
3	[912,1166,432,911,1213-681]
4	[582,85-82]
5	[59,765-116]
6	[116,1014-831]
7	[131-227]
8	[644-600]
9	[600-604]
10	[604,633-472]
-	-
52	[1131,1072,131]

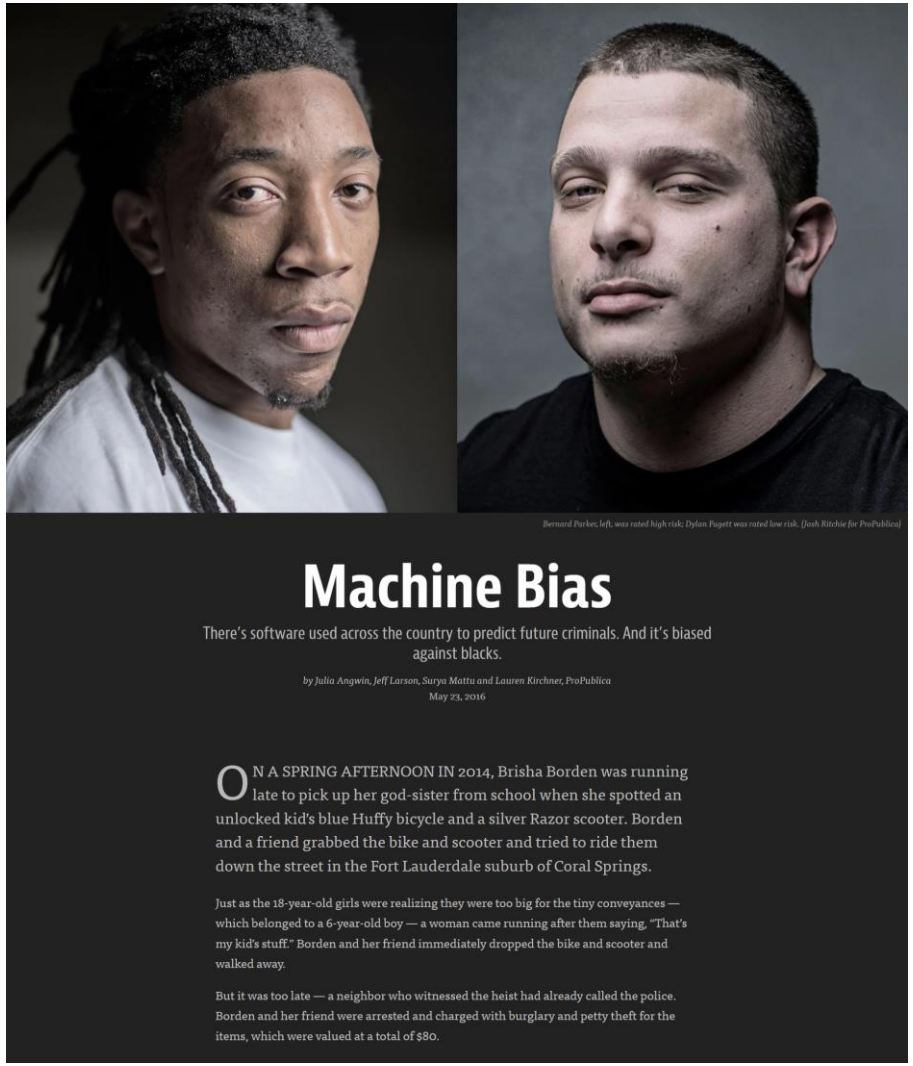
Sr.No	Cycles	Units	Buildings	PD/Roads	PP	S_P
1	Cycle 2	239	127	112	7.9	5237
2	Cycle 3	217	122	95	6.9	4527
3	Cycle 4	206	115	91	6.1	4112
4	Cycle 5	205	103	102	5.2	3601

DIORAMA architecture



BeFairest

- *Fairness* represents one of the critical quality attributes for ML systems
- Fairness ensures that ML systems are *unbiased* and do not apply discrimination among groups in the input dataset.
- The importance of fairness motivated the joint effort of studying the Bias and Fairness problem in ML through the approach named *BeFairest*.



Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

ON A SPRING AFTERNOON IN 2014, Brisha Borden was running late to pick up her god-sister from school when she spotted an unlocked kid's blue Huffy bicycle and a silver Razor scooter. Borden and a friend grabbed the bike and scooter and tried to ride them down the street in the Fort Lauderdale suburb of Coral Springs.

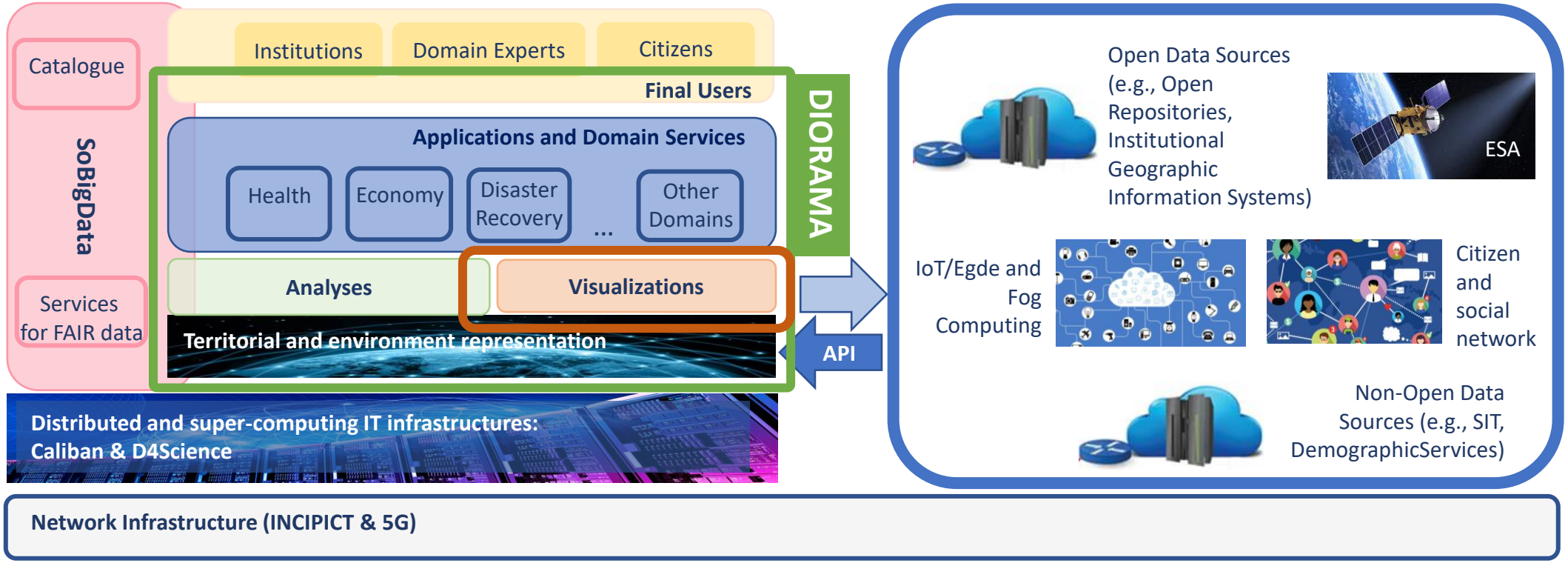
Just as the 18-year-old girls were realizing they were too big for the tiny conveyances — which belonged to a 6-year-old boy — a woman came running after them saying, "That's my kid's stuff." Borden and her friend immediately dropped the bike and scooter and walked away.

But it was too late — a neighbor who witnessed the heist had already called the police. Borden and her friend were arrested and charged with burglary and petty theft for the items, which were valued at a total of $80.

Debiaser for Multiple Variables

- We developed and tested the *Debiaser for Multiple Variables (DEMV)*
- DEMV improves fairness in binary and multi-class classification problems with any number of sensitive variables.
- DEMV efficiently manage multiple sensitive variables and is model-agnostic, making it highly flexible for any use.
- DEMV is a great improvement concerning the baselines described in Fairness literature, as they are often limited to one sensitive variable or only binary labels.

DIORAMA architecture



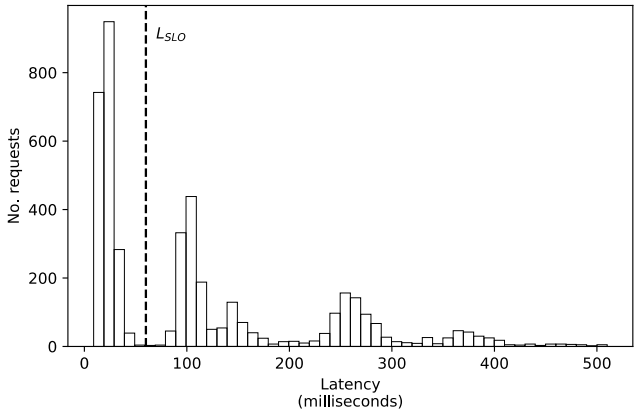
Visual Analytics

- Data visualization plays a crucial role in digital twin analysis and interpretation (Fuller *et al.,* 2020)
- In *Visual Analytics*, the phases of querying, exploring and visualizing data come together in a single process, helping to interpret data more easily and thus make analytics easier for non-experts.
- *DIORAMA* will implement a novel *Visual Analyt*ics that supports human thinking, fast data exploration and iteration, stakeholders' collaboration and insights sharing.

A. Fuller, Z. Fan, C. Day and C. Barlow, "Digital Twin: Enabling Technologies, Challenges and Open Research," in *IEEE Access*, vol. 8, pp. 108952-108971, 2020 doi: 10.1109/ACCESS.2020.2998358.

Visual Analytics: case study

- We target direct acyclic graphs (DAGs), in the context of workflow analysis
- Our current use case involves workflows that represent chains of Remote Procedure Calls (RPCs) in microservice-based systems
- We aim to leverage *Visual Analytics* techniques to support diagnosis and resolution of performance issues



Conclusions

- Diorama is a very complex and challenging project
  - Many research challenges coming to solve practical problems
  - Data Integration, data quality assessment and improvement, efficient and accurate data analysis, data visualization to efficiently and effectively communicate to several end-users, requirements elicitation of not-expert users
- We are working with municipalities, Uffici speciali per la ricostruzione and civil protection
- Several associations of citizens are involved in the requirements elicitation and assessment
- This is a joint effort of many partners and it is an open project, who can and wants to collaborate is very welcome!

Thank you for your attention!