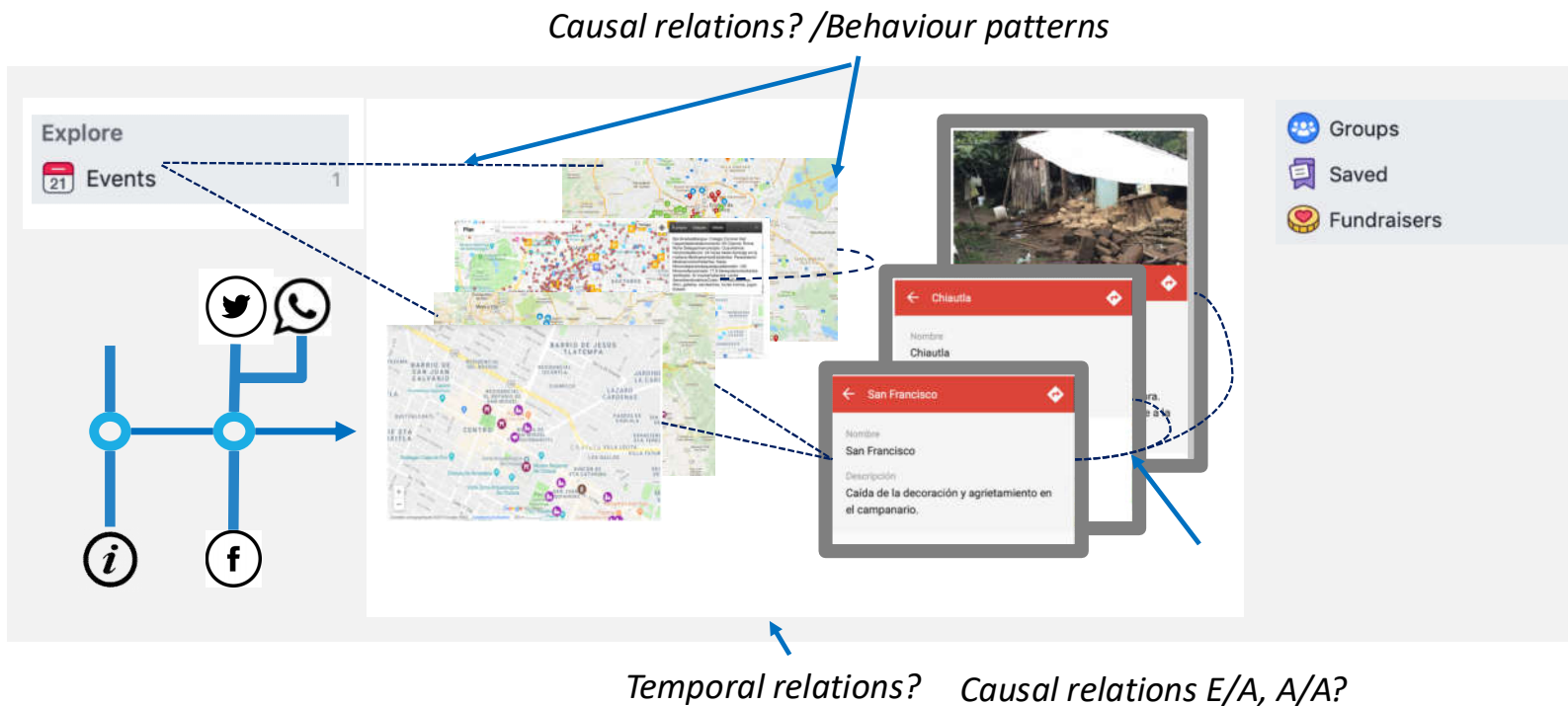# HUMAN-GUIDED DATA EXPLORATION

NOT ALWAYS SURE WHAT WE ARE LOOKING FOR... UNTIL WE FIND IT

Mehrdad Farokhnejad, UGA, LIG
Genoveva Vargas-Solar, CNRS, LIG-LAFMIA
Javier A. Espinosa-Oviedo, TU DELFT, LAFMIA

DATA CENTRIC DISASTERS MANAGEMENT

Guiding raw data exploration to define the type of questions that can be asked on top of it

Causal relations? /Behaviour patterns

KEY MOTIVATIONS

- ✓ Helping to select the right tool for preprocessing or analysis
- ✓ Making use of humans' abilities to recognize patterns

Temporal relations?   *Causal relations E/A, A/A?*

OBJECTIVE

Help data scientists express queries that can help them understand the content of data collections

2

EXPLORING DATA COLLECTIONS

DATA EXPLORATION TECHNIQUES

Query expression [guidance | automatic generation]

- Multi-scale query processing for gradual exploration
- Query morphing to adjust for proximity results
- Queries as answers: query alternatives to cope with lack of providence

Results filtering, analysis, visualization

- Result-set post processing for conveying meaningful data

Data exploration systems & environments

- Data systems kernels are tailored for data exploration: no preparation easy-to-use fast database cracking
- Auto-tuning database kernels : incremental, adaptive, partial indexing

SYSTEMATIC REVIEW STATISTICS

Machine learning & visualization are very popular for exploring data
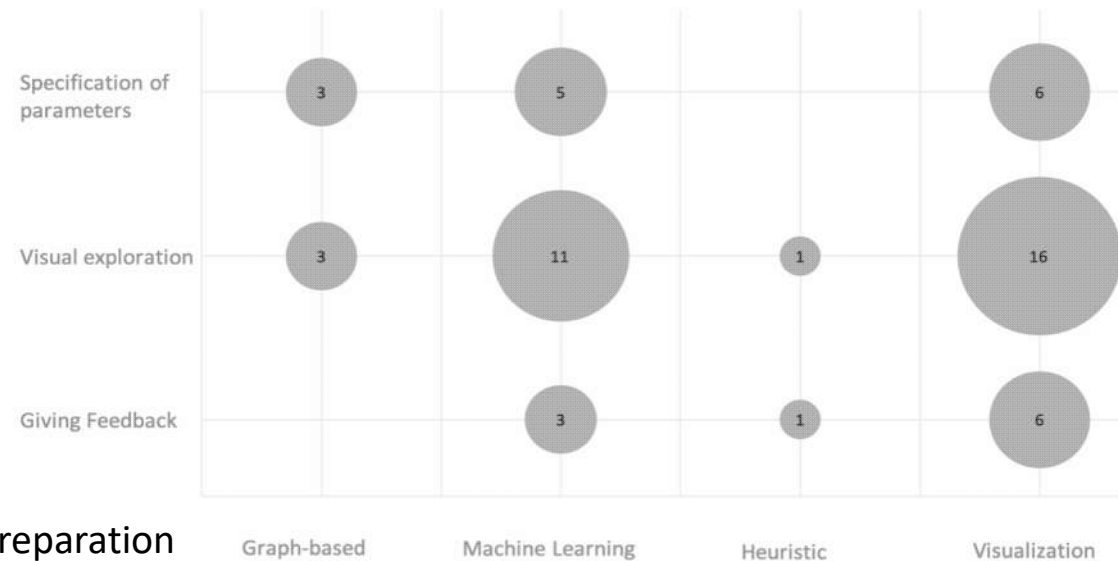
Few proposals consider human in the loop for guiding data exploration

QUERIES AS ANSWERS

What questions can be asked given a data collection?

Which data analytics technique is best adapted to be run on top data collections?

*How to prepare data collections according to the data analytics technique to be applied?*

How can data scientists express queries that can help them understand the content of data collections?
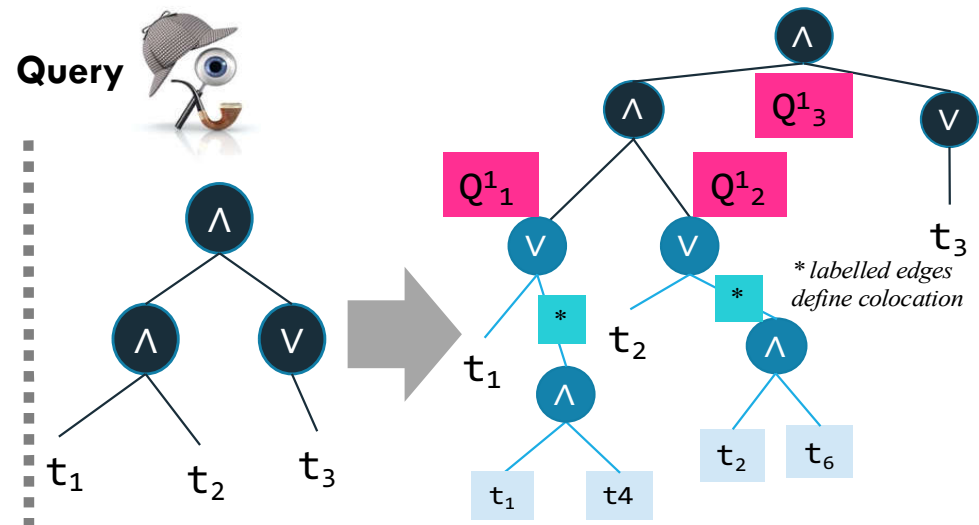
APPROACH: QUERY REWRITING

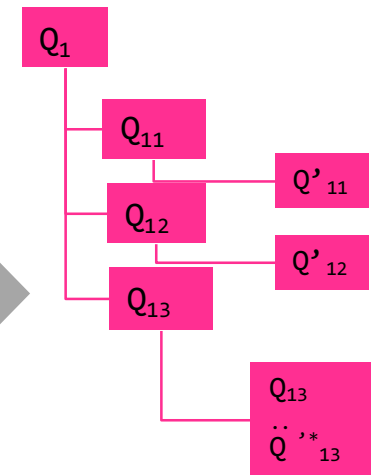*"Given an initial query provide sets of queries that can help data scientists to better exploit data collections"*

Classification

Events

Actions

Vocabulary
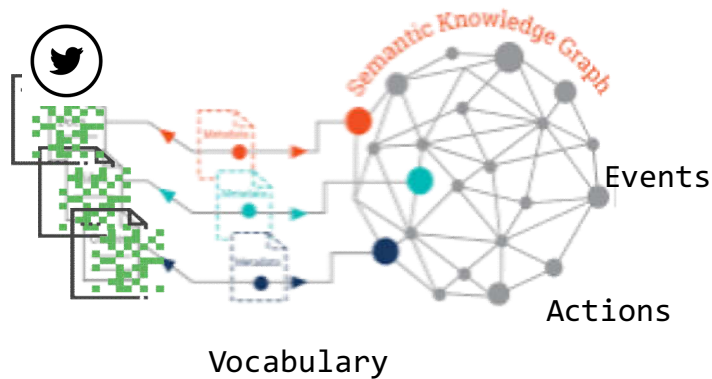
Query

Exploitability of a query

*labelled edges define colocation

Queries as answers

HUMAN-GUIDED DATA EXPLORATION

Mehrdad.**Farokhnejad**@univ-grenoble-alpes.fr