

DATA2B

WE MAKE DATA PRODUCTS

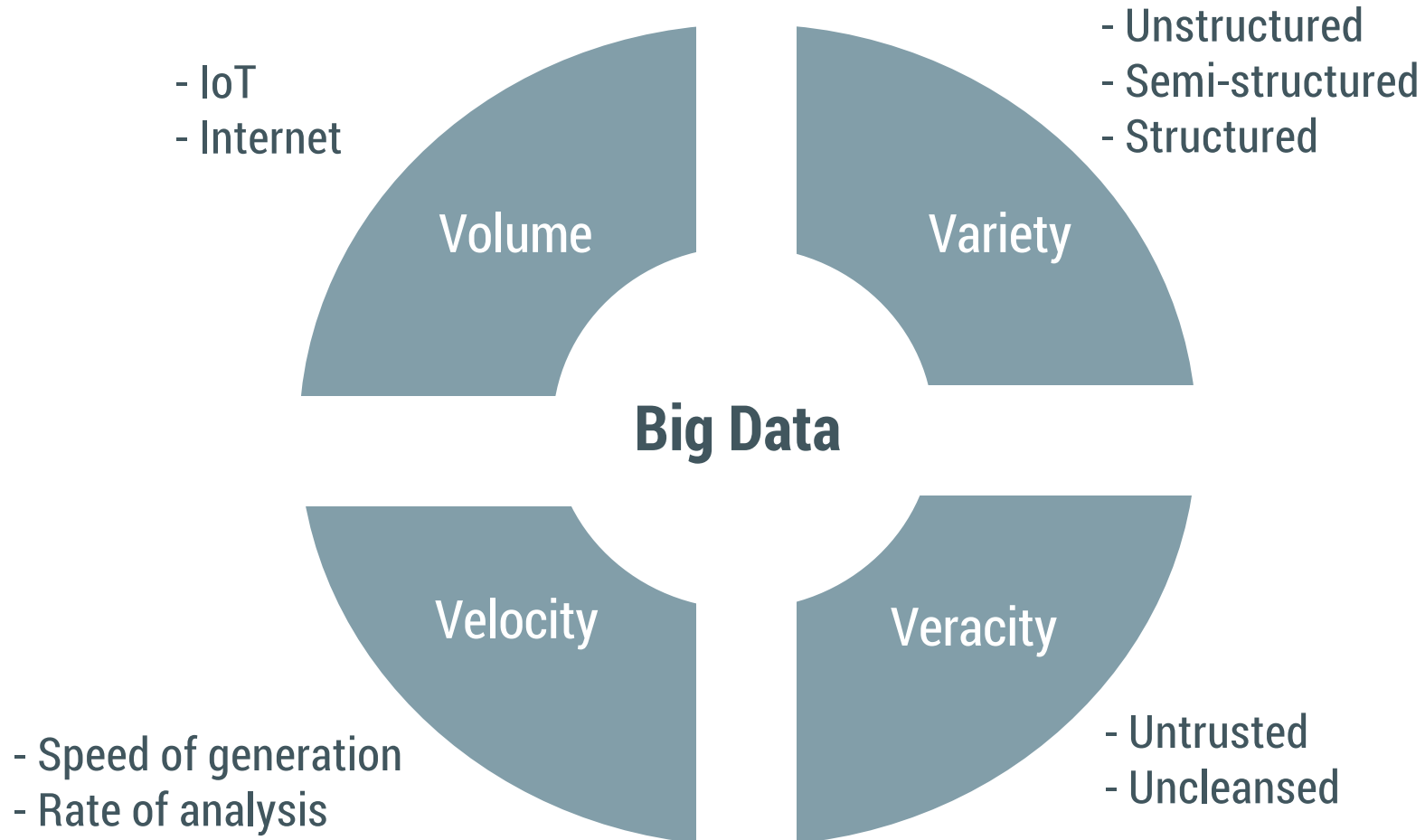
Zinnya DEL VILLAR & Christophe THOVEX

Approaching **Big Data** from a **business perspective**

What is Big Data?



What is Big Data?



Big Data eliminates intuition

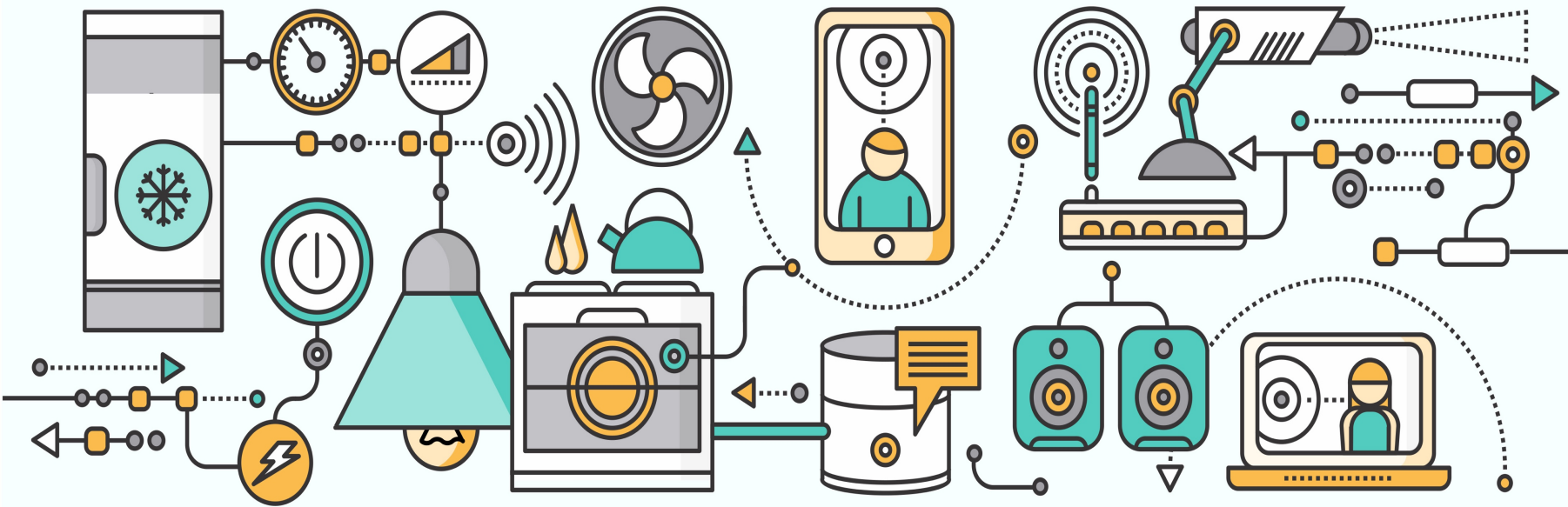
Big Data eliminates intuition

How?

Decisions can be made with a structured approach, through data driven insight.

Big Data life cycle

Creation



Big Data life cycle

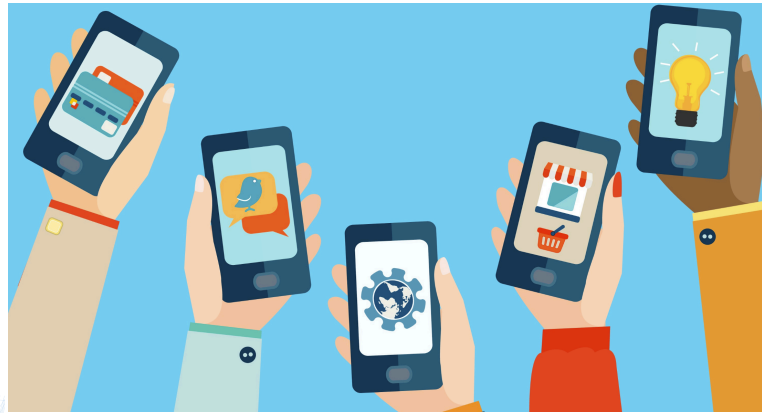
Creation

Processing



Big Data life cycle

Creation
Processing
Output



Options

datatype http

Node value

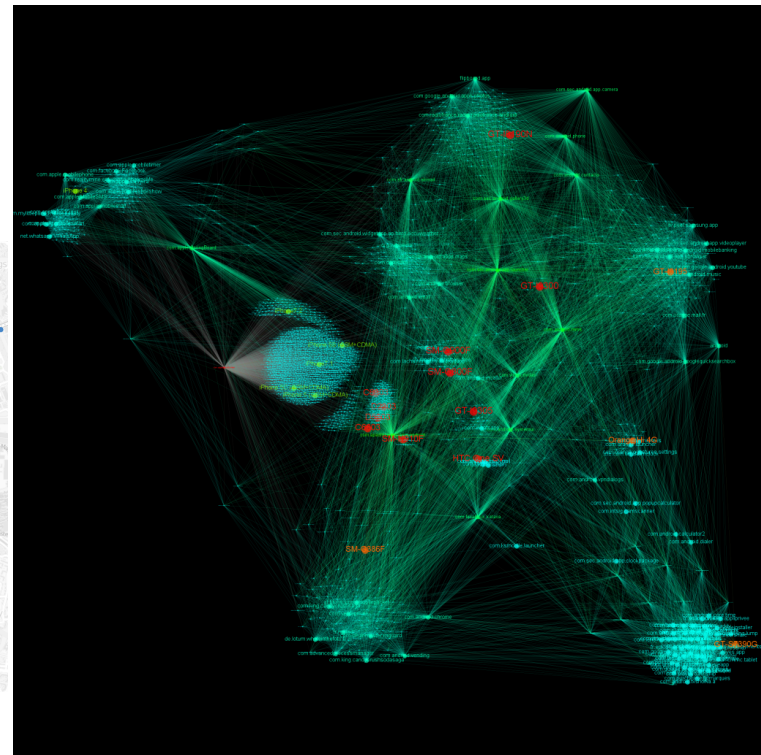
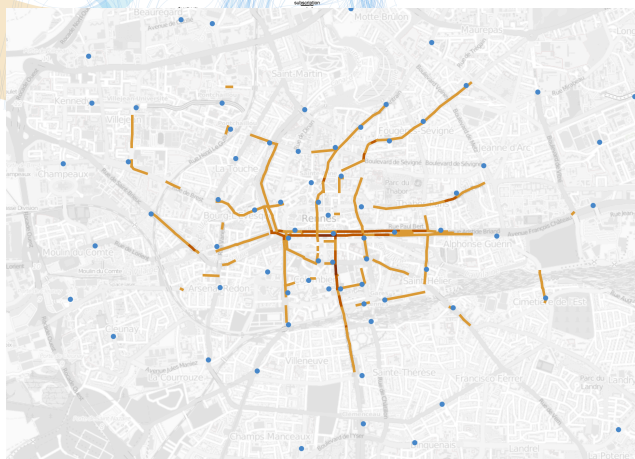
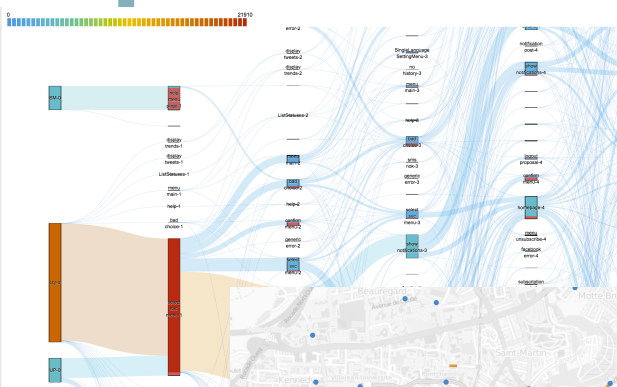
Min 0

Max 21910

Node depth

Min 0

Max 111



Big Data life cycle

Creation

Processing

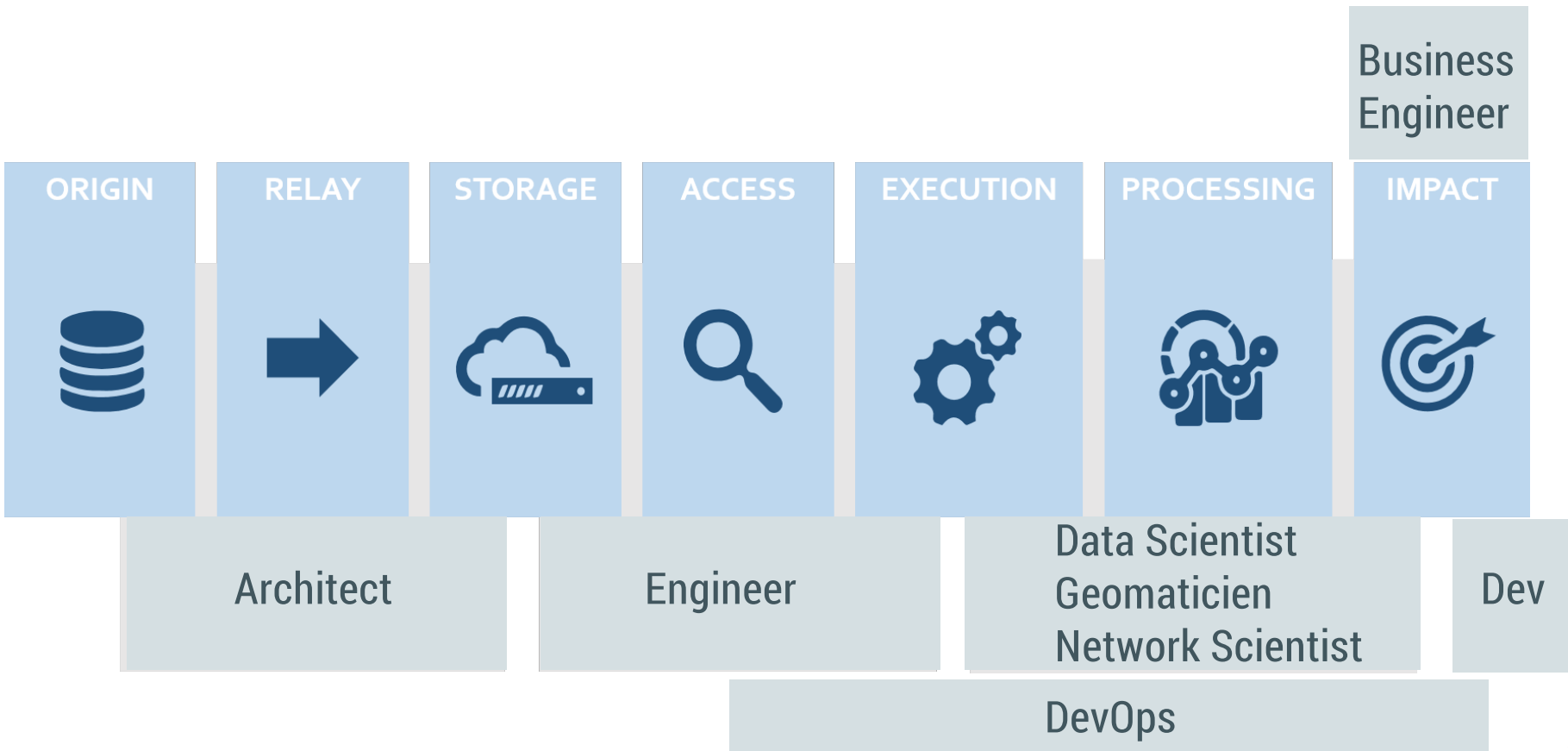
Output

Ressources and processes



Big Data life cycle

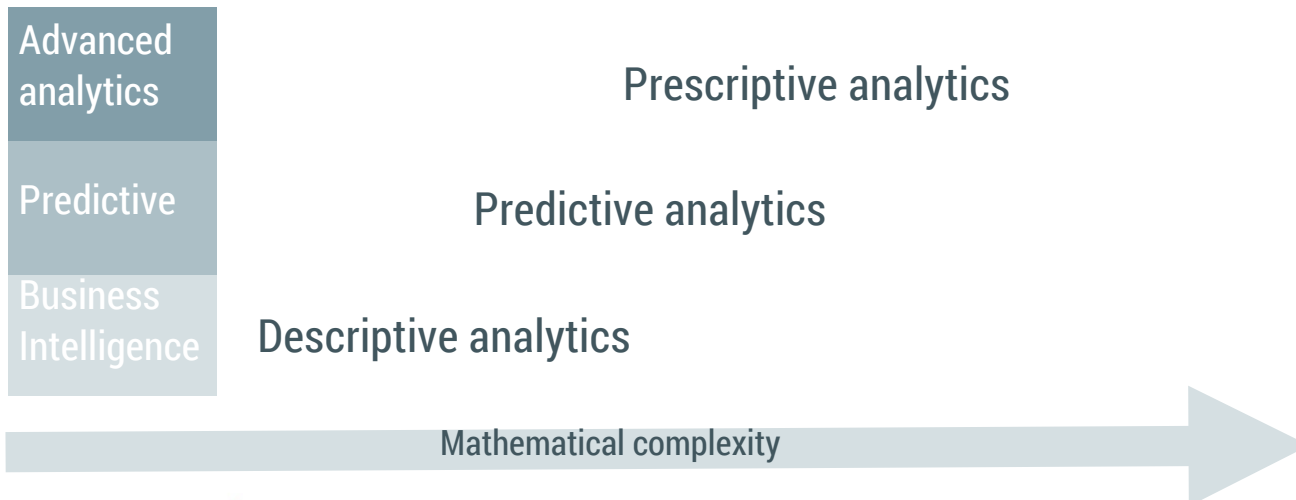
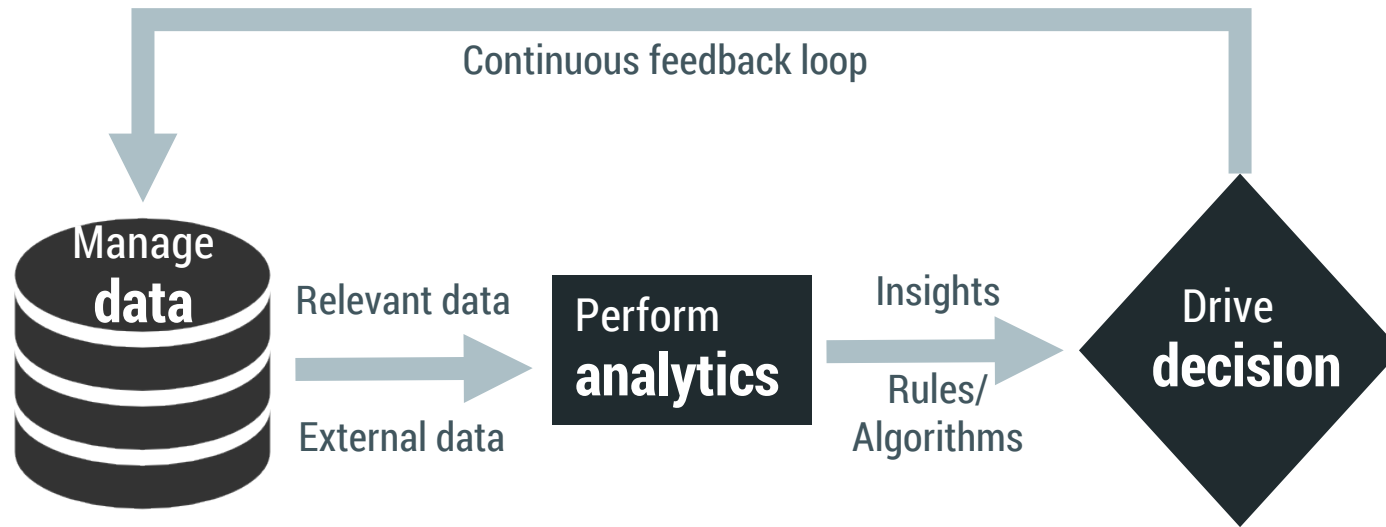
Ressources and processes



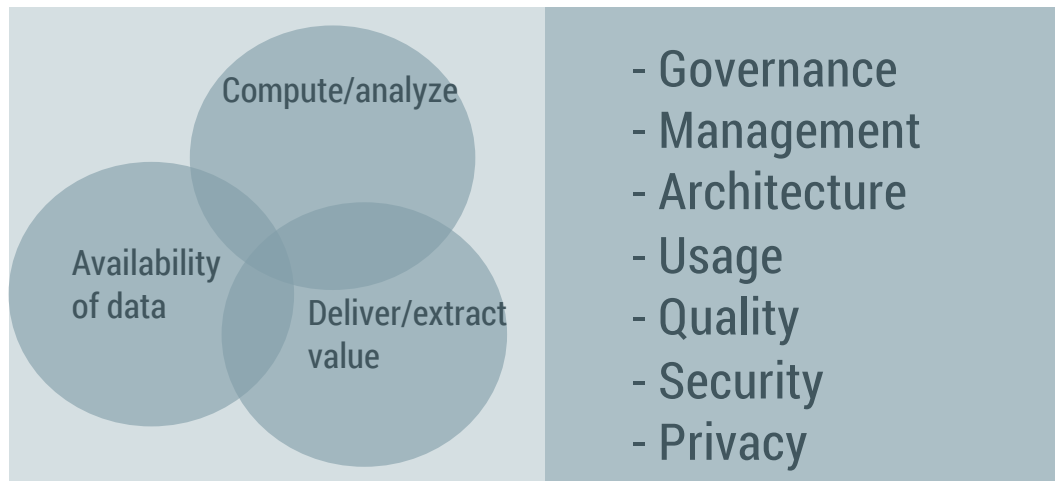
Big Data value chain

Improve the
efficiency and effectiveness
of every
decision and/or action

Big Data and analytics



Benefits and risks of Big Data



**Big Data
Success**

DATA PRODUCTS



Product improvement



Increase the production performance



Help human decisions



FOOD INDUSTRY

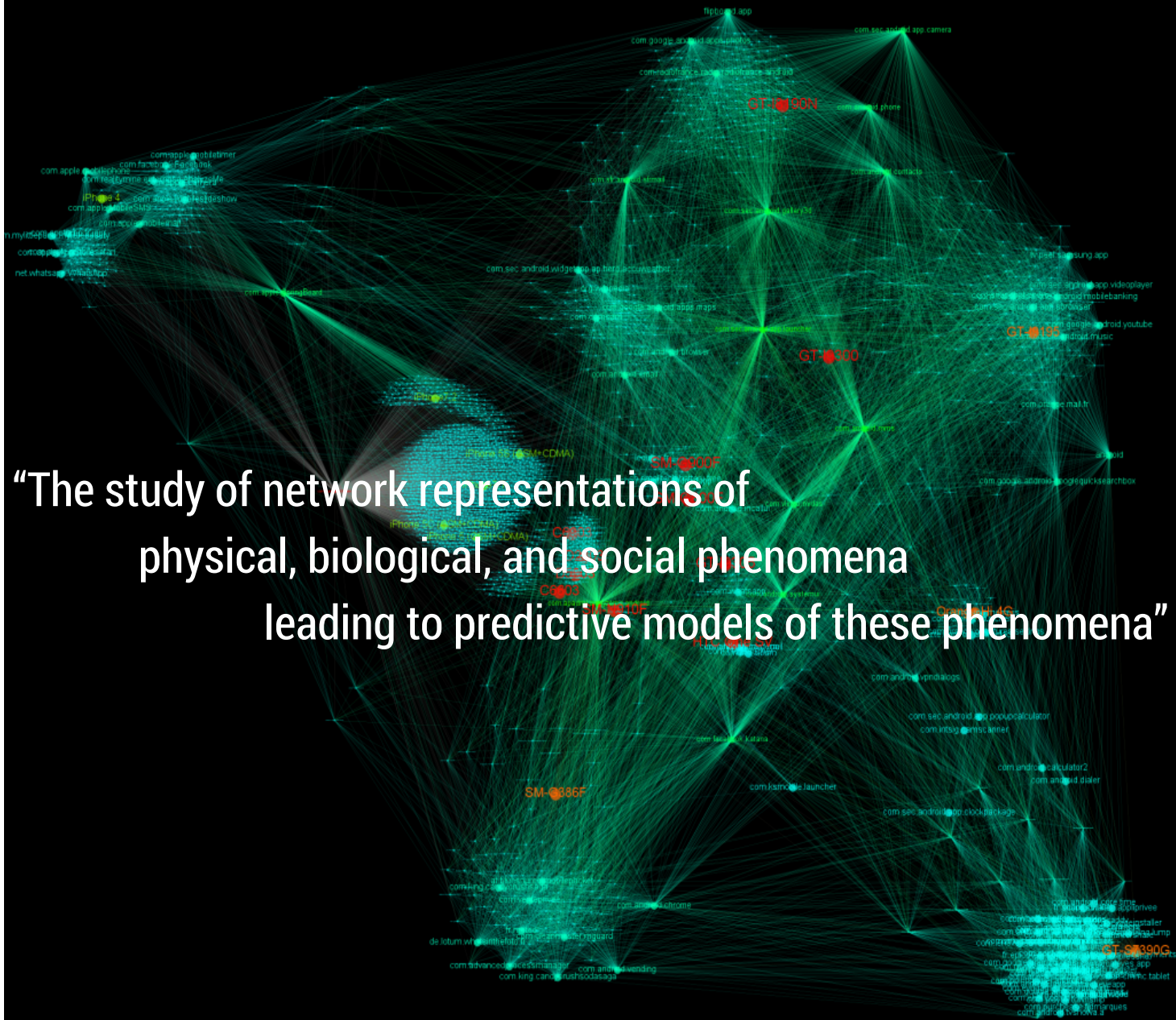
Logistic optimisation



Understand customer behavior



Network Science



“The study of network representations of physical, biological, and social phenomena leading to predictive models of these phenomena” [1]

Network Science and Big Data

Processing flows through hundreds of thousands of ties

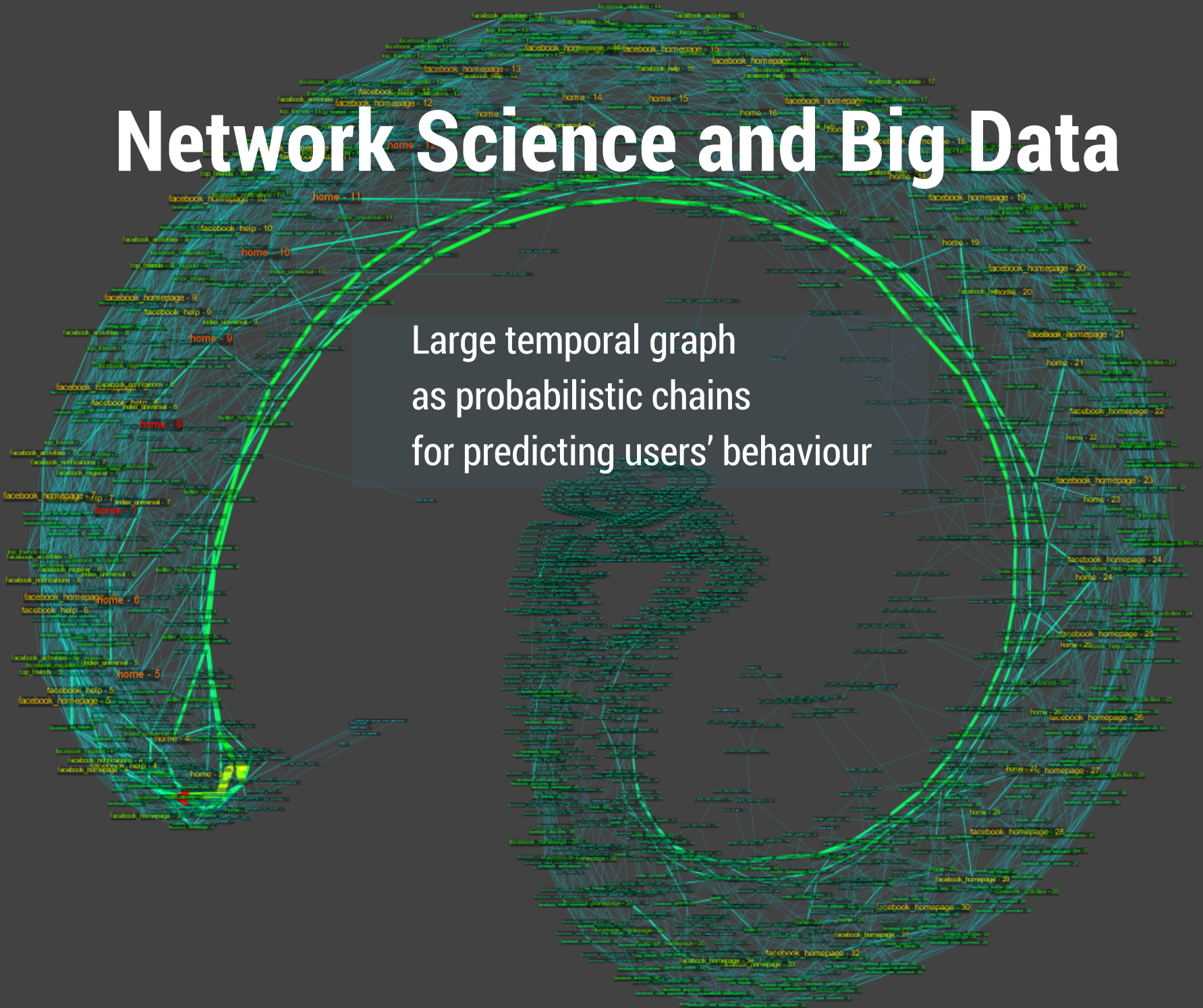
Network Science and Big Data



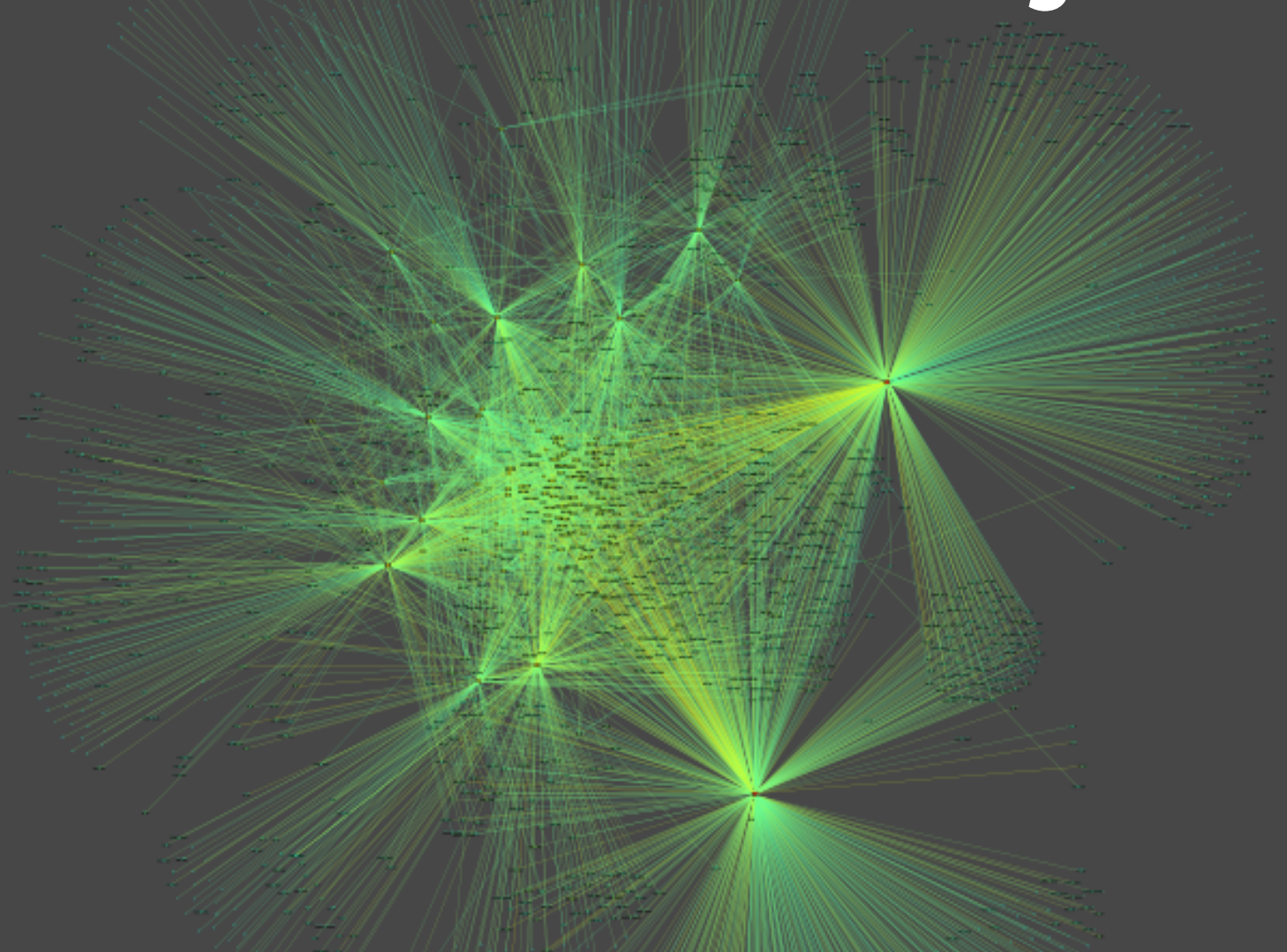
Processing statistics and/or probabilities as weights/flows in
Bayesian/Markovian networks,
Convolutional Neural Networks (Deep Learning)
Social and Semantic Networks (SSN)

Network Science and Big Data

Large temporal graph
as probabilistic chains
for predicting users' behaviour

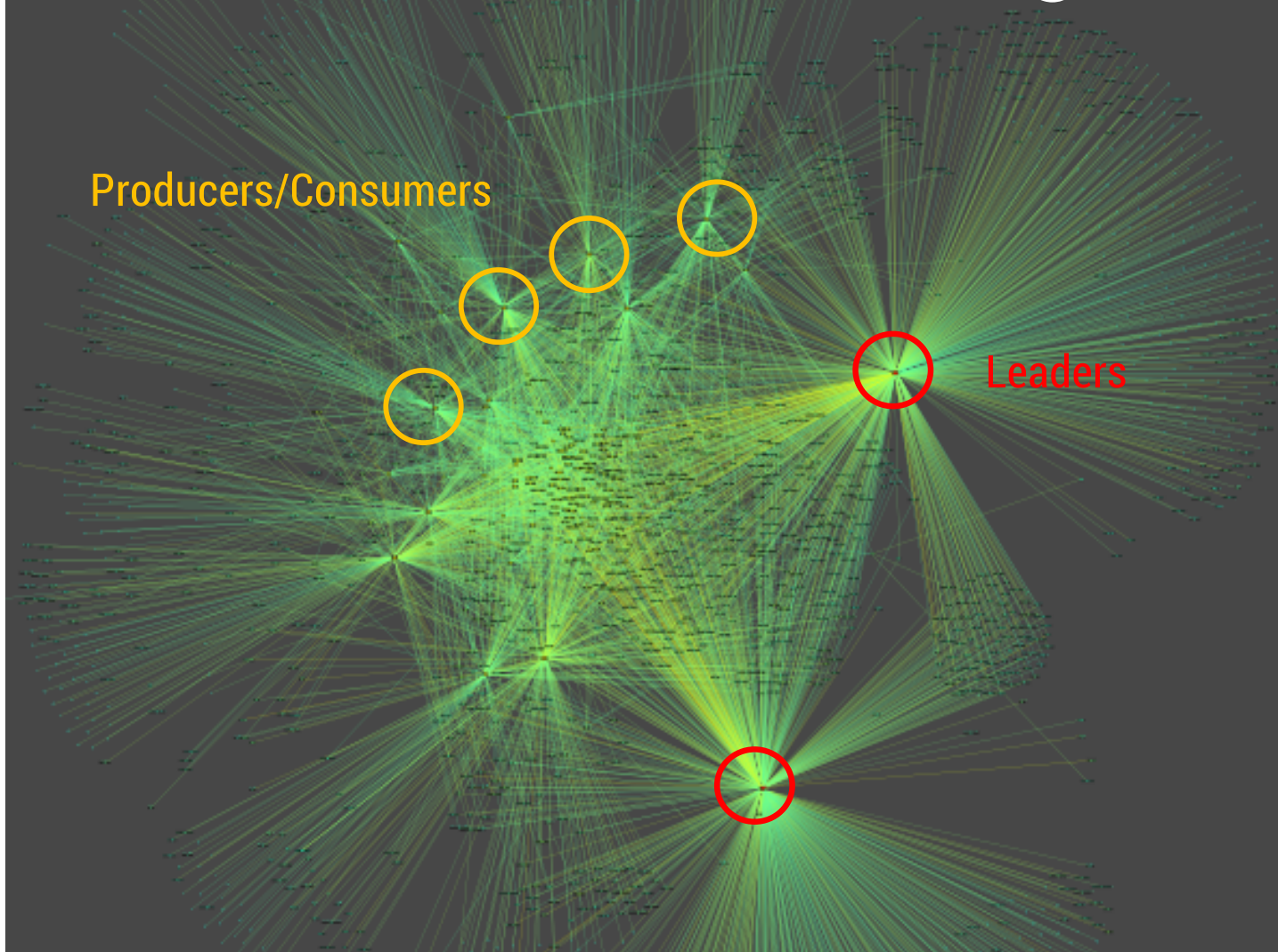


Network Science and Big Data



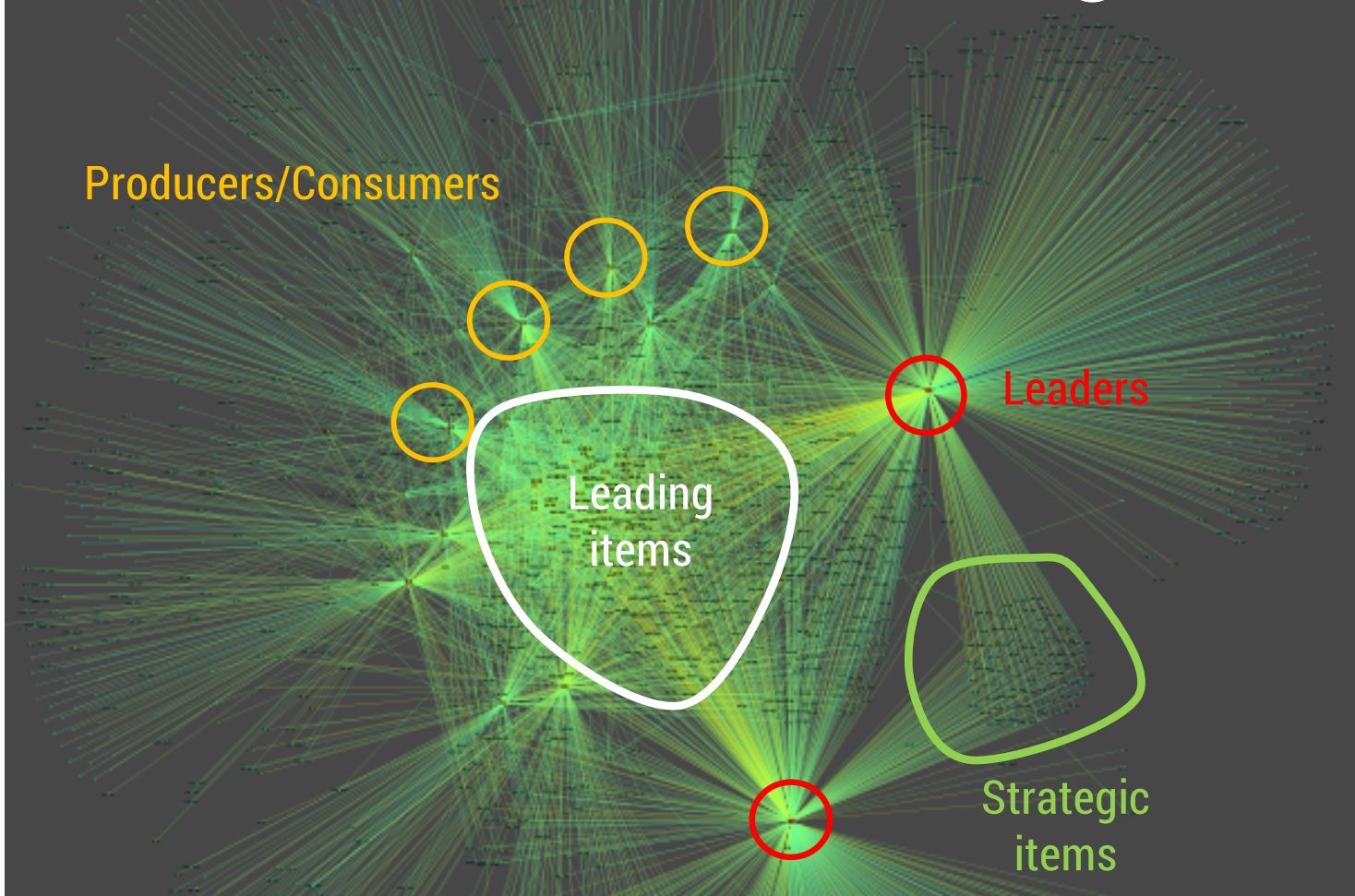
Community clustering and characterizing

Network Science and Big Data



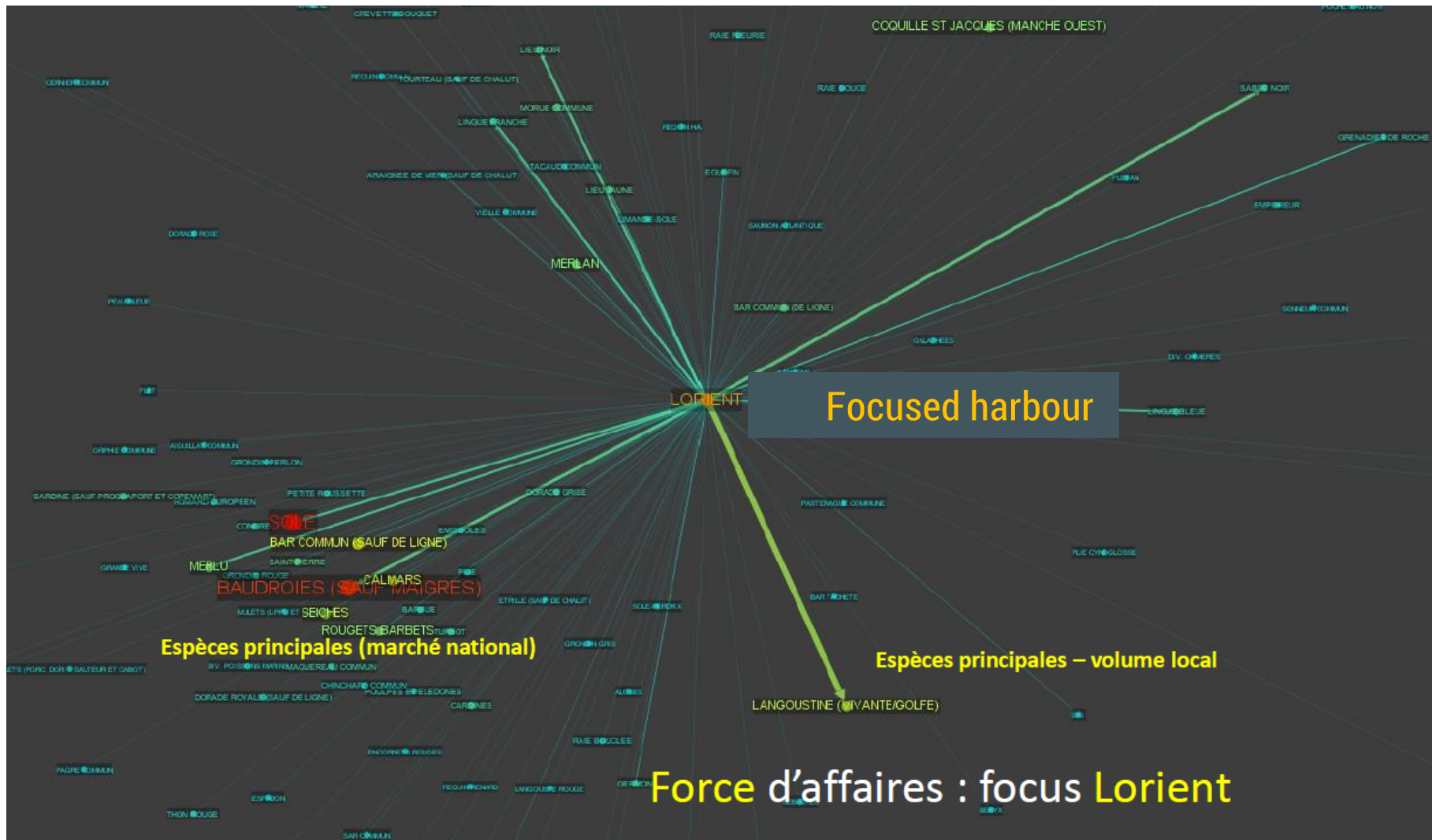
Community clustering and characterizing

Network Science and Big Data



Community clustering and characterizing

Stocks and sales by fish species – multiscale visualization



Stocks and sales by fish species – multiscale visualization

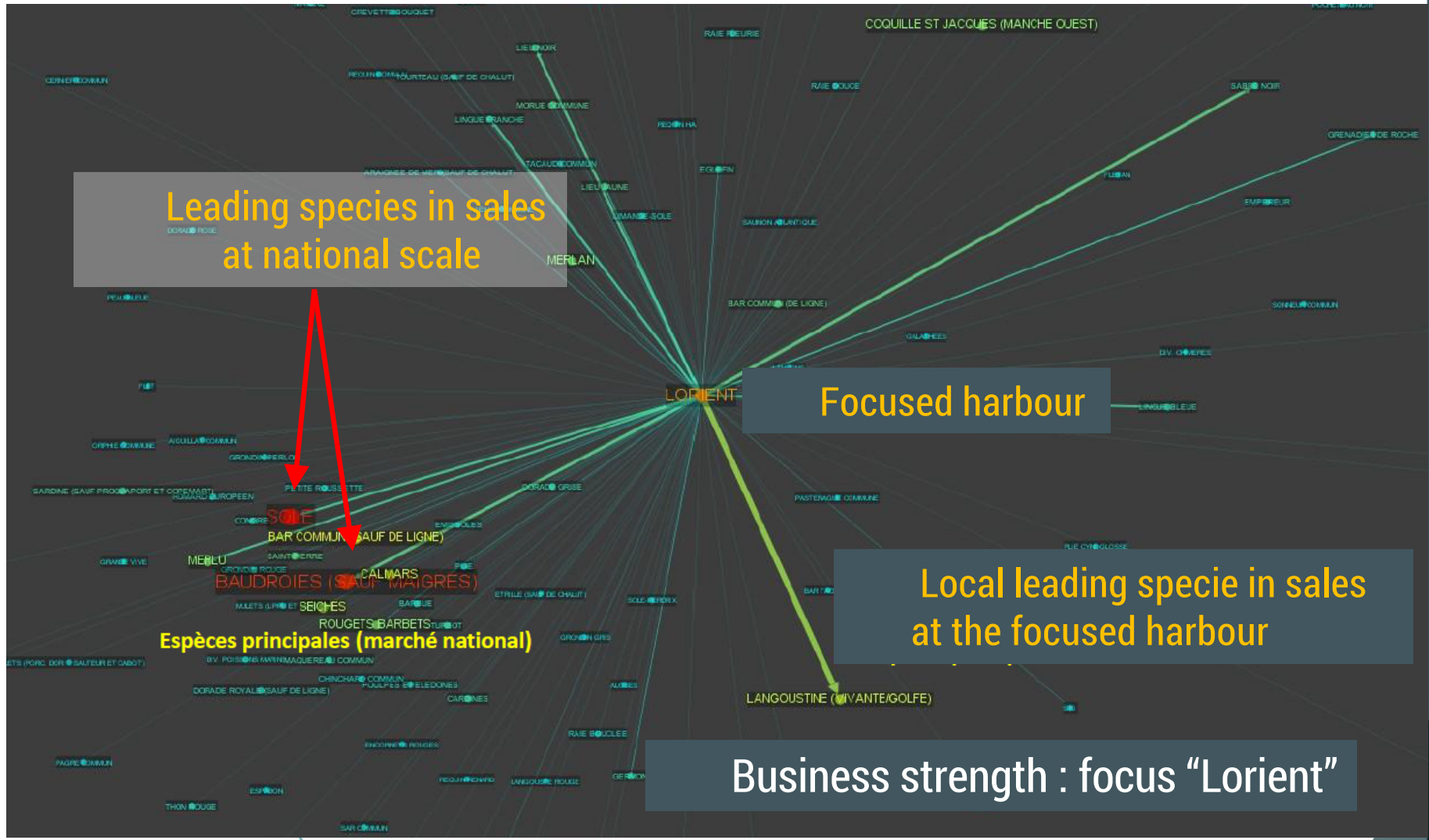
Leading species in sales
at national scale

Focused harbour

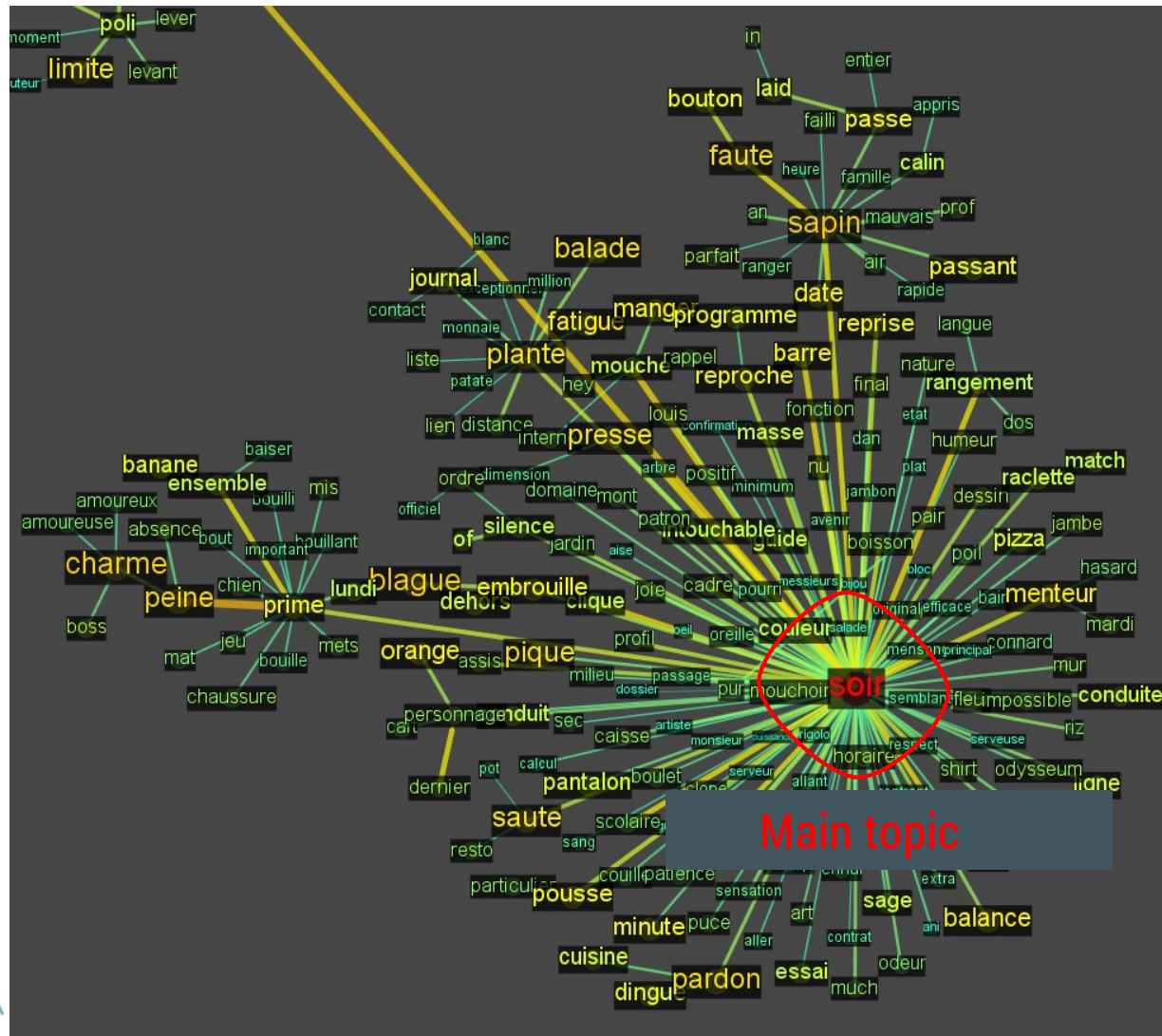
Local leading specie in sales
at the focused harbour

Espèces principales (marché national)

Business strength : focus "Lorient"

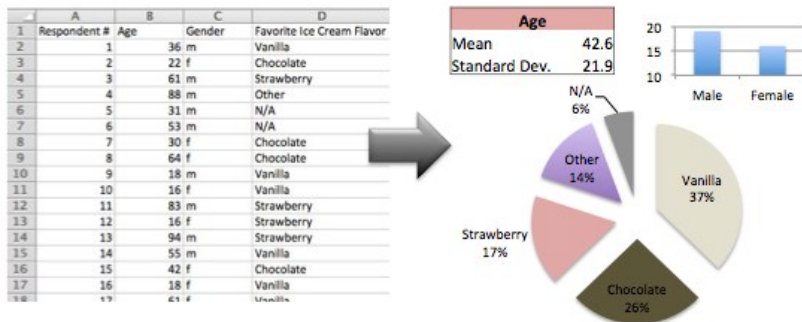


Hidden sentiment extraction from the talk of crowds



Data Science & Networks

Quantitative analysis :
Descriptive statistics, inferential statistics...

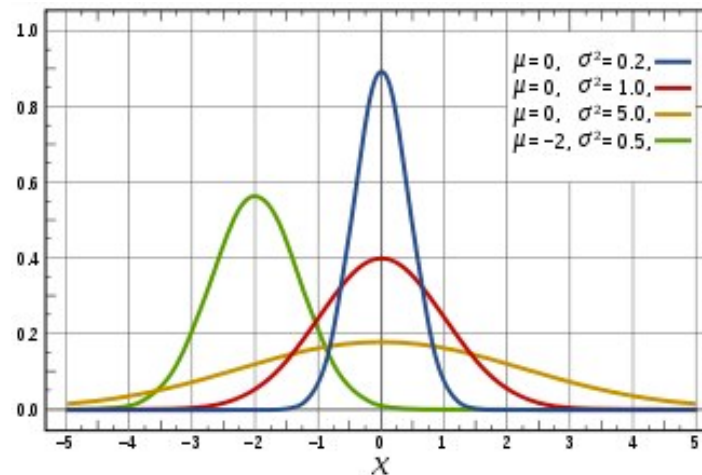


Raw Data

Descriptive Statistics

Descriptors

Estimators

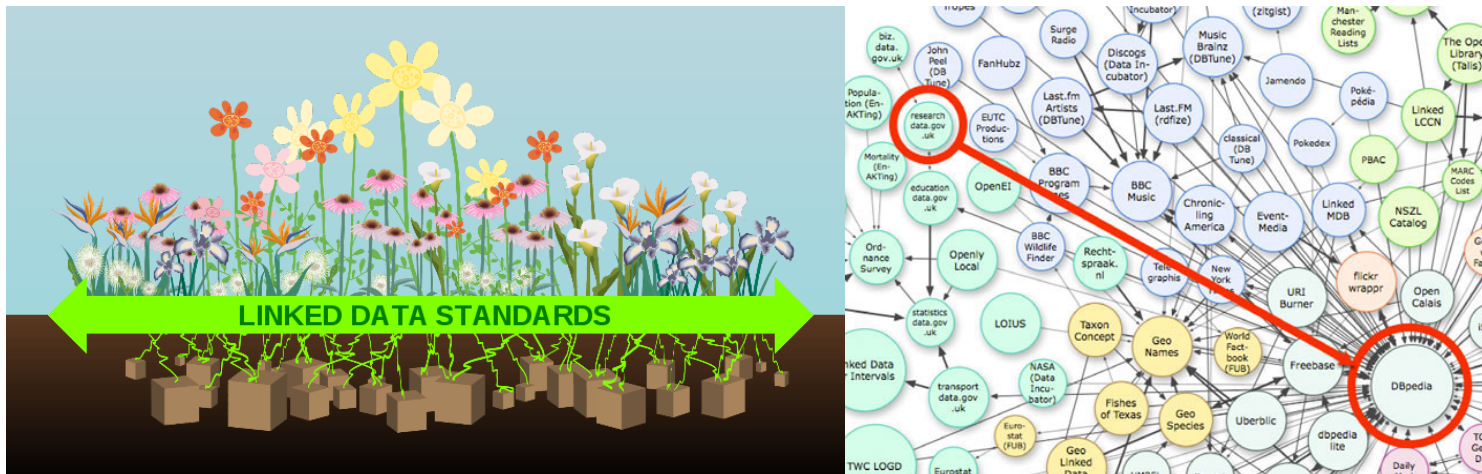


Data Science & Networks

Qualitative analysis :

To uncover and understand the “big picture”, using the data to describe the phenomenon and what this means.

Semantic and Social Web, Linked Data, Ontologies for information retrieval



Internet *is_a* communication network = true

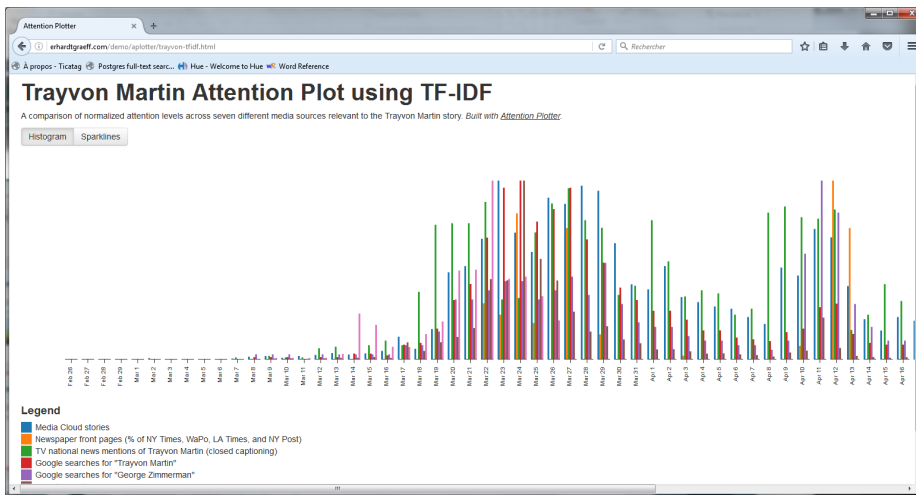
Linked Data *is_a* data network = true

Ontology *is_a* semantic network = true

Data Science & Networks

Qualitative analysis :

From linguistic statistics towards semantic inferences and fuzzy reasoning



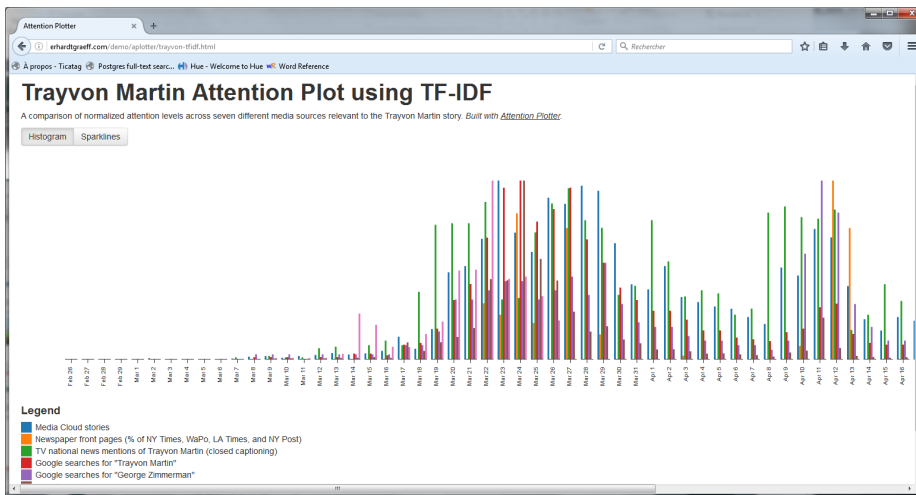
Axiom : Birds \Rightarrow to fly



Data Science & Networks

Qualitative analysis :

From linguistic statistics towards semantic inferences and fuzzy reasoning



Axiom : Birds \Rightarrow to fly



$p(\text{Birds} \Rightarrow \text{to fly}) = 0.99$



Data Science & Networks

Qualitative analysis :

From linguistic statistics towards semantic inferences and fuzzy reasoning



Axiom : Birds \Rightarrow to fly



$$p(\text{Birds} \Rightarrow \text{to fly}) = 0.99$$



Data Science & Networks

Qualitative analysis :

Fuzzy reasoning and Analytic Intelligence



Axiom : Birds \Rightarrow to fly



Discrimination
of lexical ambiguities
in semantic networks

$$p(\text{Birds} \Rightarrow \text{to fly}) = 0.99$$

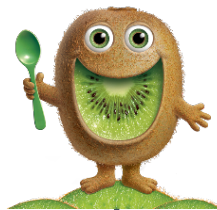
$$p(\text{Fruit} \Rightarrow \text{to fly}) = 0.001$$



Big Data + Data Science + Network Science

→ From machine learning towards machine reasoning

$$p(\text{Kiwi} \Rightarrow \text{is_bird}) = 0.5 \quad ? \quad p(\text{Kiwi} \Rightarrow \text{is_fruit}) = 0.5$$



wiseGEEK

$$p(\text{Birds} \Rightarrow \text{to fly}) = 0.99$$

$$p(\text{Kiwi} \Rightarrow \text{to fly}) = 0.01$$

$$p(\text{Kiwi} \Rightarrow \text{to fly}) = 0.001$$

$$p(\text{Fruit} \Rightarrow \text{to fly}) = 0.0001$$



Big Data + Data Science + Network Science

→ From machine learning towards machine reasoning

$p(\text{Kiwi} \Rightarrow \text{is_bird}) = 0.5$

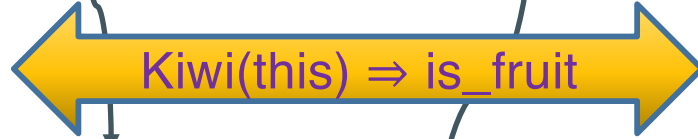
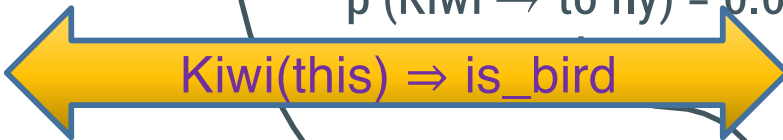
 $p(\text{Kiwi} \Rightarrow \text{is_fruit}) = 0.5$



wiseGEEK

$p(\text{Birds} \Rightarrow \text{to fly}) = 0.99$

$p(\text{Kiwi} \Rightarrow \text{to fly}) = 0.01$



$p(\text{Kiwi} \Rightarrow \text{to fly}) = 0.001$

$p(\text{Fruit} \Rightarrow \text{to fly}) = 0.0001$



Big Data Science & Networks

Other Works :

- Daily recommendations for high stock availability reducing distribution costs in round trips with stock return from delivery points and constrained transportation capacity.

Perspective example :

- Fuzzy reasoning on the Game Theory for Trading and Marketing

THANK YOU

zinnya.delvillar@data2b.net

christophe.thovex@data2b.net