

The Inria logo is written in a stylized, cursive font with a color gradient from red to orange to yellow.

# *Algorithmic systems Transparency and Accountability*

**Nozha Boujemaa** - Director of Research  
Advisor to The CEO of Inria in Big Data  
Member of the Board of Directors of BDVA

**Net Futures- Brussels 28th June 2017**

Carte de pollution urbaine : assimilation des mesures de concentration en NO<sub>2</sub> dans le modèle ADMS Urban/Numtech. Équipe CLIME.

↑ Nord



[Nozha.Boujemaa@inria.fr](mailto:Nozha.Boujemaa@inria.fr)

# Data & Algorithms

*« 2 sides of the same coin »*



- Data are everywhere in personal and professional environment
- Algorithms making sense from these data are pervasive in more and more digital services.
- Algorithmic-based decisions are embedded in several domains ranging from the processing of personal data to sensitive data in critical industrial systems (autonomous cars, conversational agents, health-care and well-being or public services etc).

# Data & Algorithms

« 2 sides of the same coin »

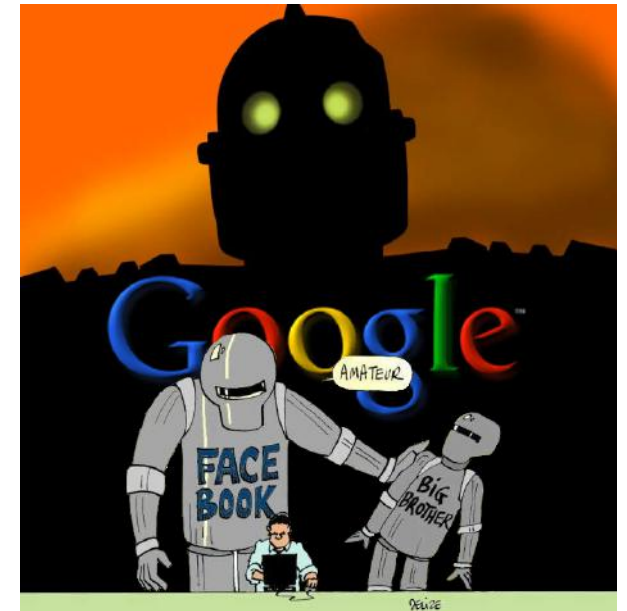


- Rising benefits from Big Data and AI technologies have wide impact on our **economy** and **social organization**.
- **Transparency** and **ethics** of such **Algorithmic Systems** (**data & algorithms**) become increasingly important properties for **trust** and appropriation of digital services
- Data analytics is changing from description of past to **predictive** and **prescriptive analytics** for decision support
- According to the countries, a vigilance has been created around **data (privacy)** and / or **algorithms (prescriptive)**

# Algorithms in every day life

- Some dominant platforms on the market play a role of "prescriber" by directing a large share of user traffic:
  - Ranking mechanisms (search engine),
  - Recommendation mechanisms and content selection

Product or service recommendation: is it most appropriate for the consumer (personalization) or the most appropriate given the stock of the seller?



- What guarantees of **impartiality to commercial factors** of classification software?
- Possible **illegal tariff** practices and **discrimination**
- Opacity** of the **use** made of the **personal data** of the users, what about the **consent**?  
Is it always **respected**? Experience Mobilitics CNIL-Inria
- Importance of remedying **the information asymmetry** between **the producer of the digital service** and its **consumer**, be it citizen or professional (B2C or B2B)

# Data Science Challenges: Accountable and Ethical Data Management and Analytics

- **Decision explanation and tractability:** Trust and Transparency of computer-aided decision-making process (**decision responsibility**): what are the different criteria/data/settings that have led to the specific decision in order to understand the global path for the reasoning?
- “How Can I trust Machine Learning prediction?” it happens to build the model of the object context rather the object itself
- **Robustness to bias/diversion/corruption**
- **Careful software reuse**

# Data Science Challenges: Accountable and Ethical Data Management and Analytics

**Mastering Big Data Technologies:** **Bias** problems could impact data technologies **accuracy** and people's lives

## Challenges 1: Data Inputs to an Algorithm

- *Poorly selected data*
- *Incomplete, incorrect, or outdated data*
- *Data sets that lack disproportionately represent certain populations*
- *Malicious attack*

## Challenges 2: The Design of Algorithmic Systems and Machine Learning

- *Poorly designed matching systems*
- *Unintentional perpetuation and promotion of historical biases*
- *Decision-making systems that assume correlation implies causation*

## **Data Science Challenges: Accountable and Ethical Data Management and Analytics**

- Opening the black box of Deep Learning
- Data provenance and usage monitoring
- Progressive user-centric analytics
- New paradigms for information flow monitoring
- Fact-checking requiring explicit, verifiable argumentation integrating heterogeneous data sources and explainable reasoning

# Relevant French Programs

## **TransAlgo** – Dec 2016, Minister of Digital Economy

- National Scientific **Platform** for **Transparency & Accountability** Tools and Methods for Data and Algorithms (Fairness, Neutrality, Loyalty); **b2b & b2c**
- Support of The new “Law for Digital Republic” after CGE report,
- Contributors: CNNum, DGCCRF (French FTC) besides academia and associations,
- 3 Objectives: \* Resource center (reports, publications, software, initiatives), \*Research & Dev. programs, \* Best practices & Moocs

## **I2-DRIVE** : Interdisciplinary Institute for **Data Research: Intelligence, Value and Ethics** (Convergence Institute)

- **4 Overarching Challenges:** *From Data to Knowledge, from Data to Decision, Deep learning toward Artificial Intelligence, Digital Trust and Appropriation, Data economy and regulation*
- **Scientific and disciplinary foundations:** Data Science, Management and Economy, Social Sciences, Legal Sciences
- *Roadmap for 10 years, 200 M€ Budget, 14 academic institutions*



The Inria logo is written in a red-to-orange gradient cursive font on a white rounded square background.

*Inria*

**Need for Interdisciplinary  
European effort!**

**Thank you for your attention**



## Trusted AI in Smart Industry