

EXPONENTIAL | BY RAVENPACK

# THE NEXT GENERATION OF QUANT RESEARCH



**Charles-Albert Lehalle**

Professor in the Dept of Applied Math (CMAP),  
Ecole Polytechnique

# The Next Generation of Quant Research

## From Data to the Supply Chain, to Narratives

Charles-Albert Lehalle



Full Professor, Applied Maths, at  
CMAP, CNRS, Ecole Polytechnique, IP Paris  
Fellow of Institut Louis Bachelier



Co-Director of the "AI for Markets and Quantitative Investment" (MaQI) Master of  
Science and Technology

Ravenpack's Exponential - June 3, 2026 - London

# The travel for today

- Prediction → forecasting / nowcasting / foresight
  - Datasets: everything, everywhere, and ...
    - The Supply Chain as a central point of recording
      - Narratives and their link with World Models
        - Biases and the need (or not) to remove them

# What to predict ?

Anticipation, prediction, speculations, ...

I am old enough in the industry to have witnessed sliding concepts:

1. I started at the era of **forecasting**: trying to continue a time series using its past
2. I saw alternative data pushing the concept of **nowcasting**: getting a grasp on what is occurring now on the economic ground
3. Now it the realm of **foresight**: forming out of distribution anticipations, with a come back of narratives

Nevertheless, it does now say **what is the target** of these efforts: future risks? future prices, or price moves? or the explanation of these moves (fundamental or due to liquidity tensions)? It requires some knowledge of price formation.

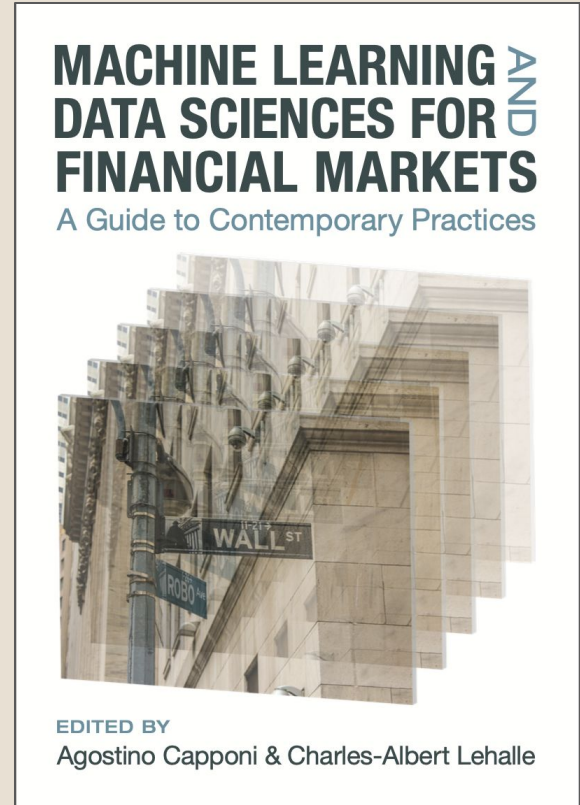
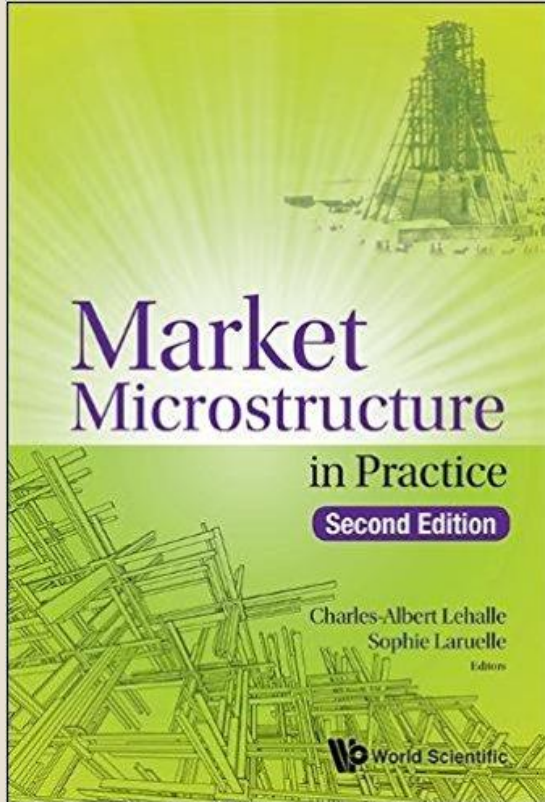
# Price formation



# Data and learning

This is the neverending story of the Grossman –Stiglitz paradox:

Information moves prices, such a way that the potential gain to understand information copes with the cost of processing it.



# The joined story of Quantitative Investment and Datasets

It started with market prices and volume only

With "Technical Signals" (mean-revert., momentum, etc), CTA captured:

- bad liquidity synchronisation, business cycles, long term economic cycles

Then Fundamentals (balance sheets of companies)

- allowed to build "factors" (quality, value, etc) that are close to "approximate relative microeconomic value", inertia of governance

Now, after the "big data era" (during which any company recorded a lot of information, "just in case"), we have information flowing from everywhere:

- retail behaviours, transportation, production, flows of capital, etc

# Different nature of information

**Pictures** (satellites, drones, mobile phones, glasses, etc)

**Positions** (of ships, trucks, car, people, etc)

**Interactions** (trade, capital flow, unload/load, sell, buy, etc)

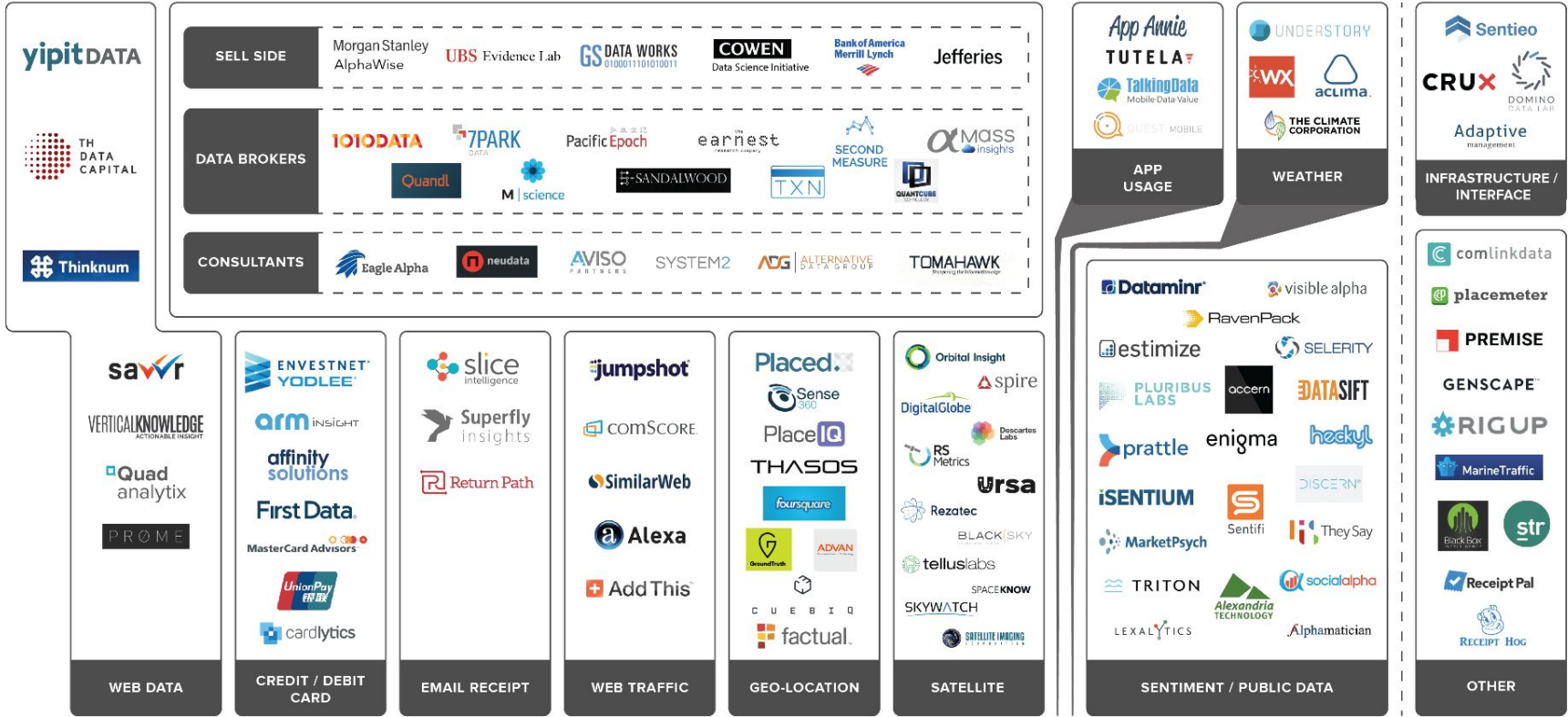
**Texts** (company filings / reports, central banks, legal documents, analysts, journalists, people, etc)

=> They are **not synchronized** (neither in space nor in time), **not structured**

=> They give you a **speed advantage** (for HFT)  
or a **breath advantage** (for large firms).

## INSTITUTIONAL INVESTORS

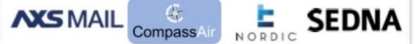
ANALYSIS  
CLEANING  
COLLECTION



Zooming on  
maritime traffic  
only

...

**Commercial Operation Software**
**Voyage Management**

**Shared E-mail**

**Pre-Fixture**

**Post-Fixture**

**Vessel Management Software**
**Technical Management**

**Vessel Performance Monitoring**

**Weather & Ship Routing**

**Maritime Intelligence & Analytics**
**AIS Tracking & Satellite Data**

**Satellite AIS Services**

**Market Data/Forecasts**


# Is there an ideal target structure to store and update?

A nowcasting project would be to build a model of **the whole Supply Chain**

- suppliers, clients, competitors, commodities...
- that should be localized (where are your warehouses?)

This model should have flexible, but realistic relations

- to identify **substitutability** (if wildfires start, or if a strait is closed)
- to enable **foresight** (what is the next event this portfolio of underlyings is exposed to?)

The "novelty" is the need to be able to form foresights, to **think out of distribution**, out of the box.

The way human do it is via **narratives**: a collection of coherent assertions affecting linked economic entities and their links (*"AAA can no more get gas or fertilizer via the strait of Hormuz"*).

# The special role of narratives

Narratives are made of **coherent sentences**

=> the notion of coherence of narratives is a "**world model**"  
(needs the support of an **ontology**)

Narratives are formed by opinion leaders, listening to the crowd/companies/countries opinion

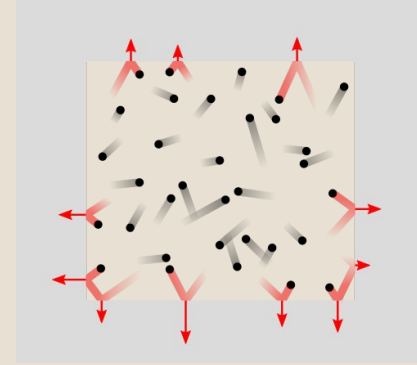
=> the crowd/companies/countries listen to the narratives and react

=> it pushes leaders to review narratives

=> Narratives **form a Mean Field** (like the atmospheric pressure)

Narratives trigger decisions (central banks wait a narrative to be dominant to lever on it)

=> **there is a closed loop**: this is a Mean Field Control



Pressure is a Mean Field  
(Wikipedia)

# Exploring the current dominance of narratives

The most natural way: "listen" to the text in real time



Overview Luxury Screener Conflict Timeline **Scenarios** Evidence Taxonomy

## Base Case: Conflict Contains & Partially Resolves (Q2 2026) 50% probability

RBC base case: sharp Q1 hit concentrated in March, followed by stabilisation as conflict largely contained by late Q1 or early Q2. Most stores reopen; footfall recovers gradually. ME revenue drops 45-20% for full year vs. prior trajectory. Luxury stocks partially recover as earnings visibility improves. LVNH, Hermès and Richemont see biggest bounces given overselling.

IMPLIED STOCK RETURN VS. CURRENT PRICE



## Bear Case: Sustained Conflict Through 2026 30% probability

RBC bear case: sustained decline in Middle East revenues throughout 2026, combined with margin pressure from logistics disruption, higher costs, and continued Gulf tourist absence from European flagships. Luxury profit scenario (H1 reporting) reveals deep margin erosion. Dubai's structural advantages (zero tax, low rent) become a lasting vulnerability for H1 earnings. Swatch and Zegna most at risk.

IMPLIED STOCK RETURN VS. CURRENT PRICE



## Bull Case: Rapid De-escalation + Pent-Up Demand 20% probability

Conflict de-escalates rapidly in Q2; Dubai reopens fully; Gulf HNW clients display strong pent-up demand; Chinese recovery accelerates simultaneously. Deutsche Bank 'sharp reversal' scenario plays out. Luxury stocks rally 20-30% from April lows as H1 2026 sales prove resilient. Saudi Arabia Vision 2030 demand resumes. Richemont and Zegna (most oversold on ME exposure) lead the recovery.

IMPLIED STOCK RETURN VS. CURRENT PRICE



# Of course it is not trivial (Remind Grossman-Stiglitz)

But this is difficult

- they are unbalanced
- their update frequency is not the same for any KPI

One good way is to **map some structure on the top of texts** (yes: LLMs are structuring texts via projecting them on embeddings that satisfy algebraic semantic relations, but it is not enough).

If you reduce texts to a short list of metadata, it will not be enough, you need to **navigate simultaneously in both metadata and semantics spaces.**

# Can it be that simple?

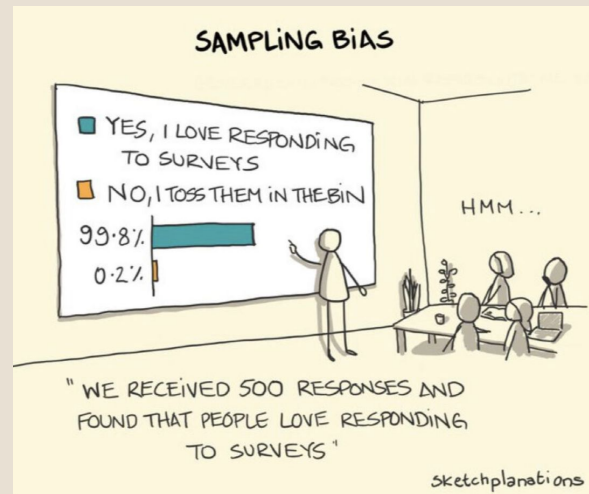
Updates are **not simultaneous** (daily, monthly, partial, sector, country, company level, etc)

=> you can only have a probabilistic view, and if you simply update what you have you will be **submitted to biases**:

- sampling / collection bias
- targeting bias

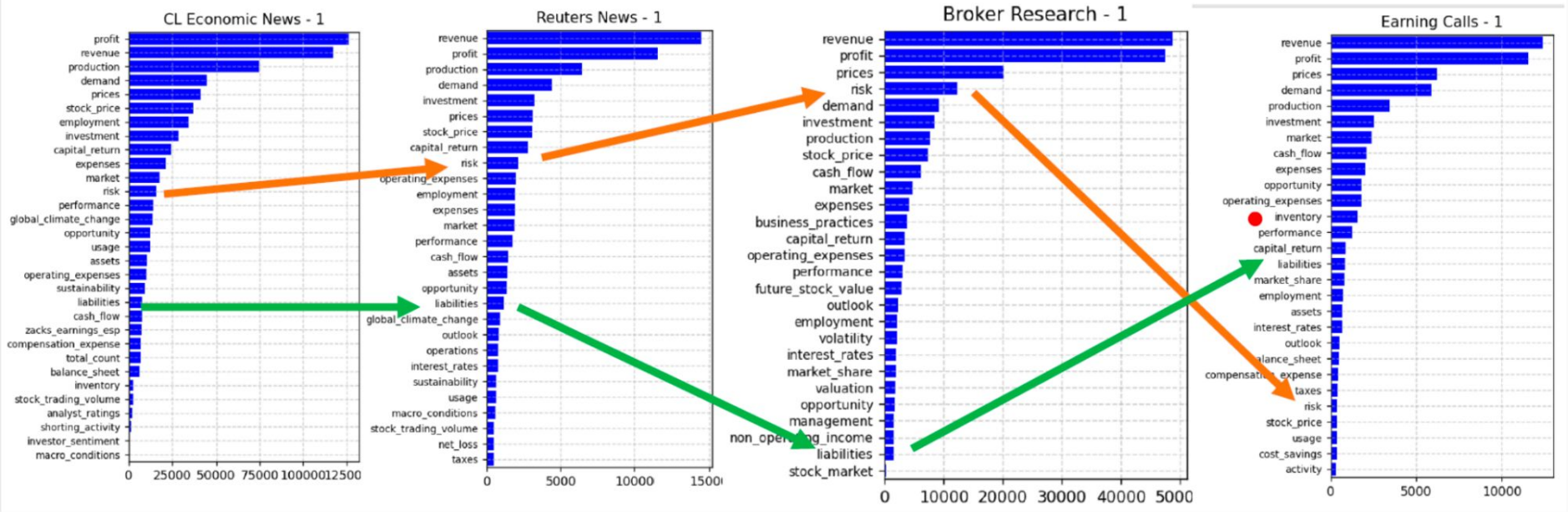
You need to learn to correct them, that is the topic of **active post-stratification**; you need reference data, metrics, and advanced statistics / ML

(call that optimal transportation, covariate shift, etc)



# Different sources: different topics, different update frequencies

Now that LLMs are a commodity: focus on the inputs



Joint work with CausalityLink.

# Do you really need to remove biases?

Think about **patents**



If you want to build a systematic investment strategy on the cross section of equities, you need to account for the fact that some sectors (and some countries) have a natural tendency to patent more than others.

=> You will normalize the intensity of patents by sector.

If you want to build a systematic investment strategy that is based on detecting when a sector will start to significantly patent more than before because it will be the sign of its impact by technology.

=> the previous normalization factor is part of your signal, you want to keep it, not cancel it.

The correct debiasing depend on the story / narrative reflected by your strategy.

# We reached the destination: Time to conclude

- Prediction → forecasting / nowcasting / foresight
  - Datasets: everything, everywhere, and ...
    - The Supply Chain as a central point of recording
      - Narratives and their link with World Models
        - Biases and the need (or not) to remove them

# What is quant research today?

- More creativity is required since there is no limit in the data you can get
- The algo is no more the start of the reasoning
- Because the structure of an investment strategy is a **narrative**
- Narratives are special set of sentences, close to '**World Models**'
- Having access to a proper ontology allows regularisation and structure, that are needed to **identify biases** and to build models of the **Supply Chain**
- They are important to project yourself outside of past data, to **foresight** scenarios

A lot of remaining open questions, like defining diversification along a Supply Chain.



## The Alternative Data Podcast

### The Charles-Albert Lehalle Episode

Season 1, Ep. 149 • Monday, March 2, 2026

In this episode I speak to Charles-Albert Lehalle, a professor at Ecole Polytechnique formerly of ADIA and CFM.



Thank you for your attention  
[charles-albert.lehalle@polytechnique.edu](mailto:charles-albert.lehalle@polytechnique.edu)