BANK OF CANADA
BANQUE DU CANADA

FEBRUARY 2024

TABE/CABE WEBINAR

Leveraging alternative data and AI in central banking analyses

Maryam Haghighi
Director, Enterprise Data Science and Insights
Bank of Canada

mhaghighi@bankofcanada.ca

The views expressed in this presentation are those of the speaker and do not necessarily reflect the official views of the Bank of Canada.

Agenda

- The data revolution
- Tools, Techniques, Use cases
- Risks
- What would it take to fully leverage all of this?

Data and Decision Making at Central Banks

Data is at the heart of decision-making at a central bank.

Statistical surveys, regulatory submissions.

Those datasets are highly reliable, but they are typically:

1) Aggregated
2) Lower frequency
3) Non agile

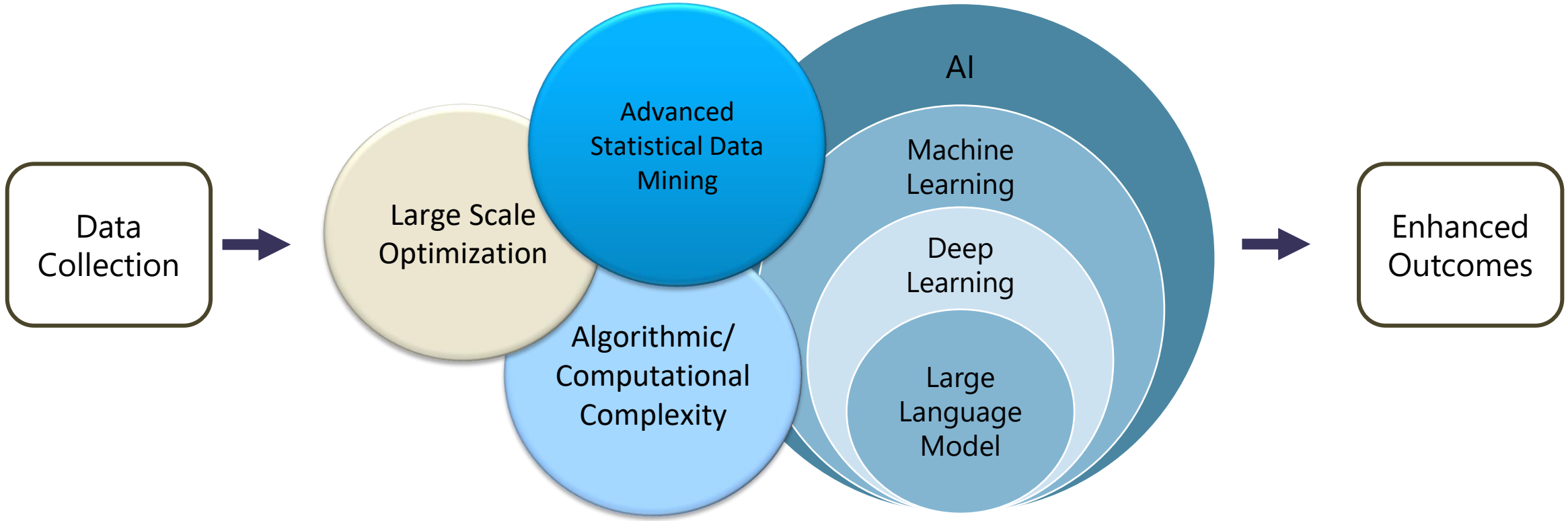
The Data revolution over the past 15 years

New devices that collect data, new tools for analysis and better data storage.

These new “non-traditional” data:

- Don’t necessarily come from surveys (e.g. satellite imagery, payment transactions, social media data, electricity usage)
- Can be in very large **volumes** (terabytes) → **Volume**
- Can be in different **formats**: not just figures and tables, but also text, videos etc. → **Variety**
- Can be obtained at much **faster** rates and higher frequencies → **Velocity**

New tools and techniques to deal with the 3V's



Need for bespoke development in a central banking context

data noun

1 : factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation

> the *data* is plentiful and easily available
> – H. A. Gleason, Jr.

> comprehensive *data* on economic growth have been published
> – N. H. Jacoby

~1600s

Dictionary

Definitions from Oxford Languages

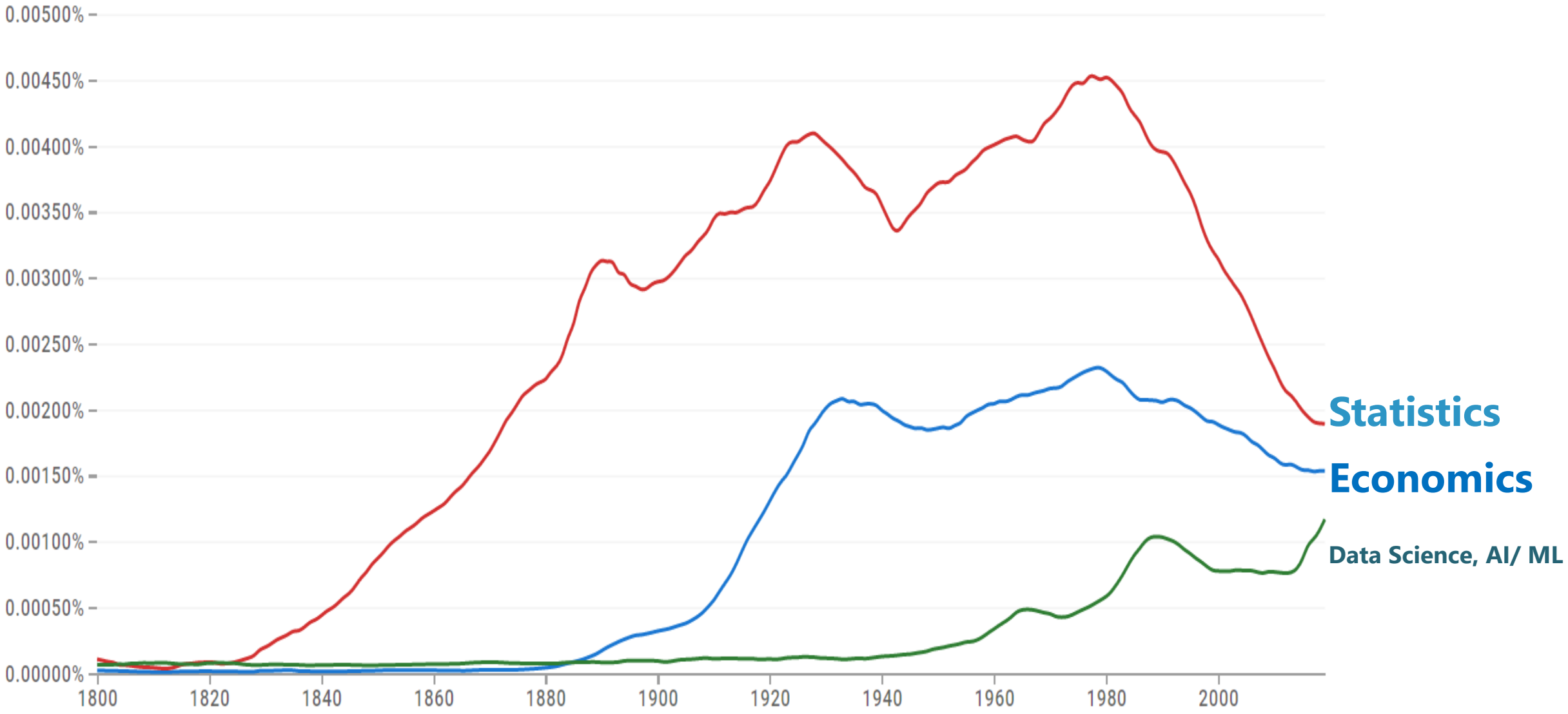
sci·ence

noun

1. the systematic study of the structure and behavior of the physical and natural world through observation, experimentation, and the testing of theories against the evidence obtained.

Much older
~1300s or before


Data Science
2010
0.0000550%
0.0000500%
0.0000450%
0.0000400%
0.0000350%
0.0000300%
0.0000250%
0.0000200%
0.0000150%
0.0000100%
0.0000050%
0.0000000%
1800
1820
1840
1860
1880
1900
1920
1940
1960
1980
2000


0.00500%
0.00450%
0.00400%
0.00350%
0.00300%
0.00250%
0.00200%
0.00150%
0.00100%
0.00050%
0.00000%
1800
1820
1840
1860
1880
1900
1920
1940
1960
1980
2000
Statistics
Economics
Data Science, AI/ ML

A data science ecosystem:

Governance
Prioritization
Risk Management

Delivery
Scalable
Technology Stack
Build once- Use across

Build Capabilities
Staffing Strategy, Training
External collaboration
Seminars and Community of Practice

The essence of successful data science is in multidisciplinary teamwork.

Examples from the experience of others

- ECB: Michele Lenza et al (2023) built a new AI model for forecasting to capture non-linearities that is informative to policy.
- Norges Bank: Are Aastveit et al. (2020) use debit card data to nowcast household consumption and find sizeable improvements over their existing nowcasting methods.
- San Francisco Fed: Shapiro et al. (2020) use major US newspapers to construct a daily measure of news sentiment. Their index fell sharply in January of 2020, nearly two months earlier than existing survey-based sentiment measures.
- BBVA Research: Buda et al. (2022) use transaction data to understand the distribution of consumption in real-time, as well as form measures of aggregate consumption in existing national accounts with a higher degree of accuracy.

Overview of some use cases

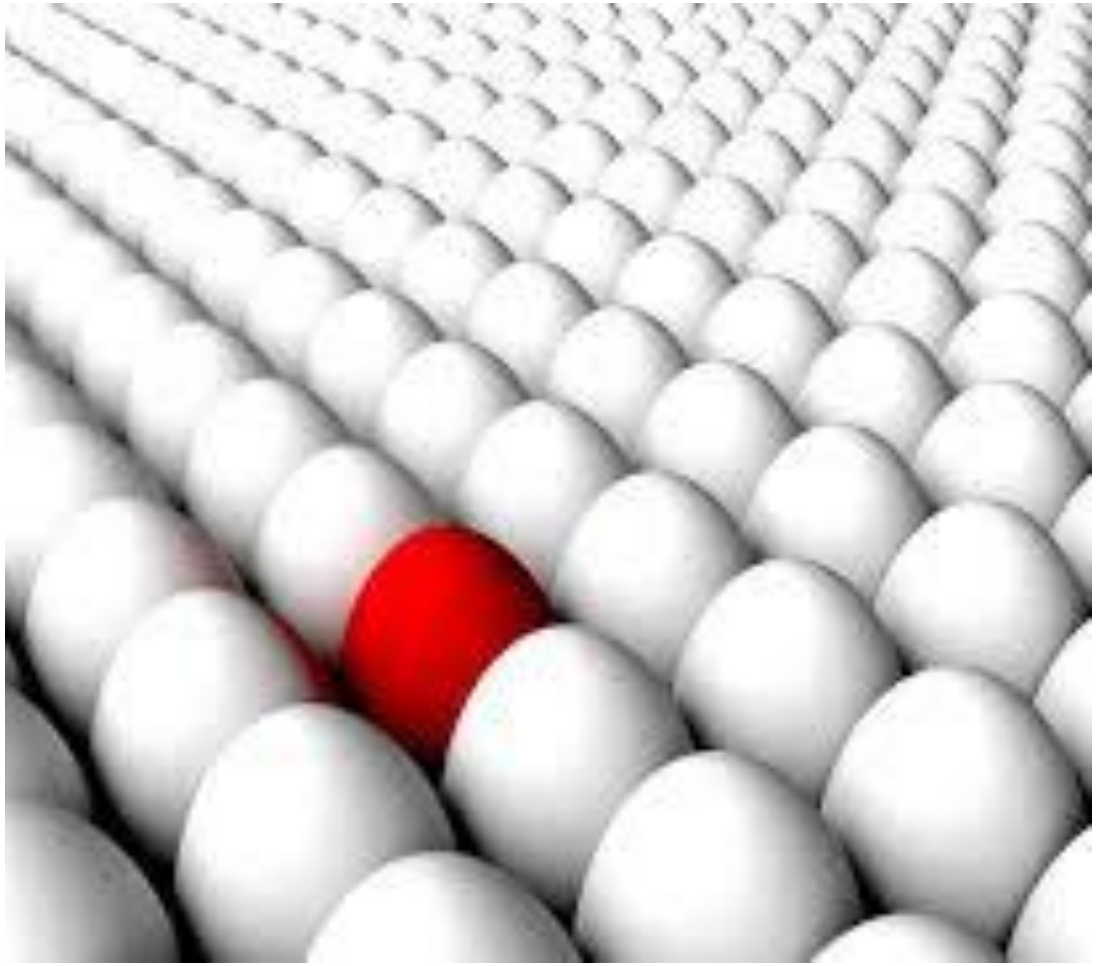
Use Case 1: Power of Operationalized Machine Learning

Millions of datapoints every month
Foundation for many of our models and analysis

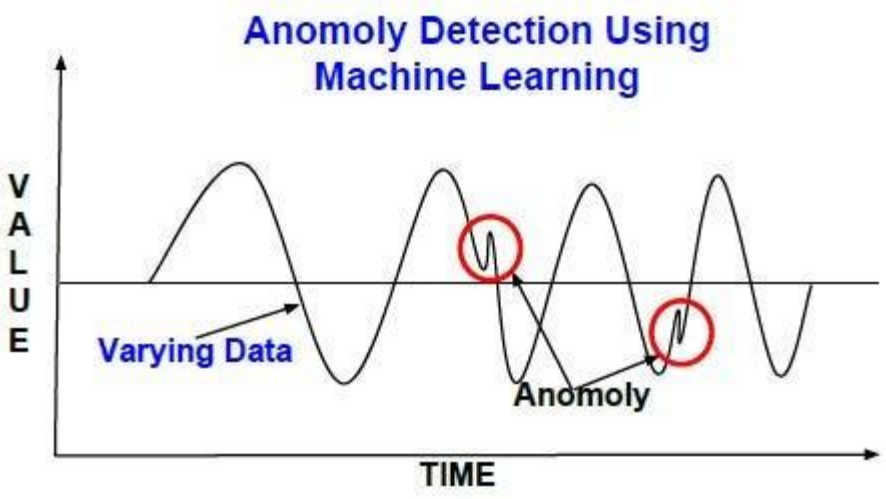
Impossible to examine every single one with traditional ways.

Challenges with **traditional rule-based approaches for anomaly detection**:

- **Significant time** using rule-based approaches.
- **Risk** that critical anomalies may be missed



Use Case 1: Power of Operationalized Machine Learning



Developed a ML model for anomaly detection.
Assessing the entire data+ Saving significant time

Now operationalized, running daily.

Formalized with modern standards, reliable, scalable, fully explainable.

Use Case 2: Advanced Visualization: Interactive Insights

Wide range of visualization tools: Excel, Tableau, Power BI etc. They have their usage, excellent for reporting and analytics.

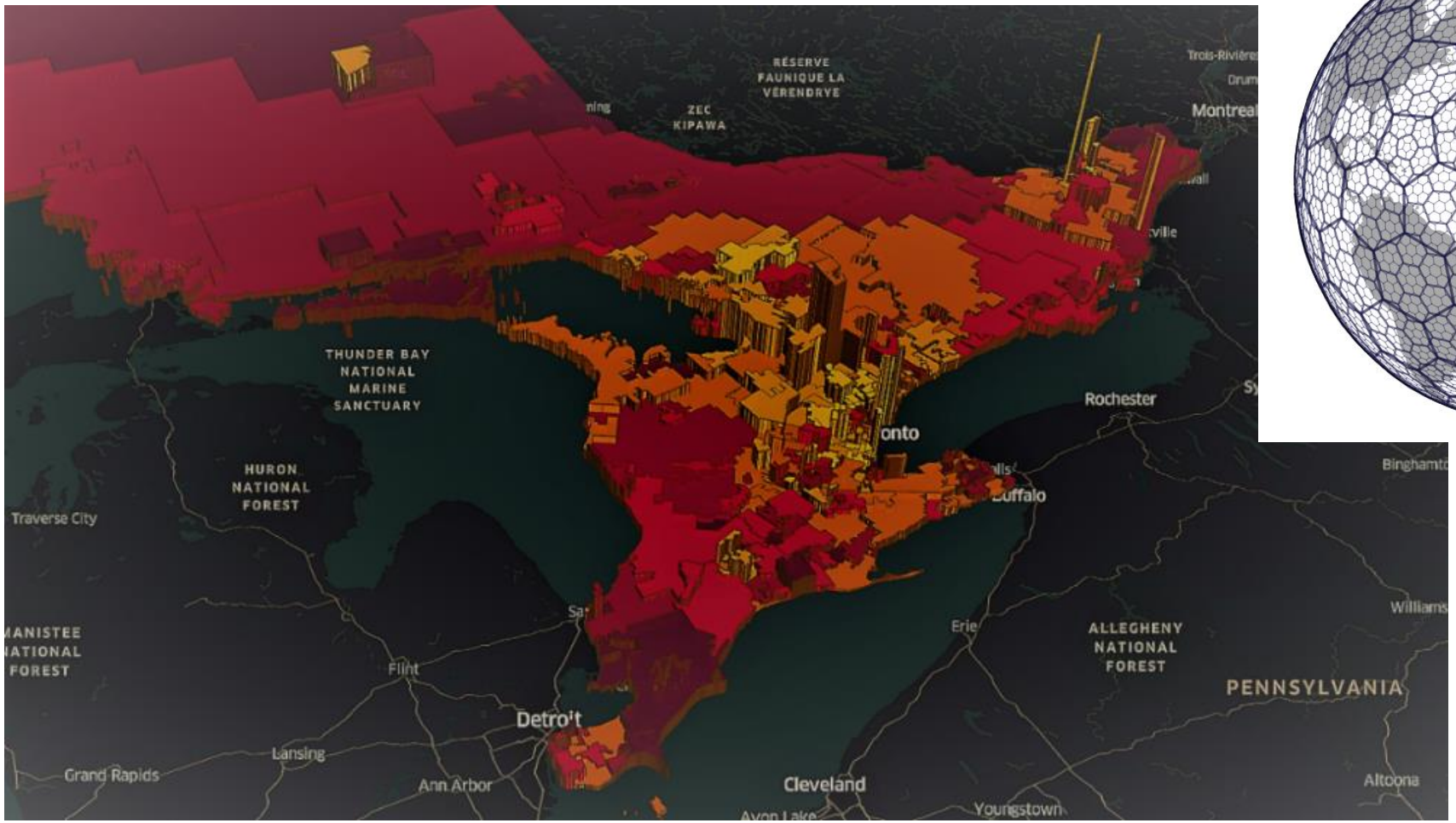
But….

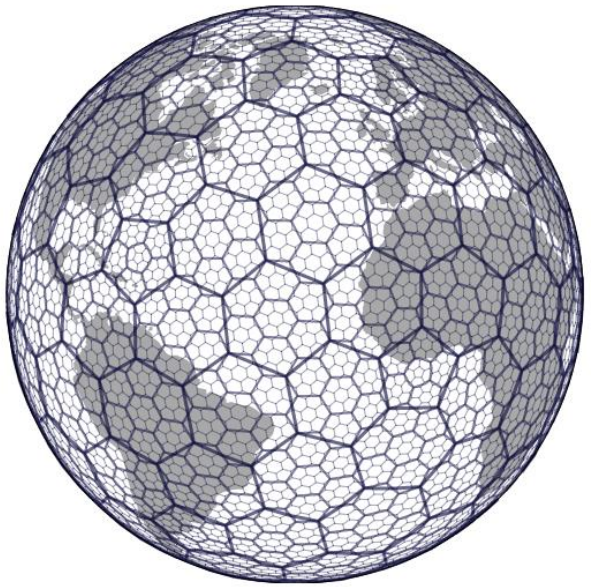
How can we look at multi-dimensional high-volume data, on-demand, interactive (not static)?

“High resolution” views of complex models to be able to drill down or aggregate up?

Example: Tiling the earth- Hexagonal geo-indexing library

Multi-dimensional datasets of high-volume, aggregate up or drill down:






RISK

What would it take to more fully leverage nontraditional data and techniques?

Tech

Modern data and tech stack, platforms and architecture
Classical AI versus Generative AI and LLMs

Talent

Multidisciplinary collaborations, internally and with peer organizations
Research intensive
Upskilling

Risk

Balancing value creation and adequate risk management
Strong governance practices

Culture

Taking a coordinated approach to leverage advances.
Scalable, repeatable, responsible innovation