Seeing the Invisible: AI-Driven Transparency in Global Supply Chains

Why our current tools cannot see risk – and what AI can do about it

Amir H. Keshavarzzadeh
Alfred Sun
Ramit Debnath
Kamiar Mohaddes
Jonathan Cullen

University of Cambridge

December 15, 2025

UNIVERSITY OF CAMBRIDGE

Why supply chains matter

- Every product in this room rests on thousands of upstream decisions.
- A single bottleneck can ripple across continents – as we saw during Covid.
- Example: **global dependence on Congolese cobalt**. In our previous work, we estimated that a disruption in the DRC could impose around **US$140 billion** in value added losses – yet we still do not see how this spills over into **EVs, smartphones, defense and other sectors** because supply chains are **not transparent**.

Global cobalt-based battery supply chain. Blue nodes show downstream firms supplying cobalt from the Democratic Republic of the Congo; red nodes show cobalt mines in the Democratic Republic of the Congo, which accounts for roughly 70 % of global cobalt mine production. Arcs indicate shipment links, highlighting how a highly concentrated upstream resource underpins a globally distributed manufacturing base.

Today’s headlines are still about chokepoints

Supply chains are front-page news

Financial Times
“Red Sea crisis pushes up delivery times for European manufacturers”

Houthi attacks force rerouting around Africa, raising freight and input costs.

Financial Times
“Gas and grain ships shun Panama Canal after drought disruption”

Low water levels severely limit transits; traders avoid the canal.

Financial Times
“Three lessons from the US port crisis”

Pandemic and the Baltimore bridge collapse stress-test US logistics.

UNIVERSITY OF CAMBRIDGE

Today’s headlines are still about chokepoints

Supply chains are front-page news

We keep reading these stories, but do we actually know **who** ships **what** to **whom**, beyond a handful of tier-1 suppliers?

Financial Times
“Red Sea crisis pushes up delivery times for European manufacturers”

Houthi attacks force rerouting around Africa, raising freight and input costs.

Financial Times
“Gas and grain ships shun Panama Canal after drought disruption”

Low water levels severely limit transits; traders avoid the canal.

Financial Times
“Three lessons from the US port crisis”

Pandemic and the Baltimore bridge collapse stress-test US logistics.

UNIVERSITY OF CAMBRIDGE

Our measurement toolkit is stuck in the 1970s

Great for macro statistics, blind to firm–firm links

- National and multi-regional Input–Output (IO) tables
- Trade statistics (customs data), sectoral surveys
- Work at **sector–country** level, not at **firm–firm** link level

Formal IO model (Leontief):

$$x = A\,x + y \quad \Rightarrow \quad x = (I - A)^{-1}\,y$$

where A is the matrix of technical coefficients, x total output, and y final demand.

IO matrices aggregate millions of firm–firm links into a handful of sector–country blocks.

Why this is not enough

We see sectors and countries, not the actual points of failure

What IO actually tells us

IO tells us that **Manufacturing in Country A** buys from **Metals in Country B**.

Why this is not enough

- Aggregates **millions** of firm–firm links into **a few** sector–country averages.
- Misses real-world **dynamics**: inventories, capacity limits, rerouting of orders, delivery delays.
- No explicit **infrastructure**: ports, plants, warehouses, logistics hubs, transport routes.
- Hides **single points of failure** and how shocks propagate through specific firms and facilities.

Why this is not enough

We see sectors and countries, not the actual points of failure

$\sim$ 97% **information loss**

Based on information-loss theory and empirical data, our comparison of IO tables with firm-level shipment networks shows that standard IO views lose almost all of the information of supply chains.

Tossing a fair coin:
50% chance of a correct guess.

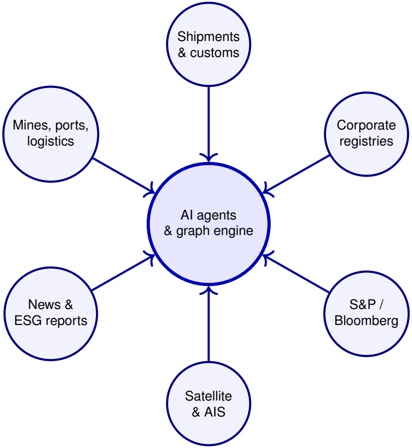
UNIVERSITY OF CAMBRIDGE

The new ingredients: compute, data, and AI agents

From web-scale digital exhaust to a live supply chain graph

AI supply chain data pipeline

- **Web scraping / APIs** for customs records, bills of lading, shipment logs, corporate registries, financial filings, news and ESG reports.
- **Satellite imagery** and **AIS / port tracking** to observe physical flows of vessels, stockpiles and mines.
- **AI agents** orchestrating scraping, parsing, entity resolution, de-duplication and **cross-validation** against S&P Capital IQ Pro and Bloomberg.
- **Graph / representation learning** to infer missing links, rank chokepoints and detect anomalies in the network.



AI agents integrate heterogeneous data sources into a unified, continuously updated supply chain graph.

UNIVERSITY OF CAMBRIDGE

Supply chains are not tree-like graphs

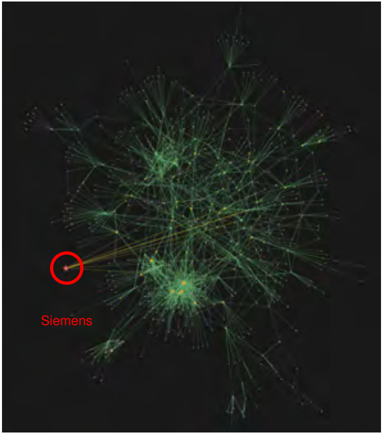
From textbook tiers to real-world graphs

As a **proof of concept**, our AI-enabled framework reconstructs a multi-tier view of the **Siemens** global supply chain.

Siemens supply chain

- A **bushy graph** with shared suppliers, cycles and cross-sector dependencies.
- **5** supplier tiers, **797** firms and **2,270** relationships across **48** countries.
- Nodes: suppliers, subsidiaries, logistics hubs, customers; edges: shipments, monetary and physical flow, ownership links.

Understanding this structure requires **graph-based methods and AI**, not just spreadsheets or simple tier diagrams.

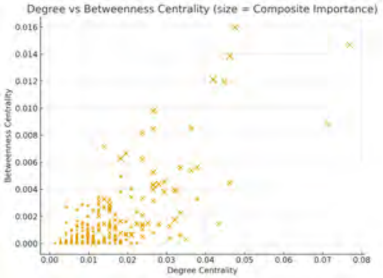


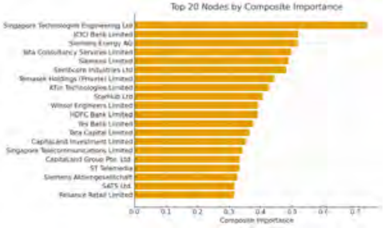
Siemens supply chain graph (5 tiers).

UNIVERSITY OF CAMBRIDGE

From network maps to measurable chokepoints

Classic graph-level metrics for identifying critical firms





Step 1: locate structural chokepoints

Each point in the top scatter plot is a firm.

- **Degree centrality** — how many direct partners a firm has.
- **Betweenness centrality** — how often a firm lies on the shortest routes between other firms.

Firms in the upper–right combine **many links** with **high control over flows**: these are natural candidates for chokepoints.

Step 2: rank critical firms

We merge degree, betweenness and exposure to flows into a single **composite importance index** and rank firms. The bar chart lists the top nodes by this index — the firms that are most structurally important for holding the network together.

UNIVERSITY OF CAMBRIDGE

Classical network metrics are not enough

Degree and betweenness still miss key supply-chain questions

Even with a full network map, **classical graph measures** (degree, betweenness, eigenvector centrality) do **not** give a complete picture of supply-chain risk.

What these metrics miss

- They treat links as static and homogeneous: no **volumes**, capacities or multi-commodity flows.
- They cannot tell whether a firm has **true substitutes** in the network, or is effectively irreplaceable.
- They ignore **infrastructure and geography**: ports, plants, corridors, regulatory and physical constraints.
- They do not integrate **rich data** such as financials, text (news, ESG, filings), or operational signals.

To really understand such graphs, we use **AI and graph neural networks**: combining shipment and customs data with firm-level, financial and textual information to learn which nodes are truly substitutable and which are systemic chokepoints.

Siemens wind-turbine supply chain graph.

UNIVERSITY OF CAMBRIDGE

Multi-layer risk engine for compound supply-chain risk

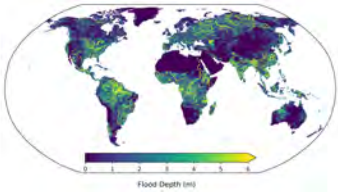
Geography + hazards + AI graph = where risk actually hits

What the multi-layer engine does

- **Geocodes facilities and routes:** mines, smelters, ports and transport corridors are located using satellite imagery and GIS layers.
- **Overlays in-house hazard databases:**
  - natural hazards (river flooding, drought, cyclones, heat),
  - anthropogenic hazards (armed conflict, sanctions, political instability).
- These hazard scores are pushed back through the **AI-derived supply-chain graph** to obtain compound risk metrics and to search for:
  - viable **substitute suppliers / routes**, and
  - **optimal inventory dynamics** to buffer shocks.

Example: we can link the Los Azules copper mine's exact location and elevation to projected flood depths, then quantify how often turbine generator copper might be disrupted and what substitutes exist.

Los Azules copper project, Andes (satellite view).



Example hazard layer: global river flood depth used in the in-house natural hazard database.

UNIVERSITY OF CAMBRIDGE

Thank you & Q&A

Thank you!

Questions & discussion

Amir Keshavarz
University of Cambridge
ak2495@cam.ac.uk

UNIVERSITY OF CAMBRIDGE