Tackling supply chain ‘blind spots’ via end-to-end visibility, traceability, and transparency: Conceptualising the role publicly data and big data analytics

Author: Wei Nie, Mariel Alem Fonseca, Dr Naoum Tsolakis, Dr Minette Bellingan, Dr Mukesh Kumar

Date: 4 July 2023 | Presentation for 2023 EurOMA Conference

UNIVERSITY OF CAMBRIDGE
Department of Engineering

IfM MANAGEMENT TECHNOLOGY POLICY

Agenda

01 Background

02 Question & Objectives

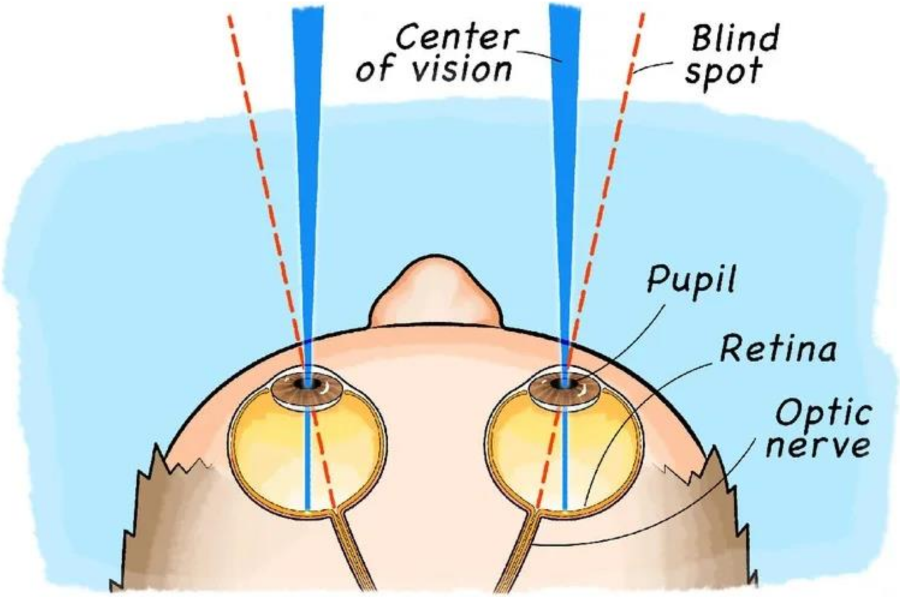
03 Method

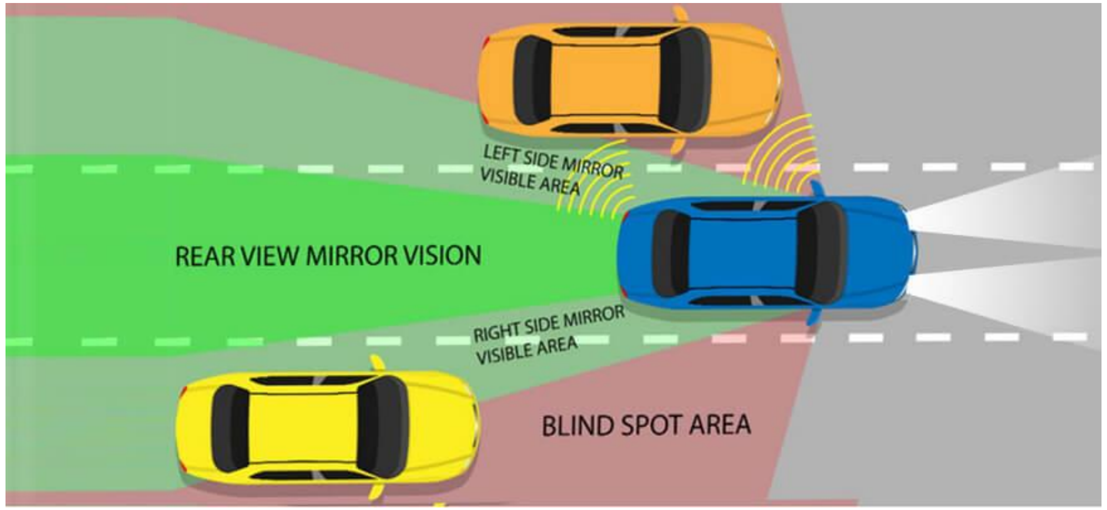
04 Findings

05 Discussion

06 Conclusion

Background – Blind Spots

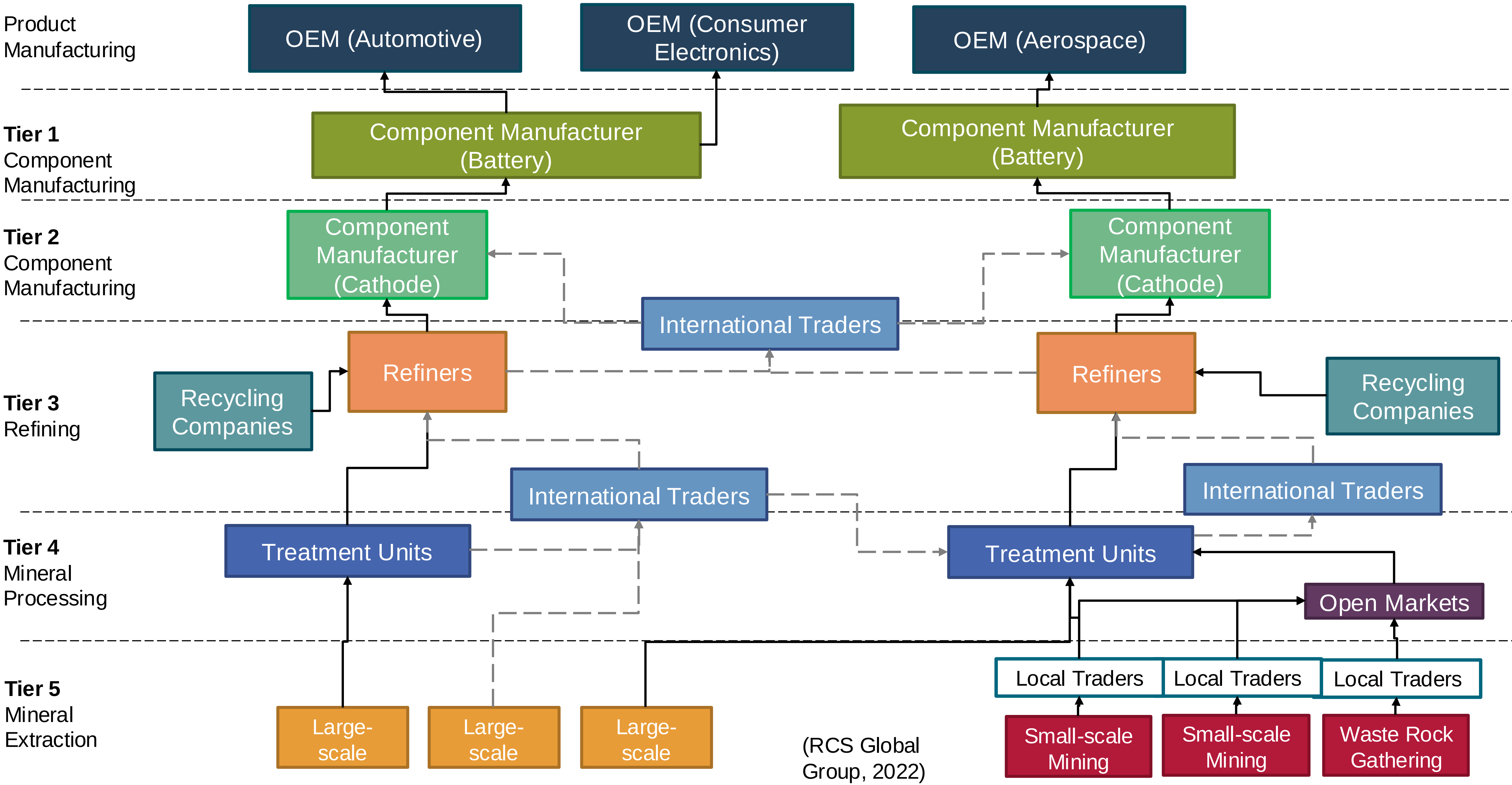




“A blind spot is **an area in your range of vision that you cannot see properly but which you really should be able to see”**
(Collins Dictionary, 2023)

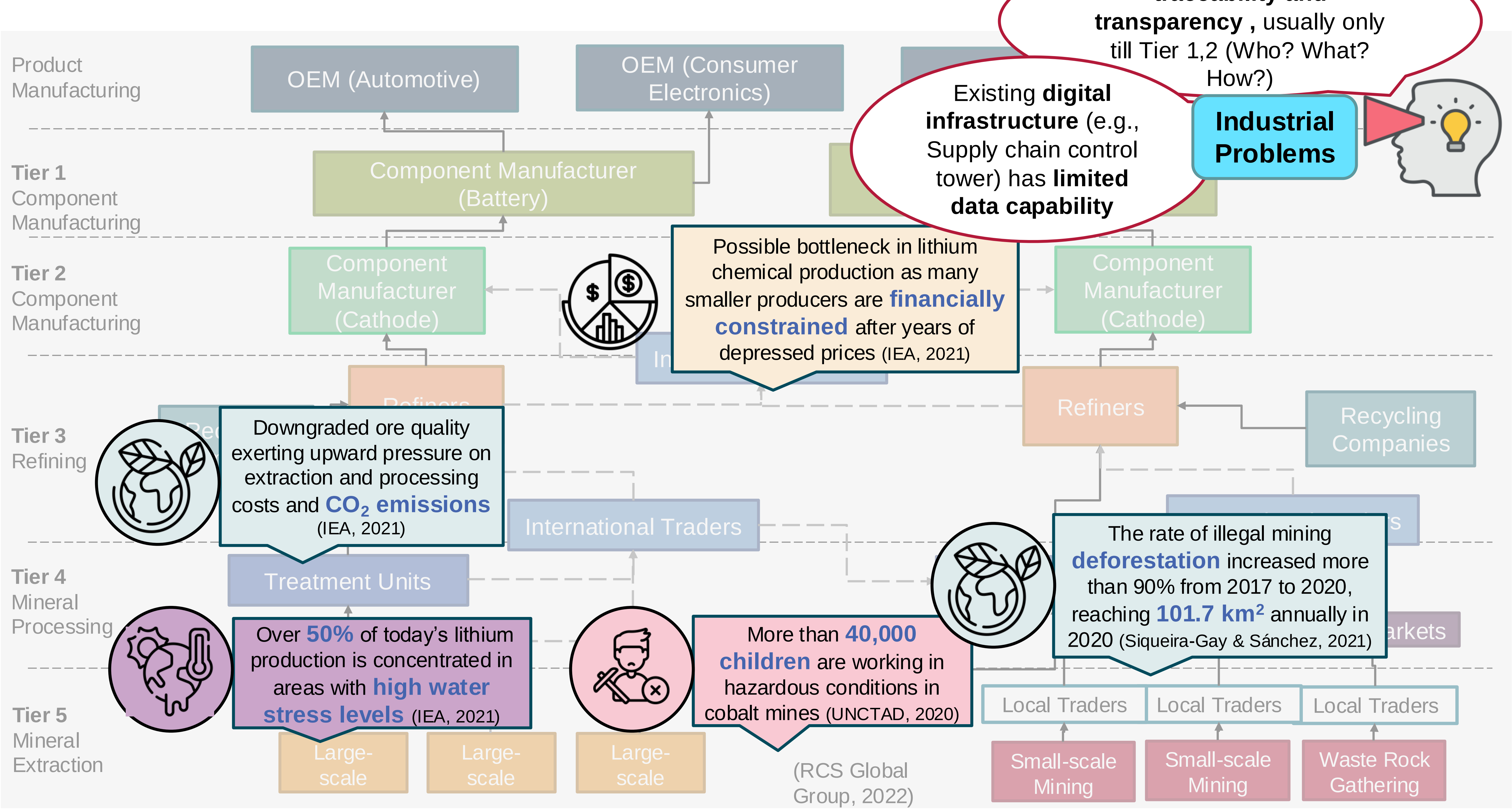
Photo credit to: https://www.scientificamerican.com/article/find-your-blind-spot/, https://www.haloview.com/blog/car-blind-spot, https://www.linkedin.com/pulse/main-blind-spots-supply-chain-management-oren-simon/

Background - Global Supply Chain Issues



(RCS Global Group, 2022)

Background - Global Supply Chain Issues



Background – Publicly Available Data & Big Data Analytics

Easily accessible data including:
- open data
- commercially available data
- data that can be obtained upon request
- public-domain data
- data with limited accessibility

(Cooper & Coetzee, 2020)

- Heterogeneous data types and formats
- High volume
- Generate supply chain knowledge

Publicly Available Data

Finance Data

Economics Data

Climate Data

Search Trend Data

Trade Data

Satellite Image

……

Spot illegal deforestation and forced labour in the Amazon rainforest (Andrews, 2022)

Use **import and export data** from UN Comtrade database for visualising the **geographical dispersion** of the commodity supply chain (Helbig et al., 2016)

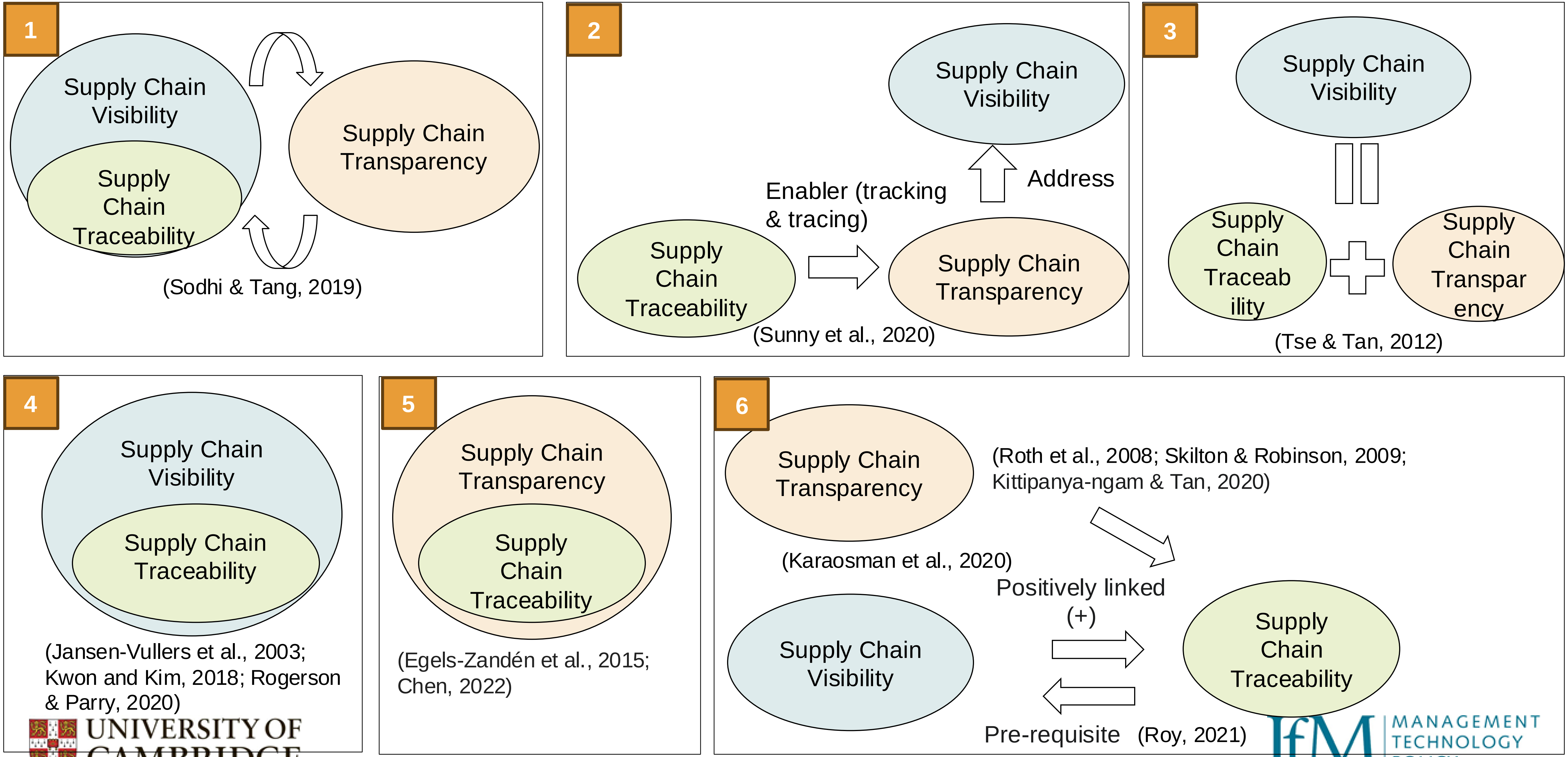
Use **Google Trends** data estimate product and service demand (Nikolopoulos et al., 2021)

Big Data Analytics (de Assis Santos & Marques, 2022)

UNIVERSITY OF CAMBRIDGE
Department of Engineering

IfM MANAGEMENT TECHNOLOGY POLICY

Background – Supply Chain Visibility, Traceability & Transparency



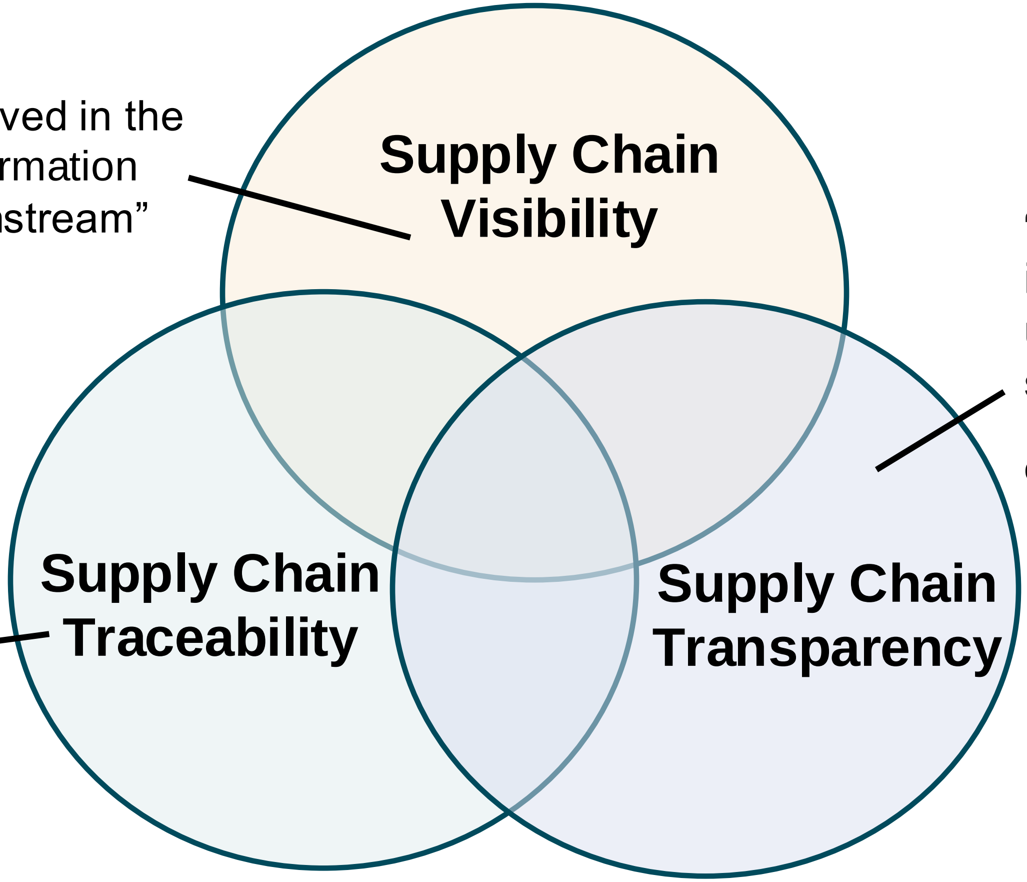
UNIVERSITY OF CAMBRIDGE
Department of Engineering

IfM MANAGEMENT TECHNOLOGY POLICY

Background – Supply Chain Visibility, Traceability & Transparency

"pertains to the degree to which entities involved in the supply chain possess timely and precise information regarding the operations upstream and downstream" (Barratt and Oke, 2007)

e.g., Demand



"company disclosing information to the public, including consumers and investors, about upstream operations and about the products it sells to consumers" (Sodhi & Tang, 2019)

e.g., Sustainability practice

"…the capability of a company for ascertaining provenance" (Sodhi & Tang, 2019)

e.g., Product journey

Considering data points related to supply chain nodes and the environments affecting them

Research Questions & Objectives

RQ: How can publicly available data be used with big data analytics (BDA) to develop the end-to-end supply chain visibility, traceability and transparency (SCVTT)?

Objectives

- Identify existing publicly available data that can be **linked** to SCVTT
- Understand the **exploitation** of publicly available data using BDA for the development of SCVTT, and their **challenges**

Research Method

A scoping review, 56 out of 2380 papers were selected for full review

Database	Web of Science
Search field	Topic
Search string	"supply chain*" AND ("simulation" OR "statistic*" OR "big data" OR "data-driven" OR "analytics" OR "machine learning" OR "data science" OR "data analy*" OR "operations research" OR "optimization" OR "optimisation" OR "open data" OR "public* data*" OR "database" OR "public* available data*" OR "data mining" OR "business intelligence")
Filters	Publication time: Last 5 years Document type: Articles Language: English Journal: Only ABS listed and with 3/4/4* scores
Title, Abstract and Full-text screening	**Inclusion criteria:** •Application paper rather than review or conceptual notes •Applied in supply chain operation context •Used publicly available data as the data source •Applied any of the mentioned techniques/models •Illustrated what and how publicly available data is used **Exclusion criteria:** •Studied the importance/capability/influence/culture of the mentioned techniques (especially using some qualitative research methods like survey and interview)

Findings (1) – Publicly Available Data related to SCVTT

1.Theme	*2. Type & Data Used*	*3. Platforms*	*4. Intended Uses*	
Traceability	Information linking supply chain places and actors **(import and export data)**	OECD, UN Comtrade	Risk management	(Nikolopoulos et al., 2021)
	Life cycle inventory information **(material & energy flow)**	GaBi, EcoInvent, Eora	Policy design & assessment	(Ingrao et al., 2019)
Supply Network Structure	Information on downstream and upstream supply chain actors & their relationships **(firms' main customers)**	Bloomberg's SPLC, FactSet	Supply chain structure identification	(Geng et al., 2022)
Supply Chain Activity	Activities of supply chain actors **(production, sales, processing data)**	FAOSTAT, British Geological Survey	Sourcing decisions, supply chain design	(Nagendra et al., 2022)
Financial Information	Firms' patterns of financial performance and transactions **(COGS, total assets)**	Bloomberg, Eikon, Compustat	Used as control variables	(Baker et al., 2022)
External Environments	Information on the earth environment **(precipitation, terrain elevation data)**	WorldClim, NOAA, EU Sentinel Satellites	Sourcing decisions, risk management	(Srinivasan et al., 2019; Nakano, 2021)
	Information on the socio-economics environment **(Population intensity, unemployment)**	Eurostat, US Department of Agriculture Market News		
ESG Policy & Commitment, Activities & Effectiveness	Supply chain actors' policy commitments to increase the transparency **(Global slavery index)**	Modern Slavery Registry, World Bank	Risk management, social responsibility and sustainability measurement	(Doni et al., 2019; Geng et al., 2022)
	Environmental impacts & performance of supply chain activities **(Emissions data)**	World Resource Institute, Bloomberg's ESG Database		

Findings (1) – Characteristics of Publicly Available Data related to SCVTT

Publicly available data used in existing research present great heterogeneity

Data Type

- Structured Data *(Tabular data)*
- Unstructured Data *(Satellite image)*

Data (File) Format

CSV, PDF, TIFF, JPEG, TXT, Webpage

Data Unit Standard

- International System of Units
- British Imperial Units
- US Customary Units

Data Provider

- Private providers *(Bloomberg)*
- Public institutions *(OECD)*
- NGOs *(Transparency International)*

Data Frequency

- Historical data *(Crop production)*
- Real-time data *(Google Trends)*
- Projected data *(Weather forecasting)*

Data Update Frequency

- (Near) Real-time
- Hourly, Daily, Monthly, Annual
- Static *(Distance between ports)*

Data Size

- Single data point *(Port distance)*
- Traditional data *(Annual crop yield)*
- Big data *(Satellite data)*

Data Geographical Coverage

- Regional *(State, Province)*
- National
- Global

Data Resolution

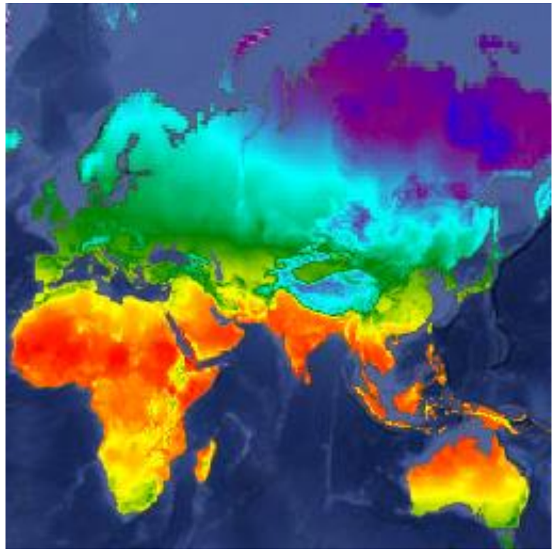
- Product level *(Commodity price)*
- Company level *(Company's' financial performance)*
- Sector level *(Pollutant emission inventory)*

(Ingrao et al., 2019; Srinivasan et al., 2019; van den Brink et al., 2020; Nikolopoulos et al., 2021; Taghizadeh et al., 2021; Baker et al., 2022; Geng et al., 2022; Nagendra et al., 2022; Ostroski et al., 2022)

Findings (1) – Publicly Available Data related to SCVTT

- **Intentions** of using publicly available data as the data source are different:
  - (1) used as the **sole** data source (Baker et al., 2022)
  - (2) used to **complement** the primary data (Ingrao et al., 2019)
- **Integrating multiple** publicly available data to generate meaningful information is a common practice (Srinivasan et al., 2019)
- Data are from various disciplines and need to be **translated** into the supply chain contexts

Example: Using publicly available data for food sourcing decision making



Historical & projected **temperature and precipitation data** from WorldClim database

+

	Optimal		Absolute	
	Min	Max	Min	Max
Temperat. requir.	15	20	2	40
Rainfall (annual)	500	1000	200	2000
Latitude	30	-	55	70
Altitude	---	---	-	4400
Soil PH	6.5	7.5	6	8

Crop growth suitability data from FAO ECOCROP database

+

Candidate sourcing locations

Estimate the impact of climate change on food sourcing decisions

(Srinivasan et al., 2019)

UNIVERSITY OF CAMBRIDGE
Department of Engineering

IfM MANAGEMENT TECHNOLOGY POLICY

Photo credit to: https://developers.google.com/earth-engine/datasets/catalog/WORLDCLIM_V1_MONTHLY
https://gaez.fao.org/pages/ecocrop-find-plant

Findings (2) – Exploitation of Publicly Available Data in Current Research

Data Source & Collection → Data Pre-processing → Data Storage & Management → Data Analysis & Use

Data Source & Collection

Extracting data from
- Online database
- Reports
- Webpage
- Social media monitoring

Tools
- API (Nagendra et al., 2020)
- Software, e.g., LCA Software GaBi (Suer et al., 2021)

Selection Reasons
- Desired data update frequency (Barker et al., 2022)
- Comprehensiveness (Dooley et al., 2019; Ingrao et al., 2019 ; Geng et al., 2022)
- Free (Costa et al., 2022)

Data Pre-processing

Cleaning the Missing Data
(data points or entire category)
- Remove the entry with missing data points (Geng et al., 2022)
- Interpolation (Taghizadeh et al., 2021)
- Supplement with additional sources (Ostroski et al., 2022)

Data Storage & Management

Big satellite data stored in **cloud platform** and fetched through API for analysis (Nagendra et al., 2020)

Data Analysis & Use

- Regression analysis (Barker et al., 2022)
- Forecasting (Yu et al., 2019)
- Optimisation (Flores & Villalobos, 2020)
- Simulation (Taghizadeh et al., 2021)
- Life cycle assessment (Ingrao et al., 2019)
- Multi-criteria decision making (Kougkoulos et al., 2021)

Summary
- Only 5 papers mentioned the data source selection reasons
- Data quality and reliability, data validation are rarely considered
- Mainly focus on the stages of data processing & analysis
- Data storage and management processes are missing from the reviewed research
- No standard set of criteria for analysis techniques/models selection with specific purpose

UNIVERSITY OF CAMBRIDGE
Department of Engineering

IfM MANAGEMENT TECHNOLOGY POLICY

Findings (2) – Key Challenges of Exploiting Publicly Available Data

Data Source & Collection

- Data definition (precipitation vs rainfall)
- Source selection
- Data heterogeneity (formats, resolutions, units, frequency, size)
- Data inconsistency (e.g., measuring method)
- Data accuracy, quality &validity
- Data source monitoring (communication, different record/update frequency)

Data Pre-processing

- Missing data
- Data conversion & fusion
- Data triangulation
- Data translation

Data Storage & Management

- Database model selection (SQL vs NoSQL)
- Storing high-dimensional & heterogeneous data

Data Analysis & Use

- Data analysis technique/ model selection
- Results scalability
- Computational time of large dataset

UNIVERSITY OF CAMBRIDGE
Department of Engineering

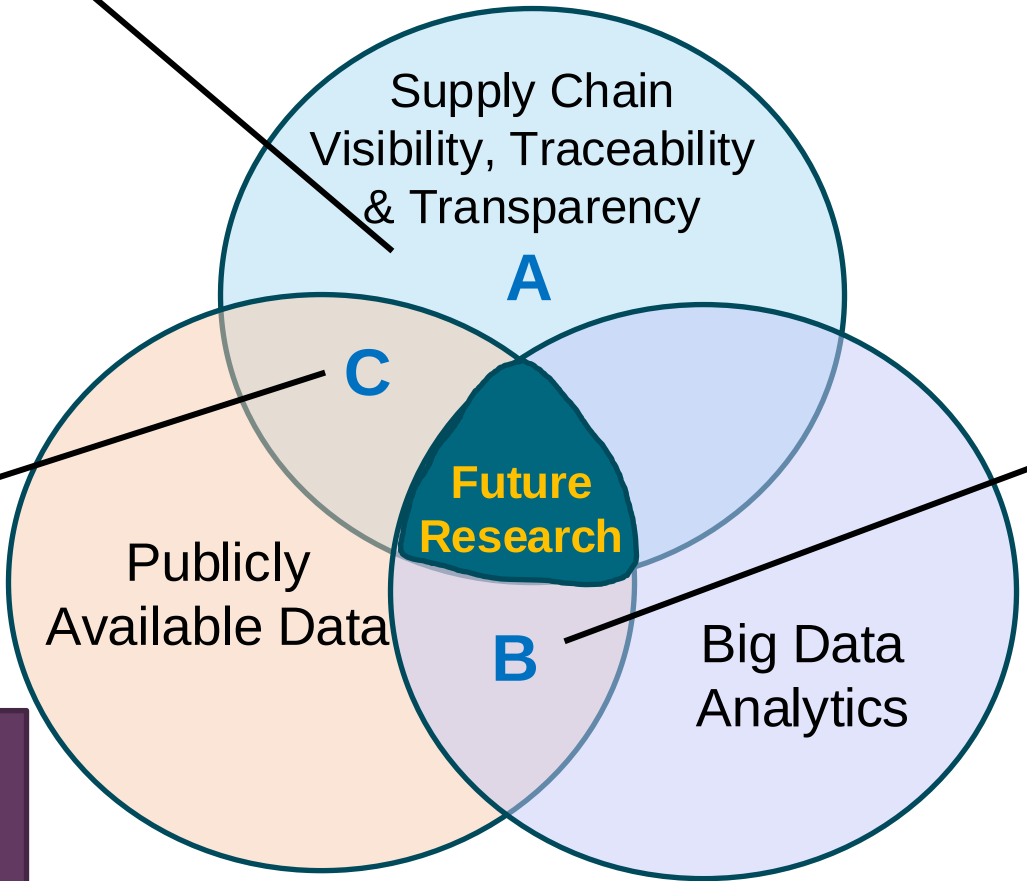
(Ingrao et al., 2019; van den Brink et al., 2020; Nikolopoulos et al., 2021; Geng et al., 2022; Nagendra et al., 2022; Ostroski et al., 2022)

IfM MANAGEMENT TECHNOLOGY POLICY

Discussions – Research gaps

Definitions & Requirements of SCVTT

- Differentiation between Visibility, Traceability & Transparency
- Agreed and standardised data requirements for SCVTT => **What data is needed?**



Implementation of big data analytics in research using publicly available data

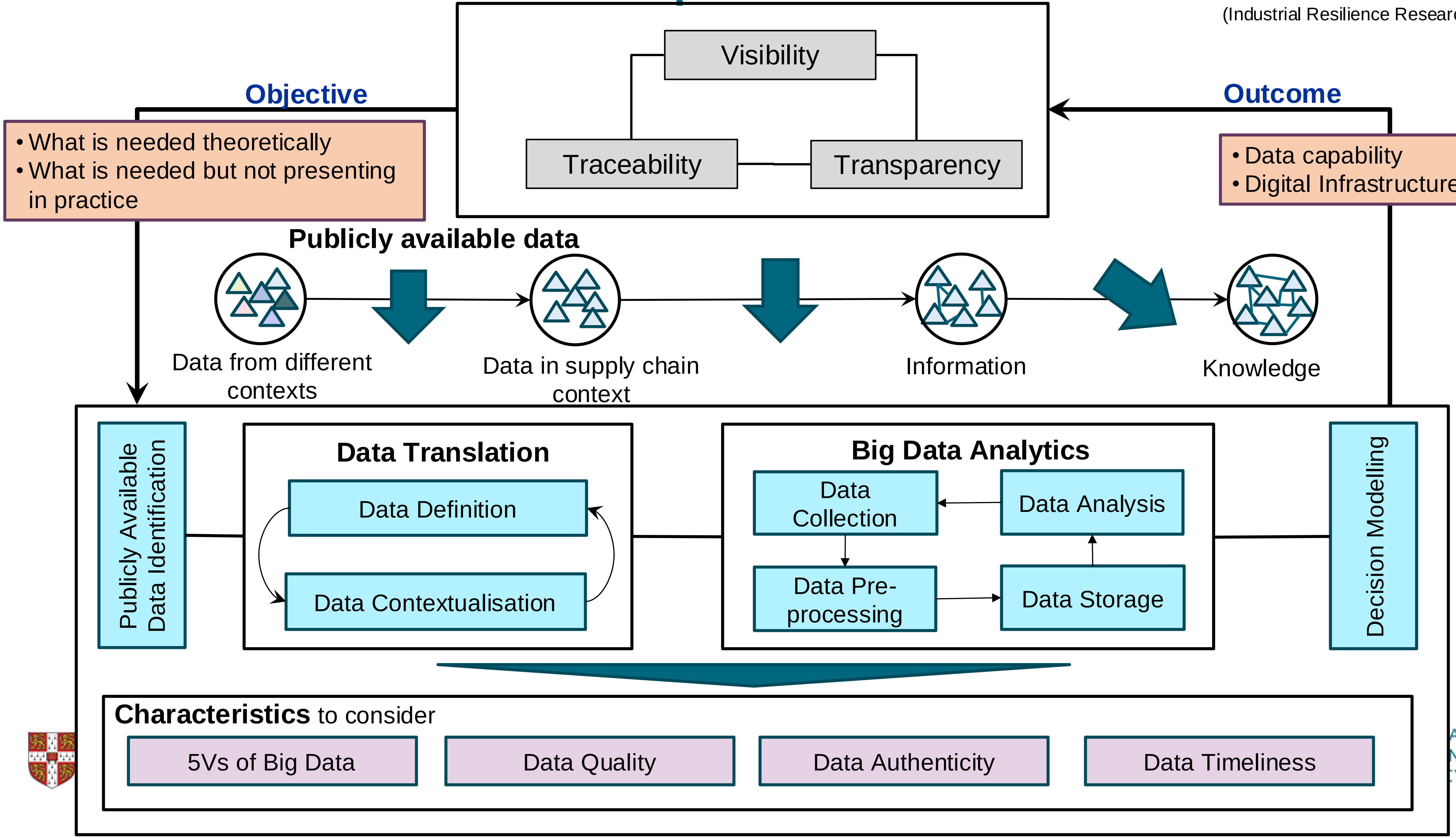
- Needs of **empirical case** which illustrates and guides the **implementation** of big data analytics (Constructs, Tools & Platforms)

Use of publicly available data in SCVTT research

- **Data definition** and **data source selection**
- More attentions should be paid on data storage and management for **sustained use**
- A standardised framework to **translate** the data from various contexts to supply chain context

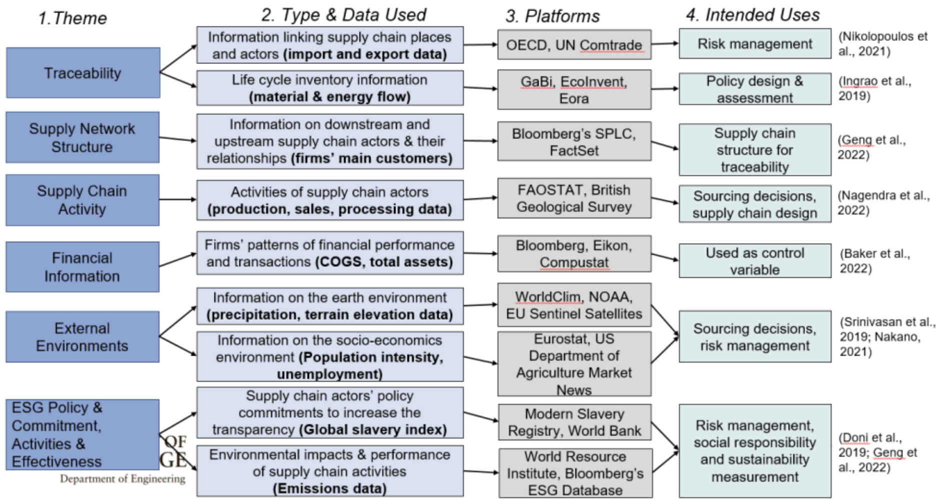
Discussions - Research Conceptual Framework

(Industrial Resilience Research Group, 2023)

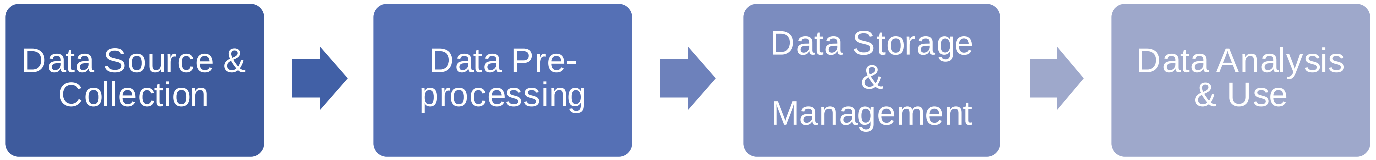


Conclusions

Objective 1: Identify existing publicly available data that can be linked to SCVTT



Objective 2: Understand the exploitation of publicly available data using BDA for the creation of SCVTT, and their challenges



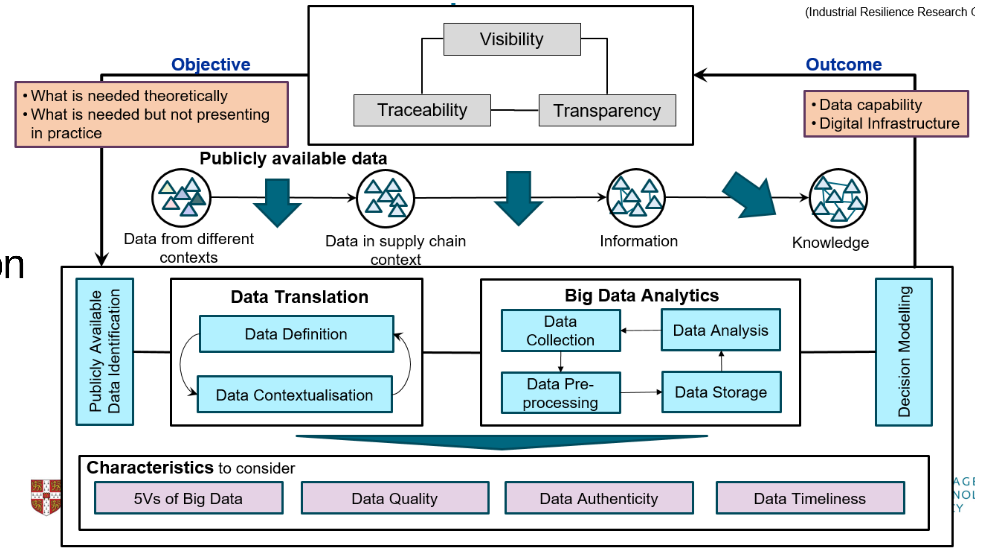
Methods, tools, techniques, challenges

Contributions

- Provide a holistic view of the use of publicly available data in current supply chain research
- Summarise themes related to SCVTT and corresponding data sources, **increase the data capability** of future research in the field of SCVTT & practical use
- Identify future research opportunities and propose a research framework

What's Next

- What data is needed to develop SCVTT
- Data definition & translation
- Digital infrastructure development



UNIVERSITY OF CAMBRIDGE
Department of Engineering

References

Andrews, E. L. (2022, April 6). *Detecting modern-day slavery from the Sky*. Stanford Human-Centered Artificial Intelligence. https://hai.stanford.edu/news/detecting-modern-day-slavery-sky

Barker, J. M., Hofer, C., Hoberg, K., & Eroglu, C. (2022). Supplier inventory leanness and financial performance. *Journal of Operations Management*, *68*(4), 385-407.

Barratt, M., & Barratt, R. (2011). Exploring internal and external supply chain linkages: Evidence from the field. *Journal of Operations Management*, *29*(5), 514-528.

Barratt, M., & Oke, A. (2007). Antecedents of supply chain visibility in retail supply chains: a resource-based theory perspective. *Journal of operations management*, *25*(6), 1217-1233.

Christopher, M., & Lee, H. (2004). Mitigating supply chain risk through improved confidence. *International Journal of Physical Distribution & Logistics Management*, *34*(5), 388–396.

Cooper, A.K., Coetzee, S. (2020). On the Ethics of Using Publicly-Available Data. In: Hattingh, M., Matthee, M., Smuts, H., Pappas, I., Dwivedi, Y.K., Mäntymäki, M. (eds) Responsible Design, Implementation and Use of Information and Communication Technology. I3E 2020. Lecture Notes in Computer Science, vol 12067. Springer, Cham.

Costa, D., Quinteiro, P., Pereira, V., & Dias, A. C. (2022). Social life cycle assessment based on input-output analysis of the Portuguese pulp and paper sector. *Journal of Cleaner Production*, *330*, 129851.

de Assis Santos, L. and Marques, L. (2022), "Big data analytics for supply chain risk management: research opportunities at process crossroads", *Business Process Management Journal*, Vol. 28 No. 4, pp. 1117-1145.

Dooley, K. J., Pathak, S. D., Kull, T. J., Wu, Z., Johnson, J., & Rabinovich, E. (2019). Process network modularity, commonality, and greenhouse gas emissions. *Journal of Operations Management*, *65*(2), 93-113.

Flores, H., & Villalobos, J. R. (2020). A stochastic planning framework for the discovery of complementary, agricultural systems. *European Journal of Operational Research*, *280*(2), 707-729.

Galbraith, J. (1973). Designing complex organizations. *Reading, Mass*.

Helbig, C., Gemechu, E. D., Pillain, B., Young, S. B., Thorenz, A., Tuma, A., & Sonnemann, G. (2016). Extending the geopolitical supply risk indicator: Application of life cycle sustainability assessment to the petrochemical supply chain of polyacrylonitrile-based carbon fibers. *Journal of Cleaner Production*, *137*, 1170-1178.

IEA (2021), The Role of Critical Minerals in Clean Energy Transitions, IEA, Paris https://www.iea.org/reports/the-role-of-critical-minerals-in-clean-energy-transitions, License: CC BY 4.0

Kaipia, R., & Hartiala, H. (2006). Information-sharing in supply chains: five proposals on how to proceed. *The International Journal of Logistics Management*.

Kim, K. K., Ryoo, S. Y., & Jung, M. D. (2011). Inter-organizational information systems visibility in buyer–supplier relationships: The case of telecommunication equipment component manufacturing industry. *Omega*, *39*(6), 667-676.

Kougkoulos, I., Cakir, M. S., Kunz, N., Boyd, D. S., Trautrims, A., Hatzinikolaou, K., & Gold, S. (2021). A multi-method approach to prioritize locations of labor exploitation for ground-based interventions. *Production and Operations Management*, *30*(12), 4396-4411.

McCrea, B. (2005). EMS completes the visibility picture. *Logistics Management*, *44*(6), 57-61.

Nagendra, N. P., Narayanamurthy, G., & Moser, R. (2020). Management of humanitarian relief operations using satellite big data analytics: The case of Kerala floods. *Annals of operations research*, 1-26.

References

Nikolopoulos, K., Punia, S., Schäfers, A., Tsinopoulos, C. and Vasilakis, C. (2021), “Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions”, *European Journal of Operational Research*, Vol. 290 No. 1, pp. 99–115.

Ostroski, A., Lagos, T., Prokopyev, O. A., & Khanna, V. (2022). Consumption-Based Accounting for Tracing Virtual Water Flows Associated with Beef Supply Chains in the United States. *Environmental Science & Technology*, *56*(22), 16347-16356.

Pero, M., & Rossi, T. (2014). RFID technology for increasing visibility in ETO supply chains: a case study. *Production planning & control*, *25*(11), 892-901.

RCS Global Group. (2022). *COBALT SUPPLY CHAIN MAPPING REPORT*.

Rogerson, M., & Parry, G. C. (2020). Blockchain: case studies in food supply chain visibility. *Supply Chain Management: An International Journal*, *25*(5), 601-614.

Sodhi, M.S. and Tang, C.S. (2019), Research Opportunities in Supply Chain Transparency. *Production and Operations Management*, 28: 2946-2959. https://doi.org/10.1111/poms.13115

Srinivasan, R., Giannikas, V., Kumar, M., Guyot, R., & McFarlane, D. (2019). Modelling food sourcing decisions under climate change: A data-driven approach. *Computers & Industrial Engineering*, *128*, 911-919.

Suer, J., Traverso, M., & Ahrenhold, F. (2021). Carbon footprint of scenarios towards climate-neutral steel according to ISO 14067. *Journal of Cleaner Production*, *318*, 128588.

Sunny, J., Undralla, N., & Madhusudanan Pillai, V. (2020). Supply Chain Transparency through Blockchain-Based Traceability: An Overview with Demonstration. *Computers & Industrial Engineering,* 106895. doi:10.1016/j.cie.2020.106895

Taghizadeh, E., Venkatachalam, S., & Chinnam, R. B. (2021). Impact of deep-tier visibility on effective resilience assessment of supply networks. *International Journal of Production Economics*, *241*, 108254.

UNCTAD. (2020). *Commodities at a Glance: Special Issue on Strategic Battery Raw Materials*. United Nations. https://www.un-ilibrary.org/content/books/9789210048293

Williams, B. D., Roh, J., Tokar, T., & Swink, M. (2013). Leveraging supply chain visibility for responsiveness: The moderating role of internal integration. *Journal of operations management*, *31*(7-8), 543-554.

Yu, L., Zhao, Y., Tang, L., & Yang, Z. (2019). Online big data-driven oil consumption forecasting with Google trends. *International Journal of Forecasting*, *35*(1), 213-223.

Wei Nie
Institute for Manufacturing
Department of Engineering
17 Charles Babbage Road
Cambridge, CB3 0DS
United Kingdom

T: +44 (0) 7425 864061
E: wn230@cam.ac.uk

Thank you!
Any questions or comments are welcome

UNIVERSITY OF CAMBRIDGE
Department of Engineering

IfM MANAGEMENT TECHNOLOGY POLICY