UC Berkeley School of Information

Summer 2024

ENERGY OWNERSHIP

Analyzing Global Energy Ownership Patterns

DATASCI-205

Anna Mowat, Jason Dong, Ryan Powers, Tara Verma

Source & Overview

- **Global Energy Monitor: Global Energy Ownership Tracker**
- Chain of ownership for energy projects
  - Lowest-level owner to highest-level parent
  - Percentage ownership
- **# of Entities: 12,520**
- **# of Energy Projects: 25,587**
- **# of ownership relationships: 41,583**

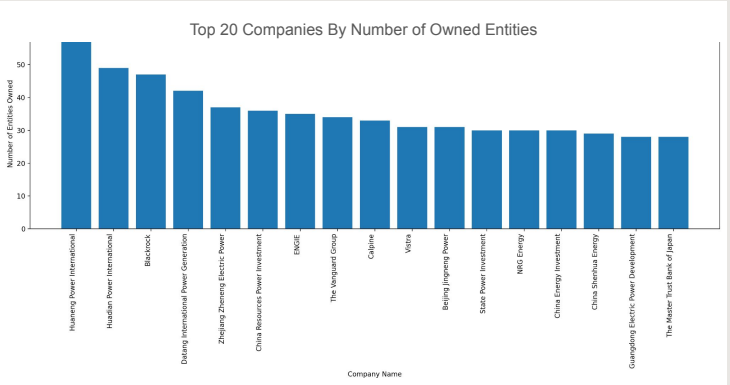
Global Energy Monitor

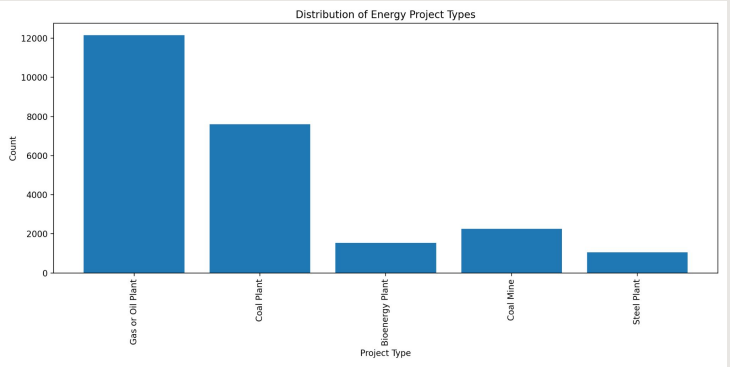
Key Definitions

- **Immediate Owner:** the entity (person, company, state, state body, etc.) that directly owns the project
- **Projects:** coal plants, oil plants, gas plants, steel plants, coal mines, and bioenergy plants
- **Parent Entities:** largest entities in the chain
- **Capacity:** maximum output the plant can produce (MW)
- **Emissions:** total greenhouse gases emitted by the plant

Cleaning & EDA

- **Cleaning:**
  - Dataset started as 8 different tabular files
  - Cleaned, reformatted, and combined files to load into a graph network
- **Feature Engineering:**
  - Function to calculate summed plant capacity for each entity (i.e. Blackrock owns X mega watts in coal plants)





Business Case

- Analytics from multiple points of view
  - **Project:** understanding the entire ownership structure for a specific energy plant
  - **Entity:** examining all of the projects owned in any part by a single entity
  - **Fuel type:** analytics on projects of a single type (coal, oil, etc.)
  - **Country:** understanding all of the entities and projects within a single country

Data Structure & Implementation

- **Flexibility with Dynamic Schema:** Stores diverse data types (JSON) for energy projects/entities without predefined schemas..
- **Scalability:** Handle large volumes of data through horizontal scaling across multiple servers.
- **Querying & Indexing:** Perform complex queries, filter, sort, and aggregate data for detailed analysis.
- **Geospatial Capabilities**: Analyze on location-based data to identify regional and opportunities.

Business Case

- **Real-Time Analytics**: Monitoring energy projects status, real-time changes, instant updates to stakeholders.
- **Session Management**: Tracking and maintaining energy project dashboards adding personalized experiences.
- **Caching Layer**: Reducing database load and improving performance, access to ownership data, and project details.
- **Multi-Database Integration**: Storing real-time updates while maintaining Noe4j complex ownership structure.

Data Structure & Implementation

- **Strings:** basic information about entities and energy projects
- **Hashes:** field-value pairs for entities and projects
- **Lists:** ownership relationships (such as the sequence of owners in a chain)
- **Sets:** unique members of all entities or projects
- **Sorted sets:** percentage of ownership or production capacity
- **Stream:** updates to ownership information and project statuses

Business Case : Challenges

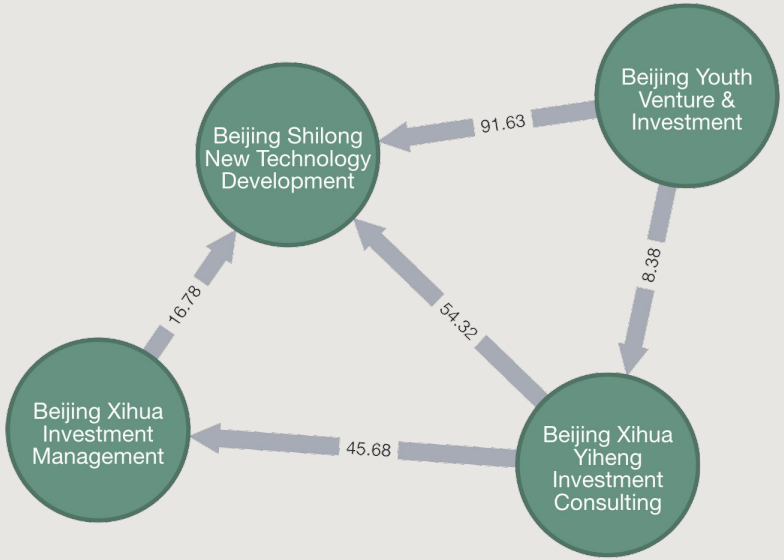
- **Complex Ownership Structures**
  - Energy projects often have intricate ownership chains involving multiple entities such as corporations, investment firms, and government bodies.
- **Hidden Influencers**
  - Identifying the majority stakeholders and decision-makers within these ownership chains may be difficult.
- **Risk Assessment**
  - Assessing the risks in any business opportunity is crucial for making informed investment decisions. Misinformed and biased data associated with indirect ownership present potential liabilities to future earnings.

Business Case : Opportunities

- **For corporations**
  - Make well-informed investment decisions if attempting to understand the complex ownership structures of potential projects, identify key players, and assess risks associated with ownership chains
    - Useful when building an ESG fund, or identifying their own climate-related risk exposure
- **For Journalist, NGOs, and shareholders**
  - Use algorithm to identify which companies are underreporting their carbon emissions
  - Identify areas where current carbon accounting frameworks have gaps

Data Structure & Implementation

- **Nodes:** entities and energy projects
  - Production capacity
  - Fuel type
  - Status
  - Emissions
  - Country of origin
- **Edges:** ownership relationships
  - Share percentage
- **Queries:** Cypher



Graph Algorithms: Centrality

Harmonic

Measures closeness of entities within the network

Entity	Closeness
Blackrock	0.017
The Vanguard Group	0.016
Blackrock Advisors	0.012
The Master Trust Bank of Japan	0.011
Custody Bank of Japan	0.001

Betweenness

Identifies key intermediaries in ownership chains

Entity	Betweenness
The Master Trust Bank of Japan	1525
Custody Bank of Japan	1174
Chongqing Energy Investment Group	809
Blackrock	655
Mitsui & Co	484

Graph Algorithms: PageRank

Unweighted

Measures overall entity influence in the network

Entity	PageRank
Huaneng Power International	2.98
Huadian Power International	2.54
Calpine	2.49
Blackrock	2.32
China Energy Investment	2.27

Weighted

Captures influence with shares as weights

Entity	PageRank
Huaneng Power International	3.20
Huadian Power International	2.73
Calpine	2.49
Datang International Power Generation	2.33
China Energy Investment	2.31

Graph Algorithms: PageRank

Personalized: **ENGIE**

Captures importance and influence to a single entity

Entity	PageRank
ENGIE	0.15
Government of France	0.06
The Capital Group Companies	0.06
(Hong Kong) Dewei Industrial Investment	0.00
100 Thuwanon	0.00

Data Visualization

- Steel Plants
- Blackrock

Entire Graph

Entity	Betweenness
The Master Trust Bank of Japan	1525
Custody Bank of Japan	1174
Chongqing Energy Investment Group	809
Blackrock	655
Mitsui & Co	484

Steel Plant Subgraph

Entity	Betweenness
The Master Trust Bank of Japan	516
Custody Bank of Japan	424
Nippon Steel	334
Blackrock	246
Tianjin Jianlong Iron & Steel Industrial	228

Database Evaluation

Database	Strengths	Weaknesses
Neo4j	• Dynamic queries for networks • SQL-like language • Data integrity and consistency	• Limited base visualization support • Must load directional relationships
MongoDB	• Different points of view • Scalability	• Limited network and graph statistic capabilities • New query language
Redis	• Deliver rapid real time results	• Lack advanced query capabilities

Insights

- Graph algorithms and subgraphs
  - Centrality
  - Key Intermediaries
  - Influence

Decision-Making

- Ownership chain analytics inform:
  - Resource allocation
  - Risk
  - Legal and regulatory action

Questions?

https://github.com/mids-w205/project-3-energy-ownership

Anna Mowat, Jason Dong, Ryan Powers, Tara Verma