

Photon



A data driven energy market opportunities map

LAPD 2020/21 - Prof. Liliana Ferreira

Eduardo Ribeiro

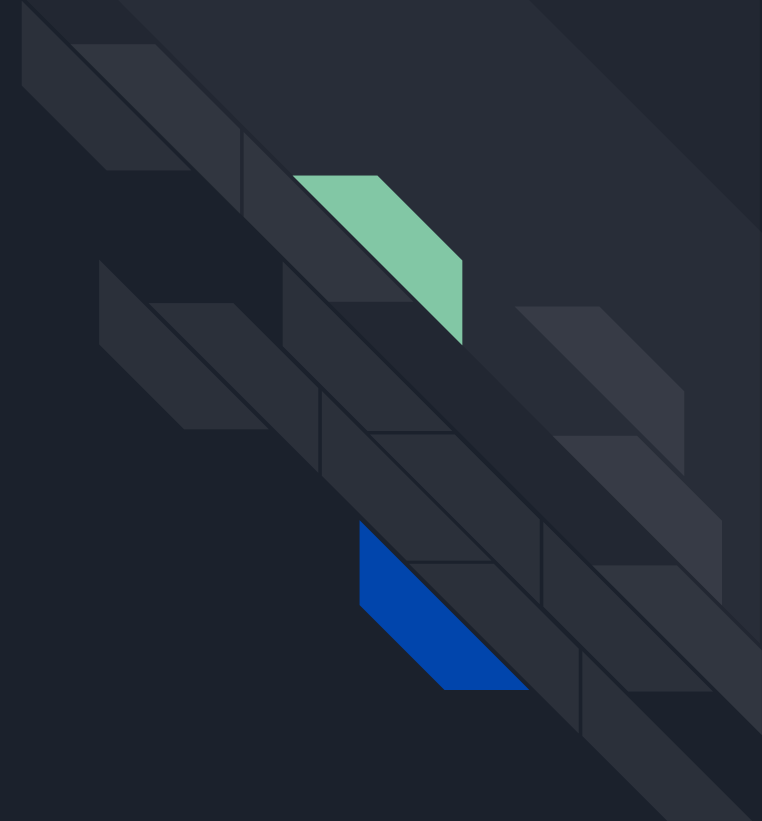
Martim Silva

Miguel Pinto

Nuno Cardoso

Table of Contents

1. Understanding the problem
2. Solution: Photon
3. Steps and Goals
4. Data Sources and Data Models
5. Tech Stack
6. Requirements
7. Prototype
8. Architecture
9. Development Roadmap





Understanding the problem

- 01 Every day, we learn about emerging technologies and developments that have the **potential to be groundbreaking**. But how do we **detect** the early proof-of-concept, non-obvious opportunities with real growth potential?
- 02 This leads us to the broad topic of **Energy**. It's one of the biggest drivers for global issues like climate change, and when starting new projects and companies, it's important to **make sure the problem is relevant**.
- 03 The process of determining whether a problem is promising is not trivial. How can we **gather** and **process** the vast amount of data revolving around energy to detect the most promising, emerging and non-obvious problems that need to be solved?



Solution: Photon

Photon is an application capable of:

- ❄ Extracting and analyzing energy related data from various sources
- ❄ Detecting and identifying real, high growth opportunities within the energy market and industry
- ❄ Showcasing that information to the user in an easy to use graph-based visual interface



Steps and Goals

Data Extraction

Development of modules that communicate with external APIs and sources in order to extract relevant, energy-related data.

Knowledge Graph

Parsing and interpretation of the collected data, in order to build knowledge graphs containing energy-related information.



ML Based Prediction

ML based screening and prediction of opportunity growth potential (can be done possibly with NLP and proximity analysis)..

Visual Interface

Creation of a visual interface, that will present the knowledge graph to the user, to have a sound source of insight into the "problem area" of Energy.

Data Sources

Social Media APIs

- Reddit API
- Twitter API

Social Media APIs can tap into and extract **public conversations** as a way to understand what's **trending**, **discover insights** and listen to events.

With these APIs, it's possible to gather different metrics data and search for specific topics using **keywords** to analyze related conversations and get popular searches in the platforms.

These 2 social media platforms are specially relevant given their abundance of cutting edge discourse.





Data Sources

News APIs

→ Usearch API

News APIs will have a great importance in data collection. This API aggregates **news from multiple worldwide sources** and utilizes different features to distinguish them.

By using it, it will be possible to retrieve news by keywords, phrases, countries and publishers.

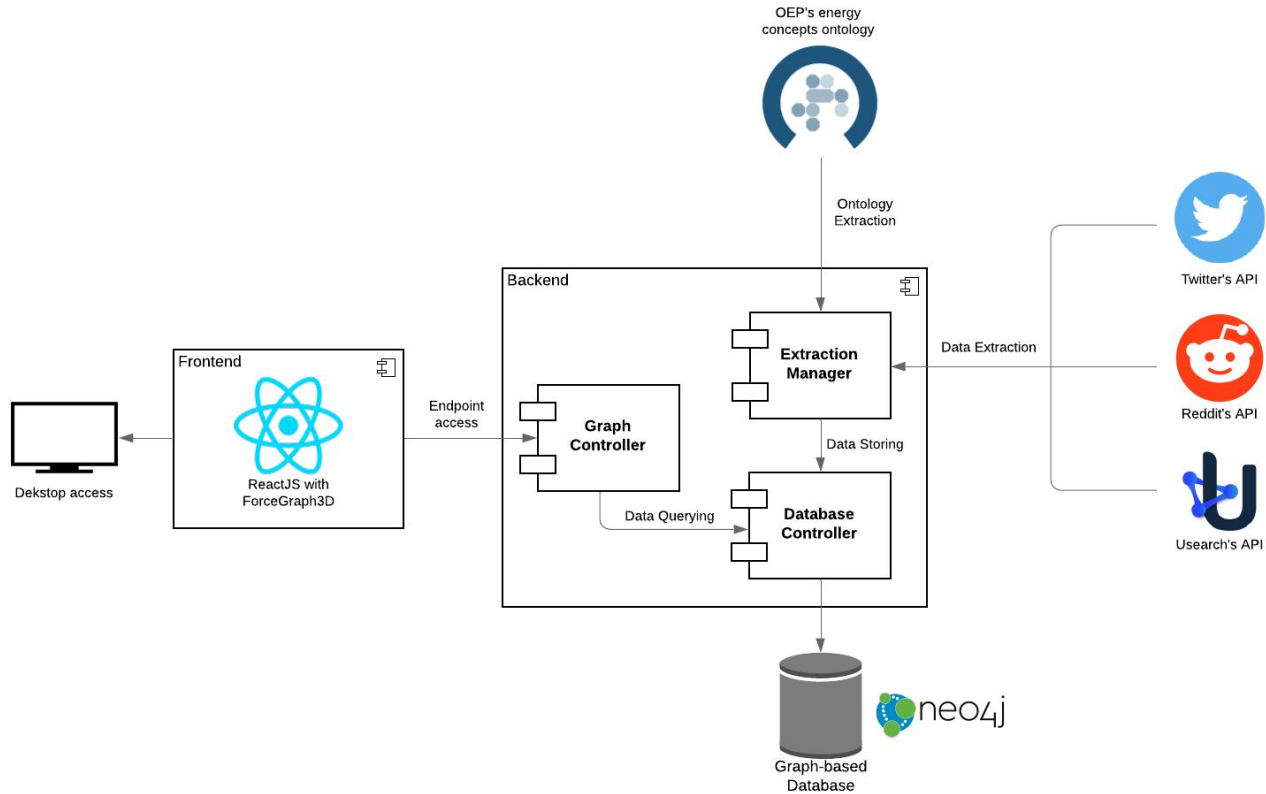




Requirements

ID	Description	Priority	MVP	Done
R01	Visualize and navigate through the graph, to be aware of the different opportunities.	High	x	x
R02	Have information about a specific graph node (energy sub-area), to allow a further analysis for a specific opportunity or area.	High	x	x
R03	See all of the attributes of the nodes and their relationships, so that it's possible to understand the connections and what makes a certain sub-area promising or not.	High	x	x
R04	Have filtering options based on the number of articles and posts, to see only relevant nodes.	Medium		x
R05	Have information about the specific articles/posts or have some links to pertinent pages, so that it's possible to analyse, in first-hand, the posts and articles that were extracted for a certain sub-area or topic.	Medium		x
R06	See distinct node sizes, proportional to a sector's disruptiveness, so that I can understand where the best opportunities lie.	Low		

Architecture

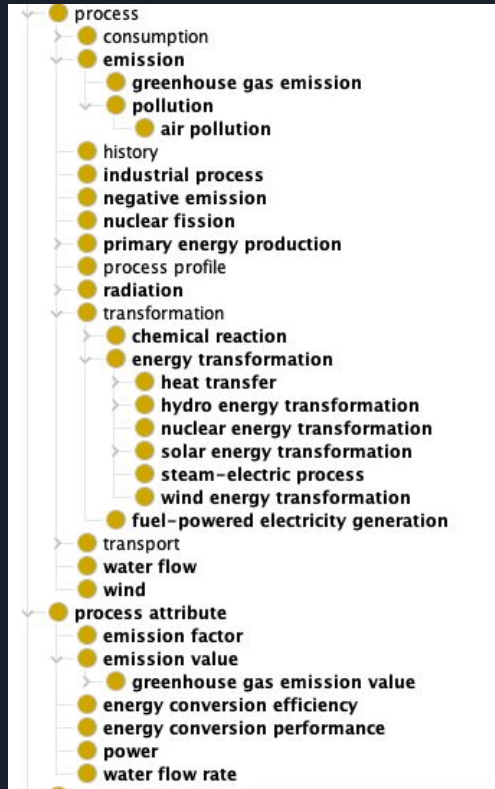




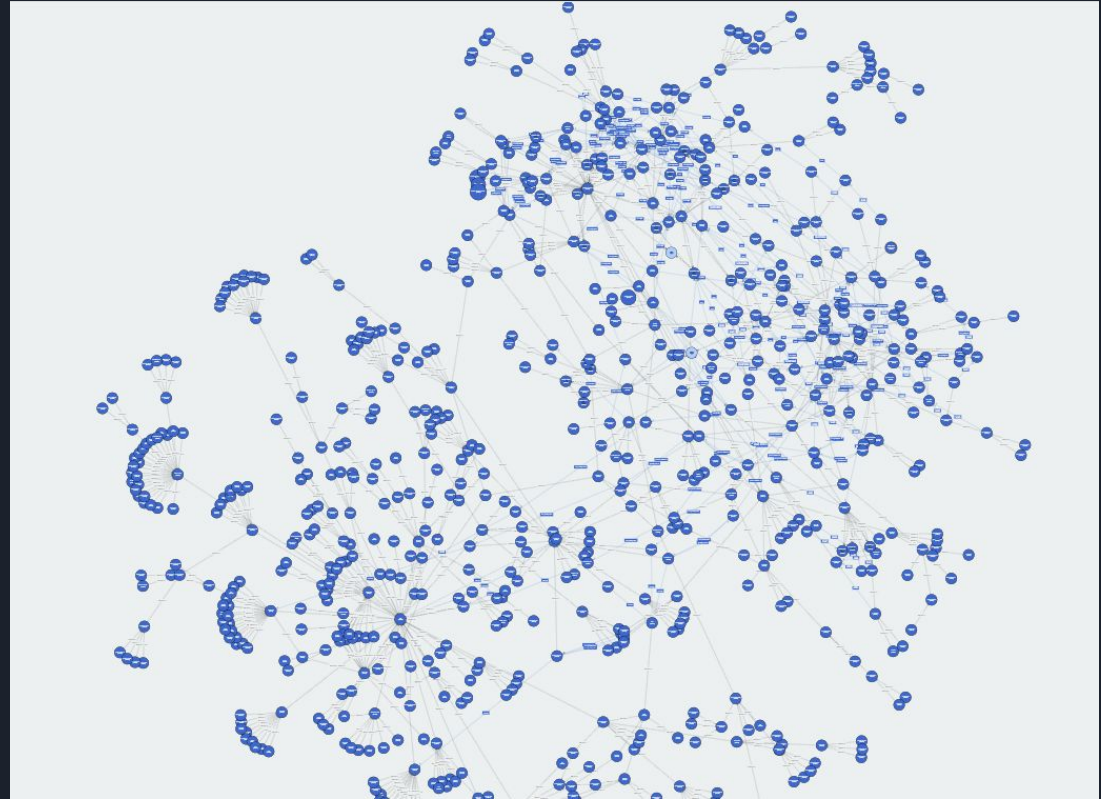
Open Energy Ontology (OEO)

- ❁ An ontology is a **formal collection of terms** and relationships between those terms, with clear definitions and a logical interpretation.
- ❁ Photon makes full use of the Open Energy Ontology. It serves as the base of the **knowledge graph** and the source of the data to query the APIs.

Open Energy Ontology



Ontology viewed in Protégé



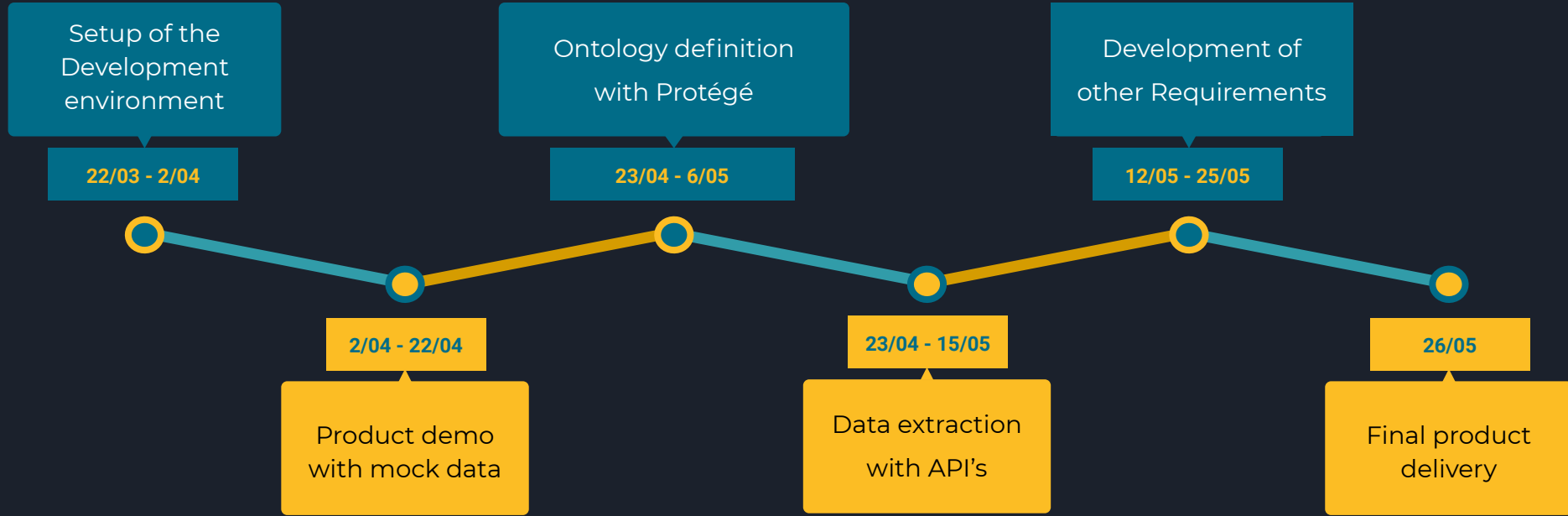
Ontology viewed in WebVOWL



Data Flow - Extraction and Knowledge Graph

- 01 Loading the database with the energy concepts from the OEO
- 02 Query the APIs using keywords related to the energy concepts
- 03 Create nodes for every tweet, reddit post and news article and link them to the corresponding energy concepts, making sure not to repeat nodes
- 04 Calculate and store node attributes for the energy concepts, allowing filtering options (eg. only displaying energy nodes with a high number of articles)

Development Roadmap





Future Work

Some features that could be added to Photon, that would improve its usability and the insight provided to the users, are:

- ❁ Changing node sizes proportionally to a sector's popularity;
- ❁ Implementing a search functionality that allows users to search for a specific sector.



Future Work

A solid base is provided to Fraunhofer in order to implement the last step: **ML Based Prediction**. Some possible tasks consist in:

- ❁ **Dynamically expanding the ontology** by detecting new energy domains/concepts, through the use of ML and NLP.
- ❁ **Improving metrics** to define the size of the nodes (that symbolizes **disruptiveness** and **potencial**), making use of sentiment analysis.
- ❁ **Automatically generate a report** with the top most disruptive energy topics, to view results quickly.



Photon



Thank you!
Questions?