



the LORD OF THE RINGS
the FELLOWSHIP
OF THE RING

Team OpenEdu B



Ceylan
Scrum Master



Mascha
Backend



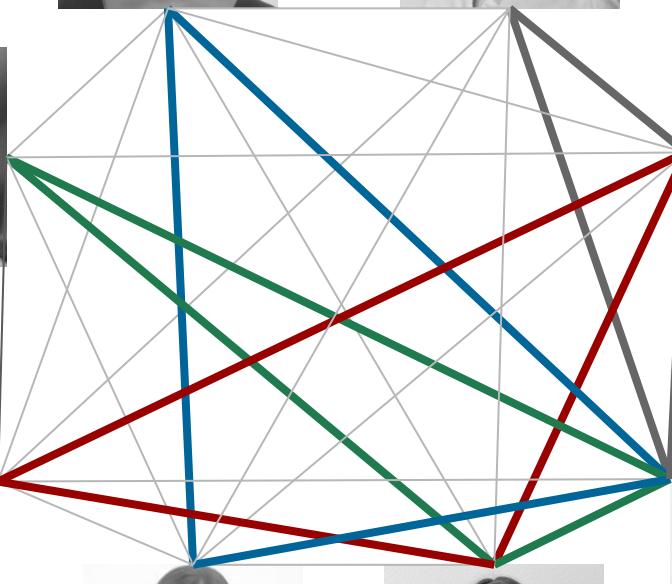
Elena
UX/UI Design



Michal
Data Scientist
(Tech Lead)



Noëline
Data Scientist
(NLP)



Andrina
Product Owner



Maria
Team Satellite

Sofia
Data Scientist

How can we bring the Wiki Universe into every classroom & learning space?



How can we bring the Wiki Universe into every classroom & learning space?



OpenEdu

Where educators find out how to bring the Wiki universe into the classroom

User Experience

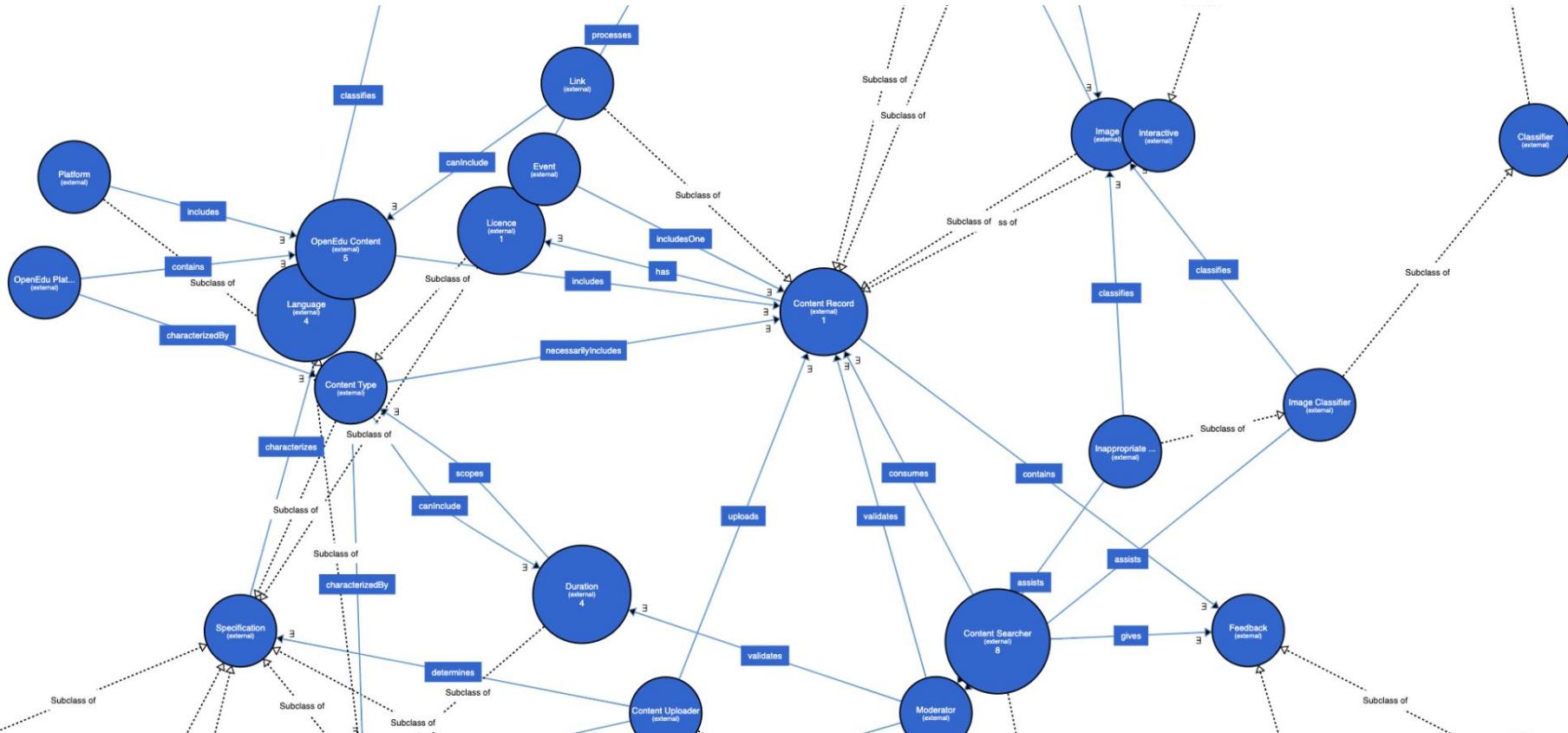
Ontology

**Automation
(NLP)**

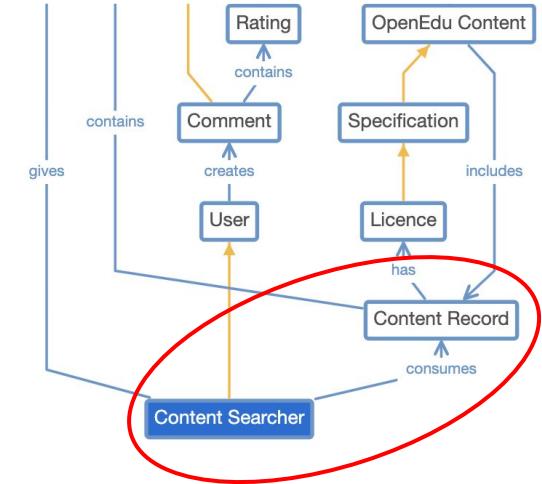
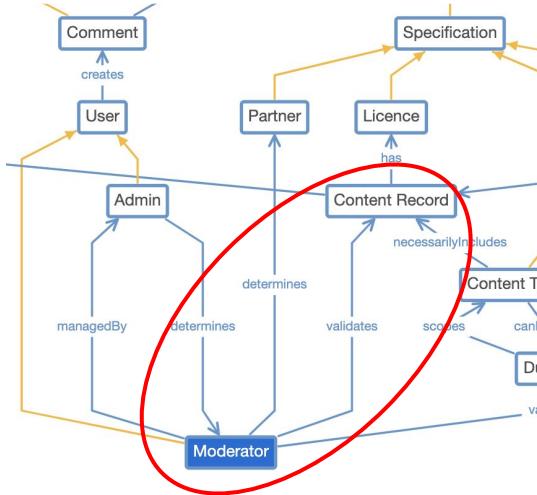
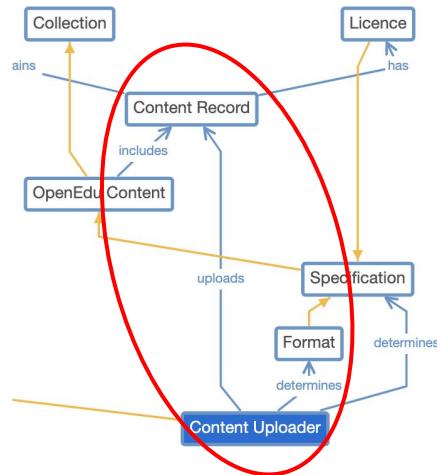
Crawlers & Scapers

Architecture

Ontology



Ontology: Users



Content Uploader
Creates Content



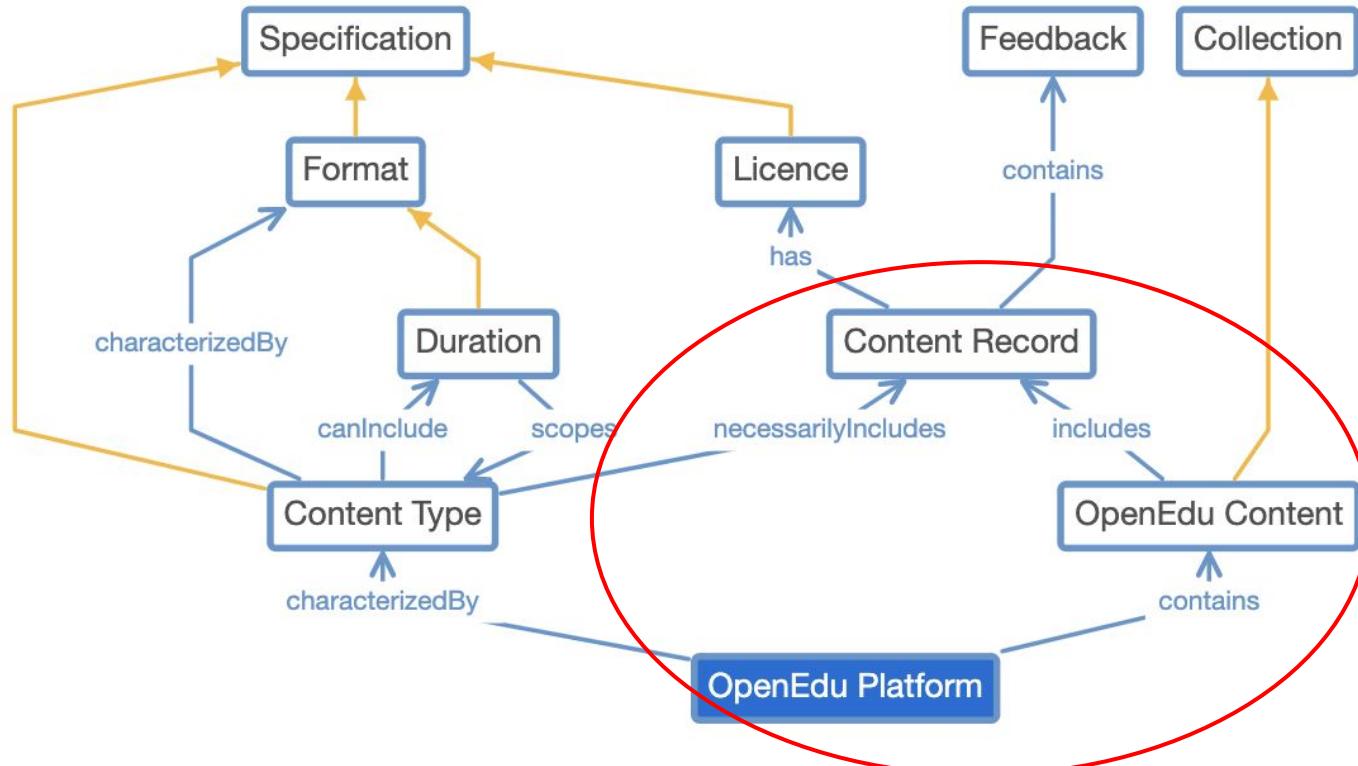
Moderator
Validates Content



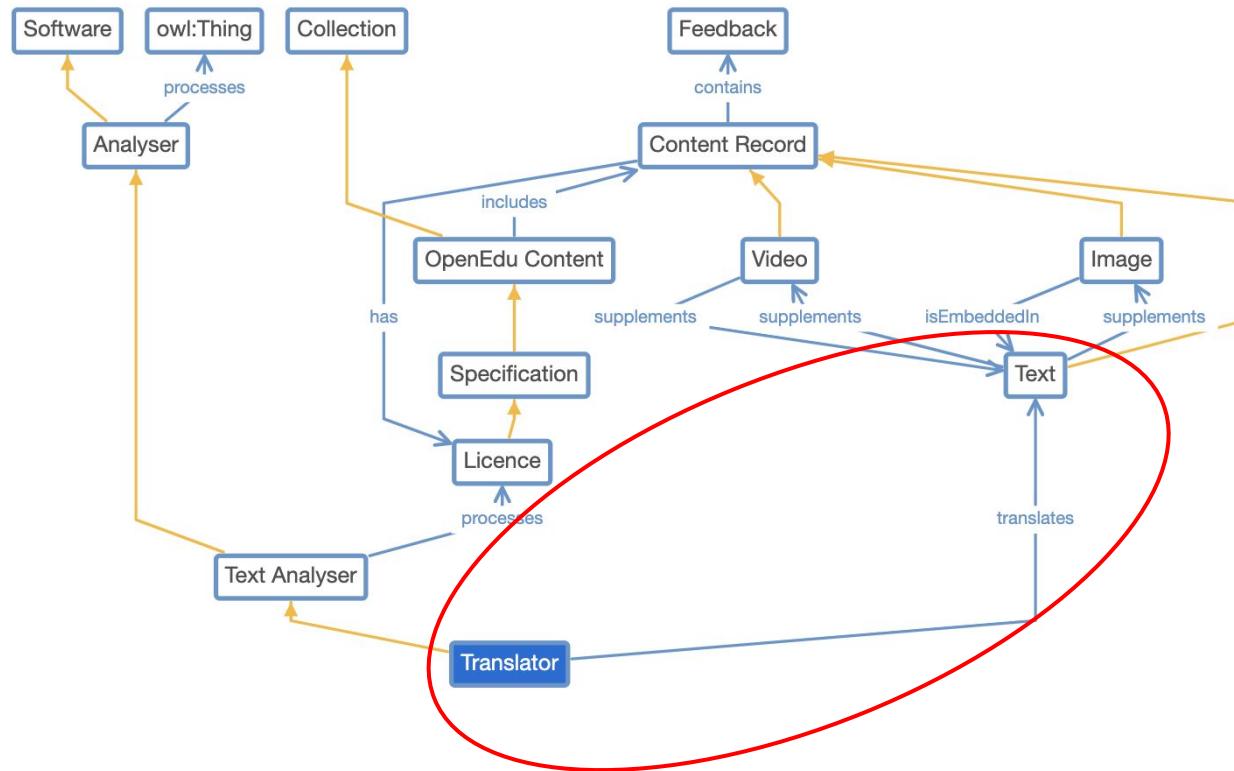
End User
Consumes Content



Ontology: Overview



Ontology: Data Science



Natural Language Processing (NLP)

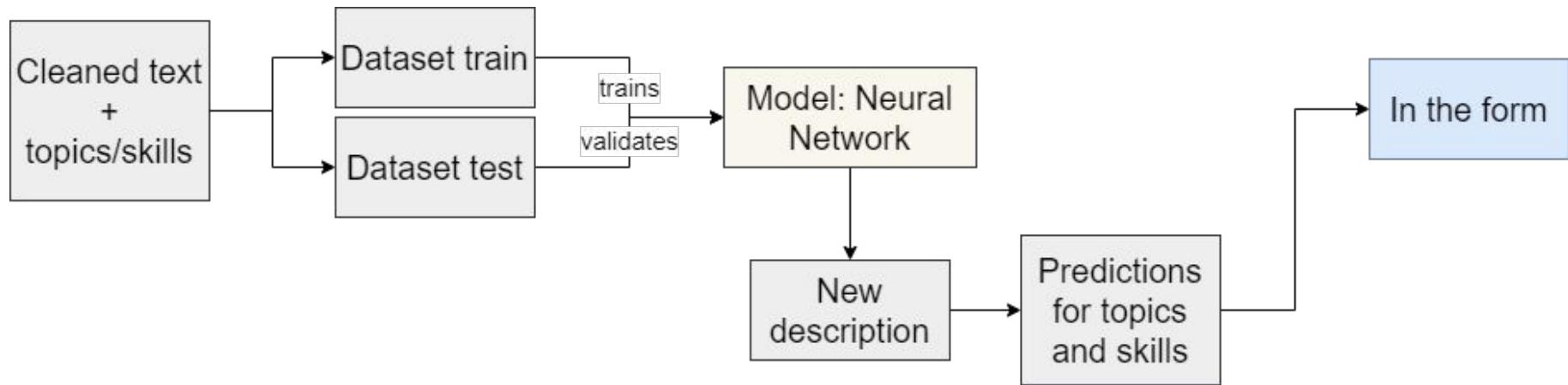
The Content uploader writes the title and a description.



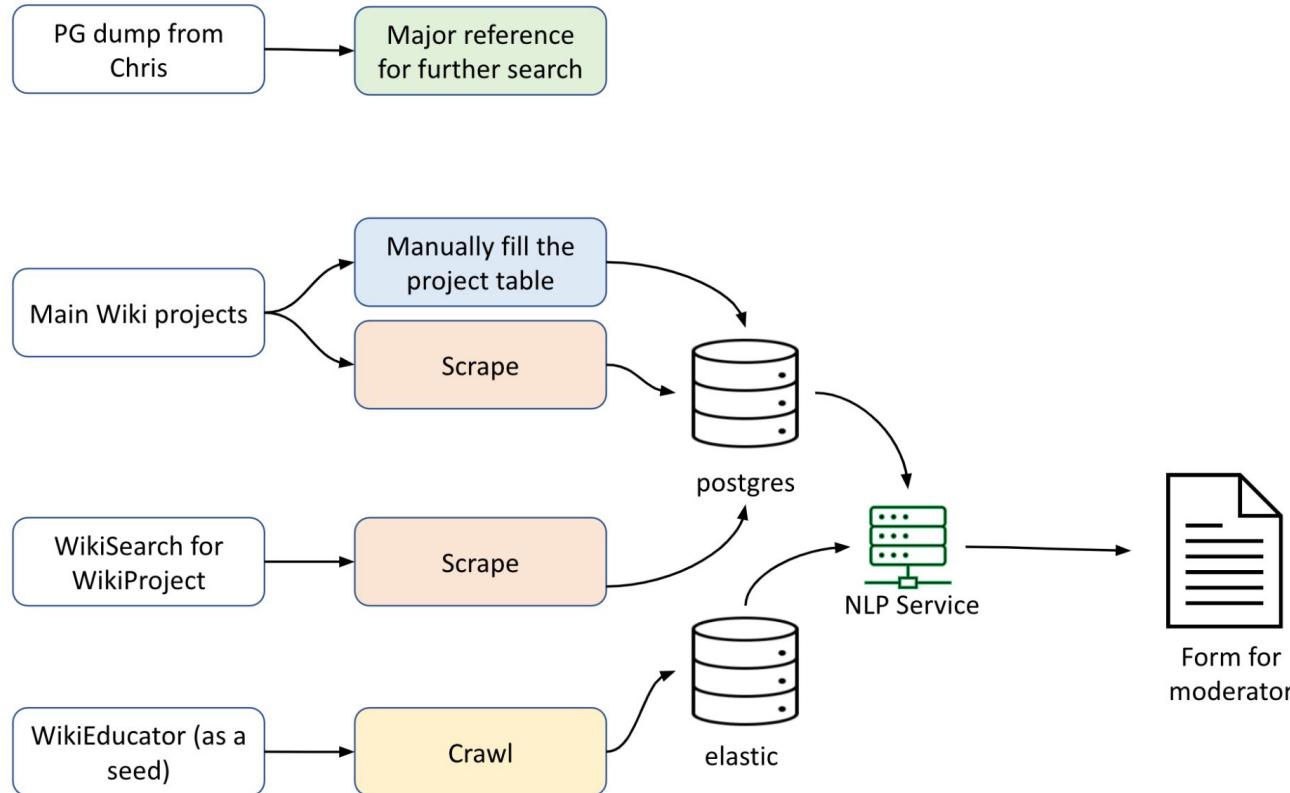
Title	language	text_en	text_fr	text_de	text_it	flesch	similar1	similar2	similar3
Commons	en	Wikimedia Commons, or simply Commons, is an on...	Wikimedia Commons, ou simplement Commons, est ...	Wikimedia Commons oder einfach Commons ist ein...	Wikimedia Commons, o semplicemente Commons, è ...	[college_graduate]	Evaluation Dashboard	WIKIDDHH	Wiki Loves Monuments
Wikipedia	en	The free encyclopedia that anyone can edit. Wi...	L'encyclopédie gratuite que n'importe qui peut...	Die kostenlose Enzyklopädie, die jeder bearbei...	L'encyclopédia gratuita che chiunque può modif...	[college_graduate]	WIKIDDHH	Wikipedia in Your University	Wiki Loves Monuments
Unesco OER Recommendation 2019	en	On 25th of November 2019, the 40th General Con...	Le 25 novembre 2019, la 40e conférence général...	Am 25. November 2019 verabschiedete die 40. Ge...	Il 25 novembre 2019, la 40a Conferenza general...	[college_graduate]	Wikiversity	Open Science Fellows Program	WikiEducator
Language		Translations			Readability Level		Similar contents		

Machine Learning: “Advanced NLP”

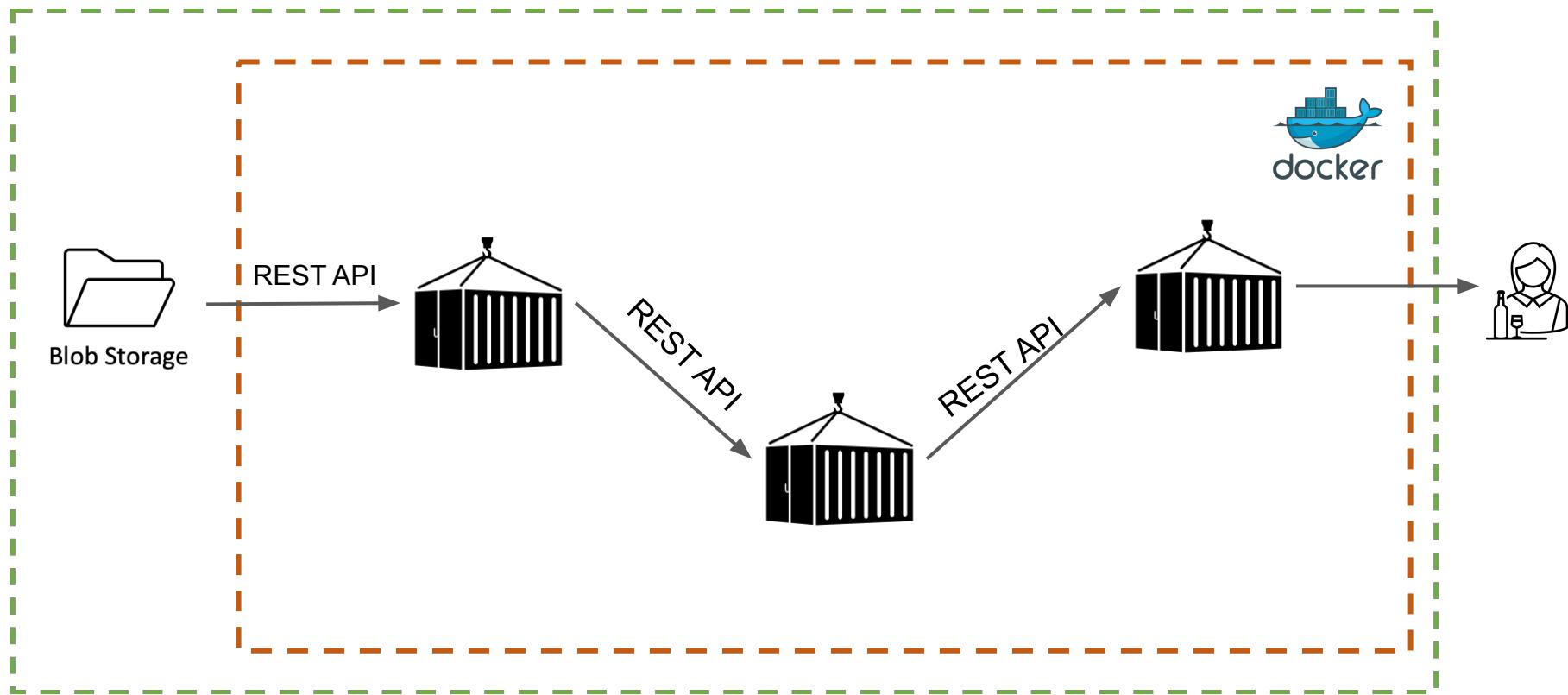
Future automation : Topics and Skills



The Flow of Data

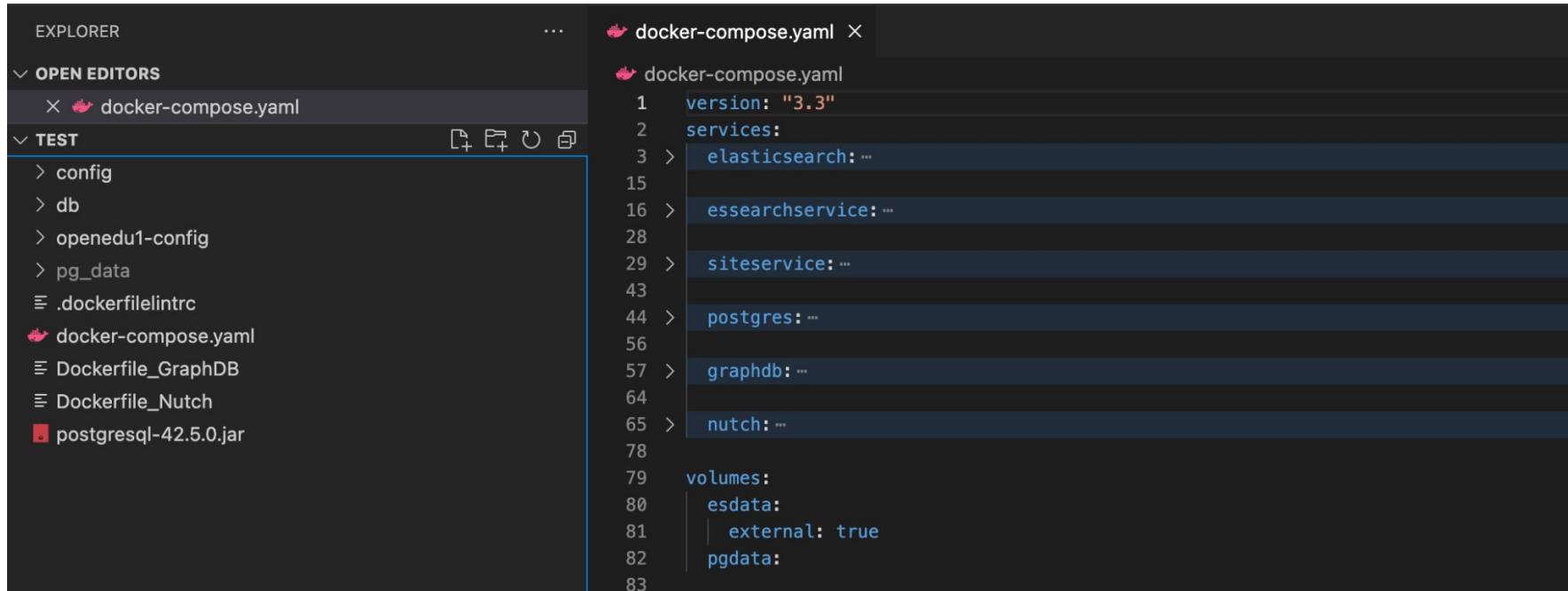


Architecture: The Big Picture



Architecture: The Containers

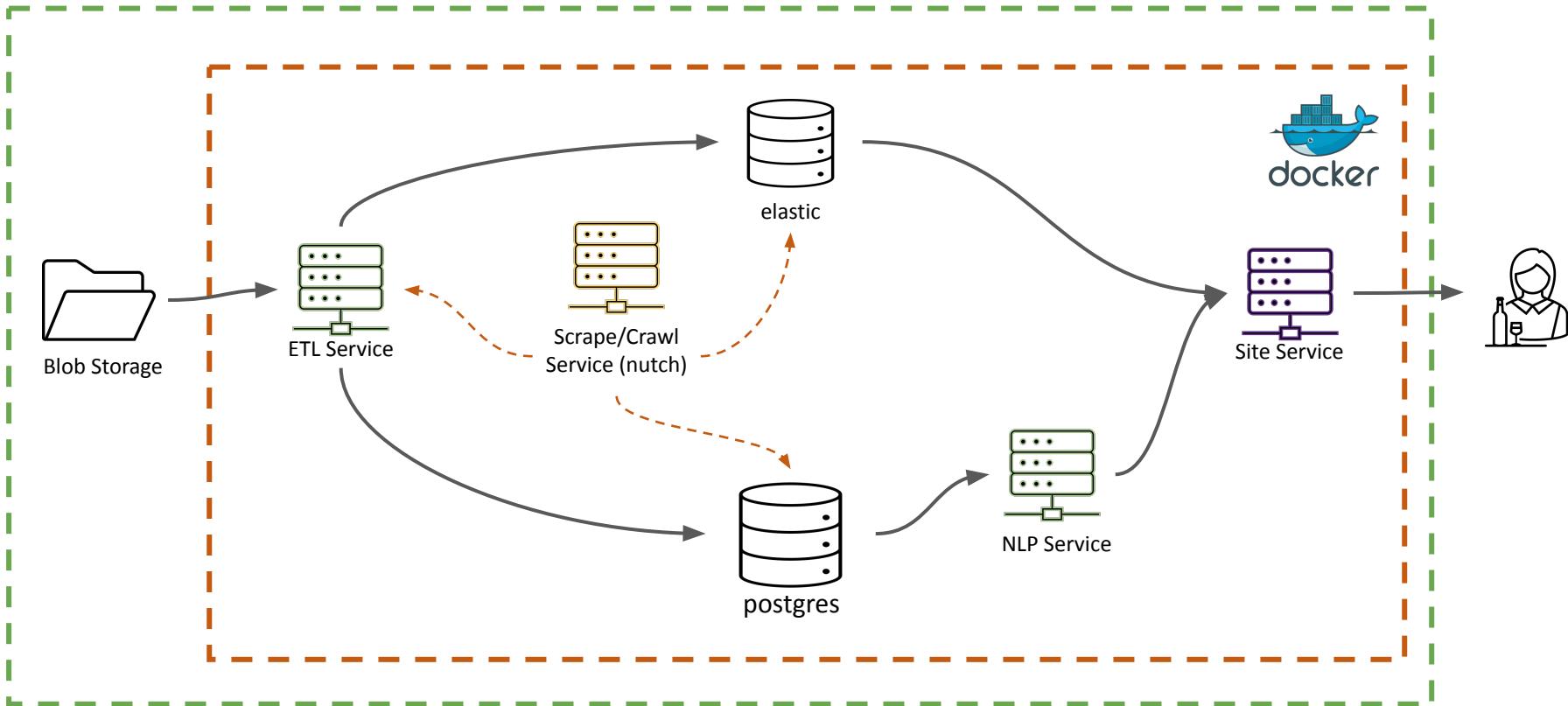
Microservice (MS) architecture packed into docker-compose



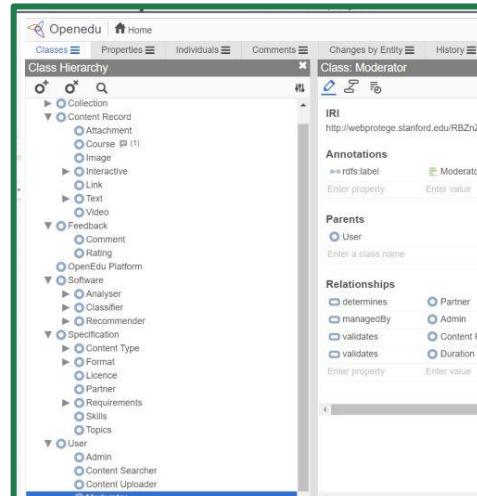
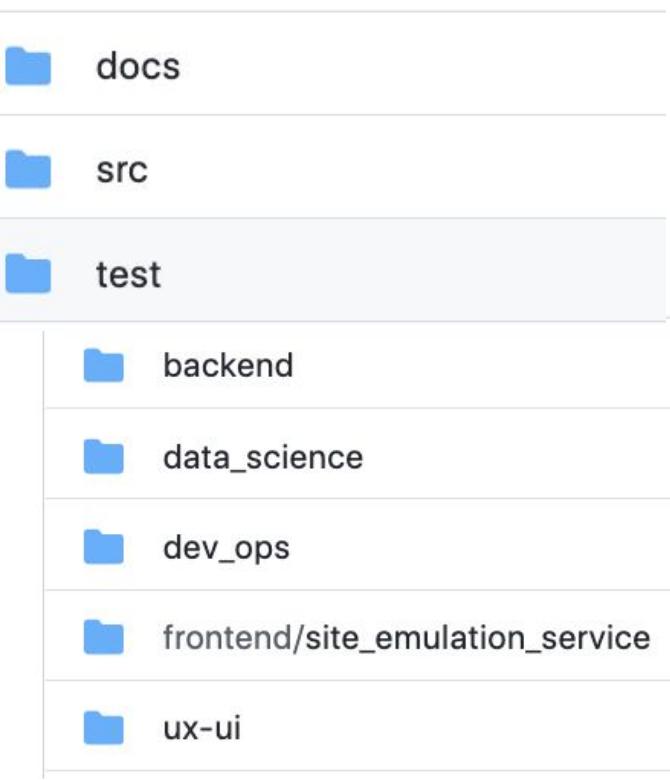
The screenshot shows a code editor interface with a dark theme. On the left is the Explorer sidebar, which lists several files and folders under the 'OPEN EDITORS' section, including 'docker-compose.yaml'. Below this is the 'TEST' section containing 'config', 'db', 'openedu1-config', 'pg_data', '.dockerfilelintrc', and 'postgresql-42.5.0.jar'. The main area is a code editor showing a portion of a 'docker-compose.yaml' file. The file defines a stack with multiple services: elasticsearch, essearchservice, siteservice, postgres, graphdb, and nutch. It also specifies volumes named esdata and pgdata, with esdata being external. The code is syntax-highlighted with blue for keywords like 'version' and 'services'.

```
version: "3.3"
services:
  elasticsearch: ...
  essearchservice: ...
  siteservice: ...
  postgres: ...
  graphdb: ...
  nutch: ...
volumes:
  esdata:
    external: true
  pgdata:
```

Architecture: Container Connections

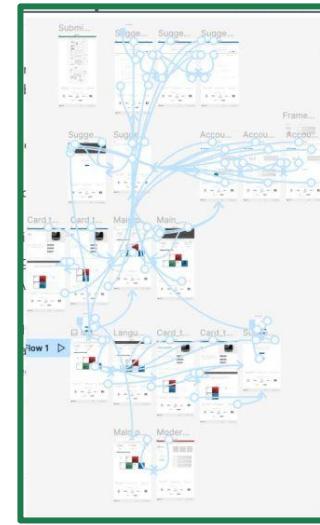


Mission accomplished! - Documentation



```
172 lines (130 sloc) 6.51 KB

1 #!pip install LangDetect
2 #!pip install googletrans
3 #!pip install py-readability-metrics
4
5 import re
6 import numpy as np
7
8 import pandas as pd
9 import nltk
10
11 from langdetect import detect
12 from googletrans import Translator
13 from readability import Readability
14 from nltk.corpus import stopwords
15 from nltk.tokenize import word_tokenize
16 from nltk.stem import WordNetLemmatizer
17 from sklearn.feature_extraction.text import TfidfVectorizer
18
19 nltk.download('stopwords')
20
21
22 # # Plan :
23 # * [1. Dataset]
24 # * [2. Language recognizer]
25 # * [3. Translation]
26 # * [4. Readability]
27 # * [5. Clean the text]
28 # * [6. Determine the topic/skills]
29 # * [7. Search similar content]
```



Bonus Suggestions

Using More Data
for Training

Detect Topics &
Skills Automatically

Better
Recommender
System

Rating System

User Dashboards

Our Learnings as a Fellowship



Questions?

Figma Prototype:



Demo Website:





Ceylan
Scrum Master



Mascha
Backend



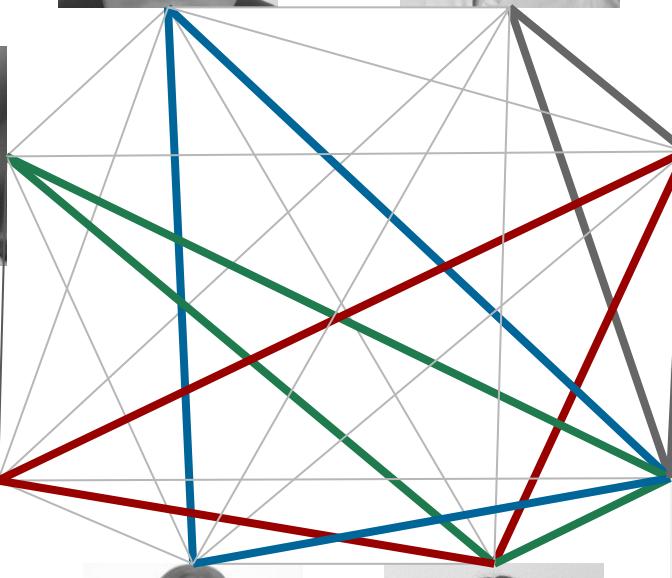
Elena
UX/UI Design



Michal
Data Scientist
(Tech Lead)



Noëline
Data Scientist
(NLP)



Andrina
Product Owner



Maria
Team Satellite

Sofia
Data Scientist