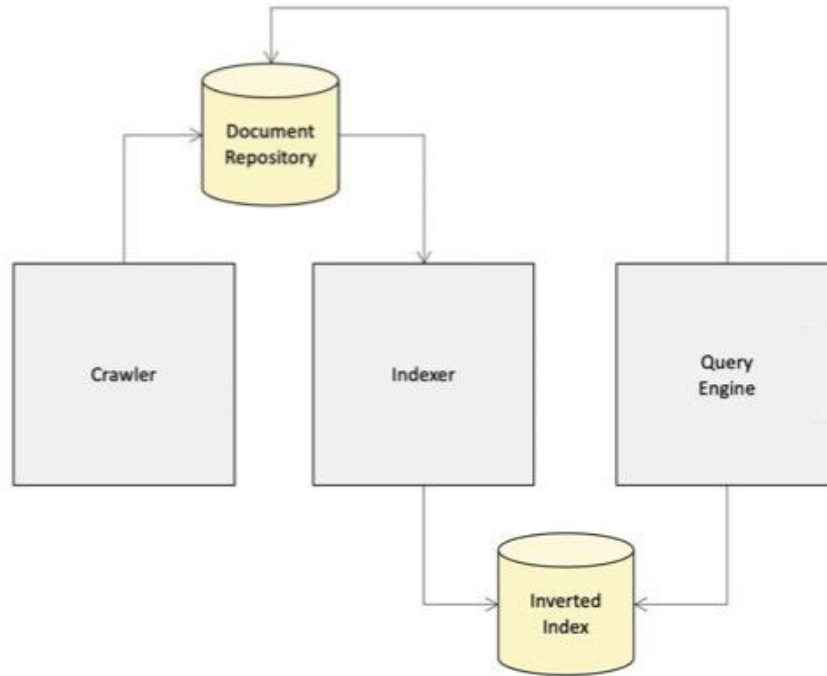


Search Engine 1

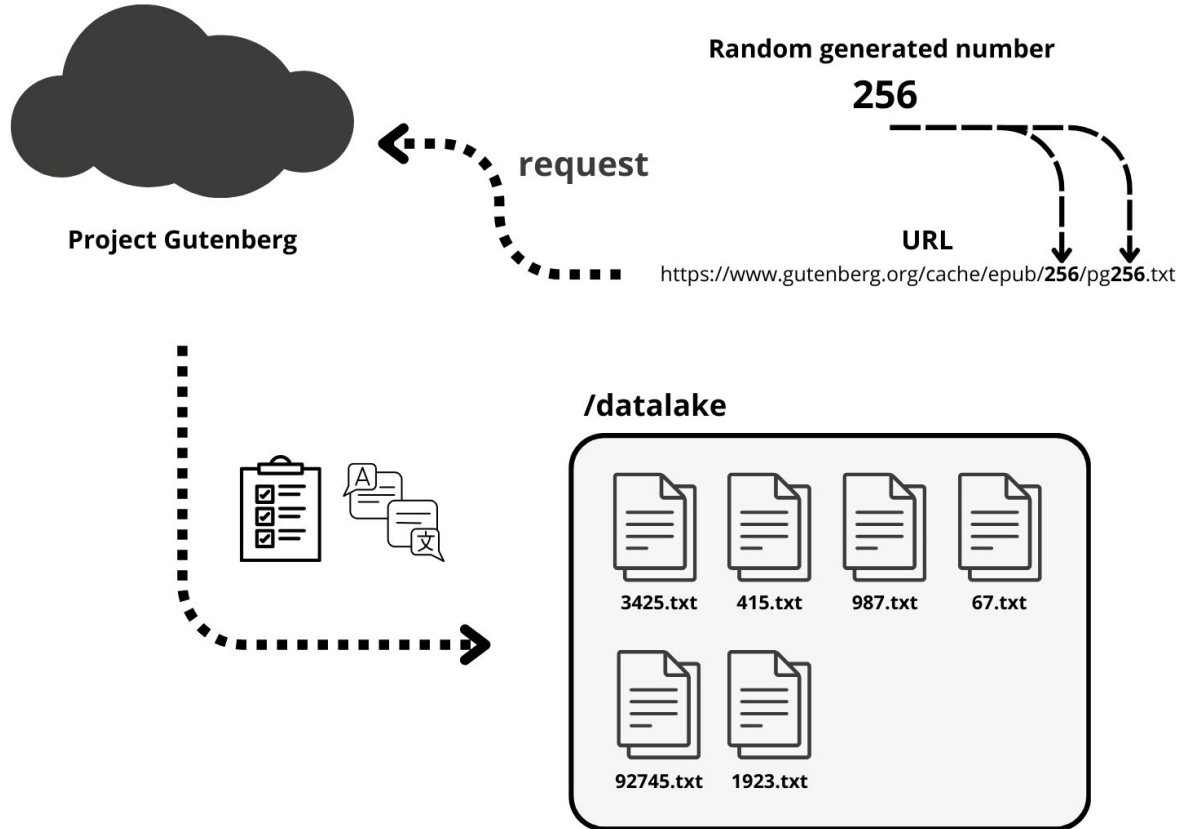
Search Oracle:

Álvaro Rodríguez González, Lucía Afonso Medina, Alejandro Alemán Alemán,
Farid Sánchez Belmadi, Néstor Ortega Pérez.

Structure of modules

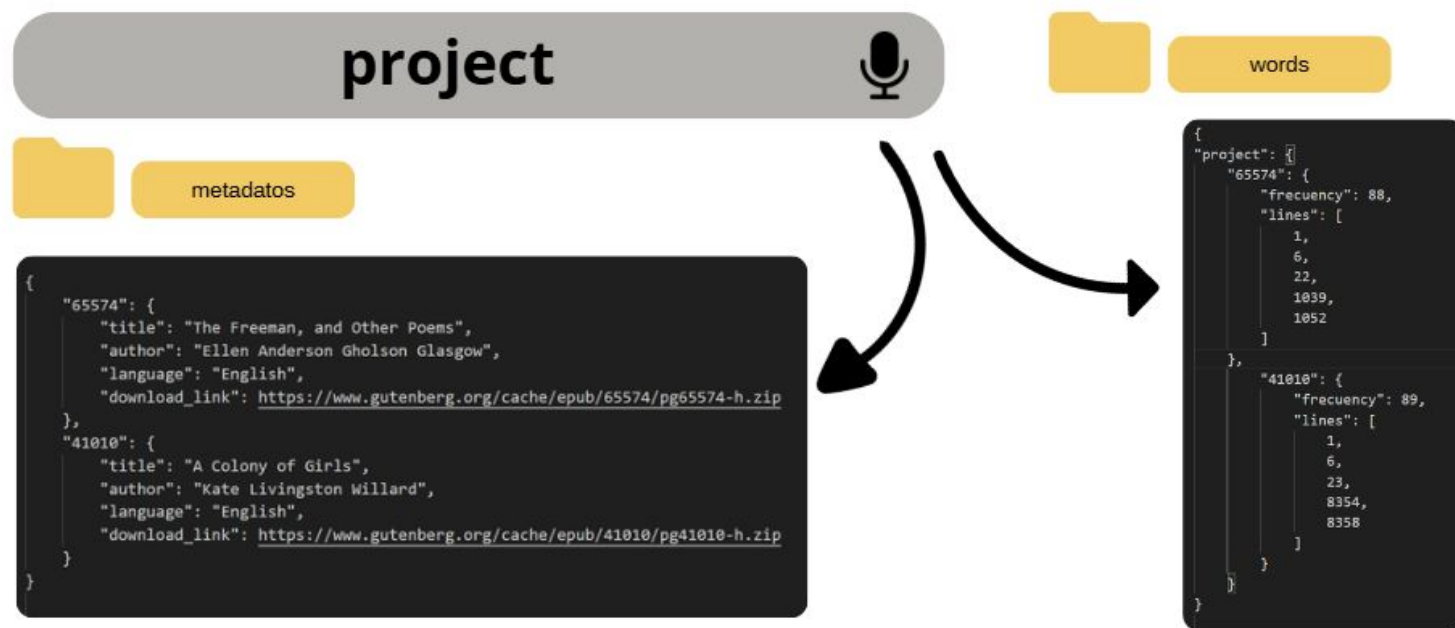


Crawler module



Data Structure 1

Indexer/Query Engine Modules



Data Structure 2

datalake

- 2565.txt
- 4134.txt
- 7665.txt

datamart

- metadatos
 - metadatos
- words
 - ab
 - abated
 - abbess
 - abbey
 - abbeys
 - abbot
 - abjured
 - able

```
{  
  "12345": {  
    "title": "Pride and Prejudice",  
    "author": "Jane Austen",  
    "language": "English",  
    "download_link": "https://www.gutenberg.org/cache/epub/12345/pg12345.txt"  
  },  
  "67890": {  
    "title": "Moby Dick",  
    "author": "Herman Melville",  
    "language": "English",  
    "download_link": "https://www.gutenberg.org/cache/epub/67890/pg67890.txt"  
  },  
  "11223": {  
    "title": "The Great Gatsby",  
    "author": "F. Scott Fitzgerald",  
    "language": "English",  
    "download_link": "https://www.gutenberg.org/cache/epub/11223/pg11223.txt"  
  }  
}
```

```
{  
  "12345": {  
    "frequency": 3,  
    "lines": [5, 18, 42]  
  },  
  "67890": {  
    "frequency": 2,  
    "lines": [10, 25]  
  }  
}
```

Query Engine Module

able

datamart
metadatos
words
ab
abated
abess
abbey
abbeys
abbot
abjured
able

```
{  
  "12345": {  
    "frequency": 3,  
    "lines": [5, 18, 42]  
  },  
  "67890": {  
    "frequency": 2,  
    "lines": [10, 25]  
  }  
}
```

```
{  
  "12345": {  
    "title": "Pride and Prejudice",  
    "author": "Jane Austen",  
    "language": "English",  
    "download_link": "https://www.gutenberg.org/cache/epub/12345/pg12345.txt"  
  },  
  "67890": {  
    "title": "Moby Dick",  
    "author": "Herman Melville",  
    "language": "English",  
    "download_link": "https://www.gutenberg.org/cache/epub/67890/pg67890.txt"  
  },  
  "11223": {  
    "title": "Don Quixote",  
    "author": "Miguel de Cervantes",  
    "language": "Spanish",  
    "download_link": "https://www.gutenberg.org/cache/epub/11223/pg11223.txt"  
  }  
}
```

datalake
12345.txt
67890.txt
7665.txt
11382.txt
22325.txt
23780.txt
30308.txt

Line 5: He was able to recover from the situation quickly.
Line 18: She did not expect to be able to convince him.
Line 42: The abble gentleman remained silent.

Line 10: He was able to guide the ship through the storm
Line 25: No one had ever been able to face such a challenge.

Conclusions



Individual Files for Each Word

Min	Max	Mean	StdDev	Median	IQR	Outliers	OPS
4.9209	14.9472	5.1407	0.8697	5.0446	0.0783	2;11	194.5244

Single File for All Words



Min	Max	Mean	StdDev	Median	IQR	Outliers	OPS
24.5000	98.9580	26.0669	2.7333	25.6250	0.7910	112;214	38.3628