



# Open Data Portal with Quality Control in Port Management Information Systems

André Alves, Bruno Meixedo, Diogo Costa, Francisco Pinto, Pedro Rei

Orientador: Prof. José Moreira, Paulo Pintor

Projeto em Informática, 3º ano, LEI.

2025



## Abstract

Modern systems often integrate data from varied sources with uneven quality, raising concerns about reliability especially in port management. This project delivers an Open Data Portal with provenance tracking, visualization, and quality control to support transparent, informed decision-making.

## Problems:

- Lack of clear metadata and data provenance tracking.
- Struggle to determine the origin, transformation history, and reliability of data.
- Presence of missing and incorrect data complicates analysis and interpretation

## Goals:

- Offer a platform to access and download datasets with full source and transformation metadata.
- Ensure transparency through automatic data provenance tracking.
- Detect and flag data issues, with suggestions for automatic correction.

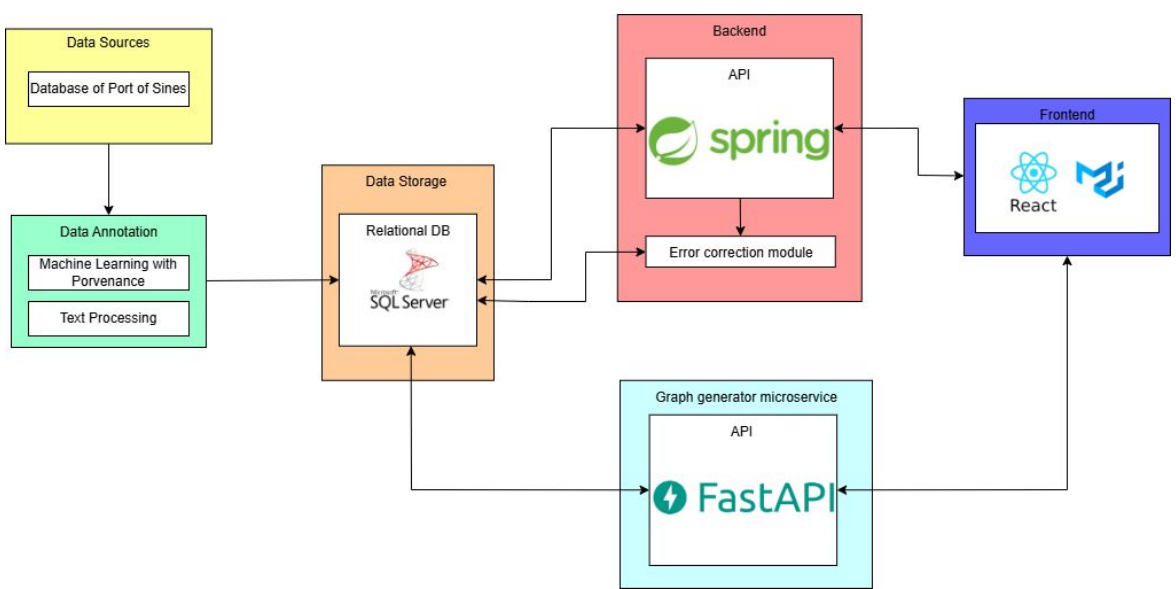


Fig 1- System Architecture Diagram

## Implemented Methods:

User-friendly interface for data access, visualization of statistics of data and metadata and download module.

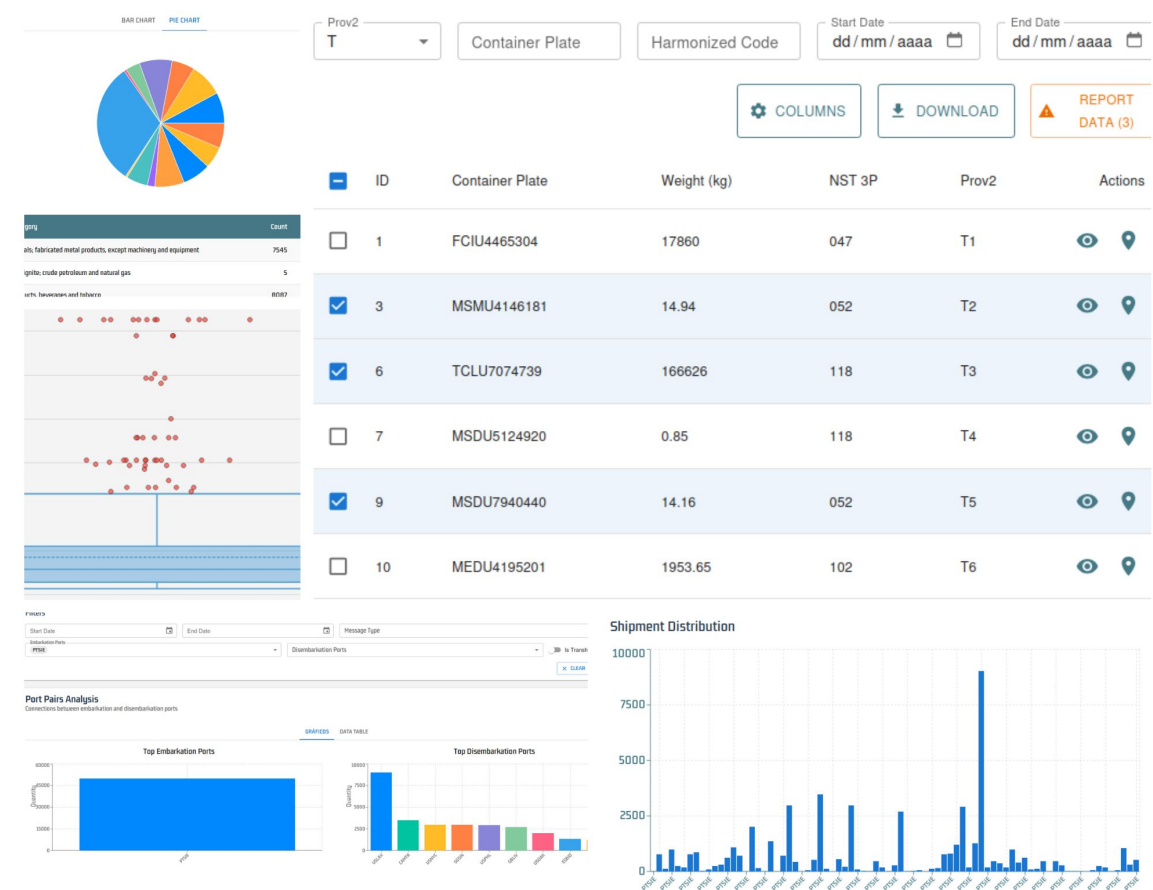


Fig 2- User Dashboard

Metadata inspection module to trace data origins interactively in a graph format.

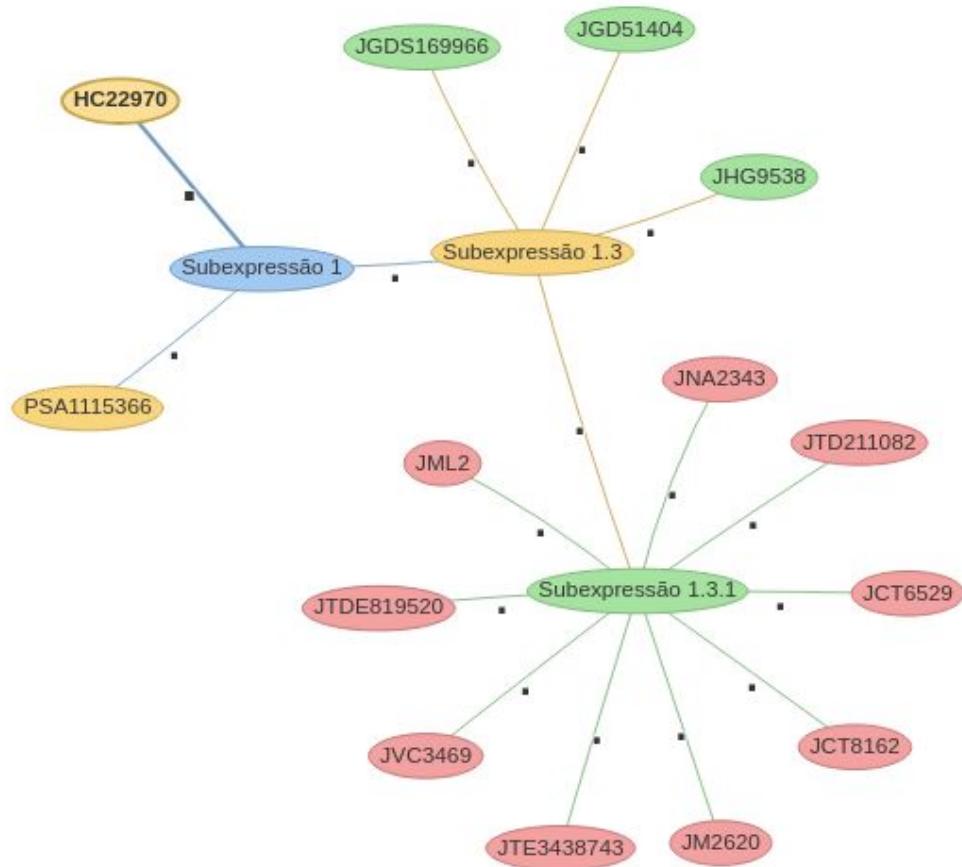


Fig 3- Provenance Graph

Data reliability module to detect inconsistencies and outliers linking issues back to the original data source.

Ship Name	Ship Type	IMO Code	Version	Version Date	Deadweight (T)	LOA (m)	Max Draught (m)	Confidence	Actions
EXPEDITO	Empurador	43	6	28/02/2020, 10:13:54	✓ 297	15.17m	2m	1	👁️ ⚙️ 🗑️
EXPEDITO	Empurador	43	7	28/02/2020, 10:14:26	✓ 297	15.17m	2m	1	👁️ ⚙️ 🗑️
EXPEDITO	Empurador	43	8	28/02/2020, 16:51:25	✓ 297	15.17m	2m	1	👁️ ⚙️ 🗑️
EXPEDITO	Empurador	43	9	28/02/2020, 16:51:59	✓ 297	15.17m	2m	1	👁️ ⚙️ 🗑️
EXPEDITO	Empurador	43	10	23/03/2020, 10:10:35	✓ 297	15.17m	2m	1	👁️ ⚙️ 🗑️
DANMARK	Passageiros (mais de 12)	5086279	2	12/10/2021, 09:30:16	GT	✓ 61.5m (74.6m)	✓ 5.1m (5.1m)	1	👁️ ⚙️ 🗑️
OCEAN MAJESTY	Cruzeiro	6602898	2	20/04/2020, 10:36:17	1031T	✓ 101.5m (130.64m)	✓ 5.1m (5.42m)	1	👁️ ⚙️ 🗑️
DANMARK	Passageiros (mais de 12)	5086279	3	12/10/2021, 09:31:50	GT	✓ 61.5m (74.6m)	✓ 5.1m (5.1m)	1	👁️ ⚙️ 🗑️
OCEAN MAJESTY	Cruzeiro	6602898	3	20/04/2020, 10:36:21	1031T	✓ 101.5m (130.64m)	✓ 5.1m (5.42m)	1	👁️ ⚙️ 🗑️
COMANDANTE AGUIAR	Tanque / Quilómetros (mais de 12)	6729438	1	13/01/2020, 13:26:46	✓ 101.5m (130.64m)	✓ 5.1m (5.1m)	✓ 5.1m (5.1m)	1	👁️ ⚙️ 🗑️

Fig 4- Data Correction Suggestion

Users can flag a datapoint with different errors. Reports are managed on a separate page for review and handling.

Status	Error Type	Item ID	<div>Filter</div>	<div>Clear</div>			
Total Errors: 5							
<div>Download Reports</div>							
Error ID	Error Type	Description	Item IDs	Status	Created At	Severity	Actions
12	<div>data inconsistency</div>	Error in Container Plate	<div>136</div>	<div>Unresolved</div>	05/26/2025, 01:11 AM	<div>Medium</div>	<div></div>
13	<div>missing value</div>	Falta o campo do peso	<div>136</div>	<div>In Progress</div>	05/26/2025, 01:11 AM	<div>Medium</div>	<div></div>
14	<div>data inconsistency</div>	Data inconsistente	<div>4241</div>	<div>Unresolved</div>	05/26/2025, 07:30 PM	<div>Medium</div>	<div></div>
15	<div>duplicate entry</div>	Entrada dupla no peso	<div>8385</div>	<div>Unresolved</div>	05/26/2025, 07:31 PM	<div>Medium</div>	<div></div>
16	<div>incorrect format</div>	Falta uma vírgula no peso	<div>152</div>	<div>Unresolved</div>	05/26/2025, 07:31 PM	<div>Medium</div>	<div></div>

Fig 5- Error Report

## Conclusion

This solution will significantly enhance data transparency, reliability, and usability at the port management level. By integrating metadata-driven provenance tracking and robust data validation tools within an accessible platform, users will be empowered to make informed decisions based on trustworthy, well-documented data.