**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE**

**PILANI (RAJASTHAN)**

**BITS ZG628T: Dissertation**

**Extended Eye**

**Presentation By**

**D BALA VENKATA MANIKANTA**

**2015HT13078**

**Project Outline**

* Background of Application
* Objectives
* Functional Requirement
* Use Case Diagram
* Process workflow Diagram
* Architecture of the Project
* Activity Diagram
* Sample Vessel Tracker
* Implementation Modules
  + IText
  + Natural Language Processing
    - Parts Of Speech
  + Java Script
* Result & Analysis
* Conclusion & Future Work

**Background of Application**

For cargo shipping tracking where dealers and shipping companies are involved internationally, to know their Activities are legal as per the UN / US / operating countries regulations.

Some of the dealers might be black listed and can be some embargos / sanctions on counties and some black listed shipping routes to calculate the risk associated with this deal.

Today the customer must visit many websites to collect information and embargos and fraudulent deals Involving various shipping agencies, dealers and few blocked countries to calculate the risk.

This entire data is dynamic and static document / list will not help and in the volatile environment, Customers must be careful about the sanctions and black listed companies to avoid the fraudulent deals.

**Objectives**

System shall be able to scan the physical deal document.  It shall then be able to parse the deal document using NLP or any equivalent system to identity key candidates for decision making analysis.

Key candidates could be dealer names, shipping corporation names, type of ships, ship id, vessel id, shipping route etc. Using these keywords system shall be able to search authentic public websites such as <https://www.vesseltracker.com> to get the information about these keywords.

It shall be able to parse the HTML output that is given by the search release. It shall then present user with the probable decision.

**Functional** **Requirement**

A functional requirement will explain the behavior of the software and in built third parties will give some specific functionalities for the inputs. Based on the issue functional requirement will change.

Based on previous results of those trackers will provide some recommendations. According into the Extended Eye Functional requirement will split below tasks. Most of the functional requirement will determine using Use cases or BDD scenarios.

Getting the customer data, which consists dealer names, shipping corporation names, type of ships and ship id, vessel id and shipping route etc.

Convert the PDF or Excel or CSV any format of customer data and Search the nouns and verbs in that document.

Search those data in existing government or any third-party websites and come up some user interface because of that user not able to check all website to conform those dealers are black listed or some sanctions are involved or not.

**Use Case Diagram**

Use case will capture the list of events or actions regarding the application overview. Client will act as the actor and detail business case will capture the using this use case.

Dictionary

Parts of Speech Engine

Parser Engine

**Process workflow Diagram**

Project Model and Design is very key to application which tells how much heavy or light is our project. Normally we will draw the work flow diagram to tell the process of the application.

Customer Data Collection

Parts of Speech Library (NLP)

Convert the Data into Plain Text

JavaScript(Search With Param)

User Interface

**Architecture of the Project**

Low Level design is the step by step process application. In this process, we will discuss about Project structure, software architecture, data design and implementation code etc.

Database

Data Base Server

Oracle

JavaScript

Open NLP

IText PDF Parser

Tomcat Server

HTML

Browser

Client

Service Layer

Client Layer

**Activity Diagram**

Activity Diagram which will tell the dynamic action of the application. It will tell in each activity Business functionality.

Import the customer Data

Parse the customer Data into Plain Text

NO

Is Parse Successful

Y

Yes

Search the nouns and alphanumeric words

NO

Found the Nouns

*Yes*

Yes

NO

Browse in the Third parties

No

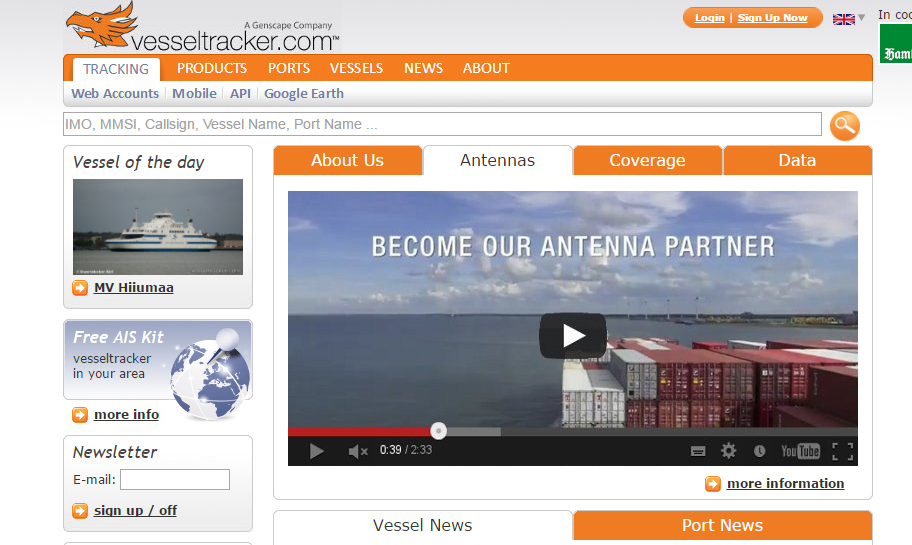
Found the fraud data

Yes

Prepare HTML Page

**Sample Vessel Tracker**

This is one of the sample vessel tracker and we can add so many websites in future



**Implementation** **Modules**

Project development heart is implementation. In this world, there is no one will be without heart the same way without implementations there is no project.

**ITEXT**

iText is open source library to create or manipulate the PDF documents in java. In this iText we will use the SimpleTextExtractionStrategy.

This renderer keeps track of the current Y position of each string. If it detects that the y position has changed, it inserts a line break into the output. If the PDF renders text in a non-top-to-bottom fashion, this will result in the text not being a true representation of how it appears in the PDF.

**Natural Language Processing**

Natural Language Process is small component of the Artificial intelligence and it can understood the human speech and it is based on machine learning.

If you have lot of data that is Structured or unstructured and we want some data extract from that then These Natural Language Processing Techniques will be helpful.

**Parts of Speech**

In Our Project, we will mainly Parts of Speech library. Parts of Speech is a piece of software that reads the text from language and split each word into the Noun, verbe, adjective, etc.

In Parts of speech classification techniques.

1. Supervised Parts of speech:

It is a Structured collection of text it will learn based on already existing data.

1. Unsupervised Parts of Speech

It does not require any exist learning experience based on assumption it will give output.

**Java Script**

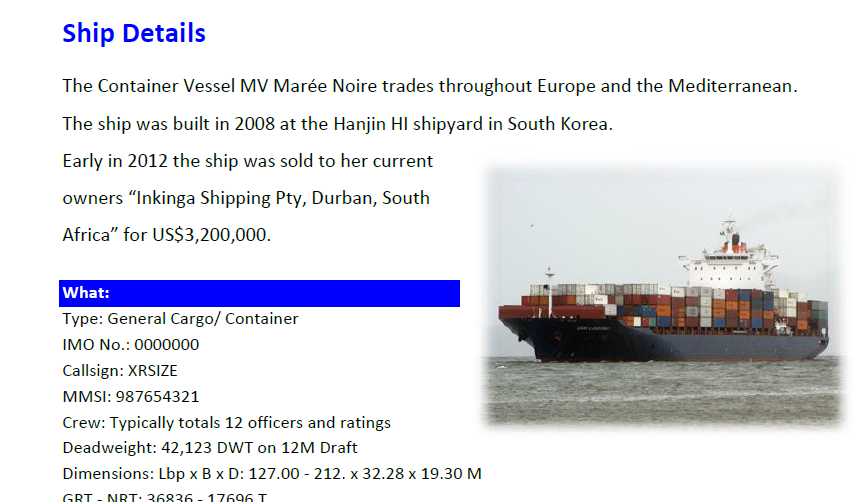
**J**ava Script is a dynamic programing scripting language which will add the interactivity and behavior to the applications.

There are two main operation we will use in this Project. The parameter which we want to search First populate the parameter into the search box and second step will click the submit button but these actions has to done automatically. So application will pass the param and search those parameter.

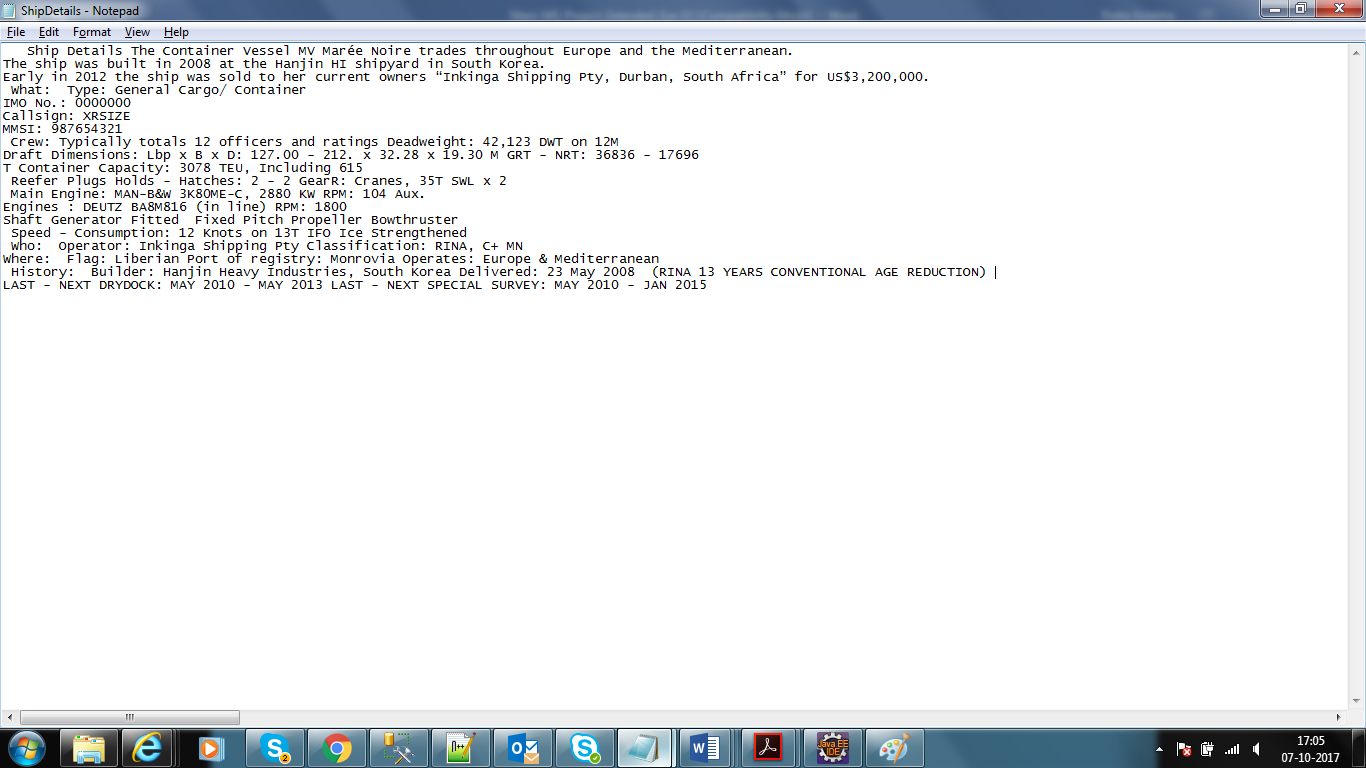
**Results and Analysis**

The subsequent snapshots characterize the outcomes or yields that we will attain after orderly implementation of the considerable number of modules of the framework.

Input : Example of Customer Data in PDF



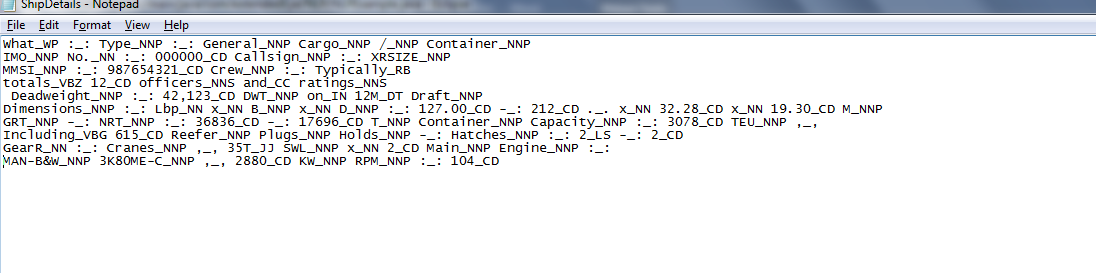
Output : Example of Customer Data in PDF



**Figure 14. Vessel Output file**

* 1. **Result of the NLP Data**

After getting the customer data in plain text and we will apply NLP Techniques which will the all nouns and verbs and alphanumeric words.

****

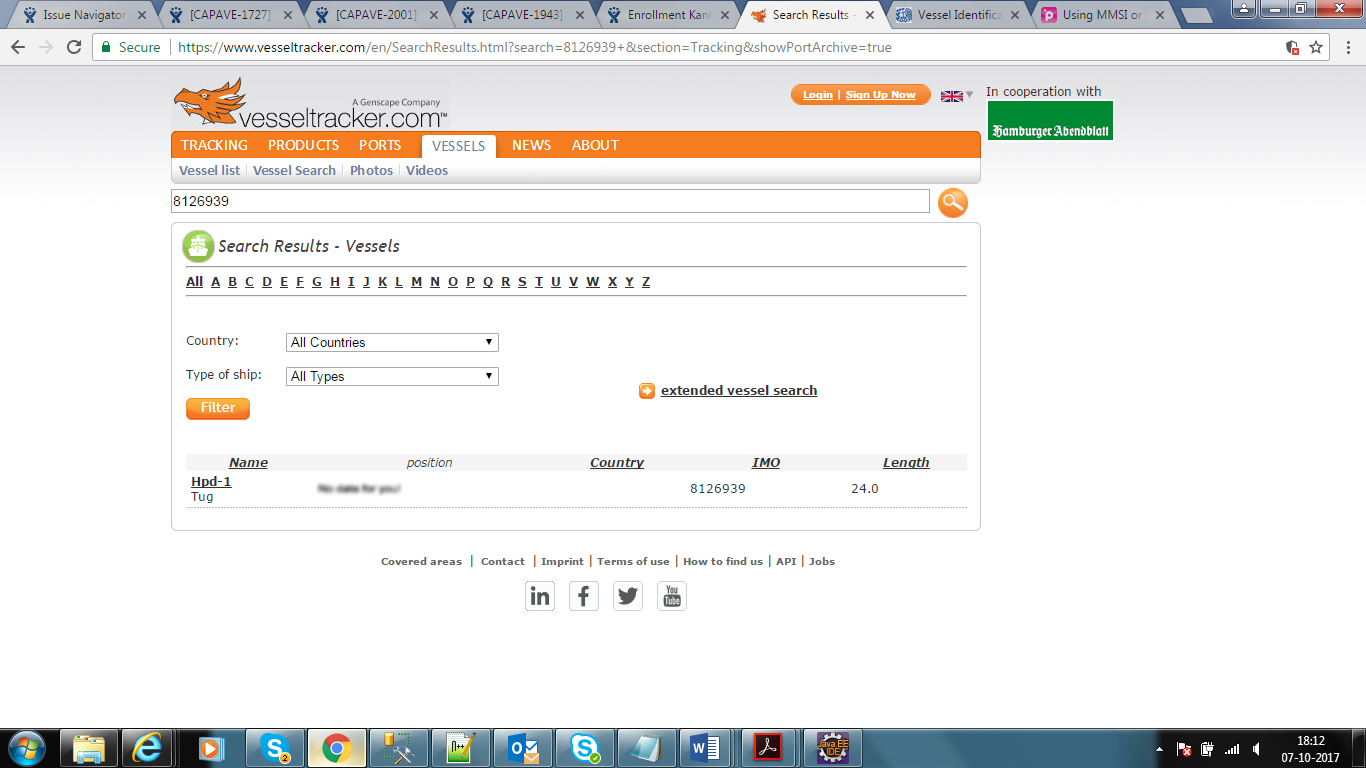
**Figure 15. After Apply NLP Technique**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Page | 25

* 1. **Result of search Parameters.**

In this project, mainly we will focus on nouns because IMO, MMIS, Callsign, Vassal name and port are nouns. we will extract these words and search in the vessel tracker website.



**Conclusion and Future Work**

The thesis project achieved its goal and met the set requirements of the developed application. So, based on this application we can reduce the effort of customer to identify the about the sanctions and black listed companies to avoid the fraudulent deals.

Now customer no need to search to visit many websites to collect information and embargos and fraudulent deals Involving various shipping agencies. Based on this application we can easily find out.

In Future, we will work on the performance and complexity and we can extend this project in serval ways. We will configure a properties file for all third-party web site and we will take one parameter and search all third-party websites.

Second, we will send the mail to customer after finding fraud or sanctions involved in that vessel, in such way we can still reduce the effort as well as we can send the notification alert messages also.

Finally, in Future it very scalable once we configure the all third-party websites where user want to search. It we should this performance and complexity also.



Thank you