Exploiting Data from Govt. Data Portals
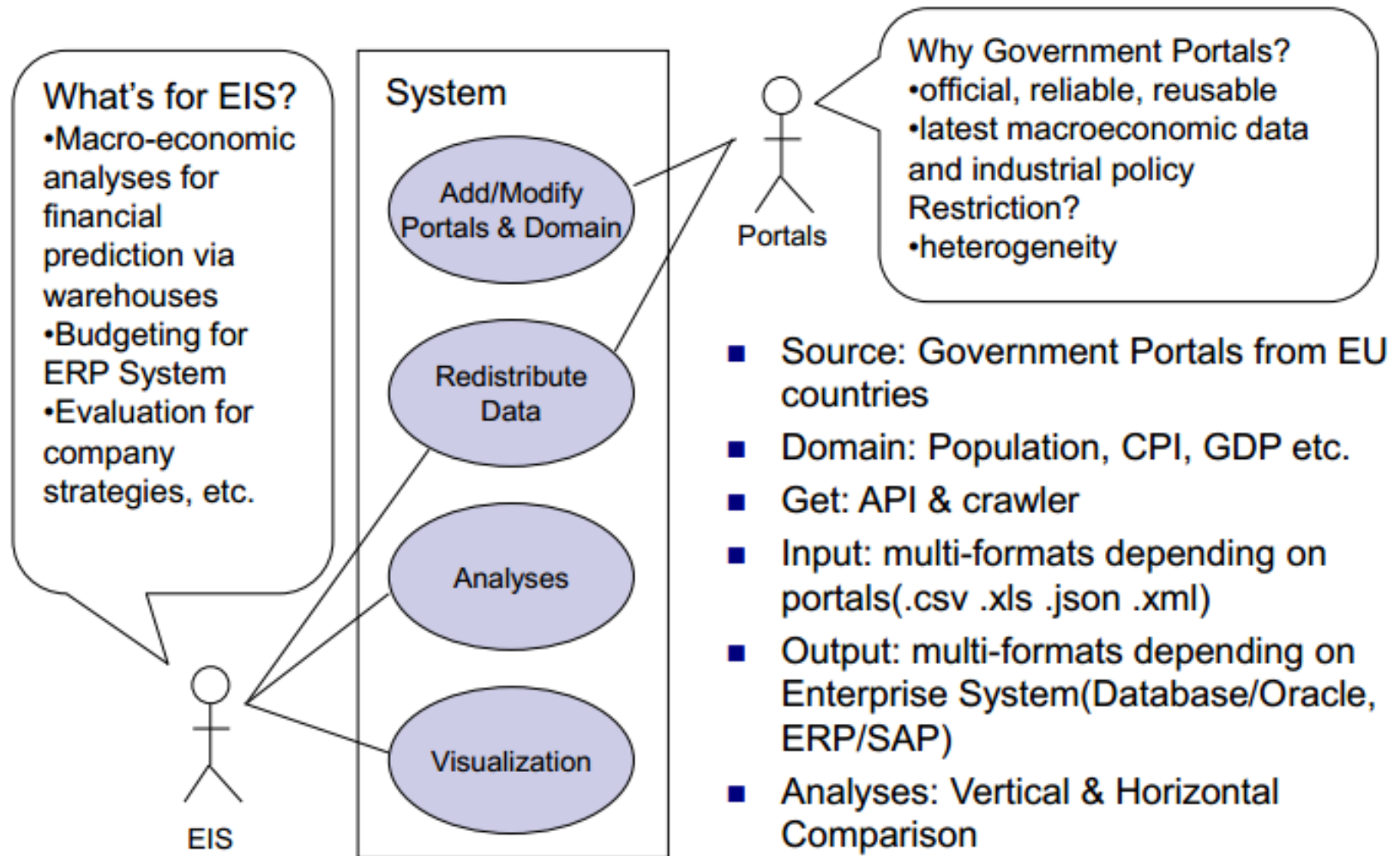
Muhammad Usman

Introduction

- **Goal :**
  - Identify and compare a number of EU Government data portals.
  - Exporting data from these portals and importing it into an Enterprise System.
  - A comparative analysis of this data.
  - Benefits:

System Overview

What's for EIS?
- Macro-economic analyses for financial prediction via warehouses
- Budgeting for ERP System
- Evaluation for company strategies, etc.

System
- Add/Modify Portals & Domain
- Redistribute Data
- Analyses
- Visualization

EIS

Portals

Why Government Portals?
- official, reliable, reusable
- latest macroeconomic data and industrial policy Restriction?
- heterogeneity

- Source: Government Portals from EU countries
- Domain: Population, CPI, GDP etc.
- Get: API & crawler
- Input: multi-formats depending on portals(.csv .xls .json .xml)
- Output: multi-formats depending on Enterprise System(Database/Oracle, ERP/SAP)
- Analyses: Vertical & Horizontal Comparison

DATA
Exploiting Data from Government Data Portals

Gov. Data Portals

- To make Gov. processes more transparent
- To help people to understand how government works?
- How policies are made ?
- Easy decisions and suggestions by public about government policies based on detailed info
- Hear more about the Government's Transparency agenda
- Pilot Data Portals: UK, Germany and France
- Pilot Domain: Population

Gov. Data Portals

	UK	Germany	France
Website	www.data.gov.uk	www.govdata.de	www.data.gouv.fr
Data Access	Data Set , Map Search, Data Request	API, Combining Search Terms, Specific Fields Search, Fuzzy Search	Manual Download
API	CKAN	SCKAN, Liferay	EtaLab
Data Provision	CKAN	Manual: Web Forms, Automated: SKAN API, JSON Dump, CSW	EtaLab
Licensing	Open Gov. License	Open GL	Open GL
Data Usage	Free and Commercial under Open GL	Free and Non Commercial	Free and Non Commercial
Data Formats	XLS, CSV, HTML	XLS, CSV, HTML	XLS, CSV, HTML

Testing

- Unit Testing

- Integration Testing

- System Usability Testing
  - Hallway Testing
  - Load Testing
  - Module Testing

Documentation

- Documentation
  - Introduction
  - Requirement Specification
  - Use Case Specification
  - Architecture
  - Gov. Data Portals
  - Issues in Data Portals
  - API's
  - Future Work and Recommendations
- User Manual
  - Data Export to EIS
  - Data Analysis and Visualization
- Poster

Technical Specification

- Data Export to EIS

    - C#, Windows Forms, MS SQL etc.

- Data Analysis and Visualization

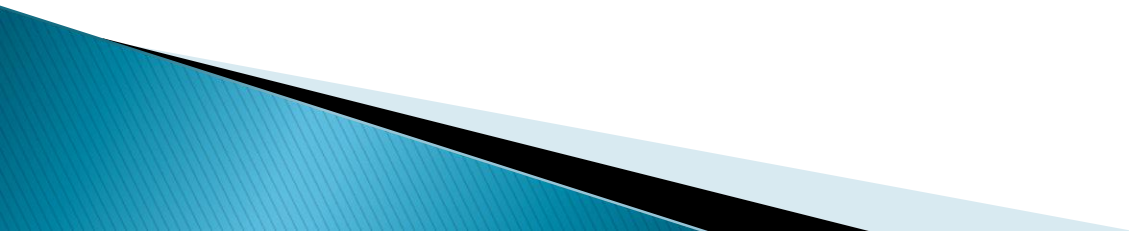   - C#, ASP. Net, HTML, Java Script , Ajax etc.

- Database: MS SQL

- Visual Studio 2010

Issues in Gov.Data Portals

❖ No. Multi Lingual Data Portals

❖ No Single and Easy Data Access Point

❖ Non Availability of Open Machine Readable Data Set Formats

❖ Heterogeneity of Available API for Each Portal

Future Work and Recommendation

- ❖ Multilingual Data download and Comparison Support

- ❖ Single and Easy Data Access Point

- ❖ Availability of Open Machine Readable Data set formats like RDF , JSON Etc.

- ❖ Homogeneous API for Data Access

- ❖ Multi Format ,Multi Lingual and Platform Independent Data Export to EIS and Data Visualization Tool

Demo

Thanks