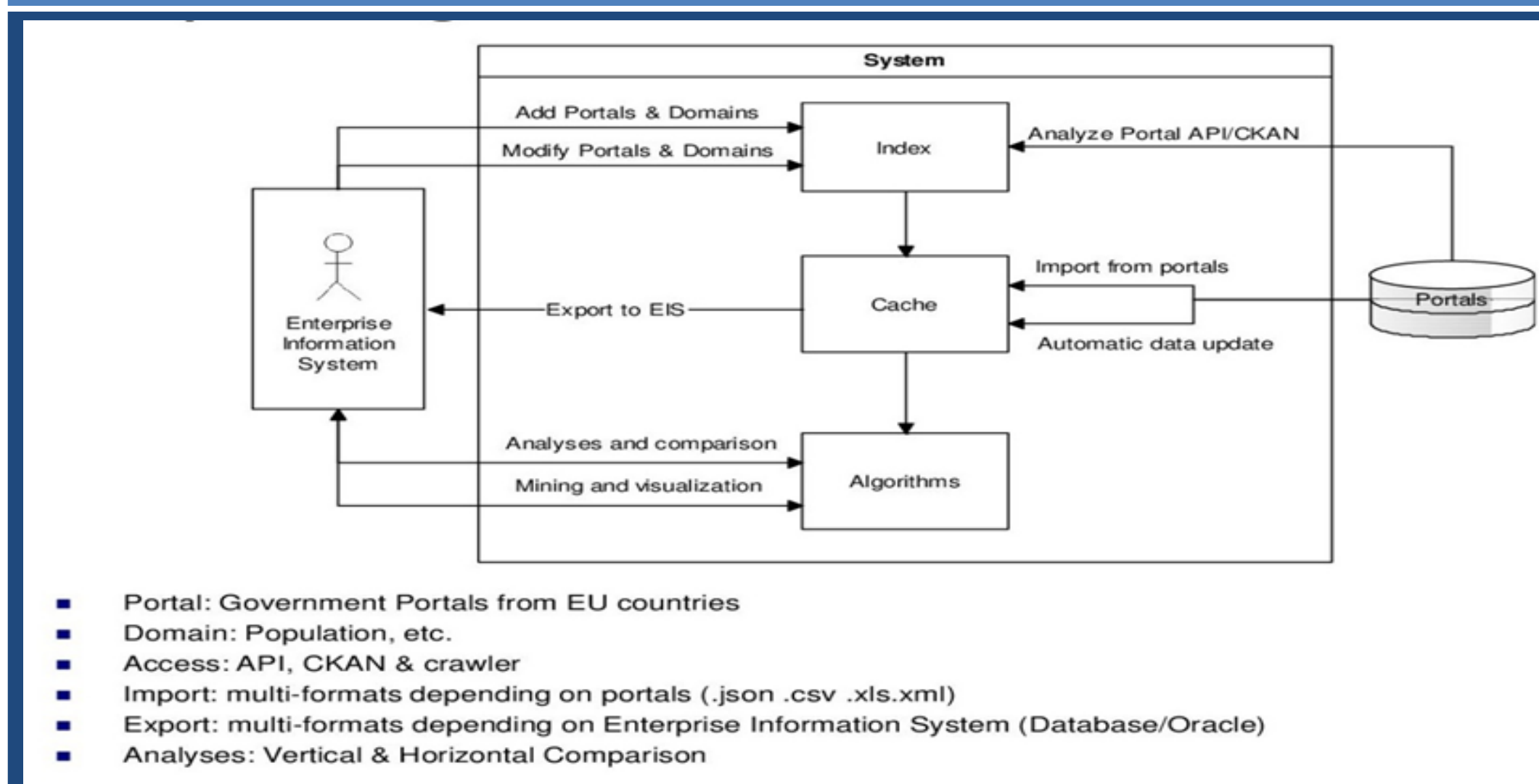# Exploiting Data From Govt. Data Portals

Compare Portals, Download & Export Data to EIS, Analysis and Visualise Data

Muhammad Bajwa, Matriculation:2383436, Bajwa@informatik.uni-bonn.de, Group F
Judie Attard: attard@iai.uni-bonn.de, Prof. Dr.Sören Auer

Introduction

The aim of this lab was to identify and compare a number of EU Government data portals. This included exporting data from these portals and importing it into an Enterprise System. A comparative analysis was then led on this data.

System Architecture

- Portal: Government Portals from EU countries
- Domain: Population, etc.
- Access: API, CKAN & crawler
- Import: multi-formats depending on portals (.json .csv .xls.xml)
- Export: multi-formats depending on Enterprise Information System (Database/Oracle)
- Analyses: Vertical & Horizontal Comparison

Govt. Data Portals

The Government is releasing public data to make its processes more transparent and to help people to understand how government works and how policies are made and to make easier for people to make decisions and suggestions about government policies based on detailed information and hear more about the Government's Transparency agenda from the Government. **UK, Germany** and **France** data portals was our pilot data portals from EU and **Population** as pilot data set domain.

	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,PDF	XLS, CSV, HTML,PDF	XLS, CSV, HTML,PDF

Issues in Data Portals to Access Data

- ❖ **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**

Data Analysis and Visualization

This tool helps the user not only to analyze and visualize data per Portal but also compare data between the portals both in Graphical as well as in tabular forms.

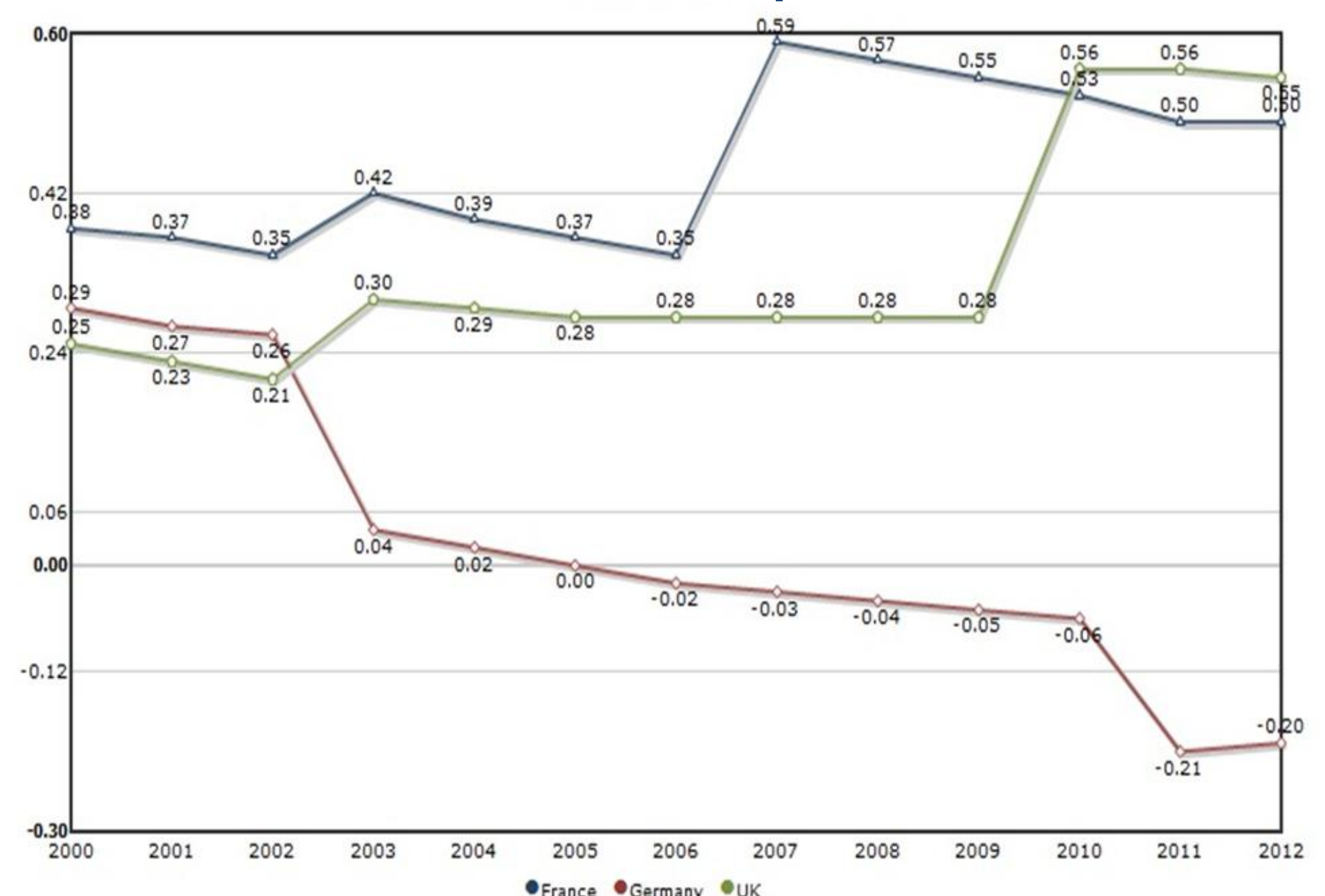**Population:** The total number of persons inhabiting a country
Population Density: # of people /kilometer Sqr. (Unit Area)
Growth Rate: The average annual % change in Population
Birth Rate: Birth/1000 Population
Death Rate: Death/1000 Population

Growth Rate Comparison

Future Work and Recommendations

- ❖ **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**

<Virtual Machine Description>

Hostname: Windows-ExpGovDataGroupF

Software Installed on EIS3
Visual Studio.net 2010: MSSQL 2008,ASP.Net, .Net Framework 4.5, IIS, EXDataExcel2EIS, Data Analysis and Visualization 1.0
Code Access: https://github.com/EIS-Bonn/MA-INF3232
Lab/tree/master/SomSem2014/Group%20F%20(GovData_Attard)

Local Access:	Remote Access:
OS User: EIS-User	Partner ID: 902677001
Password: GovData	Password: i34g3b

EIS
Enterprise Information Systems

universität bonn