# Appendix 4   Details and data related to Asian Regional Integration Observatory with Greater Mekong Subregion/Association of Southeast Asian Nations Economic Community Cluster Demonstration (also Central Asian Regional Economic Cooperation Demonstration)

1   SYSTEM DESIGN

The platform is built on a general Web-GIS structure. Thematic socioeconomic datasets are organized as GIS layers and stored in a PostgreSQL database. The database has a GIS application layer, and PostGIS, as extension of the standard database to process spatial queries. GeoServer is used as GIS server to receive hypertext transfer protocol (HTTP) queries from the browser. It decodes the queries for vector or image layers, requests the data from the database, and passes the result back to the browser. The Web server is the component that deals with the HTTP query on the interface of the website. It returns the webpage in HTML format as well as the style file in cascading style sheets (CSS) format and browser-side code in JavaScript. The JavaScript code helps to implement the interactions and to compose a correct query for the thematic maps (see Figure A4.1).

144
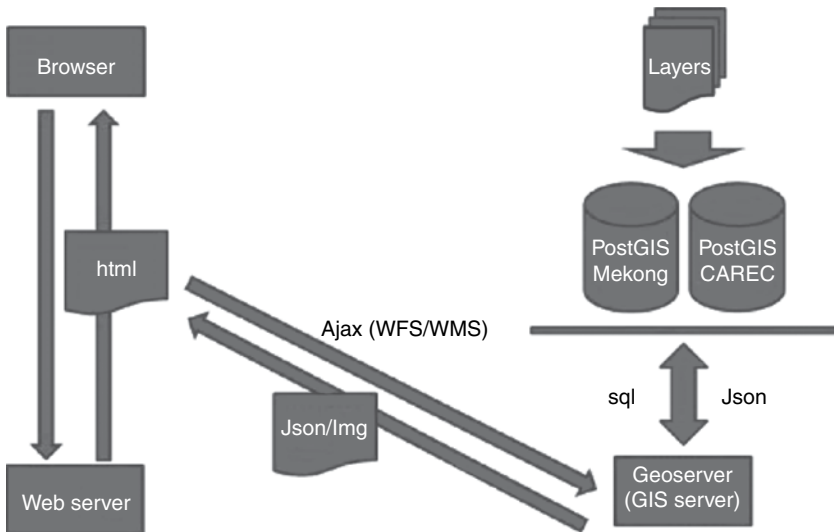
Notes: CAREC = Central Asian Regional Economic Cooperation; GIS = geographic information systems; html = hypertext markup language; Img = image; Json = Java Script Object Notation; sql = structure query language; WFS = Web Feature Service; WMS = Web Map Service.

Source: Fu (2015).

Figure A4.1   System structure of observatory

2   INDICATORS

Due to the expected data availability issue, the project confines itself to 14 critical indicators: population, GDP, employment, wage levels, consumption, productivity, poverty level, transportation accessibility, transportation cost, land use, product space, trade and value-added flows, sector cluster location, and R&D expenditure.

3   DATA SOURCE

The main data sources of socioeconomic data are government statistics agencies (see Table A4.1).

As socioeconomic data is not detailed enough to identify cluster locations, information on industrial zones in GMS and CAREC was also collected as a complementation source. Survey reports from Japan

Table A4.1   Government statistics agencies

Country	Agency
Cambodia	National Institute of Statistics
PRC	National Bureau of Statistics
Lao PDR	Statistics Bureau
Myanmar	Central Statistical Organization
Myanmar	Myanmar Industries Association
Thailand	National Statistical Office
Viet Nam	General Statistics Office

Notes: PRC = People's Republic of China; Lao PDR = Lao People's Democratic Republic.

Source: Author.

External Trade Organization and Vietnam Investment Network are major sources.

Detailed trade flows in specific sectors and products can also help to narrow down the searching scope on industrial clusters. The UN Commodity Trade Statistics Database (COMTRADE) and Corridor Performance Measurement and Monitoring database are used. Corridor Performance Measurement and Monitoring is hosted by ADB. The project samples cargo flows among CAREC countries every month to record the type of commodity, origin, destination and intermediate cities, total weight of commodity, reported value to customs and so on. By identifying the origin cities of one commodity, we can assume that the commodity is produced in the city or its nearby area to narrow down searching scope.

4   DATA AVAILABILITY

For the data intensive knowledge platform as an observatory, data availability is affected in three dimensions: indicator, geographic unit, and time.

Data accessibility is critical in both GMS and CAREC countries. In practice, there are several data gaps:

1.   No indicator data at a certain administrative level. Usually countries have a value at country level but no more detailed values for its provinces. For instance, Myanmar is still in the process of conducting the

2014 census, which is the first census in the past 30 years. Thus no reliable fundamental demographic indicators are available.

2. Different countries have different statistics definitions and/or standards on an indicator. For instance, the Bureau of Statistics in PRC only counts firms with RMB 5 million ($807 000) annual turnover, above which firms are surveyed and sampled.

3. Monetary indicators in different countries are usually presented in own currency. The value is also not adjusted to constant price.

4. Language is a challenge as GMS and CAREC countries do not use English as the official language. Although yearbooks usually have an English version, some surveys and reports are written in the local language.

The administrative unit is also a critical issue. Usually, socioeconomic indicators are aggregated by administrative boundaries. A small geographic unit is better to gain finer spatial knowledge. However, data is also more difficult to access for smaller geographic units. At present, the target administrative and/or geographic unit for presenting data is set at provincial level for balancing data accessibility and spatial detail. Table A4.2 shows detailed information about the geographic unit selected for each GMS country.

Time is the third critical dimension. To collect evidence on the process of clustering, time series data is required. As most socioeconomic indicators are reported annually, a year is the basic time unit in the observatory. The time span that this observatory tries to cover is between 2007 and 2012. The observatory also adds the most recent

Table A4.2   Geographic unit in Greater Mekong Subregion countries

| Country	Unit	Administrative level
Cambodia	Province	Province
PRC	Prefecture	Between province and county
Lao PDR	Province	Province
Myanmar	State/Division	Province
Thailand	Changwat	Province
Viet Nam	Province	Province

Notes: PRC = People's Republic of China; Lao PDR = Lao People's Democratic Republic.

Source: Author.

survey on GMS and CAREC key industrial zones conducted by ADB in 2014/15.

5   INTERFACE

The user interface, as the Web page, is composed of the digital map on the left and the indicator selection column on the right (see Figure A4.2). The map has basic GIS functions, including zoom in and out, pan, and interaction with markers. Data of the selected indicator will be visualized as circle markers in the map. The size of the circle represents the value. Once data is
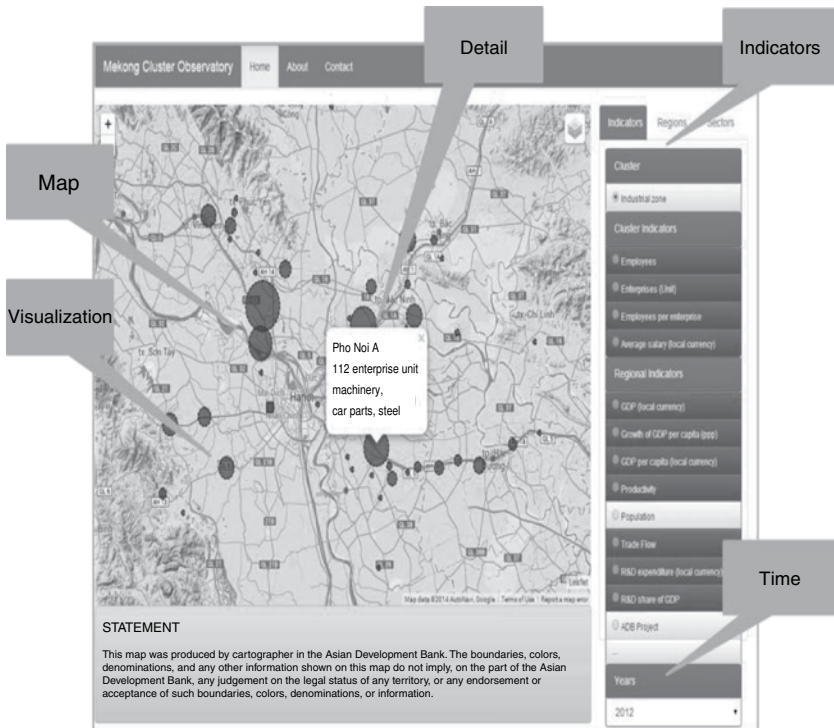
Notes: ADB = Asian Development Bank; GDP = gross domestic product; R&D = research and development.

Source: Fu (2015).

Figure A4.2   User interface of Mekong Cluster Observatory

Notes: ADB = Asian Development Bank; APAN = Asian Pacific Advanced Network; cont'd = continued; MoU = memorandum of understanding.

Source: Author.

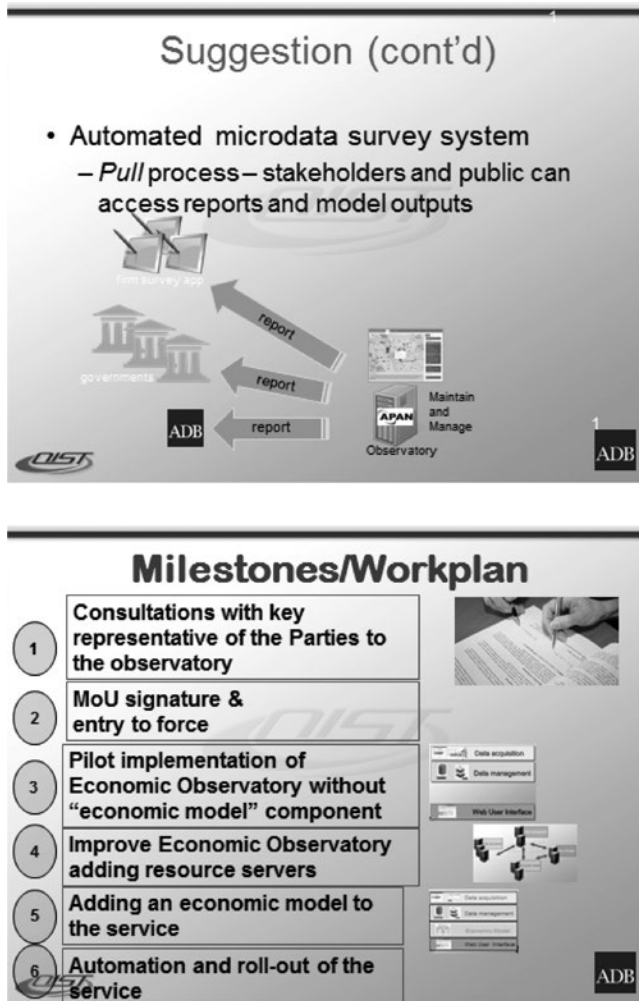*Figure A4.3 Asian Development Bank–Asia Pacific Advanced Network Memorandum of Understanding outline*

Figure A4.3 (continued)

loaded and visualized, users can click the markers to access the exact value and other relative information in the pop-up window.

For the indicator column, the visualization design helps to identify data availability: a dark-shaded bar (for example 'Employees') means that data for the selected indicator at the current year is not available for all countries. A light-shaded bar (for example 'Population') means that data for the selected indicator is available for at least one country.