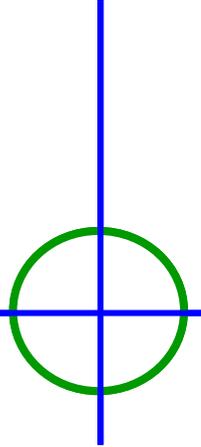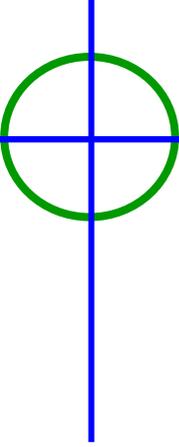# The 11th Meeting of the Heads of National Statistical Offices of East Asian Countries (6-8 November 2006, Tokyo, Japan)

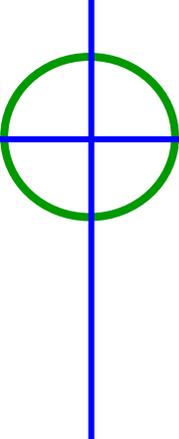## The Design and Development for Business Common Sampling Frame of Establishments in Thailand

Overview

❑ Introduction & Objectives

❑ Phase I: System Design and Prototype

❑ Phase II: Database System and Usage

❑ Future Operational Plan

❑ Problems encountered and lessons learned

Introduction& Objectives

❑ Current establishment sampling frame obtained from the establishment censuses: costly ,time consuming, out of date.

❑ NSO tries to construct a Common Sampling Frame of Establishment (CSF) using administration data from various government agencies: reduce cost and time, easy to update.

❑ Promote the CSF to be used by other SUs: reduce provider burden, able to link data across surveys done by diff. SUs, improve est. survey data accuracy.
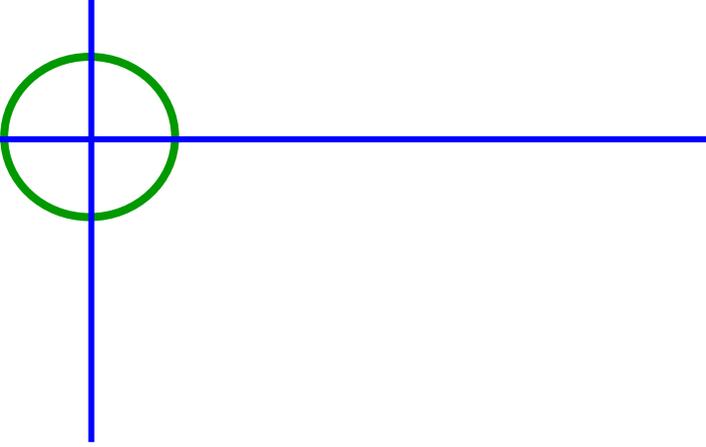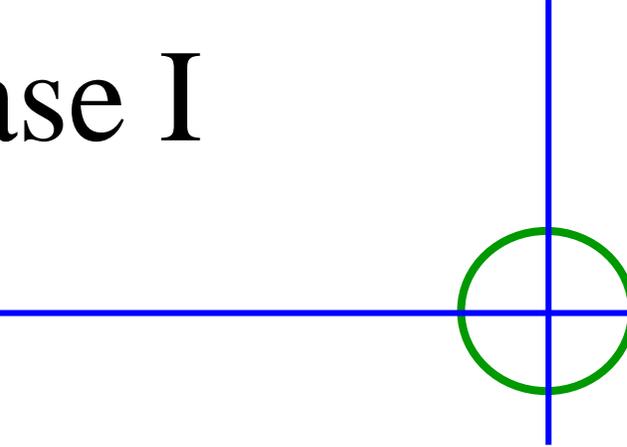
Process

❑ A consultant team from Chulalongkorn Uni. worked with NSO team: focus on learning process.

❑ CSF Project is divided into 2 Phases

– Phase I: System Design and Prototype

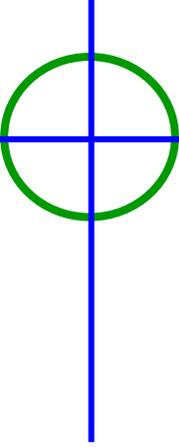– Phase II: Database System and Usage.

Detailed Process: Phase I
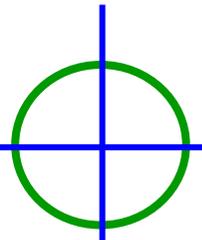
System Design and Prototype

Phase I: System Design

❑ Database Analysis

– Review establishment data bases of 19 related government agencies.

– Review meta data of each data base: type of data stored, format, legal support, time to update, concepts & definitions used etc.

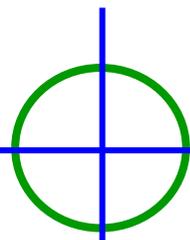– Consult with 20 Ministries: establishment data producers and users.

Owners of Establishment Registration& Related Agencies

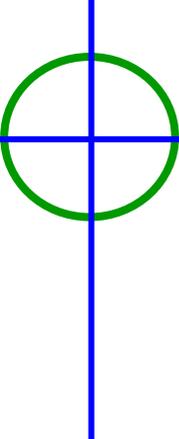| Agencies	Departments
Ministry of Finance	Customs Dept., Excise Dept, and Revenue Dept.
Ministry of Transport and communication	Dept. of Waterways Transport and Maritime, Dept, of Land Transport and Dept. of Air Transport
Social and Security Development	Off. Of Investigation and Legal Affairs and Bureau of Provincial Affairs
Ministry of Commerce	Dept.of Business Development and Dept. of Export Promotion
Labour	Labour Protection and Welfare and Social Security Off.
Education	Dept. of the Private Education Commission

Owners of Establishment Registration & Related Agencies

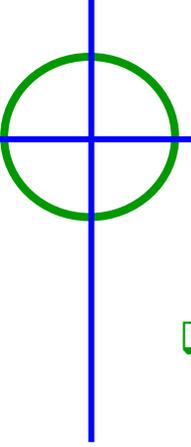| Agencies	Departments
Ministry of Public Health	Dept. of Medical Science and the Food and Drug Administration
Ministry of Industry	Dept. of Industrial Works, the Board of Investment of Thailand, Off. of Industrial Economics and Off. of SMEs Promotion
Office of the Prime Minister	The Royal Thai Police
Min. of ICT	NSO
Private bodies	the Thai Chamber of Commerce and Federation of Thai Industries

Fundamental Items

- Legal Name
- Address, Postal code, Telephone number, Fax number, E-mail address
- Date of operational, Date of close down
- Period of operation
- Registered capital
- Form of legal organization
- Form of economic organization

- Proportion of foreign investment or share holding
- Status
- TSIC/ISIC Economic activity,
- Number of person engaged
- Number of employees
- Remuneration
- Wages and salaries
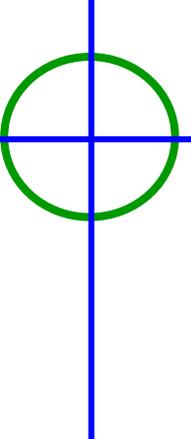- Total Expenses
- Total Sales and Book value of fixed assets

Results from Data Base Analysis

❑ Each government agency collects different detail of data to serve different objectives.

❑ Necessary to compile and store metadata to be able to route to owner of data.

❑ Define Business Establishment :a location or a part of single physical location that does one or more type of economic activities under a single ownership, no matter if the activities are performed by the owner himself or are controlled by the registered company.

❑ The industry classification:
  – United Nations: ISIC Rev.3
  – Thailand Standard Industrial Classification: TSIC 2001
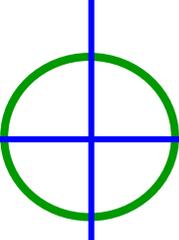
System Requirements

❑ System requirements :

– To support input data from every sources in text file and relational database format

– To retrieve, convert, and transfer data from many sources

– To upload data from all sources into file sever through web pages, file transfer, or web services

– To verify constraints in updating data including basic data items, details data items, and demographic data from each data sources

Output Design

❑ Output design of the systems :

– Report showing list of establishments according to search condition and their fundamental data items

– Output screen showing number and/or details of business establishments according to search criteria

Input Design

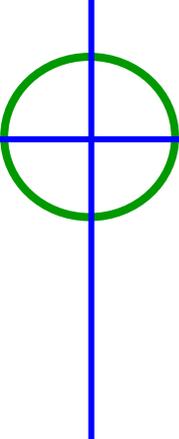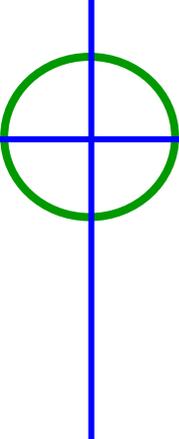| Source of Data	Detail
Revenue Department	Use to update attribute items and status of business establishments.
Department of Business Development	Use to update the list of business establishments, fundamental items, attribute item and status of business establishments.
Social Security Office	Use to update attribute items of business establishments
The National Statistical Office	Use to update the list of business establishment, fundamental items, attribute item and status of business establishments.

Database Design

❑ Data components include business establishment data in two categories.

– Business Establishment Data: Business establishment data includes fundamental items, attribute items, and establishment demographic items.

– Metadata of business establishment: Metadata includes data of attribute items data of business establishments that each government agencies recorded.
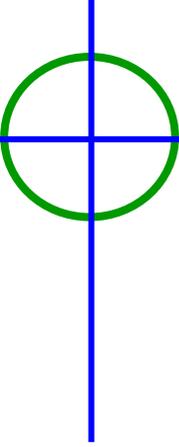
Process Design

❑ Extraction, Transformation and Loading: ETL

- – Data Loading Process
- – Data loading to staging area process
- – Syntax Validation
- – Data Matching
- – Creating Business establishment ID
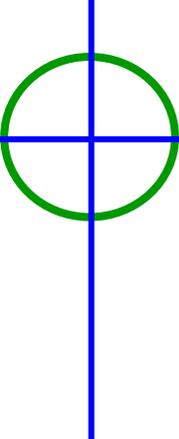- – Process of updating data in the database
- – The First Loading Process

Use of the new systems

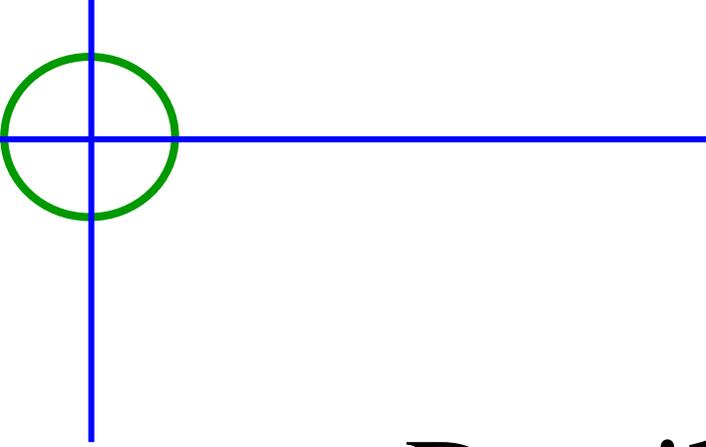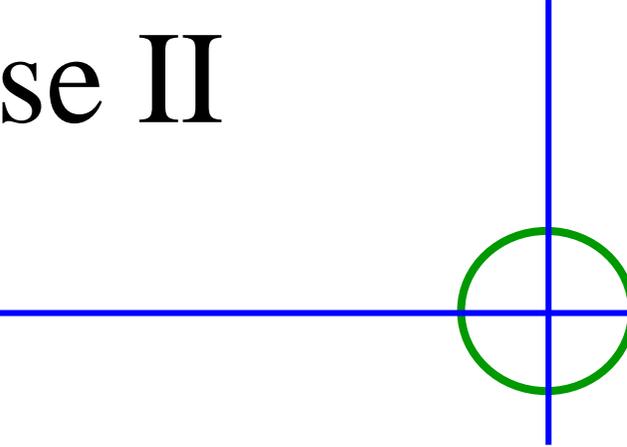❑ Business Establishment Search

❑ Metadata Search

Security Design

❑ Users in each category have different access level.

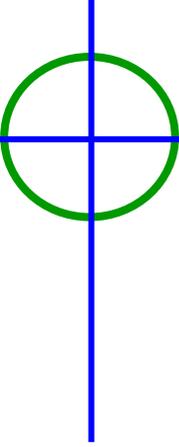– System Administrators

– Internal users

– External users

Detailed Process: Phase II

Database System for selecting the sample of establishment

Phase II: Database System

❑ Database System for selecting the sample of establishment

– Database system is developed to provide the sampling frame

– Data are collected from 4 government agencies i.e. RD, DBD, SSO and NSO

– Data go through the process of data cleaning or ETL (Extraction, Transformation and Loading)

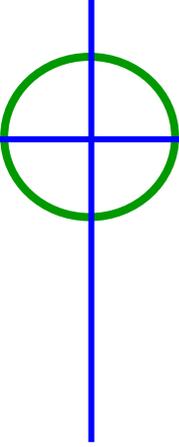– Database is ready to be used for selecting the sample

The result of Phase II

❑ The Common Sampling Frame of Establishments Database System is complete, but it is still to be improved in order to use in the framework for selecting the sample used in establishment surveys.

❑ Training programme for staff in order to understand how to utilize the establishment database.

Future Operational Plan

❑ Promoting other SUs to use this database as a sampling frame for selecting the sample of establishments: more user friendly approach: software e.g. CHORMY-GEN(EZ-sample), Excel

❑ This database will be merged to the Enumeration District System of NSO and connected to the country's map.
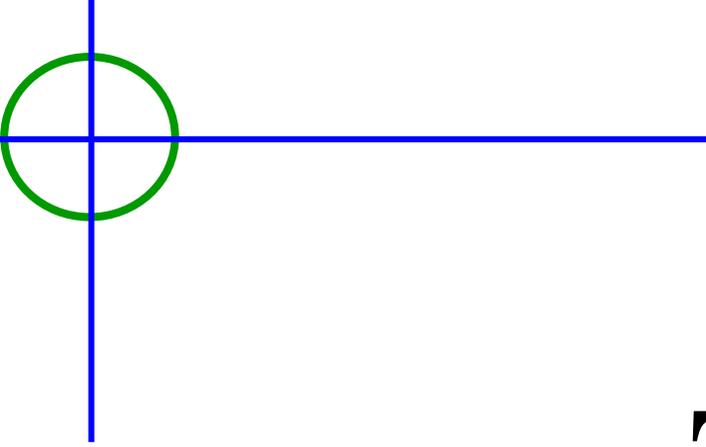
Problems encountered .

- ❑ Less cooperation from government agencies.

- ❑ Format of data is different

- ❑ Transformation and loading Thai language software is not available

- ❑ According to Thai alphabet, there are some differences in the way of recording data from the various agencies

Thank You