Meeting on the Management of Statistical Information Systems (MSIS 2014)
(Dublin, Ireland and Manila, Philippines 14-16 April 2014)

Topic (ii): Implementation of standards supporting the modernization of official statistics (Manila Session)

Toward the Modernization of Official Statistics at BPS-Statistics Indonesia:
Standardization Initiatives and Future Directions

Prepared by Said Mirza Pahlevi, BPS-Statistics Indonesia

I. Introduction

1. The Change and Reform for the Development of Statistics (STATCAP-CERDAS) is a reform strategy of BPS with the aim of modernizing the statistical system in Indonesia and, in particular, improving the quality of the statistical products produced by BPS. The STATCAP-CERDAS project is a mid-term project for 2014-2018 that provides four interrelated major goals or components: (i) improving statistical quality as well as improving statistical user satisfaction and confidence of BPS products and services; (ii) increasing the efficiency of statistics administration through application of information and communication technology; (iii) improving human resources management and development to support statistics administration; and (iv) strengthening the organizational structure.

2. At the core of the project is the intention to move the statistical program from a set of semi-autonomous surveys to a harmonized suite of surveys that draw on a common statistical infrastructure. This involves three steps. The first step is to recognize the common features that underlie all surveys in the form of a generic statistical process model. This is the starting point for the harmonization and is further described in Section 2. The second step is to build the corporate statistical infrastructure to support all the processes identified in the process model. The third step is to review and revise the entire statistical program so that the component surveys are built on the corporate statistical infrastructure and hence integrated to the fullest extent possible. The second and third steps include modernization of ICT and information management that will be described in Section 3. As the basis of the three steps, the Generic Statistical Business Process Model (GSBPM) is adopted as references in defining BPS future business processes.

3. The internationally recognized GSBPM published by the UNECE Statistical Division is a well defined international standard. It divides the survey process (cycle) into 8 phases, which are themselves further divided into 3-8 sub-processes. The GSBPM serves well to model surveys, including censuses and administrative collections. These constitute the major part of BPS statistical program. The GSBPM also covers metadata
management and quality assurance, which are viewed as over-arching activities relevant to all phases of the survey process.

4. This paper describes the first two major goals of the STATCAP-CERDAS project, namely, improving statistical quality and improving ICT and information management. It also gives a brief description of several initiatives that have been done in order to prepare for the project.

II. Improving Statistical Quality

5. This section provides general information on the improved statistical quality component of BPS. BPS has initiated a program to improve statistical quality and efficiency of statistics administration as well as user satisfaction to keep pace with the demands of a modernizing country in a rapidly globalizing world. The improvement in statistical quality will not succeed without major support from proper organizational capabilities and capacities, information and communication technology (ICT), and human resources.

6. BPS conducts data collection via three methods - census, survey, and administrative product compilation based on RI Law number 16 Year 1997. The data collections are conducted by different Subject Matter Areas (SMAs) and are supported by a methodology unit and a dissemination unit. Many of BPS surveys were developed independently of one another and still operate largely independently as so-called survey silos.

7. Based on an in-depth study of the implementation and operation of the statistical collections in BPS, in order to improve statistical quality, BPS needs to improve the following 8 important areas:
   - Statistical governance
   - Business process and system
   - Communication with stakeholders
   - Skill and knowledge of staffs
   - Integration of statistical surveys
   - Survey response rates
   - User satisfaction on statistical data
   - Data and metadata management

8. To achieve high statistical quality, Figure 2.1 illustrates the lifecycle of BPS Statistics that aligns all data collection activities in BPS. There are 7 steps in BPS Statistical Lifecycle: Planning, Preparation,
Collection, Processing, Analysis, Dissemination, and Evaluation. Planning is included in specify needs phase which determines whether there is a presently unmet demand, externally and/or internally, for the identified statistics, whether BPS can produce them and how to produce them. The preparation and collection steps are dedicated to provide effective approach in collecting data from respondents. The processing, analysis, dissemination and evaluation steps are dedicated to maximize data utilization that is able to meet the user requirement and presented in a user friendly way. Every step in BPS Statistical Lifecycle is governed with statistical governance such as ICT Strategy, ICT governance and management, etc. Moreover, every activity in BPS Statistical Lifecycle is recorded in metadata repository and data produced at the statistical production processes are maintained in databases.

9. Figure 2.2 shows the mapping between the BPS Survey Business Process and the (first level) GSBPM. From the figure, the Planning process in BPS is mapped to the Specify Needs process in GSBPM, while the Preparation process is mapped to the two processes in GSBPM, namely, Design and Build processes.

10. Figure 2.3 shows the BPS Survey Class diagram. The diagram consists of three process groups (of the BPS Statistical Lifecycle) and five common services that supported the process groups. The three process groups are collection, processing, and analysis. The five common services are Field Management, Frame and Register System, Data Management and Metadata, Input Data Warehouse, and Survey/Project Management System.

11. The collection group has five process elements which are Data Capture, Provider Management, Interview Management, Profiling, and Frame Maintenance and Sample Selection. Collection data is obtained from BPS interviewers that interview household/retail and wholesale/establishment respondents. The collection data is also obtained from third party data, such as ministries and institution. The captured data is loaded into the Input Data Warehouse (IDW).

![Figure 2.3 BPS class diagram](image)
12. The Processing Group has seven process elements which are Validation, Microediting, Automated Correction, Imputation, Derivation, Weighting, and Aggregation/Indexing, while the Analysis Group also has seven process elements which are Estimation, Macro Editing/Analysis, Output Validation, Statistical Analysis (including seasonal analysis), Microdata Confidentialization, Output Finalization, and Macrodata Loading.

13. Common services of survey class are Field Management, Input Data Warehouse, Frame and Register System, Data and metadata Management, and Survey/Project Management. Field management assists in coordinating, collaborating, and arranging enumerator/BPS field interviewer’s workload, travel and time, and input data. Input Data Warehouse is a repository of BPS data that can be accessed by authorized person with different security levels. Frame and Register System (Integrated Business Register) is a repository of samples household, business, and spatial data. Data Management and Metadata is a system to facilitate in defining and amending BPS metadata during survey cycles. Survey/Project Management provides planning, coordinating, and monitoring capability to manage end-to-end survey cycle.

III. Improving ICT and Information Management

14. ICT has been widely utilized in statistical activities within BPS from planning to dissemination. However, there are areas where ICT utilization currently is inefficient. In most cases, ICT capabilities in BPS have been developed, operated, and managed in an isolated manner by each SMA. This approach is perceived to present greater flexibility, control, and speed for SMAs in establishing ICT to support their activities. However, it also introduces barriers in statistical process integration and information sharing which in turn will affect the consistency and coherency of statistical products. This condition also introduces a complex ICT portfolio (application and platform), which raises difficulties in maintaining this portfolio.

15. Improvement of ICT should be performed as part of a whole organization considering BPS’ goals instead of looking at it individually such as by product line or by Subject Matter Area. ICT should be improved in all aspects including the alignment with BPS organization goals, statistical business requirements, information which covers data and metadata, applications and information systems including common tools like data warehouse and analytical tools for time series analysis, ICT platform and network infrastructure, information security, and governance. Improvement of ICT in all of these aspects would enable the improvement of statistical quality, in terms of accuracy, relevance, timeliness, accessibility, coherence, and interpretability.

16. ICT Enterprise Architecture would play a key role in providing pathway for ICT improvements in all those aspects. Such architecture should consider the following areas:

a. Statistical Business – statistical process requirements which drive the requirements for other components of BPS ICT Enterprise Architecture.

b. Governance and Operating Model – a view from the context of good management within broader BPS management planning, monitoring, and reporting arrangements. Management to ensure delivery of appropriate services to enable BPS staffs to do their work effectively and efficiently.

c. Information – a view of data, metadata, and information to support statistical business process and how these data flow along the process.

d. Application Systems – use and re-use of common applications systems to support statistical business processes.

e. Infrastructure – ICT platform (hardware, operating system, database management system, and other platform software), local and wide area network as well as physical computing facilities (data center and data recovery center) that support the applications.

f. Security – components required to preserve the confidentiality, integrity, and availability of information.
17. Figure 2.4 shows the overview of ICT improvement of BPS in the context of the BPS ICT Enterprise Architecture. BPS Statistical Business Process can be generalized for simple illustration in Figure 2.3. It describes data collection, processing, analysis, and publication. Survey control and monitoring process is the process to monitor and control the execution of survey from collection to publication.

18. The Governance and Operating Model has six aspects of improvements, which mainly to improve the governance and management of ICT resources and the model of interaction between ICT Support and Subject Matter Areas as well as other units.

19. Future BPS applications shall be developed based on an agile and shareable application architecture. Service Oriented Architecture (SOA) is an application design principle that emphasizes the idea to divide applications into loosely coupled services which allow use and re-use of same service components by many applications. By dividing applications into service components, it also makes the application more agile in responding to future requirements. In addition, re-using a common component will improve processing consistency among applications. This is very important in BPS to support the quality of statistics produced by the applications. In light of that, usage of SOA in the design of application is recommended. Common
applications shall be designed with flexible system interfaces to allow re-use and integration between Subject Matter Areas.

20. The data/information infrastructure consists of two separate data warehouses, the Input Warehouse and Dissemination Warehouse. Input Warehouse stores survey raw data and metadata collected and processed in all surveys, including historical survey data, while Dissemination Warehouse designed to store confidentialized version (data with no identifiable information) of survey data and metadata, mainly for publication purposes and to support other external users’ data requirements.

21. Information security is a major component of BPS ICT Enterprise Architecture. Therefore, BPS should implement security measure to preserve the confidentiality, integrity, and availability of statistical data. The goals of the security architecture are integrated access management, malware control, desktop security management, network and server security management, and security event management.

22. The last BPS ICT infrastructure is designed and developed to hosts the statistical application systems and other supporting systems and also serves as a media to connect all components of the systems. The infrastructure has several components, such as the Data and Metadata Exchange Gateway which is a system that will act as a the common gateway for data and metadata exchange with external institutions utilizing commonly used data and metadata exchange protocols such as SDMX, the Development Tools which are common and shared software development tools provided for enhancement of statistical systems, and BPS Data Center that hosts computing facilities which support survey and other statistical operations.

IV. Current Achievements and Future Plan

A. BPS Integrated Business Register (BPS-IBR)

23. BPS-IBR is an active list of Indonesian business. It will be used as a basis for business surveys intended to measure the real world Indonesian economy. Currently, BPS-IBR has around 230,000 business from 23 sectors collected from business directories maintained by SMAs in BPS. The business directories have been built and maintained by the SMAa based on Indonesia Economic Census 2006.

24. BPS-IBR provides many functionalities including business searching, browsing, matching, and profiling. The business profiling function includes adding a new business into the repository, updating existing business data, and managing relationships among businesses in term of a statistical unit model (i.e., Enterprise Group, Enterprise, and Establishment). BPS-IBR also has capability to store the history of the business updates so that it can create reports for statistics on new business and business updates for a specific period of time.

25. In near the future, BPS will initiate a formal agreement (MoU) with other government institutions that also maintain their own business directories so that the scope of BPS-IBR can be expanded. The government institutions include the Ministry of Law and Human Rights, the Ministry of Cooperatives and Small and Medium Enterprises, and Indonesia Financial Services Authority.

B. Foreign Trade Data Warehouse

26. BPS has built a data warehouse prototype for Foreign Trade Statistics. Around 400,000 records of export data are stored monthly in the data warehouse with 34 variables, while for import data there are around 850,000 records with 39 variables. Currently, the data warehouse stores Foreign Trade data from 2005 to 2010 with the total size of 400 GB and around 1.1 billion records. The data warehouse is mainly used for research and development purposes to achieve the most suitable data warehouse for BPS.

27. The data warehouse has the main functionalities as follows: (a) Storing export and import data in the same warehouse. Therefore, it can be used as a comprehensive analytic tool for foreign trade statistics; (b) Providing web based interface for drag-and-drop tabulation; (c) Providing a data editing tool to edit foreign
trade data in the warehouse; (d) Metadata driven application model; (e) Single sign on model in conjunction to others BPS application; (f) SPSS analytic tool connection.

28. In the near future we are going to expand the warehouse to house other data clusters, such as population Census 2010 and Agriculture Census 2013. Furthermore, we also extend the warehouse by adding some functionalities such as, remote execution facilities, seamless connection to BPS metadata repository, and interactive dashboarding and data visualization.

C. Data Center and Disaster Recovery Center

29. BPS has built a place for a data center with the floor area of about 379.10 m². The data center meets the Tier 2 based on TIA-942 Standard and has several room segments, such as server room, network room, utility room, and console rooms. There will be 89 Physical Servers and 7 SANs (Storage Area Networks) with the storage capacity 389 TB.

30. All critical systems, such as Survey/Project Management system, should be designed with capabilities to recover from disruption. These systems should have stand-by backup located at Disaster Recovery Center (DRC) which is ready to be activated anytime when disruption occur. BPS has planned and designed such DRC in a different island at South Borneo Province which is located approximately 921 km away from Jakarta.

V. Conclusion

31. The STATCAP-CERDAS project of BPS has just been started at the beginning of 2014. The main aim of this project is to modernizing the statistical system in Indonesia, especially in BPS. This project has four main goals, namely, improved quality of statistics, improved information and communication technology, human resources management and development, and institutional arrangement.

32. The establishment of new corporate statistical infrastructure will be based primarily on the use of ICT, and be built using knowledge of practices and experiences in other statistical offices. It will mean that common methods and tools will be used for data collection, processing, analysis, output production and dissemination. This approach will also help achieve consistent use across BPS of statistical standards and classifications. In addition to building the infrastructure and related systems, BPS will need to develop policies and processes for their use so that the goals of the STATCAP-CERDAS project can be achieved.