Data warehouse capability maturity model assessment for efficient monitoring process: a case study in National Narcotics Board

Fahmi¹, A N Hidayanto¹, Solikin², H S Indriany³, A Prastyá³ and S F Mardiansyah⁴

¹Faculty of Computer Science, Universitas Indonesia
²Universitas Bina Insani
³National Narcotics Board

Abstract. Data Warehouse/Business Intelligence nowadays has become a necessity for every organization. Data analysis supports in making decisions and policies. The problem is the data warehouse management process, both related to ETL, data modelling, and data architecture, which results in the non-optimal and inefficient dashboard that wastes energy, as a data monitoring tool in displaying important information needed by the organization. This research was conducted by providing an assessment to the working group team in the data warehouse section using the Spruit & Sacu Data Warehouse Capability Maturity Model (DWCMM) to measure the maturity of the data warehouse at the National Narcotics Board. The results showed that the DWCMM maturity level is at the repeatable level, from the measurement results it can be seen that the worst DW category is in the Service Processes component so that many improvements need to be made to increase the DWCMM maturity level. As a recommendation, it is necessary to design a virtual integrated DW architecture that has real-time updates, data quality system, automatic synchronization of data modelling, standardization of metadata, real-time ETL monitoring management, closed-loop & BI real-time BI application, and improves the development process and process services in the management of DW/BI at BNN to become more efficient.

1. Introduction

Nowadays, almost all organizations need data analysis which is one way of making decisions. Data analysis, processing, and visualization are the main components of a Business Intelligence (BI) [1]. The application of Data Warehouse and Business Intelligence (DW / BI) is a stage where an organization manages data as property or asset [2]. Huge transactional data requires a high-scale server to process it to ensure the quick output information. Data warehouse was developed to handle big size structured data that confirming efficient process to realise energy efficiency.

One Data Indonesia Portal (SDI), namely data.go.id which is the official Indonesian open data portal, through this portal all public data collections can be accessed easily and quickly. The policy on SDI has been regulated [3]. This is one form of government attention in terms of data management. As one of the government agencies in charge of the Prevention of Eradication of Narcotics Abuse and Illicit
Trafficking (P4GN) [4] the National Narcotics Board (NNB) also manages data management, which in this organization has also managed the DW/BI.

In terms of data warehouse management, there are problems related to ETL, data modeling, and data architecture, among others, namely in terms of data cleansing, data aggregation, data query, data integration, and response time which results in the dashboard being not optimal as a data monitoring tool in displaying important information needed by the organization. Based on the software maintenance report [5], real problems that often occur include inefficient queries in data warehouses that affect data staging and data marts, complex data sources that need adjustments to stage data, scheduler ETL on ODI (Oracle Data integrator) that requires improvements that affect the response time and data display on the dashboard. Based on these problems, it is important for NNB to know the current conditions and know the steps that need to be taken to make continuous improvements. As a research question, what is the maturity level of the data warehouse at NNB? This study aims to determine the level of maturity of DW in organizations to assist organizations in defining and understanding processes in the data warehouse so that strategic steps can be drawn up to improve DW and determine policies related to data warehouse/business intelligence.

From previous studies on data warehouse/business intelligence maturity, as in [6] which discusses the Business Intelligence Maturity Model (BIMM) and proposes the Service-Oriented Business Intelligence Maturity Model (SOBIMM) to be able to help find solutions to problems such as integration, and quality issues, [7] developed and evaluated an impact-oriented theoretical model of BI maturity aimed at integrating the methodical strengths of IS and BIMM theory. [8] evaluated the maturity level of Business Intelligence (BI) in Polish organizations, showing that BI systems can make more effective decisions and improve business processes/business performance and in [9] develop a Business Intelligence Maturity Model for government organizations in Malaysia. The difference between this study and previous research is that this study measures DWCMM in government agencies using Spruit & Sacu.

The structure of this article is: the first part is the introduction, and the second part is the literature review. The third part presents the research methods used in this research. The fourth part, results, and discussion. The fifth part is the recommendation, and the sixth part is the conclusion of the research and future work.

2. Literature Review

2.1. Data Warehouse (DW) /Business Intelligence (BI)

The phrase Business Intelligence (BI) has two meanings. Firstly, it means a data analysis purposed at knowing organizational occupations and occasions. The findings of the analysis are used to develop the success rate of the organization. Secondly, Business Intelligence denotes the collection of technologies that backing this data analysis. The development of decision support tools and BI tools makes a query, data mining, statistical analysis, reporting, scenario modelling, data visualization, and dashboards possible [10].

A data warehouse (DW) is a mixture of two main elements: a combined determination backing database and related application programs expended to collect, clean, transform and store data from various operations and external sources. To support historical, analytical, and BI needs, a data warehouse can also contain a data mart, which is a copy of a subset of the warehouse data. In the broadest setting, the data warehouse contains all data storage or data extraction used to support data distribution for BI determinations. Data warehousing describes the process of extracting, cleansing, transforming, controlling, and loading operations to maintain data in the data warehouse [10].

A data warehouse/business intelligence system is a query and reporting system that aims to improve operational activities in the form of decision-making support, compliance, operational reporting, performance management, and integration between analysis results and predictive modelling [11].
2.2. Approaches to Data Warehousing
There are two common approaches to data warehouses, namely Bill Inmon and Ralph Kimball. Inmon defines the data warehouse as the subject-oriented, integrated, time-varying, and non-volatile data set to aid the organization's decision-making process [12]. Kimball describes a warehouse as a duplicate of transaction data especially structured for querying and analysis [13].

2.3. Data Warehouse/Business Intelligence Maturity Model (DW/BI MM)
By measuring the DW/BI maturity, it can be seen how the DW/BI condition of an organization is. There are several maturity models to measure DW/BI. [14] this BIMM model from The Data Warehouse Institute (TDWI) focuses on technical aspects, especially on data warehouse aspects. It consists of six maturity stages and uses the metaphor of human growth: prenatal, infant, child, teenager, adult, and sage. It has eight maturity components, namely: scope, analytical structure, executive perceptions, types of analytics, stewardship, funding, technology platforms, and change and management administration.

Gartner's BI maturity model emphasizes three main areas, explicitly people, processes, and metrics or technology at five maturity stages, explicitly: unaware, tactical, focused, strategic, and pervasive. The model helps assess the business maturity and the maturity of each department. This model is a tool to evaluate the maturity of the company's BI and PM efforts and the maturity required to achieve business goals [15].

The DWCMM delivers a rapid method for companies to evaluate their DW/BI maturity and objectively [16]. The DWCMM developed by Spruit & Sacu was chosen because it offering the Data Warehouse Capability Maturity Model (DWCMM) which evaluates two sides, is technical and organizational concerned in evolving the data warehouse. On the technical aspects which are the main components in the assessment of DW technical solutions, namely architecture, data modelling, ETL, and BI application, in which this technical solution assesses the technology side more deeply in each of its characteristics compared to other models. While the organizational aspects need to be assessed because in addition to technical solutions, the procedures and functions concerned in the project, namely development processes and service processes, are supporting components that cannot be separated and are sustainable from the technical aspects. The assessment of the components of the technical aspect is in line with the technology implemented by NNB, and the components of the organizational aspect are also inseparable from the process of data warehouse management activities at NNB. In addition, DWCMM provides a detailed maturity matrix. DWCMM be able to be used to aid organizations to evaluate Data Warehouse solutions in their current state and provide guidance for future improvements. This research is expected to contribute to the maturity level data warehouse in government organizations.

Comparison between three DW/BI MM at the maturity level and the developed dimensions can be seen on Table 1.

2.4. Data Warehouse Capability Maturity Model (DWCMM) Spruit & Sacu
The Spruit & Sacu DWCMM framework developed at Utrecht University has a model that can be used to measure data warehouse maturity. This data warehouse maturity model defines five maturity levels, explicitly initial at level 1, repeatable at level 2, defined at level 3, managed at level 4, and optimized at level 5 [16]. To determine the maturity level of the data warehouse, this framework has 6 components. Of the 8 components, it is further elaborated into 50 characteristics and these characteristics are parameters or references to measure the maturity value of the data warehouse at each level. Here are 6 DW components that are in the framework developed by Spruit & Sacu:

1) General Architecture and Infrastructure: DW architecture contains: three core elements (ETL, BI, data modelling), some data storage elements (DW database, data marts, data staging area, source system, operational data store), and the fundamental components such as security, infrastructure, and metadata, that assistance the movement of data from the system to end-users. The characteristics of this component include performance optimization, infrastructure, conceptual architecture, and its
layer, DW data sources, business policies, metadata and security management, and update frequency.

2) Data Modelling: It is a collection of models that be able to define the structure and operations of a database. The characteristics of this component are design levels, tools, standards, metadata management, dimensional modelling, and synchronization between all the data models.

3) ETL (Extract Transform Load): This is the process of extracting data from external sources, changing the data somewhat according to the requirements, and loading the data into the target database. The main characteristics of this component include complexity, data quality system, management, and monitoring, tools, metadata management, and standards.

4) BI Application: Commonly called “front-end” tools/dashboard. The characteristics of this domain are types of BI application, delivery methods, tools, metadata management, and standards.

5) Development Processes: The influencing characteristics are: CMM levels, project planning, and management, DW/BI sponsor, DW project team and roles, requirements definition, testing and acceptance, development/testing/acceptance/production environments, DW quality management, knowledge management, and standards.

6) Service Processes: The characteristics of this component are the management of service quality knowledge, service level, incident, release, technical resource, availability, and change.

| Authors | Maturity Models | Maturity Levels | Component Maturity |
|---------|----------------|----------------|--------------------|
| The Data Warehousing Institute, 2007 | Business Intelligence Maturity Model (BIMM) | 1. Prenatal 2. Infant 3. Child 4. Teenager 5. Adult 6. Sage | 1. Scope 2. Analytic Structure 3. Executive Perceptions 4. Types of Analytics 5. Stewardship 6. Funding |
| Spruit & Sacu, 2015 | Data Warehouse Capability Maturity Model (DWCM) | 1. Initial 2. Repeatable 3. Defined 4. Managed 5. Optimized | 1. Architecture 2. Data Modelling 3. ETL 4. BI Application 5. Development Processes 6. Service Processes |
| Gartner & Hostmann, 2007 | Business Intelligence and Performance Management Maturity Model (BI&PMMM) | 1. Unaware 2. Tactical 3. Focused 4. Strategic 5. Pervasive | 1. People 2. Processes 3. Metrics and Technology |

2.5. Theoretical Framework
The following Figure 1 is the theoretical framework in this research. As a framework used to measure the level of DW in this research is the framework of Spruit & Sacu to determine the level of DWCMM. Recommendations are obtained from the gap between the target and the existing condition.
3. Research Methodology

The research methodology used follows the research stages as shown in Figure 2, including problem identification, framework determination, data collection, data processing, analysis, and recommendations.

3.1. Research Instrument

The research instrument is based on 50 questions from 6 DW Spruit & Sacu domains [17]. DW Technical Solution (32 questions) - involves several questions in sub-classifications:

1) General Architecture and Infrastructure, divided into 9 questions
2) Data Modelling, divided into 9 questions
3) Extract Transform and Load, divided into 7 questions
4) Business Intelligent Applications, divided into 7 questions

DW Organization & Processes (18 questions) - consists of several questions in the following sub-categories:

5) Development Process, divided into 10 questions
6) Service Process, divided into 8 questions

3.2. Data Collection

Spruit & Sacu DWCMM components as parameters to perform measurements that aim to determine the condition of the data warehouse. By gathering a working group team consisting of three computer administrators as the person in charge of DW at NNB. The three team members are members of a working group that together answer the DWCMM assessment. Furthermore, data collection is carried out by collecting and exploring data related to the condition of the organization and the data warehouse management process. The objectives are, among others, to clarify the description of the evaluation/measurement results that have been carried out [17] and to increase the cogency of the data in qualitative research by discovering the fact of specific information across various techniques and sources of data acquirement [18]. By studying Presidential Regulations, Agency Regulations, manual books, and report documents, as well as studying dashboards, data warehouses, Oracle Data Integrator (ODI), and applications at NNB, we get the current condition of NNB which has implemented DW/BI for 4 years. Three employees are responsible for DW/BI with data warehouse maintenance carried out throughout the year by a third party. The technology used is:

1) Data modelling: SQL Developer, Draw.io
2) ETL (Extract Transform Load): ODI (Oracle Data Integrator
3) BI Applications: Tableau
4) Database: Oracle

3.3. Data processing method

From the calculation results, the DWCMM Spruit & Sacu obtained guidelines for making improvements. Furthermore, recommendations will be given to NNB regarding things that must be done to increase the maturity level of DW and be able to answer the problems that have been identified previously. In this case, the characteristics with values 1 and 2 are considered weak and need to be improved.

4. Result and Discussion

DW's maturity level is based on the results of an assessment/interview on current conditions. The results of the assessment are presented in Table 2, it can be seen that of the 50 characteristics in the six maturity components of the Spruit & Sacu DWCMM, there are 20 characteristics (40.00%) that are considered weak, and 30 characteristics (60.00%) are strong enough.

| Component Maturity                  | Number of Questions | Weak Characteristics | % Weak Characteristics |
|-------------------------------------|---------------------|----------------------|------------------------|
| DW Technical Solution               | 9                   | 1                    | 11.11%                 |
| Data Modelling                      | 9                   | 3                    | 33.33%                 |
| ETL                                 | 7                   | 3                    | 42.86%                 |
| BI Application                      | 7                   | 2                    | 28.57%                 |
| Development Processes               | 10                  | 4                    | 40.00%                 |
| Service Processes                   | 8                   | 7                    | 87.50%                 |
| **Total**                           | **50**              | **20**               | **40.00%**             |

As for the calculation results for each maturity component, from five data levels of DWCMM Spruit & Sacu, namely: initial, repeatable, defined, managed, and optimized, the DW maturity level is found to be at the repeatable level. There are two maturity components at the defined level, namely Architecture and BI Application, three maturity components at the Repeatable level, namely Data Modelling, ETL, and Development Processes, and one maturity component at the Initial level, namely Service Processes. As the worst DW category is Service Processes. As shown in Table 3, the maturity level of the Service Processes component maturity level is initial with a capability value of 1.75.

| No | Component Maturity       | Capability Level | Maturity Level | Target |
|----|--------------------------|------------------|----------------|--------|
| 1  | Architecture             | 3.78             | Defined        | 5      |
| 2  | Data Modelling           | 2.89             | Repeatable     | 5      |
| 3  | ETL                      | 2.71             | Repeatable     | 5      |
| 4  | BI Application           | 3.00             | Defined        | 5      |
| 5  | Development Processes    | 2.60             | Repeatable     | 5      |
| 6  | Service Processes        | 1.75             | Initial        | 5      |
|    | **Average**              | **2.79**         | **Repeatable** |        |

Of the 6 maturity components of Spruit & Sacu DWCMM, the average overall component capability is 2.79 and with maturity at the Repeatable level. This data warehouse maturity measurement is the first DW maturity measurement carried out at NNB.
Figure 3. Radar Chart per component of DWCMM capability.

From the radar chart, it can be seen that the organization has a big target for the data warehouse, considering that NNB is a vertical organization that must present data up to the regional or district level throughout Indonesia. Some characteristics are still weak because, in general, the process management in the government is still weak. The following are the components of DWCMM with characteristics that are still considered weak:

1) Data Modelling: There is no documentation for the data model yet, several fact tables have granularity at the lowest level achievable, and only several dimensional modelling tables are proposed with replacement keys and complex hierarchies.

2) ETL (Extract Transform Load): ETL still with simple transformations such as: changing formatting, sorting, filtering, merging, and aggregating. There is no data quality system yet, management and monitoring of ETL is still in the form of simple monitoring.

3) Development Processes: DW quality management is still in the ad-hoc stage, Knowledge sharing in the DW development process is organized through written documentation (e.g., knowledge database, intranet, wiki), and besides overtraining and mentoring programs. In the DW project, the definition of requirements is still ad-hoc and has not used a specific methodology. The types of testing used include acceptance testing, regression testing, and system integration testing.

4) Services Processes: Knowledge sharing in the DW services process is organized through written documentation (e.g., knowledge database, intranet, wiki), and also across training and mentoring programs. The customer service requirements are recorded and formulated in standard procedures. Incident management is done on an ad-hoc basis without a dedicated ticket handling system or service desk to evaluate and categorize them before referring them to specialists. Change requests are created and resolved on an ad-hoc basis. Ad-hoc resource management activities (only in case of problems). Availability management is still ad-hoc.

5. Recommendation

Characteristic gaps are obtained from expectations or targets compared to current conditions based on the results of the assessment/interview. The recommendations given by Spruit & Sacu are:

1) Architecture: designing a virtual integrated DW architecture, designing a web-accessed central metadata warehouse that is integrated, standardized, up-to-date metadata, implementing combined
company-wide authorization security, increasing data source support to DW with various types of unstructured data sources and web data source, using DW specialized appliances to update DW in real-time.

2) Data Modelling: Using standardized data modelling tools to manage metadata, using automatic synchronization on all data models, designing all data models to have conceptual, logical, and physical levels, standardizing documentation for all data models, rather all fact tables have granularity at the stages as low as achievable, implementing enterprise-wide standardized data model dimensions for all business processes.

3) ETL: Using complete ETL generated form metadata, using real-time ETL, implementing a data quality system on ETL, managing and monitoring ETL in real-time, establishing, implementing, and documenting ETL management standards, managing metadata management for all ETLs.

4) BI Application: Using closed-loop BI systems and real-time BI systems, using one BI tool for mainstream BI (reporting and visualization application) and one BI tool for special BI application (data mining, financial analysis, etc), setting, implement and document standards for BI applications, all BI applications use standard objects and patterns, perform complete metadata integration with BI applications (accessed by one button push on attributes) as a way to access metadata, use highly interactive, the business process-oriented, up-to-date portal (without separation among operational and BI portals) as BI application technique.

5) Development Processes: Improve the continuous improvement process in categorizing weaknesses and strengthening processes proactively, by the aim of avoiding faults from occurring; establish/implement and document standards for development, testing, and implementation of DW functionality (ie: ETL and BI applications); implementing DW quality management by conducting causal analysis meetings to identify the causes of defects, conducting service quality management certification; perform project planning and scheduling, project risk management, project control, standard and efficient procedures and documentation, and conduct evaluation and assessment; conduct periodic evaluations also provide performance assessments for each role and responsibility; improve knowledge sharing continuously; using all major test types (regression testing, acceptance testing, unit testing by another; system integration testing); user training; standard procedures and documentation; external evaluations and reviews.

6) Service Processes: Implement DW service quality management in the organization by conducting causal analysis meetings to identify the causes of defects, implementing service quality management certification; improve knowledge sharing continuously; implement real service delivery that is constantly observed and evaluated to customers on an event-driven basis for continuous progress improvement; perform trend analysis in the incidence of incidents as well as in customer satisfaction and perceived benefit of the services delivered; perform trend analysis and statistics concerning the happening of changes, success rates, customer satisfaction, and perceived value of the services delivered; conduct analysis and monitoring of resource management trends to define the most popular bottlenecks and ensure that there is enough capability to assistance the scheduled services; perform analysis and planning of availability management trends to define the most bottlenecks and ensure that all components are available for agreed service level goals; perform announce management trend analysis, statistics and planning. Meanwhile, based on the characteristics of DWCMM, the strategies that must be carried out by the organization to improve the data warehouse are:

1) Designing a virtual integrated DW architecture that real-time updates with a web-accessed central metadata repository
2) Designing enterprise-wide standards and automatic synchronization data modelling
3) Standardize metadata
4) Create real-time ETL monitoring management
5) Using closed-loop & BI real-time BI application.
6) Create a data quality system
7) Improve knowledge sharing and DW service quality management
8) Continuously improve development processes
9) Improve service.

6. Conclusion and Future Works

6.1. Conclusion

The results of DW maturity measured using the Spruit & Sacu model show that in general, the DW maturity level is at the repeatable level with an average maturity level of 2.79. The architecture maturity component is at the managed level (3.78), the data modelling maturity component is at the repeatable level (2.89), the ETL (Extract Transform Load) maturity component is at the repeatable level (2.71), the BI Application maturity component is at a defined level (3.00), Development Processes maturity component is at repeatable level (2.60), and Service Processes maturity component is at initial level (1.75). The results of the DW component assessment are based on the current DW conditions at NNB. From the gap between the target and the existing conditions at DWCMM Spruit & Sacu, recommendations were obtained as a strategy to improve the data warehouse in the organization. To answer problems related to data warehouses, it is necessary to design a virtual integrated DW architecture that real-time updates, automatic synchronization data modelling, metadata standardization, real-time ETL monitoring management, closed-loop & BI real-time BI applications, data quality systems and improve the development process and service process.

6.2. Future Works

This study has a limited scope on DW/BI managed by NNB which focuses on DW/BI to support the main application, not including supporting applications that have not been managed. Further research is expected to be able to develop a data warehouse design and implementation that is more complex and comprehensive than the NNB system and can periodically measure DW/BI, so that with improvements and improvements that have been made it can be seen the development of DW/BI organizations.

References

[1] Naneva V and Stefanova K 2021 Thermophys. Basis Energy Technol (Thet 2020) 2337 070010
[2] Halim S, Mubarokah I and Hidayanto A N 2020 Proc. - 2nd Int. Conf. Informatics, Multimedia, Cyber, Inf. Syst. ICIMCIS 2020 313–318
[3] Anonymous 2019 Presidential Regulation of the Republic of Indonesia Number 39 of 2019 concerning One Indonesian Data (Republic of Indonesia: Deputy for Economic Affairs)
[4] Anonymous 2009 Law of the Republic of Indonesia Number 35 of 2009 Concerning Narcotics (Republic of Indonesia: Minister of Law and Human Rights)
[5] Anonymous 2020 Software Maintenance Report (Republic of Indonesia: BNN)
[6] Shaaban E, Helmy Y, Khedr A and Nasr M 2011 Int. Arab Conf. Inf. Technol. (ACIT 2011) 11–14
[7] Lahrmann G, Marx F, Winter R and Wortmann F 2011 44th Hawaii International Conference on System Sciences IEEE 11836578
[8] Olszak C M 2013 Fed. Conf. Comput. Sci. Inf. Syst. FedCSIS 2013 951–958
[9] Zahrullaili M H and M F Noordin 2018 Proc. - 2018 4th Int. Conf. Inf. Retr. Knowl. Manag. Diving into Data Sci. CAMP 2018 223–228
[10] Dama 2017 DAMA-DMBOK Data Management Body of Knowledge (Chicago and Philadelphia: DAMA International)
[11] Loshin D 2011 Master Data Management and Data Quality BT - The Practitioner’s Guide to Data Quality Improvement (Burlington: Elsevier)
[12] Inmon W H 2005 Building the Data Warehouse (New Jersey: John Wiley & Sons)
[13] Kimball R and Ross M 2013 The Data Warehouse Toolkit The Definitive Guide to Dimensional Modelling Third Edition (Indianapolis: John Wiley & Sons)
[14] Eckerson W W 2007 Beyond the Basics : Accelerating BI Maturity (The Data Warehousing Institute)
[15] Rayner N and Schlegel K 2008 Maturity Model Overview for Business Intelligence and
Performance Management (Gartner: Stamford)

[16] Spruit M and Sacu C 2015 *J. Univers. Comput. Sci.* **21** 1508–1534
[17] Adami M F and Kiger A 2005 *Nurse Res.* **12** 19–29
[18] Walidin W, Saifullah and Tabran 2015 *International Journal of Business and Management* **9** 224–233