Expanding the Discipline of Enterprise Architecture Modeling to Business Intelligence with EA4BI

Rudi Claes

Inno.com Institute, Beerzel, Belgium

Abstract. The current mainstream enterprise architecture modeling languages provide little syntax to describe the primary goal of any business intelligence (BI) project, i.e. the transformation of data into information and knowledge. As a consequence, modeling in most BI projects is done in isolation with the focus on design instead of architecture, leading to poor alignment with enterprise architecture and the goals it seeks to realize. This paper introduces a new enterprise architecture framework for BI called EA4BI. This framework addresses these issues by expanding the discipline of enterprise architecture modeling to BI. As an enterprise architecture framework and toolkit focused on BI, it facilitates and empowers end-to-end BI modeling, but also provides linkage to other enterprise architecture domains and modeling languages.

Keywords: enterprise modeling, business intelligence, enterprise architecture, BI architecture, framework, toolkit.

1 Introduction

Business intelligence (BI) is an umbrella term that includes the applications, infrastructure, tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance [Gartner IT Glossary, 2017]. In order to realize a successful BI project of any scope within the context of any organization, this requires close cooperation and mutual understanding between the organizational entities producing the information (IT), analyzing the information (Business), and taking decisions based on the information (Management). Business and Management also need to be made clear that information is not readily available as raw material, but rather the result of an often-complex production process tailored to the specific environment of the organization. In order to explain this core process without having to reveal its internal workings and practical complexities, IT professionals commonly resort to using the Data, Information, Knowledge, Wisdom (DIKW) hierarchy [Ackoff, 1989].

DIKW facilitates in delivering the key message, that adding relevant context to raw data results in information. With proper analysis of this information, and appropriate presentation and visualization, it gets the potential to yield knowledge. This in turn, has the potential to support and streamline the decision-making process.

J. Ralyté, B. Roelens, and S. Demeyer (Eds.):

PoEM 2017 Doctoral Consortium and Industry Track Papers, pp. 50-57, 2017.

When Business and Management are able to identify and acknowledge this as one of the primary goals of the BI project, IT can obtain its mandate to allocate sufficient resources on this topic, paving the way for a successful realization.

BI project members quickly face the complexity of gaining the required insights in, and actively managing the overview of, both the problem space and the solution space related to the core process of turning data into information and knowledge. This problem is not only related to BI projects though, but rather inherent to IT projects having properties such as broad organizational scope, large technical complexity, cross-functional business domains, and/or requiring a substantial amount of enterprise application integration. Large BI projects usually include all the aforementioned properties, in addition to a set of critical success factors (business driven and iterative development approach, user-oriented change management, sustainable data quality and integrity, …) unique to the realization of BI systems [Yeoh et al, 2009].

This paper introduces a new four-layered enterprise architecture framework for BI called EA4BI, that attempts to address these issues. The rest of this paper is structured as follows: section 2 gives an overview of the current usage of enterprise modeling in real-life BI projects and highlights some of the issues that are present. Section 3 introduces the EA4BI framework. We conclude in section 4.

2 Enterprise modeling in BI projects

Software development processes such as the Rational Unified Process [Rational Software Corporation, 1998] have been able to provide adequate countermeasures against the aforementioned issues present in both problem space and solution space of IT projects, of which the principle of applying visual modeling languages, more specifically the Unified Modeling Language [Object Management Group, 1997], has proven to be very effective [Kruchten, 2004]. Subsequent enterprise modeling languages such as the Business Process Model and Notation (BPMN) [Object Management Group, 2008] and ArchiMate [Open Group, 2009] have been able to successfully re-apply the same principle in their respective areas of business process modeling and enterprise architecture.

Therefore, it is not surprising that, in addition to the obvious usage of classic data modeling [ANSI, 1975], these current mainstream enterprise modeling languages have also been applied in BI projects. However, this leads to a number of interesting observations.

The first observation is the limitation of linguistic expressivity towards the BI spectrum, both in problem space as well as in solution space. As ArchiMate, BPMN, and UML do not have BI as a focus area, the number of provided language elements, capable of describing the problem space (i.e. elaborating the core BI process of turning data into information and knowledge), is limited. For the same reason, none of those languages is sufficiently capable of elaborating the solution space and identifying an appropriate BI architecture and a resulting contribution to the BI asset library of the organization.

This leads to a second observation: enterprise modeling in most BI projects is done in isolation and with the focus on design (white-box view revealing the implementation details), instead of architecture (black-box view still hiding the implementation details). As BI project members struggle with the issues above, they frequently give up early on capturing the problem space and the solution space into enterprise models and viewpoints (often stating it is too difficult and/or too time-consuming). They usually resort to performing requirements management and business process modeling framed within the scope of the BI project in the problem space, and translate the results into detailed design models in the solution space while ignoring the BI architecture.

All aforementioned issues lead to a third observation: as these BI project modeling artefacts are limited to the scope of the project, have a low strategy focus and favor a high technology depth, this leads to poor alignment with enterprise architecture and the goals it seeks to realize [Ross et al, 2006]. This affects both Management, Business and IT, as it is their combined responsibility to define a suitable BI strategy resulting into a roadmap for a consistent set of concrete BI projects supported by an appropriate BI architecture, that produce valuable additions to the BI asset library of the organization [Deloitte, 2009].

The objective of EA4BI is to address these issues by expanding the discipline of enterprise architecture modeling to BI. As an enterprise architecture framework and toolkit focused on BI, it facilitates and empowers end-to-end BI modeling, but also provides interfaces with other enterprise architecture domains and modeling languages.

3 EA4BI Framework and Toolkit

3.1 EA4BI Overview

This paper proposes a four-layered enterprise architecture framework for BI (see Figure 1), taking into consideration the three major perspectives enterprise architecture offers and adding a fourth one that allows for realizing the connection with technical design. The four layers are enterprise architecture, solution architecture, technical architecture and technical design. This paragraph positions the proposed EA4BI viewpoints in this frame of reference; the viewpoints themselves are discussed in more detail below in this section.

Enterprise Architecture (EA) takes place at a high level in the organizational structure. Strategy and concepts are strongly focused in this discipline, with very limited attention for concrete technology choices and implementation details. In order to support this conceptual level, EA4BI provides the viewpoints of Ross & Weill Core Diagram [Ross et al, 2006] and Analytical Master Data Management [Allen et al, 2015].

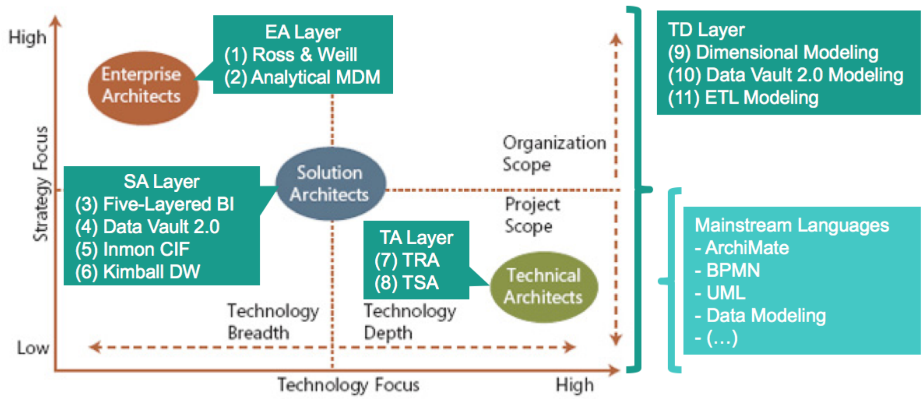

Fig. 1. EA4BI Overview

Solution Architecture (SA) is deliberately positioned at the center. This discipline translates the enterprise architecture into a time-based project portfolio and facilitates the selection of appropriate technology, thereby allowing for the implementation of the organizational strategy. In order to support this concrete level, EA4BI provides the viewpoints of Five-Layered BI Architecture [In Lih Ong, 2011], Data Vault 2.0 Architecture [Linstedt et al, 2016], Inmon Corporate Information Factory (CIF) [Inmon et al, 2000], and Kimball Data Warehouse Architecture [Kimball et al, 2013].
Technical Architecture (TA) takes a deep dive into the technology for a given project scope, as such preparing it for detailed design and implementation. In order to support this practical level, EA4BI provides the viewpoints of Technical Reference Architecture (TRA) and Technical Solution Architecture (TSA).

The Technical Design (TD) layer enables the linkage of detailed technical design with the previous architectural layers. Where the latter offers guidance and black-box views (still hiding the implementation details), TD offers a white-box view (revealing the implementation details) and actually controls the implementation done by the development team. EA4BI offers the viewpoints of Dimensional Modeling [Kimball et al, 2013] and Data Vault 2.0 Modeling [Linstedt et al, 2016], in order to provide support for both methodologies. The Extract Transform Load (ETL) viewpoint allows to describe the complexities of data retrieval, data processing and data storage.

3.2 EA4BI Framework Design Principles

EA4BI applies the framework design principles of KISS, traceability and extensibility. The popular acronym KISS should be interpreted in this paper as "Keep It Small and Simple". First, KISS can be found in the fact that the number of elements and relationships is limited to a strict minimum through "need to have" selection strategy. This has been achieved by categorizing the complete set of elements for all EA4BI viewpoints, and assigning a representative structural/behavioral element to each category. The same method has been applied for the relationships, resulting in the assignment of a relationship type and optional stereotype. Secondly, KISS manifests itself visually in the sense that the resulting elements and relationships have a recognizable design, given the fact that they are based on a UML profile (see Figure 2).

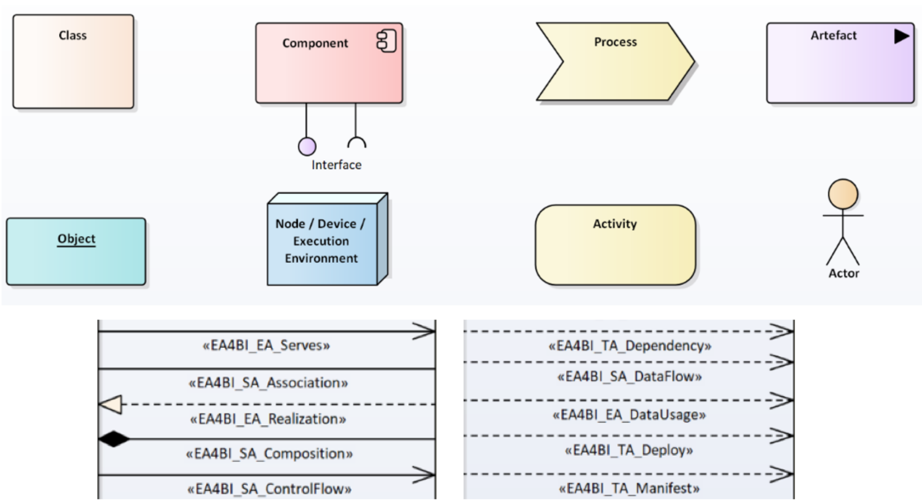

Fig. 2. EA4BI Elements and Connectors

A model is a graph of elements connected by relationships. Traceability is the ability to explore this graph from any given starting point in any direction [Aizenbud-Reshef et al, 2006]. EA4BI provides an appropriate set of relationships that offer full traceability between all the elements present on all the viewpoints that are part of the four layers. Traceability between the three architectural layers is achieved by means of the relationship types "composition" and "realization". Traceability towards the technical design layer is realized by means of a "dependency".

EA4BI offers two extensibility features: association (same level of detail) and decomposition (higher level of detail). The former would manifest itself by creating an EA4BI relationship between an EA4BI element/diagram and a non-EA4BI element/diagram. The latter would result in an EA4BI element having a non-EA4BI composite diagram depicting the substructure. As such, both extensibility features can be used to provide linkage to mainstream enterprise modeling languages such as ArchiMate, BPMN, UML, and data modeling. A coordinated application of these features can also result in extending EA4BI with organization-specific and/or industry-specific modeling policies and practices.

3.3 EA4BI Viewpoints

The first version of EA4BI offers 11 viewpoints, covering all four layers of the framework. This paragraph gives an overview and provides the rationale for incorporating the viewpoints.

Enterprise Architecture:

1. Ross & Weill Core Diagram: a simple one-page picture that provides a high-level view of the process, data, and technologies constituting the desired foundation for execution [Ross et al, 2006]. It helps Management, Business and IT to understand the required enterprise architecture for the organization. The early identification of processes and data provides a good starting point the definition of a BI architecture
2. Analytical Master Data Management: its primary purpose is to improve analytics, reports, and business intelligence in general [Allen et al, 2015]. An analytical master data capability, adequately managed by Business and IT, results in the identification and the provisioning of a single point of reference for the critical data, required by an organization in order to support decision making by Management. As such, it elaborates and details the core data present on the Ross & Weill Core Diagram, and represents an important foundational building block of the BI architecture

Solution Architecture:

3. Five-Layered BI Architecture: a reference architecture for BI [In Lih Ong, 2011]. As it is methodology-neutral, conceptual by nature, and defined at the level of solution architecture, it can serve as a good starting point for the elaboration of the BI solution space and the resulting BI architecture, without having to make an early "dogmatic" choice between the major BI methodologies represented by the other viewpoints in this architectural layer
4. Data Vault 2.0 Architecture: the solution architecture component of the Data Vault 2.0 methodology [Linstedt et al, 2016]. It allows for the creation of a BI architecture based on this methodology
5. Inmon Corporate Information Factory (CIF): the solution architecture component of the Inmon Data Warehouse 2.0 methodology [Inmon et al, 2000]. It allows for the creation of a BI architecture based on this methodology
6. Kimball Data Warehouse Architecture: the solution architecture component of the of the Kimball Data Warehouse Toolkit methodology [Kimball et al, 2013]. It allows for the creation of a BI architecture based on this methodology

Technical Architecture:

7. Technical Reference Architecture (TRA): allows an organization to define a platform and technology dependent technical BI reference architecture, as such providing an overarching guidance to all BI projects
8. Technical Solution Architecture (TSA): elaborates the platform and technology dependent technical BI architecture for a given project scope. It can be based on the aforementioned TRA or created ad hoc

Technical Design:

9. Dimensional Modeling: a database design technique for modeling data warehouses and data marts, and part of the Kimball Data Warehouse Toolkit [Kimball et al, 2013]. It allows for the creation of technical designs based on this methodology
10. Data Vault 2.0 Modeling: a database design technique for modeling data vaults, and part of the Data Vault 2.0 methodology [Linstedt et al, 2016]. It allows for the creation of technical designs based on this methodology
11. ETL Modeling: allows depicting the Extract Transform Load (ETL) processes that enable the data flows in the BI ecosystem of an organization. The purpose of this viewpoint is not to describe all ETL flows, but to select those that expose high complexity and/or have a major impact on the BI landscape

3.4 EA4BI Toolkit

The EA4BI framework is based on a UML profile and can thus be depicted as a UML profile diagram. UML modeling tools, capable of performing a Model-Driven Generation (MDG) based on UML profile diagrams, can create a tool-specific toolkit. At the time of writing this paper, this has been done with the Sparx Enterprise Architect tool (see Figure 3).

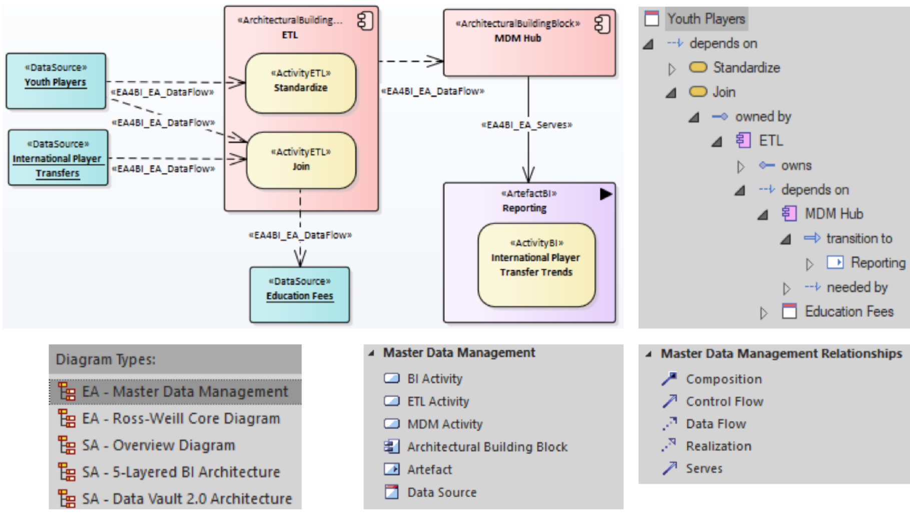

Fig. 3. Sample EA4BI MDM diagram and Sparx Enterprise Architect Toolkit

4 Conclusion

This paper has introduced EA4BI, a four-layered enterprise architecture framework for BI, taking into consideration the three major perspectives enterprise architecture offers (enterprise, solution, technical) and a fourth one providing linkage to technical design. It expands the discipline of enterprise architecture modeling to BI, by adding more linguistic expressivity towards the BI spectrum and offering better alignment with enterprise architecture. This by offering an integrated set of 8 BI architecture and 3 BI design viewpoints. Though it is built based on real-life BI cases, the framework still needs further validation by the BI community.

5 References

1. Ackoff, R.L.: From data to wisdom. Journal of Applied Systems Analysis 16, 3--9 (1989)
2. Allen, M., Cervo, D.: Multi-Domain Master Data Management. Morgan Kaufmann, San Francisco (2015)
3. Aizenbud-Reshef, N., Nolan, B.T., Rubin, J., Shaman-Gafni, Y.: Model traceability. IBM Systems Journal, vol. 45/3 (2006)
4. American National Standards Institute: ANSI/X3/SPARC Study Group on Data Base Management Systems; Interim Report (1975)
5. Prashant, P.: Business intelligence (BI) How to build successful BI strategy. Deloitte (2009)
6. Retrieved from: https://www.gartner.com/it-glossary/ on 08/11/2017
7. In Lih Ong: A Five-Layered Business Intelligence Architecture. Communications of the IBIMA vol. 2011 (2011)
8. Inmon, W.H., Imhoff, C., Sousa, R.: Corporate Information Factory, Second Edition. John Wiley and Sons, New York (2000)
9. Kimball, R., Ross, M.: The Data Warehouse Toolkit, Third Edition. John Wiley and Sons, New York (2013)
10. Kruchten, P.: The Rational Unified Process: an introduction. Addison-Wesley Professional, Boston (2004)
11. Linstedt, D., Olschimke, M.: Building a Scalable Data Warehouse with Data Vault 2.0. Morgan Kaufmann, San Francisco (2016)
12. Object Management Group: Unified Modeling Language (UML), http://www.uml.org
13. Object Management Group: Business Process Model and Notation (BPMN), http://www.bpmn.org
14. Open Group: ArchiMate, http://www.opengroup.org/subjectareas/enterprise/archimate
15. Rational Software Corporation: Rational Unified Process. Rational Software White Paper TP026B, Rev 11/01 (1998)
16. Ross, J.W., Weill, P., Robertson, D.C.: Enterprise Architecture as Strategy. Harvard Business Review Press, Boston (2006)
17. Yeoh, W., Koronios, A.: Critical Success Factors for Business Intelligence Systems. Journal of Computer Information Systems 50, 23--32 (2009)