THE ROLE OF DATA ARCHITECTURE AS A PART OF ENTERPRISE ARCHITECTURE*

Cristian Etegan Ph. D. Student
University of Craiova
Faculty of Economics and Business Administration
Craiova, Romania
Assoc. Prof. Anca Mehedintu Ph. D
University of Craiova
Faculty of Economics and Business Administration
Craiova, Romania
Prof. Cerasela Pirvu Ph. D
University of Craiova
Faculty of Economics and Business Administration
Craiova, Romania

Abstract: : In the early days of computing, technology simply automated manual processes with greater efficiency. The new organizational context provides input into the data architecture and is the primary tool for the management and sharing of enterprise data. It enables architects, data modelers, and stakeholders to identify, classify, and analyze information requirements across the enterprise, allowing the right priorities for data sharing initiatives. Data architecture states how data are persisted, managed, and utilized within an organization. Data architecture is made up of the structure of all corporate data and its relationships to itself and external systems. In far too many situations, the business community has to enlist the assistance of IT to retrieve information due to the community's inconsistency, lack of intuitiveness, or other factors. The goal of any architecture should illustrate how the components of the architecture will fit together and how the system will adapt and evolve over time.

JEL classification: M15, M10

Key words: data architecture, enterprise architecture, business process planning, databases, business objects.

1. INTRODUCTION

In the early days of computing, technology simply automated manual processes with greater efficiency. As technology evolved, new innovations enabled new capabilities and processes in the enterprise that were driven by IT. Gradually, IT changed the business but not necessarily in alignment with the business strategy. This lack of alignment resulted in significant waste of resources and missed opportunities, and placed the organization in a competitive disadvantage in the market.

In the case of an enterprise, the Information Technology term means the entire corporate activity and function related to computing, networking or storage. This

* This work was supported by CNCSIS –UEFISCSU, project number PNII – IDEI code 378/2008

includes the people, the assets, the systems, the software, the applications, the practice and the principles related to this function.

So, it is obvious that information technology became critical to organizations, depicted whose have claimed that IT is no longer a strategic asset.

Information technology (IT) issues are a major enabler of the new logistics concepts necessary to manufacture and deliver a built-to-order car in within 5 days. Unfortunately, the current IT infrastructure and systems applied within the automotive industry do not fulfill the requirements posed by new collaborative processes. The existing legacy systems were originally built for a "different world" of IT development, specific tasks (not integrated) and where technology was associated with central control. Thus, supply chain partners currently depend on software applications exchanging information mostly on the basis of proprietary data schemes and interfaces using non-standard transportation and application protocols hampering customized system integration. New collaborative interaction approaches, however, require the flexible integration of IT systems from different organizations. The resulting IT infrastructure must be interoperable and has to follow a distributed architecture. On the basis of a common understanding of content and the meaning of information transferred by computer systems, intelligent IT technologies such as ontology-based data interoperability and service-orientated design principles support stakeholders in coping with the increased run-time complexity of supply chain management amongst integrated network partners.

2. OBJECTIVES

Managing the product development process from design through manufacturing is a complex task that is further complicated by the need to incorporate sub-tier support and low-cost country operations. Today's automotive and transportation companies need to ensure that their suppliers, customers and business partners keep their information synchronized throughout the entire process regardless of organizational or geographic boundaries.

Since the rules of competition have changed, organizations are forced to eliminate the process of reinventing the wheel each time a new business problem arises that forces the enhancement of an existing system or implementation of a new one.

To align the strategies of business with IT, a new approach for managing IT has been developed called Enterprise Architecture. Just as architecture provides a blueprint for constructing a building, Enterprise Architecture provides a blueprint and roadmap for aligning business strategy with IT.

3. METHODOLOGY

The methodology approached involves the framework of enterprise architecture. Enterprise architecture is one of the most challenging roles in information technology today. Many aspects of the role are technical while much more of it is about interaction. Many people who have this position have significant responsibility but do not have authority or control. Enterprise architecture as an area of study allows one to focus on interesting, complex problems, to advance on the corporate ladder, and to maintain technical competencies while making a colossal difference to the enterprise.

Enterprise architecture is a plan of record, a blueprint of the permitted structure, arrangement, configuration, functional groupings/partitioning, interfaces, data,

protocols, logical functionality, integration, technology, of IT resources needed to support a corporate or organizational business function or mission. Typically, resources that need architectural formulations include applications, security subsystems, data structures, networks, hardware platforms, storage, desktop systems, to name just a few.

To understand the need of a framework regarding the enterprise architecture, we must follow the "white rabbit" – systems architecture paradigm.

In this case, a system is an interconnected set of machines, applications, and network resources. Systems architecture unifies that set by imposing structure on the system. More importantly, this structure aligns the functionality of the system with the goals of the business.

4. ANALYSES

The basic purpose of systems architecture is to support the higher layers of the enterprise architecture. In many companies, the software and hardware represent a significant portion of the enterprise's total assets. It is important that enterprise architects do not equate their duties with the objects, the applications, or the machines that comprise their domain. The fundamental purpose is to support and further the business objectives of the enterprise. Hardware and software objects are fundamentally transient and exist only to further the purposes of the business.

4.1. Framework

Like the **Open System Interconnection Reference Model** (OSI Reference Model or **OSI Model)** of the layered communications and computer network protocol design, we could notice a layered model for systems architecture in the case of an enterprise.

(Source: A Practical Guide to Enterprise Architecture Prentice Hall PTR, 2008)

Figure no. 1. Enterprise architecture model

Systems architecture is also used as part of the process of keeping the enterprise architecture aligned with the business goals and processes of the organization. It is important to understand the technical details of the infrastructure and the applications

running within it but also to have the knowledge to participate in the process of architectural change with the enterprise architectural team. That involves the following:

Defining the structure, relationships, views, assumptions, and rationales for the existing systems architecture and the changes in relationships, views, assumptions, and rationales that are involved in any changes required for moving from what is to what is desired.

Creation of models, guides, templates, and design standards for in use in developing the systems architecture.

Enterprise architecture terminology carries many variations within each organization and in the vast array of literature. As described in American National Standards Institute/Institute of Electrical and Electronics Engineers (ANSI/ IEEE) Std 1471-2000, an architecture is "the fundamental organization of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution." A metaphor can be drawn by thinking of a corporate/IT blueprint for the planning of a city or a large development. Specifically, then, the blueprint provides the macro view of how elements (roads, lots, utilities—read: platforms, networks, applications, applications' logical components) fit, particularly in relation with one another [2].

Enterprise Architecture (EA) is a method and an organizing principle that aligns functional business objectives and strategies with an IT strategy and execution plan. The Enterprise Architecture provides a guide to direct the evolution and transformation of enterprises with technology. This in turn makes IT a more strategic asset for successfully implementing a modern business strategy [5].

An Enterprise Architecture typically produces deliverables such as:

- Current State Enterprise Architecture model
- Future State Enterprise Architecture reference model that is needed to execute on the proposed business strategy
- Gap analysis that identifies the shortfalls of the current state in terms of its ability to support the objectives and strategies of the business
- Architecture Roadmap that defines the initiatives required to migrate from the current state into the future state.

Enterprise architecture at any instance n is an exact description of the functionality, the interfaces, the data, and the interface protocols supported by the (partitioned) set of functional elements in the environment, as of time instance n. Call that simply "enterprise architecture, the current blueprint," or just "enterprise architecture." One is then able to provide a description of a desired target state at time n + 1 or n + j in terms of possibly new functionalities, new interfaces, new data, new interface protocols, and possibly new partitioning of the set of functional elements.

Architecture is a concept with many facets and can almost be all things to all people. Our interpretation, perception, and understanding of architecture will not necessarily be the same as that of our colleagues. A unique, unilaterally accepted definition of architecture has not yet been established. This need not be a problem, as long as we are aware of the differences in interpretation and make them explicit in our communication.

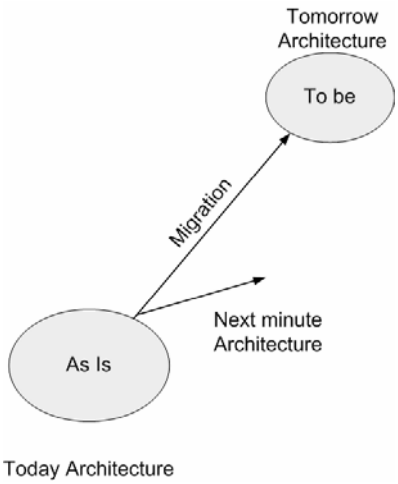
When we discuss architecture, there are three aspects that must be clarified before - otherwise, confusion and misunderstanding can result [3].

1. Chronology
2. Context or subject matter
3. Level of abstraction

There is a chronological aspect to architecture. We have encountered the following (chronological) definitions of architecture:

- A description of the current situation
- A blueprint for a desired future situation
- A set of guidelines for carrying out changes

The Gartner Group makes this even more explicit (in terms of time) by identifying the above definitions as three forms of architecture: the today architecture, the tomorrow architecture, and the next minute architecture. These three forms are represented in Figure 2.



(Source: Roel Wagter, Martin van den Berg, Joost Luijpers, Marlies van Steenbergen Dynamic Enterprise Architecture: How to Make It Work, John Wiley & Sons, Inc, 2005)

Figure no. 2. Chronological Aspects of Architecture

It is easy to imagine that a great deal of confusion can arise when the type of architecture is not made explicit in a specific situation. The aspect of context, or subject matter, also needs to be explained. Architecture, as a concept, can be applied to different contexts or subject matter. Architecture can be developed for products and services, processes, organizational structures, information, applications, middleware, platforms, and networks. This creates product-and-services architecture, process architecture, organizational architecture, information architecture, and so on. The various architectural domains are often grouped together to form three main types of architecture:

1. Business architecture
2. Information architecture
3. Technical architecture

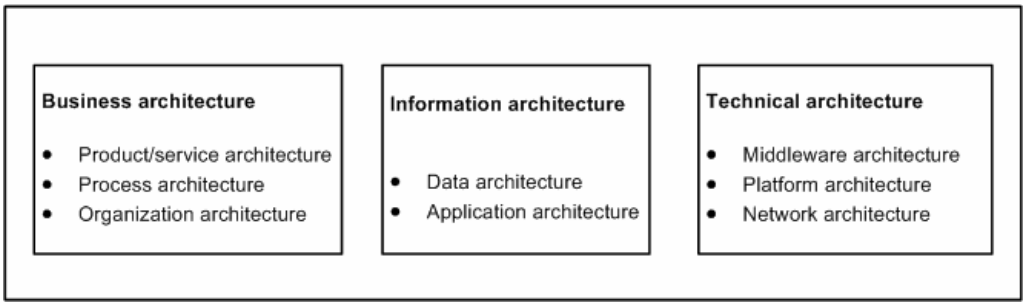
Business architecture sketches the contours for the way in which an organization can be structured to effectively pursue its business objectives. Business architecture consists of three domains: (1) the products and services offered; (2) the

processes responsible for producing these products and services; and (3) the organizational structure required to carry out these processes.

Information architecture sketches the design contours for the provision of information within an organization. It consists of two domains: (1) the data that is important for the correct functioning of the organization; and (2) the applications that ensure that this information is correctly distributed within the organization.

Technical architecture sketches the contours of the technical infrastructure necessary to support the organization. It consists of three domains: (1) the hardware platforms; (2) the network components; and (3) the software required for information sharing between applications (also known as middleware).

Figure 3 illustrates the constituent parts—that is, the architectural domains—of the three types of architecture.



(Source: Roel Wagter, Martin van den Berg, Joost Luijpers, Marlies van Steenbergen Dynamic Enterprise Architecture: How to Make It Work, John Wiley & Sons, Inc, 2005)

Figure no. 3. Architectural Types and Domains

The next step in developing enterprise architecture is considering it as a dynamic and / or agile one. And from this point of view, let's look upon the Data architecture as a part of information architecture.

4.2 The Data Architecture

Data is the lifeblood of an organization and a valuable enterprise asset. It provides the foundation on which to base critical and everyday business decisions. It is, therefore, essential that decision makers can access and depend on quality data to operate confidently in a high-performance environment.

Data Architecture in enterprise architecture is the design of data for use in defining the target state and the subsequent planning needed to hit the target state. It is usually one of several architecture domains that form the pillars of an enterprise architecture or solution architecture [9].

A real enterprise has very complicated data architecture. Most of the data will be held in large legacy or package systems, for which the details of data structure may be unknown. Other data will be held in spreadsheets and personal databases (such as Microsoft Access), and may be invisible to the IT department or senior business data administrators. Some key data may reside in external systems maintained by service

providers or business partners. As we explore enterprise data architecture, we come to accept two realities:

1. We have little control over the way high-level business data concepts are realized. Data is likely to be highly dispersed, often without adequate controls on quality.
2. Most data is duplicated across a number of systems, with significant variations in quality, format, and meaning. Some of the copies, maintained by Enterprise Application Integration technology (EAI) or careful business processes may be good (but probably not perfect). Most are very poor, maintained only by occasional batch transfers and stressed or broken manual processes. Organizational and business process conflicts, or simple failures of trust, may get in the way of common sense improvements.

The poor copies may be causing business problems. Furthermore, initiatives such as Customer Relationship Management (CRM) and Business Intelligence will need to merge data from various sources. Some organizations work to harness various legacy systems in end-to-end processes. Either the business or IT may be driving changes to simplify business processes, streamline data flows, and reduce duplication. Modeling can be of great benefit in meeting these challenges. But most traditional modeling approaches don't address these requirements.

They produce models which are either too detailed to be of use, or not detailed enough, and typically fail to focus on the difficult issues of the enterprise data architecture and the integration of its various components.

We believe it is important to create powerful, simple, but effective models of the data structure from an enterprise viewpoint - a set of models known as the "Enterprise Data Architecture."

The core of any enterprise data architecture is an enterprise data model (EDM). No enterprise data architecture exists without an enterprise data model. The EDM is an integrated subject-oriented data model defining the essential data produced and consumed across an entire organization [6].

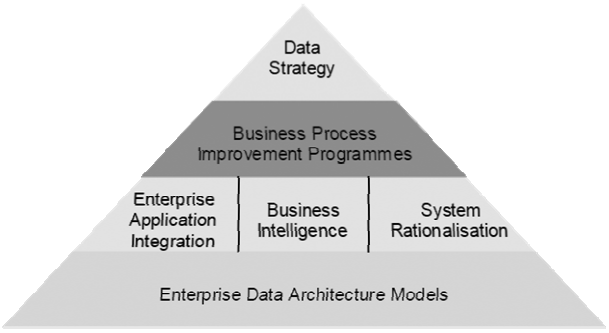
- **Essential** means the data critical to the effective operation and decision-making of the organization. Few (if any) enterprise data models define all the data within an enterprise. Decisions must be made (and revisited) about the scope of enterprise data modeling efforts. "Essential" does not mean "common" or "shared." Essential data requirements may or may not be common to multiple applications and projects. Some data defined in the enterprise data model may be shared by multiple systems, but other data may be critically important yet created and used within a single system. Over time, the enterprise data model should define all data of importance to the enterprise.
- **Integrated** means that all of the entities attribute and rules in the model are defined once, without redundancy. The concepts in the model fit together as the CEO sees the enterprise, not reflecting separate and limited functional or departmental views. There is only one version of the Customer entity, one Order entity, etc. Every data element also has a single name and definition. The data model may also identify

common synonyms and important distinctions between different sub-types of the same common business entity.

- **Subject-oriented** means the model is divided into commonly recognized subject areas that span across multiple business processes and application systems. Subject areas are focused around the most essential business entities.

An enterprise's information systems architecture has many interrelated aspects, including applications, hardware, networks, business processes, technology choices, and data. As shown in Figure no. 4, the data architecture is a layered set of models which provides a solid foundation for strategic initiatives such as:

- A Data Strategy, outlining the business's aims and objectives for improved collection and use of data,
- Business process improvements,
- Decisions on the future of new and changed systems,
- Integration, data warehousing, and reporting initiatives.

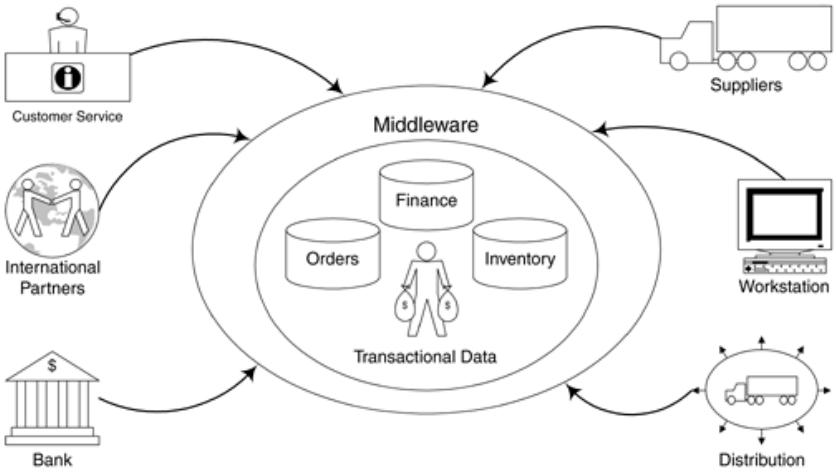


(Source - Andrew K. Johnston and Richard Wiggins - Modelling the Enterprise Data Architecture, 2003)

Figure no. 4. Enterprise data architecture models support a variety of common IT and business improvement initiatives.

The typical organizational context considered provides input into the data architecture and is the primary tool for the management and sharing of enterprise data. It enables architects, data modelers, and stakeholders to identify, classify, and analyze information requirements across the enterprise, allowing the right priorities for data sharing initiatives. Data architecture states how data are persisted, managed, and utilized within an organization (as shown in Figure no. 5). Data architecture also describes the following:

- How data are stored in both a transient and permanent manner
- How components, services, and other processes utilize and manipulate the data
- How legacy systems and external business partners access the data
- How common data operations (create, read, update, delete) should occur in a consistent manner



(Source: A Practical Guide to Enterprise Architecture Prentice Hall PTR, 2008)

Figure no. 5. A typical Organizational context

The baseline data architecture defines the business information needed to support the business functions that are encompassed in the baseline business model. It essentially contains information on high-level data areas and their classes. Additionally, it identifies the need for data warehouses and data marts. Data architecture is not a data model (entities, relationships, classes, elements, etc.). Data models are best created as part of a departmental project where business requirements are best understood. The data architecture cannot always anticipate all implementation details.

Data architecture defines how data is stored, managed, and used in a system. It establishes common guidelines for data operations that make it possible to predict, model, gauge, and control the flow of data in the system. This is even more important when system components are developed by or acquired from different contractors or vendors.

It is also necessary for an enterprise software architect to identify candidates for business objects.

4.3. The Business Object

According to the Object oriented Paradigms, a business object is a code construct that corresponds directly to a *thing* in the actual business the software is meant to represent. The object encapsulates the business logic related to the *thing,* and encapsulates the data that is required by the logic and also describes, defines, makes up, is contained by, and or is associated with the *thing.* These things should be recognizable to a non-technical person familiar with the business, like the users, business analysts, etc. [4]. For example, in a video rental store business we can assume there would be DVDs, console games, customers, rental orders, late fees, etc. In software representing that business there would be business objects that represent DVDs, console games, customers, rental orders, late fees, etc. Each object has data that describes or is attributed to the object and methods that make decisions based on that data. For

example, a DVD may have a Title and DateReleased attributes (data), and may have the method *calculate rental price* to determine how much it costs to rent the DVD.

Business architecture (conceptual architecture) identifies processes at the enterprise level and outlines the desired outcome the business seeks, as well as the steps necessary to accomplish it. In a vast majority of organizations, it will be relatively easy to find documentation on the solutions that have been implemented to date.

Identifying business objects and the data that need to be associated, classifying business objects and their relationships to each other allows the architect to develop classification schemes that are necessary for creating metadata. In this step, it is a good idea to assign a unique identifier to each object, as well as to capture a description of the business processes the business object fulfills.

One should also consider attaching a representative sample of the data, as well as attaching a statement for the original motivation behind collecting this piece of data.

At this stage, one should make sure that processes and their associated data are consistently named throughout the enterprise. In all but the smallest organizations, it will be difficult to find the same process referred to with two different names. Usually this occurs across different business units. Inconsistencies in processes and their associated data may include assigning different labels to refer to the same type of process/data, associate the same label with different types of process/data, representing the same process/data using different formats, and so on. Eliminating these types of inconsistencies will remove hurdles now and in the future when integrating disparate business systems. It is important that one not get hung up on getting it perfect at this step.

Identifying the major business units within the enterprise and the relationships among them should be considered thoroughly at this stage. One will find that the identified business processes within each business unit should ideally map to the identified business objects.

In the case of an automotive industry data architecture concept, the business objects should be defined in the following manner:

Table no 1. Partial list of Business objects definitions in an automotive assembly structure.

Name	Definition	Identifiers	Characteristics	Notes
Agency	Department or administrative unit of a government entity	Title, type, ID number	Description, role, address, contact information (telephone, fax, Web address)	Legislative branches (Ministry of Transports, Ministry of Economy)
Agreement	An arrangement between two or more parties as to a course of action that identifies the roles and responsibilities of each	ID, title, dates, parties	Subject, scope, period of performance, funding, terms and conditions	Contracts, permits, collective bargaining, certifications and labor management agreements
Authority	Specific rights and permissions assigned to an employee, person, and job	Type, data, name	Source, duration, level, scope	Rights, warrants, delegation orders, access rights to information, clearances, digital signatures, authentication, credit card

				information
Benefit	Pack of services and programs provided to employees	Date, name, type, grade	Description, duration, terms and conditions	Health insurance, counseling, massages, free tickets
Budget	Estimated funds required for goods and services	Year, type, organization	Amount	Planning estimates, budget recommendations
Compliance	Conformance with legal, HR, and regulatory requirements	Date, type, mandate, department, name	Methodology, result	Audit reports, certifications, accident reports
Cost	Value of goods and services	Data, name, type	Amount, fiscal year, period covered, currency	Salary, benefits, products
Departments	Organizational units	Names, codes, acronyms	Mission, vision, function, head	Locations in each country
Dispute	Controversy among parties	Type, date, number, title	Description, resolution, status	Security clearances, HR policies, whistle blowers, union cutbacks, etc.
Document	Data that are contained within a physical or electronic media	Name, number, type, date	Content, description, format, media, sensitivity	Media can be hard copy, analog, digital, and can come in many forms
Employee	A person who works for the enterprise	SSN, employee ID	Education, name, type, grade, job title, skill, location	Does not include contractors or consultants
Goal	Desired outcome	Title, type, date	Content, event, source	Settled, closed, cancelled, etc.
Incident	Significant event where disclosure is required	Date, location, type	Sensitivity, severity, duration, department	Accident, injury, chemical release, security breach
Intellectual property	Rights owned by enterprise	Name, date, type	Concept, description, formulae, value	Copyrights, patents, trademarks, secrets
Job	Responsibilities of an employee	Title, position, job ID, date of creation, job grade	Description, job family, type, role, status	Enterprise Architect, Technology Director, Thought Leader
Mandate	Specific instruction	Identifier, date, topic	Authorization reference, context, type, approval	Law, regulation, policy, requirements (e.g., Healthcare Portability Act)
Measure	Standard of evaluation	Type, organization, date	Unit, scale, dimension, period, description	HR performance standards, miles per gallon
Product	Items that are created, manufactured, or distributed by enterprise	Name, date	Description	Wheels, metal detectors for finding unicorns, etc.
Program	Grouping of related projects that have common attributes	Name, acronym	Scope, subject, context	financial management, customer relationship management
Project	Set of activities that are managed and funded together	Name, cost center, acronym, type	Scope, description, schedule	World domination, human genome
Proposal	Offer to provide product or services	Title, date, type	Value, approach, key personnel, deviations	May result in agreement

			from norm	
Resource	Physical property used to accomplish a goal	Identifier, type	Description, location, age, quantity, availability, value	Equipment, space, supplies, buildings, inventory
Risk	Possibility of loss, failure, or other undesirable outcome	Name, type	Probability, impact, description, outcome, scenarios	Market, financial, litigation, security, safety, environmental, etc.
Task	A discrete unit of work	Title, number, date	Description, schedule, milestones	The lowest level of work, such as evaluate software for ERP

4.4. Metadata

Another important aspect is the use of metadata concepts in defining data architecture. Since the metadata is information about data, the new approach is the use of federated metadata, witch describes data the same way in each physical table in which the data are stored.

A federated metadata approach can be implemented even when information is physically partitioned across disparate data sources. In the ideal world, all related data would exist within a single physical location, making application development easier. Reality dictates that information will be spread across multiple heterogeneous platforms and topologies. Defining this up front will simplify application integration. Data sharing can be further extended through reusable services using a service-oriented approach or by using a broker-based architecture. To access data, each application would send a request to the service. The service would execute the request and return the appropriate response to the application. This approach protects the integrity of data and guarantees that the data retrieved are both accurate and consistent.

To be successful in defining metadata, the first step is to define the appropriate owner of each data element. Usually data elements can be classified by who is the authoritative source. For example, one enterprise assigns each employee an employee number. The employee number is used by more than one department, including security, manufacturing, and legal, but the human resources department issues the employee number and is, therefore, the authoritative source. Sometimes, data elements are used by more than one department but are not identified as the authoritative source. For example, each employee has a CNP number that is used in multiple applications, including payroll, employee benefits, and security.

Table no. 2 – Use of metadata in the case of CNP Social security number.

Field	Values
Data Element Name	Social Security number
Element Definition	A 13-digit number assigned to an individual by the Social Security Administration
Business Format	9-999999-99-999-9
Business Length	17 positions

Field	Values
Business Type	Number
Exchange Format	9999999999999
Exchange Length	13 bytes
Exchange Type	Character
Storage Format	9999999999999
Storage Length	13 bytes
Storage Type	Character or variable character
Scope	Romania

Using the metadata repository to store data element definitions has the result of building the enterprise data model. In this usage, if the repository is kept up to date, the resulting enterprise data model is also up to date. By being able to start from either direction, the repository can serve as the tool that will empower an enterprise to determine how changes in data affect other processes. It will also help realize the goal of data reliability, reusability, and sharing across organizational boundaries.

5. CONCLUSIONS

It is rare for all but a few privileged organizations to have complete data architectures. Data architecture is made up of the structure of all corporate data and its relationships to itself and external systems. In far too many situations, the business community has to enlist the assistance of IT to retrieve information due to the community's inconsistency, lack of intuitiveness, or other factors. The goal of any architecture should illustrate how the components of the architecture will fit together and how the system will adapt and evolve over time.

REFERENCES

1. McGovern, J., Ambler, S. W., Stevens, M. E., Linn, J., Sharan, V., Jo, E. K. — A Practical Guide to Enterprise Architecture Prentice Hall PTR, 2008
2. Minoli, D. — Enterprise Architecture A to Z, Auerbach, 2008
3. Wagter, R., Van den Berg, M., Luijpers, J., Van Steenbergen, M. — Dynamic Enterprise Architecture: How to Make It Work – John Wiley and Sons, 2005
4. Reddy, J. — A definition of business object, 2006
5. Covington, R., — The Oracle Enterprise Architecture Framework, October 2009
6. Mosley, M. — What’s In YOUR Data Architecture? - EIMI Archives, Volume

2, June 2008 Edition
7. * * * http://www.agiledata.org
8. * * * http:// www.gartner.com
9. * * * http:// www.wikipedia.org