NOVA IMS
Information Management School

MAA

Mestrado em Métodos Analíticos Avançados
Master Program in Advanced Analytics

The ENTERPRISE DNA: static and dynamic digital representation of organizations.

Roman Vladimirovich Trotsyuk

Dissertation presented as partial requirement for obtaining the Master's degree in Advanced Analytics

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

2019

The ENTERPRISE DNA: static and dynamic digital representation of organizations.

Roman Vladimirovich Trotsyuk

MAA

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

The ENTERPRISE DNA: static and dynamic digital representation of organizations

by

Roman Vladimirovich Trotsyuk

Dissertation presented as partial requirement for obtaining the Master's degree in Advanced Analytics

Advisor: Vítor Manuel Pereira Duarte dos Santos, PhD

February 2019

DEDICATION

I dedicate this master thesis to my wife and mother. The support of my family during this two-years NOVA IMS educational journey made my dream come true.

ACKNOWLEDGEMENTS

I would like to express appreciation to Professor Leonardo Vanneschi. ~~for m~~My interest in ~~the topic of~~ Evolutionary Computation ~~, that~~ emerged by virtue of his ~~gorgeous~~ excellent lectures, and related subjects included in the study plan for Advanced Analytics master program. ~~This~~ His passion, theoretical knowledge, and hands-on experience~~,~~ acquired during the study, inspired me to learn Genetic Algorithms further.

I also want to thank Professor Vitor Santos for suggesting this engaging topic, and supervising my research ~~brilliant and very interesting topic proposed for this master~~. His idea of applying evolutionary computation for modelling the business entities' behavior, and considering organizations as living organisms ~~and~~ in a competitive environment, as is in ~~a~~ nature, is very innovative ~~and perspective~~ due the complexity and ~~the~~ novelty of the subject. His patience, involvement, tips, availability for frequent meetings and ~~spent~~ his previous time helped me ~~to~~ perform this research.

Finally, I want to express appreciation for NOVA IMS Professors Mauro Castelli and Flávio Pinheiro for their very warm and productive feedback interview that brought a lot of valuable insights, ideas and remarks that helped me to improve the paper.

ABSTRACT

The contemporary business environment ~~become~~ is becoming more complex and fast-changing ~~during processes of business~~due to globalization, technical progress, competition, accumulation and sharing of knowledge. Companies arise, grow, exchange ideas, merge, interact with each other and adapt to the market environment, and~~,~~ finally, some of them disappear. Their behavior determined by sophisticated ~~laws, that~~laws that cannot be discovered in isolation, because existence of each entity depends on its specific, internal properties and the interplay with other participants of the market. In a way, ~~The~~ the~~complex interactions and~~ enterprise life cycle ~~reminds~~ resembles the life of species in ~~the~~ nature and thus, the mimic of biological entities can be applied as a modelling tool ~~for~~ to better understand~~ing~~ how a company works and what makes it successful.

The broad purpose of this work is to create a foundation for applying Artificial Life simulation ~~for~~ to business analysis. Artificial Life is a concept that ~~allow to~~ mimics biological evolution and behavior of living creatures ~~for~~ in modelling complex systems, forming a specific environment with interacting and evolving agents. Thus, ~~the~~ Artificial Life can be applied ~~for~~ to the analysis of the enterprise on the competitive market for studying its success factors. Possible combination of factors and their value are company- specific and represent properties that affects performance of the organization. The enterprises can exchange their characteristics with others by means of stuff swap, consulting, merging etc., acquiring best practices and becoming more adapted for specific challenges. The main goal of this paper is ~~the~~ to research and ~~suggestion~~ explore~~of~~ such characteristics and their representation~~,~~ which, by analogy with biology, will constitute Enterprise DNA.

In this thesis, the digital representation of the Enterprise DNA inspired by the biological notion of living organism's DNA is proposed. As the foundation of important company's features, the Enterprise Architecture concept was applied. Despite the fact~~,~~ that it was previously used for an Information Technology architecture, this discipline ~~was~~ has evolved to more broad science and became a tool for describing the business architecture of ~~the~~ a company. The Zachman Enterprise Architecture framework is used as a basis of enterprise representation. Regarding this tool, the artifacts for phenotype representation are proposed and then~~,~~ their digital XML representation found. The DNA digital representation model (genotype) for artifacts is proposed, which can be used for further evaluation of fitness of the specific company to the competitive environment on the specific market. This representation can be used by means of Genetic Algorithm for further implementation of Artificial Life simulation on real company's data. The evaluation by the experts showed ~~perspectivity~~perspective of the idea of applying ALife modelling for solving business problems. As a result, some ideas for model improvement are discovered.

KEYWORDS

Artificial Life, Corporate DNA, Competitive Strategy, Enterprise Architecture, Enterprise DNA, Evolutionary Algorithms, Genetic Algorithms, GA Representation

INDEX

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AI	Artificial Intelligence
APQC	American Productivity & Quality Center
ALIfe	Artificial life
BPEL4WS	Business Process Execution Language for Web Services
BPM	Business Process Modeling
BPML	Business Process Modeling Language
BPMN	Business Process Model and Notation
BPMS	Business Process Management System
BSC	Balanced Scorecard
CRM	Customer Relationship Management
DNA	Deoxyribonucleic acid
DSR	Design Science Research
EA	Enterprise Architecture
ER	Entity-Relationship Diagram
ERM	Entity-Relationship Model
eTOM	Enhanced Telecom Operations Map
FEA	Federal Enterprise Architecture
GA	Genetic Algorithm
IDEF 3	Integrated DEFinition for Process Description Capture Method
ITIL	Information Technology Infrastructure Library
MU	Mutation genetic operator
PEST	Political, Economic, Socio-cultural and Technological analysis
PMXML	Project Management eXtensible Markup Language
SCML	Supply Chain eXtensible Markup Language
TOGAF	The Open Group Architecture Framework

UML	Unified Modelling Language
UNSTAT	United Nations Statistical Comission
W3C	The World Wide Web Consortium
WSCI	Web Services Choreography Interface
WSFL	Web Services Flow Language
XLANG	Web Services for Business Process Design language (Microsoft)
XML	eXtensible Markup Language
XO	Crossover genetic operator
XPDL	The XML Process Definition Language
ZF	Zachman Framework

1. INTRODUCTION

We live in the fast-changed era when environment, people, companies and other entities are changing with tremendous pace, ~~become~~ and are becoming more complex and less understandable. Interactions of companies within environments they operate ~~became~~ has become more and more complex and it is hard to create a particular ~~to~~ model that can describe an enterprise, its interactions with competitors and market where it is presented.

The way the enterprises operate, when they born, grow up, interact and compete, survive and die ~~remind~~ is similar to the way ~~as~~ natural organisms behave. Thus, we can see an analogy of companies with existing living species that live by laws of evolution and natural selection. Technology transfers, stuff exchange or even reading business books by mangers and inviting consultants leads to some specific knowledge, ideas, structure, systems etc. interchange that reminds the exchange of genes when the living species mate.

Company's internal efforts ~~as a~~ in discovering a new strategy, business processes reengineering, and other internal improvements lead to better performance and survival on the market. We can consider such changes, not inspired by external knowledge but generated inside the enterprise as biological mutations. Some of them will improve company's market position, and some we lead to failure. Only markets and competition ~~will show~~can determine which of them were positive.

Thus, the idea to~~of~~ mimic ~~of~~ the biological behavior of ~~the~~ living organisms, their natural selection and adaptation to specific environment with respect to real business ventures arises. Considering each company as an individual with specific set of genes – good and bad, we can simulate the life cycle of the organization, ~~–~~its competition with rivals, head hunting, development, used strategy and other efforts of managers that lead the company to the a specific position on the market, and its success.

When we mimic an organism (or companies) and we can compare some properties. Number of legs, or color, size, speed or brain size – all of these are characteristics with describes specific ~~spicy~~ specie and can be extended (with adaptation) to the business entity. Thus, every organization as a living organism have some visible signs, that are defined by its phenotype and their DNA encoding that is represented by special code, or gene.

1.1. BACKGROUND AND PROBLEM IDENTIFICATION

The main idea is to mimic company's behavior by means of biologically inspired simulation models. Such models ~~usually~~ are called Artificial Life simulation (Langton, 1989) by analogy with real life, and was developed for modelling complex~~,~~ evolutional systems. The ALIfe conception implies modelling of research entities by means of Genetic Algorithms (Holland, 1975), which can model every individual within a specific population, their interaction, mating, mutations and adaptation to the environment. The fitness function defines how an individual ~~is~~ fits ~~to the~~ specific conditions and ~~is~~ if ~~there~~ it is a success or failure. The population is evolving, exchanges its genes, mutates and adopts to the environment ~~and~~ by law of natural selection - only the best individuals survive. ~~Best means the fitness to the environment.~~

Although the author ~~have~~ has some industry experience relevant to the topic, more scientific and systematic approach must be used. Thus, the idea is not to create Enterprise DNA "from head" but use some widely used framework, which will simplify the design and collecting the data in the future. The Enterprise Architecture frameworks are considering as a basement for projecting the Enterprise's DNA structure.

The problem is the modelling of the enterprises and their interactions by analogy with living species for better understanding of how they behave, ~~which~~ and what components make them successful. This kind of modelling is called Artificial Life simulation and is usually implemented by means of evolutionary algorithms. Genetic algorithms implementation implies three main steps:

- Development of genotype representation of the individual (representation problem)
- Design of fitness function~~,~~ that will evaluate successfulness of the individual
- Simulation by means of Artificial Life algorithm

Thus, the representation problem is the first on the way to implement evolutionary competition model of the enterprises, and~~,~~ thus~~,~~ is the objective of this research. Other steps considered as a suggestion for further development of the topic. The representation of the individual for Genetic Algorithm is not straightforward and depends on the goals, so it is not a trivial task and can be implemented by several ways. The possible implementations are discussed in literature review part of this work.

1.2. STUDY RELEVANCE AND JUSTIFICATION

As far as we are aware, there were no similar researches proposed before. The novelty is in ~~the~~ applying Artificial Life simulations to the business competitiveness. Another innovation proposed ~~are~~ is the representation of the enterprise by its DNA like a living organism and use the Enterprise Architecture framework ~~for that~~is used for that purpose ~~the Enterprise Architecture framework~~ because it is a modern approach to describe the business entity. Each company existing mimic living organisms with their DNA – specific structure and properties sets, behavior, reactions and relevance to the environment. Thus, the idea of application of Alife simulation to the business competitiveness looks promising.

This work is theoretical one and its evaluation is a challenge. For this purpose, experts in evolutionary algorithms should be attracted for validation of developed model. Further works will show in practice value of ideas proposed.

1.3. RESEARCH OBJECTIVES

The goal of this work is to propose the representation of Enterprise's DNA that can be used for further modelling of organizations by means of ~~for~~ Artificial Life simulation. This representation must cover the main aspects of the company and ~~reflex~~ reflect its constitution. Thus, the research objective is the design of Enterprise's DNA.

The representation problem is one of the main challenges when Genetic Algorithms are applied. There are several kinds of genome representations and their choice can affect results of simulation. The main objective is to find such representation that can be used for Artificial Life simulation with respect to proposed fitness function.

As a foundation, ~~for~~ individual representation with Enterprise Architecture frameworks are considered, which are used for description of company's business architecture. The idea is new regarding goals stated, nonetheless, it seems very promising to use this approach for achieving better results in understanding how business operates, what makes it better and what "Genes" lead to success.

2. THEORETICAL FRAMEWORK

2.1. ENTERPRISES

The main goal of this work is to create a basement in finding features of the enterprise that distinguishes it from others. For this purpose, companies can be seen from points of view: internal and external. From external viewpoint there are number of classifications of enterprises that discern them by different typology. Belonging to each class or group carry some specific characteristic. From internal point of view companies differ by their Enterprise Architecture.

2.1.1 Typology of enterprises

There are number of enterprise classifications that examine companies from different sides and serving for specific purposes. There are two main groups: Academic approach and applied International Organizations and other institutions classifications. Scientific study usually concentrated on finding general methods and frameworks for classifications while other institutes develop taxonomy for use for their operational purposes (Governments, Banks, Statistical agencies etc.)

2.1.1.1 Academic Approach

Researchers are working on the development of scientific methods of organizations classification. The 10 guidelines for the Empirical Classification of Organizations for exploring of multivariate approaches were

*Possible Kinds of Taxonomic Characters**

1. Morphological Characters
   a. General formal structural (formalization, specialization, levels, etc.)
   b. Special structures (technical, accounting, control, planning systems, etc.)
   c. Internal morphology (workflow configuration, div. of labor, staff groups, etc.)
   d. Subunit characteristics (kinds of subunits, formal-informal nature, etc.)
   e. Variance characteristics (variance in subunit size, formality, etc.) [38]
   f. Interdependency networks (coordination structures, etc.)
2. Physiological (process and functional) Characters
   a. Metabolic flows (personnel, communications, workflow, rates, etc.) [48]
   b. Managerial functions and processes (decisionmaking, conflict handling, etc.)
   c. Adaptation and change characteristics (managerial succession, changes in influence postures, etc.) [38]
   d. Workplace throughput and conversion processes (assembly lines, work stages, etc.)
3. Ecological Characters
   a. Environmental (physical, cultural, economic, social, technical, etc.)
   b. Epiphysical (buildings, layouts, personnel characteristics, etc.) [48]
   c. Dependency networks (on others, by subcontractors, etc.)
   d. Environmental variances (diversity, dynamism, uncertainty changes, etc.)
   e. Input-output characteristics (supplies, products, information, energy, etc.)
4. Behavioral Characteristics (nontransient)
   a. Attacker, avoider, achiever styles, etc. [5]
   b. Competitive posture (monopolistic, oligarchic, etc.)
   c. Human resource posture (conserver, user, developer, of people, etc.)
5. Geographic Characters
   a. Location patterns (local, national, multinational, etc.)
   b. Product distribution patterns
   c. Employee recruitment patterns
   d. Variance in cultural-social forms dealt with

Table 2.1 Possbile characters of organization taxonomy (McKelvey, 1978) proposed by Bill McKelvey (McKelvey, 1975). In a ~~further~~ newer paper, he proposed characters for organizations typology (see Table 2.1) based on work (Mayr, 1969) implies finding similarities and differences among organization applying biological taxonomy methods (McKelvey, 1978)

Other work (Rich, 1992) discusses empirical, theoretical and evolutionary perspectives of organizations typology, hierarchical taxonomy, and defines methodological consistency of classification systems.

2.1.1.2 International and other institutions classifications

Globalization of companies, international trade and capital movement requires similar view on organization for different stakeholders that lead to development special classification by many institutions. Some of meaningful examples are showed below.

United Nations developed its "International Standard Industrial Classification of All Economic Activities" (United Nations, 2008) to compile, generate and analyse a wide range of economic, social and environmental data and information on which States Members of the United Nations draw to review common problems and to take stock of policy options; to facilitate the negotiations of Member States in many intergovernmental bodies on joint courses of action to address ongoing or emerging global challenges; and to advise interested Governments on the ways and means of translating policy frameworks developed in United Nations conferences and summits into programmes at the country level and, through technical assistance, helps build national capacities.

~~Europian~~European Comission agency Eurostat provides own standard "NACE Rev. 2 Statistical classifiation of economic activites in the European Community" (European Communities, 2008) which "is a basic element of the international integrated system of economic classifiations, which is based on classifiations of the UN Statistical Commission (UNSTAT), Eurostat as well as national classifications; all of them strongly related each to the others, allowing the comparability of economic statistics produced worldwide by diffrent institutions".

In 1999, Standard & Poor's and MSCI/Barra jointly developed the Global Industry Classification Standard (GICS) (Standard & Poor, 2006) to establish a global standard for categorizing companies into sectors and industries. "The GICS methodology has been commonly accepted as an industry analysis framework for investment research, portfolio management and asset allocation which defines 10 Sectors, 24 Industry Groups, 67 Industries and 147 Sub-Industries. GICS was developed in response to the global financial community's need for one complete, consistent set of global sector and industry definitions, thereby enabling asset owners, asset managers and investment research specialists to make seamless company, sector and industry comparisons across countries, regions, and lobally" (Standard & Poor, 2006).

London Stock Exchange Group provides its own classification Benchmark (FTSE Russell, 2008) which "is a detailed and comprehensive structure for sector and industry analysis, facilitating the comparison of companies across four levels of classification and national boundaries. The classification system allocates companies to the Subsector whose definition closely describes the nature of its business as determined from the source of its revenue or the source of the majority of its revenue where available".

2.1.2 Enterprise's architecture

The subject of these research is the Enterprise's DNA which is composed from two parts: tangible and intangible. If we will draw an analogy with the computers that are "hardware" and "software". "Hardware" are assets of the company, including human capital, "software" is knowledge and culture. In the literature under term “Corporate DNA” most authors understand knowledge and the term “Corporate Architecture” means assets and their structure. For our purposes we have to consider both domains.

In Figure 2.1 enterprise architecture is positioned within the context of managing the enterprise. At the top of this pyramid, we see the mission of the enterprise: why does it exist? The vision states its ‘image of the future’ 8 Introduction to Enterprise Architecture and the values the enterprise holds. Next there is its strategy, which states the route the enterprise will take in achieving this mission and vision. This is translated into concrete goals that give direction and provide the milestones in executing the strategy. Translating those goals into concrete changes to the daily operations of the company is where enterprise architecture comes into play. It offers a holistic perspective of the current and future operations, and on the actions that should be taken to achieve the company’s goals. Next to its architecture, which could be viewed as the ‘hard’ part of the company, the ‘soft’ part, its culture, is formed by its people and leadership, and is of equal if not higher importance in achieving these goals. Finally, of course, we see the enterprise’s daily operations, which are governed by the pyramid (Lankhorst, 2017).

Figure 2.1 Enterprise Architecture as a management tool (Lankhorst, 2017)

2.1.3 Enterprise Architecture frameworks

The Enterprise Architecture is a “hardware” part of the company, which includes specific artifacts ~~which~~ that are defined by used EA framework. There are many frameworks developed for Enterprise Architecture modelling. Today, four methods are dominant: Zachman's structure for enterprise architecture, TOGAF (The Open Group Architecture Framework), Federal Organization Architecture (FEA) and Methodology Gartner (formerly called Meta Framework).

2.1.4 Zachman framework

Significant contribution to the development of the concept of the enterprise architecture was made by J. Zachman (John A. Zachman). Since the publication of the “Framework for Information Systems Architecture” (Zachman, 1987) has become the basis on which many organizations create their own methods of describing the information enterprise infrastructure. Since 1987, when ~~it was proposed~~ the first version of this model was proposed, extended subsequently in the works of 1992-96 (Zachman & Sowa, 1992), it was used quite often by world's largest corporations such as General Motors, Bank of America and others. The Zachman model also served as the basis for creating a range of other techniques and models for enterprise architecture descriptions.

Zachman's model is based on the discipline of the classical architecture, structure and provides, a common vocabulary and perspective set or structures (framework) to describe the modern complex educational systems. For convenience of description, Zachman suggested the so-called 'Zachman Framework for Enterprise Architecture. The model has two main objectives: on the one hand, logically break the entire description of the Architecture into separate sections to simplify their formation and on the other hand, to ensure that the architecture review from selected points of view or according to abstraction level (Lankhorst, 2017).

At the time when the works of Zachman were published, approach to the formation of the description of the system used the concept of "life cycle", including stages such as planning, analyzing, designing development, documentation, implementation and industrial exploitation. At each of these stages are considered issues associated with both system functions and data. Zachman suggested instead of the traditional approach associated with consideration of certain aspects of the system at different stages use the consideration of systems from various perspectives (points of view) (Lankhorst, 2017).

Historically, the Zachman model was first created for IT systems. This approach in the subsequent work was generalized to consider not only IT systems, but also to describe enterprises in general, so the proposed model can be used as a tool for describing of complex production systems of any type. The basic idea is to ensure ability to consistently describe each individual aspect of a system in coordination with others. For any a complex system the total nu0mber of connections, conditions and rules usually exceeds the capacity for simultaneous consideration. At the same time, separated, in isolation from others, consideration of every aspect of the system most often leads to non-optimal solutions, both in terms of performance and cost of implementation (Lankhorst, 2017).

The actual model is presented in the form of a table with five rows and six columns, which is shown in Figure 2.2. Note that there are five lines in the model, simply displayed in the figure, the sixth line corresponds to the level

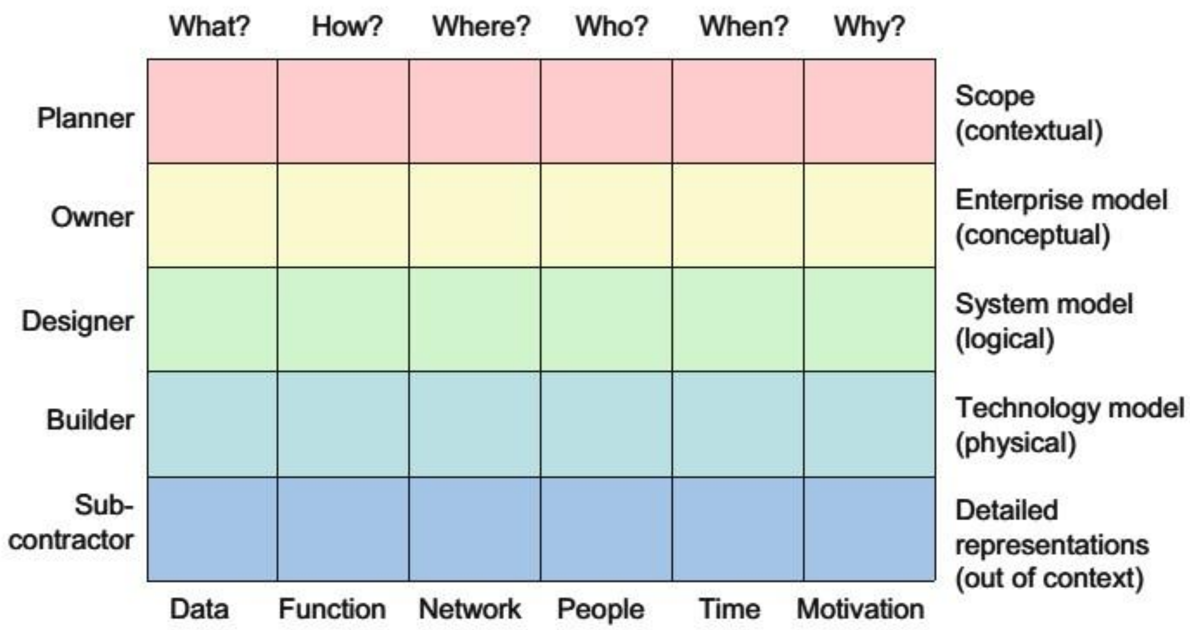


Figure 2.2 The Zachman framework, adapted from (Lankhorst, 2017)

of description architecture, and the level of the operating system or enterprise the whole event (Lankhorst, 2017).

The framework (Figure 2.2) in its most simple form depicts the design artefacts that constitute the intersection between the roles in the design process, that is, owner, designer and builder, and the produ ct abstractions, that is, what (material) it is made of, how (process) it works and where (geometry) the components are relative to one another. Empirically, in the older disciplines, some other ‘artefacts’ were observable that were being used for scoping and for implementation purposes. These roles are somewhat arbitrarily labelled planner and subcontractor and are included in the framework graphic that is commonly exhibited. (Zachman & Sowa, 1992)

From the very inception of the framework, some other product abstractions were known to exist because it was obvious that in addition to what, how, and where, a complete description would necessarily have to include the remaining primitive interrogatives: who, when and why. These three additional interrogatives would be manifest as three additional columns of models that, in the case of enterprises, would depict: who does what work, when do things happen, and why are various choices made (Lankhorst, 2017)?

Advantages of the Zachman framework (ZF) are that it is easy to understand, it addresses the enterprise entirely, and it is defined independently of tools or methodologies, and any issues can be mapped against it to understand where they fit. An important drawback is the large number of cells, which is an obstacle for the practical applicability of the framework. Also, the relations between the different cells are not that well specified. Notwithstanding these drawbacks, Zachman is to be credited with providing the first comprehensive framework for enterprise architecture, and his work is still widely used. (Lankhorst, 2017)

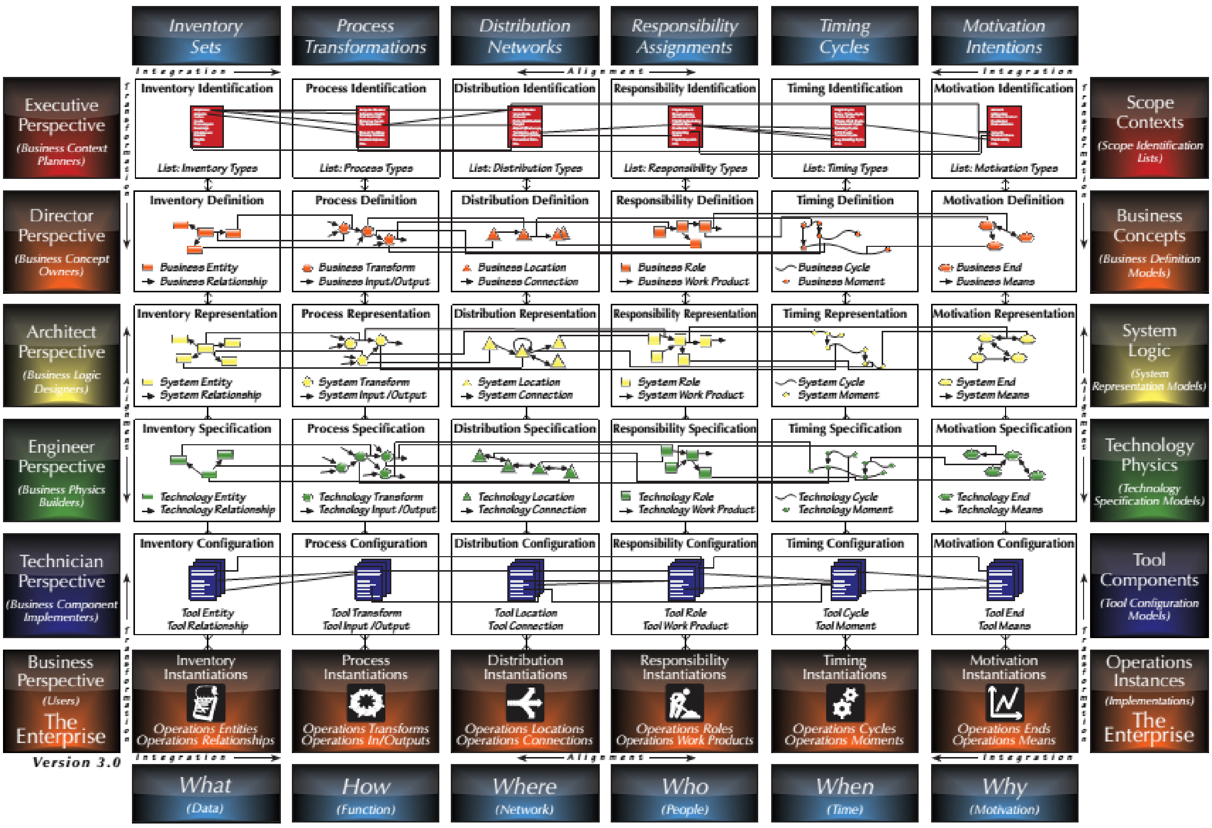


Figure 2.3 Zachman Framework v3.0 for Enterprise Architecture (Zachman, 2011)

To correct using of inconsistent terminology and clear up several misconceptions and misunderstandings that have occurred over the years with the original version of the Framework (Version 1.0) Zachman continued to work on his framework and created new versions: v2 (2008) and v3 (2011), see Figure 2.3 (Zachman, 2011).

2.1.4.1 Aspects of Zachman framework

"What"

The first column answers the question "WHAT?" and defines data used in the system. At the top level it is enough to put a simple listing of the main objects used by business. At the second level, these objects are combined into a high-level semantic model and usually described written down as an entity-relationship diagram (E-R diagram) with a reflection of the main links and the most significant business-constraints. At the third level, this model is reduced to normalized form, all attributes and keys are defined. The fourth level is a physical model of data in the system (in the object-oriented approach - class hierarchy). The next level contains a description of data management language model for the tables, new class libraries and database. Finally, the last level can describe the actual data sets, including characteristics such as access logs, the size of the actual disk space, query statistics, etc. (Lankhorst, 2017)

"How"

The function column (the answer to the question "HOW?") Is intended for detailed descriptions how the Mission Enterprise is implemented at the level of individual operations. On the first level it will be a simple enumeration of business-processes. The second level will contain del business processes, which is subsequently detailed as data operations and application architecture (level 3), class methods (level 4), program code (level 5) and, finally, executable code. Starting from the 4th level, consideration is no longer within the Enterprise as a whole, but on individual subsystems or applications. (Lankhorst, 2017)

"Where"

The next column (the question "WHERE?") defines the space distribution of system components and network organization. At the level of business planning, there is enough to define the location of all production facilities. At the next level, these objects are combined into a model with links, characterizing the interaction between entities – whether it is information or delivery of goods. At the third level of the system architecture is a binding of information system components to the network nodes. The fourth level defines a physical implementation in terms of hardware platforms, system software, and intermediate level tools (the so-called "Middleware") used to integrate various components of information systems. At the fifth level are defined protocols and specification of communication channels. Last level describes functioning of the implemented network. (Lankhorst, 2017)

"Who"

The column of the table that answers the question "WHO?" are participants of the process. At the business planning level here is showed the list of enterprise departments and their functions. The next level is organizational chart and common information security requirements. Next are defined participants of business processes and their roles, user interface requirements and access to individual objects, their physical implementation at the level of code or statements for access to tables in the DBMS. The last level describes trained users of system. (Lankhorst, 2017)

"When"

The fifth column answers the question "WHEN?" And defines time characteristics of business processes and system execution. Detailing is done from top to bottom, starting from schedule (level 1) and the main parameters of the business processes execution - for example, requirements for the time of the end of transaction (level 2). Third level is determined by events that cause a change in information objects and the initiation of operations. At the next level, these events are broadcasted as program calls (triggers) or transmitted messages. Fifth level determines the physical implementation of events processing. Finally, on the 6th level - the actual history of system functioning. (Lankhorst, 2017)

"Why"

The last column "WHY?" determines motivation and sets the paths of transition from business tasks to the requirements and elements of information systems. The starting point is the business strategy, which is consequentially translated into a business plan, then into rules and restrictions for the implementation of business processes, and at level 4, to relevant applications required for the composition of information systems and, subsequently, in their physical implementation. (Lankhorst, 2017)

2.1.4.2 Views (Perspectives) of Zachman framework

Scope Context (Executive Perspective)

This row shows Scope Context perspective – the big picture of desired architecture made by Business Context Planners - the executives or investors (Zachman, 2011) to estimate scope of the business (products, services, clients and business locations, strategy etc. (Bogomolova, 2016)), what it would cost and how it would perform. In analogy with classical architecture we can say it would be size, spatial relationships and basic purpose of the final structure (Zachman & Sowa, 1992).

Business Concepts (Director Perspective)

On this row is depicted the business concepts on the level of directors and process owners which correspond to business model, which constitutes design and shows business entities (organizational structure (Bogomolova, 2016)) and processes and how they interact (Zachman, 2011). In analogy with classical architecture it is architect's drawings that depict final building from the perspective of the owner (Zachman & Sowa, 1992).

System Logic (Architect Perspective)

This row shows system logic from the Architect Perspective (Zachman, 2011). Here are a business processes are already described in terms of information systems, including data elements, their transformation rules and processing to perform previous level 2 business-functions (Bogomolova, 2016). In analogy with classical architecture it is the architect's plans that are the translations of the drawings into detailed plan (Zachman & Sowa, 1992).

Technology Physics (Engineer Perspective)

At level 4 there is the Engineer Perspective addresses Technology Physics for Technology Specification Models (Zachman, 2011). The builders plan corresponding to the technology model, which must adapt the information system model to the details of the programming languages, I/O devices and other technology. In analogy with classical architecture it is builder's plan drown from architect's plan to consider contractor constrains of tools, technology and materials. (Zachman & Sowa, 1992)

Tool Components (Technician Perspective)

Fifth row of the framework is the Technician Perspective correspond to the detailed specification that are given to programmers (or other technicians) without being concerned with the overall context and structure of the system. In analogy with classical architecture those are the subcontractor's shop plans that specify the details or parts of subsections. (Zachman & Sowa, 1992)

Operations Instances (Business Perspective)

The last, sixth level, which is not presented in early versions of the framework (Zachman, 2011) describes the working system. At this level can described such objects as instructions for working with the system, the actual databases data from the end user point of view (Bogomolova, 2016).

2.1.5 Artefacts for the representation of Enterprise aspects

All six aspects of the Zachman enterprise framework can be represented with specific for each column artefacts. Some of possible artefacts (see Figure 2.4) but without specific standard recommendations were show by Zachman in his framework (Zachman, 2011)

	DATA *what*	FUNCTION *How*	NETWORK *Where*	PEOPLE *Who*	TIME *When*	MOTIVATION *Why*	
Objective Scope ***Contextual*** *Role: Planner*	List of Things Important in the Business	List of core Business Process	List of Business Locations	List of Important Organizations	List of Events	List of Business Goals/Strategies	**Objective Scope** ***Contextual*** *Role: Planner*
Enterprise Model ***Conceptual*** *Role: Owner*	Conceptual Data/ Object Model	Business Process Model	Business Logistics System	Work Flow Model	Master Schedule	Business Plan	**Enterprise Model** ***Conceptual*** *Role: Owner*
System Model ***Logica*** *Role: Designer*	Logical Data Model	System Architecture Model	Distributed Systems Architecture	Human Interface Architecture	Procesing Structure	Business Role Model	**System Model** ***Logica*** *Role: Designer*
Technology Model ***Physical*** *Role: Builder*	Physical Data/ Class Model	Technology Design Model	Technology Architecture	Presentation Architecture	Control Structure	Rule Design	**Technology Model** ***Physical*** *Role: Builder*
Detailed Representations ***out of Context*** *Role: Programmer*	Data Definations	Program	Network Architecture	Security Architecture	Timing Defination	Rule Specification	**Detailed Representations** ***out of Context*** *Role: Programmer*
Functioning Enterprise *Role: User*	Usable Data	Working Function	Usable Network	Functioning Organization	Implemented Schedule	Working Strategy	**Functioning Enterprise** *Role: User*

Figure 2.4 Zachman Framework artefacts (Zachman, 2011)

For purposes of this work we will consider Conceptual perspective of Zachman framework. The justification of this choice is the following: Scope level is too general and some enterprises can be described equally though there are different that we can see only on following layers. Logical and more detail perspective are too specific and can even distinguish in the same company. Moreover, a very detailed description can make representation very complex, company specific and thus, become an incomparable model.

~~More specifically a~~Artefacts for implementation of Zachman framework were proposed later (Noran, 2003). Such diversity of standards was explained by Noran: “Today, no single existing modelling language by itself is capable of modelling all necessary aspects of an enterprise” (Noran, 2003). Possible modelling languages to populate Zachman’s modelling framework are shown in the Figure 2.5. For the Conceptual level there were considered following standards: WHAT – ERM, HOW – IDEF3, WHERE – Graph, WHO – Org Chart, WHEN - Gantt chart, WHY – Structured English.

	What	How	Where	Who	When	Why
Scope (Contextual)	RP / English	RP / English	RP / Map	RP / English	RP / English	RP / English
Enterprise Model (Conceptual)	ER(M), IDEF1, UML Class	IDEF0, IDEF3 UOB, UML Act, GRAI Nets	Graph	Org Chart, GRAI Grid	GANTT / PERT, IDEF3 OSTN, Timed Petri	Struct English
System Model (Logical)	ER, IDEF1x, UML Class	UML Use Case Data Flow Diag.	UML Component	GRAI Grid, UML Use Case	Data Flow, IDEF3 OSTN, Timed / Colored Petri	FOL, Struct English, Z
Technology Model (Physical)	Relational, UML Class	UML Class, Activity, Structure Chart	UML Deployment	UML Real Use Case, UI Design	UML Sequence, Collab, State, Statecharts	FOL, Struct English, Z
Components (Out of Context)	DB Schema	Programming language	URL, IP, TCP/IP	UI Programming language	Struct English	Rule spec. In Prg Lang
Functioning Enterprise	DD L(*SQL*)	Machine code (0/1)	Address, Comm language	User / Worker	English (Schedule)	English

Figure 2.5 Possible modelling languages for Zachman framework (Noran, 2003)

2.1.5.1 Artifacts for data representation

We can represent our What (Data) perspective of Zachman Framework with ER diagrams (Noran, 2003). Entity-relationship concept was applied for data modeling (Chen, 2002) which explains how data entities interact (relate) with each other (see Figure 2.6).

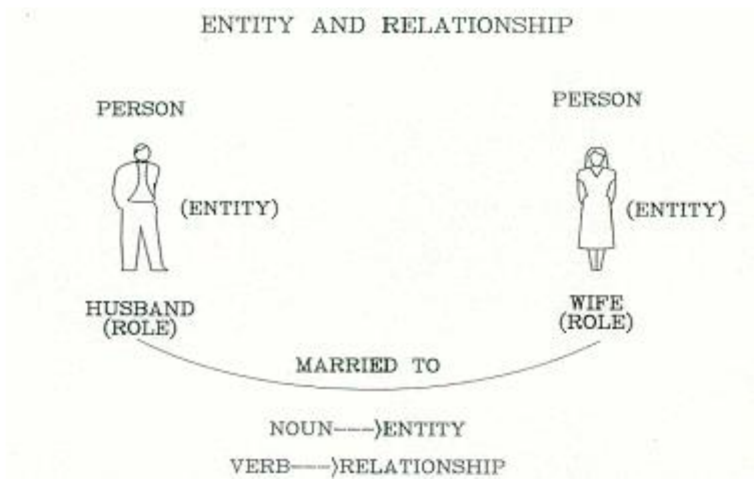


Figure 2.6 The concept of Entity and Relationship (Chen, 2002)

ER diagrams were proposed by Piter Chen (Chen, 1976) which became a standard for database development. They show how normalized data tables corresponds by key and are an outline of the database. We can model any structures data using Entity-Relationship diagram models. This representation is perceived by people very easily and is proliferated widely, so it can be used for representation of Conceptual perspective of Data aspect of Zachman framework. An ER diagram example for analysis of information in a manufacturing firm is depicted on the Figure 2.7.

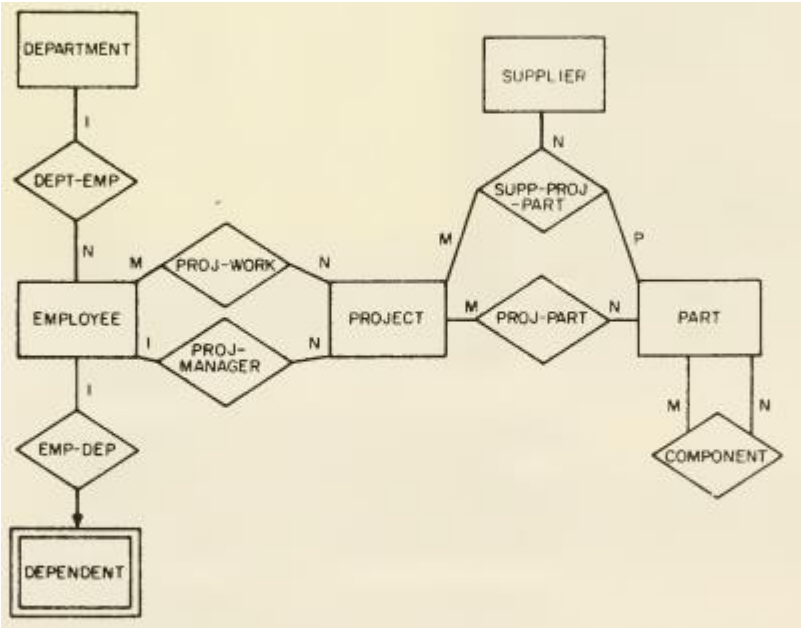


Figure 2.7 An ER diagram example (Chen, 1976)

For the purposes of this paper, we not only have to consider ~~only~~ the representation designed for human, but also one consigned for computers. For this, W3C (with Peter Chen as one of participants) (W3C, 1999a) proposed the XML format (Chen, 2002) for digital ER diagram representation. ~~There were~~They proposed several works for converting ER to XML (Sung & Kang, 2007), (Franceschet M., Gubiani D., Montanari A., Piazza C, 2009). ~~,~~ T~~t~~hus we can ~~make~~ represent~~ation of~~ our Data ~~perspective that~~ for a computer ~~can~~ to treat (see Figure 2.8).

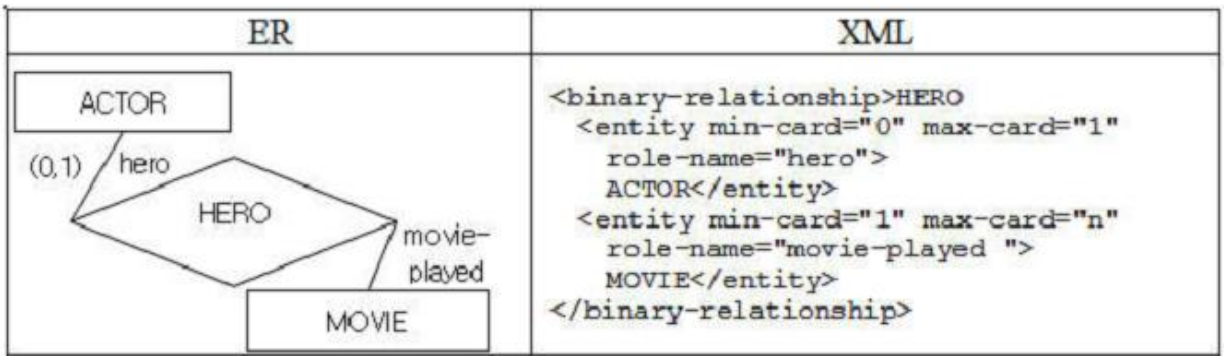


Figure 2.8 Mapping relationship from ER to XML (Sung & Kang, 2007)

2.1.5.2 . Artifacts for representation of functions at the process level

The second perspective of Zachman Framework is How (Function). It represents business processes workflow. For a Conceptual aspect, it (Noran, 2003) ~~proposed~~ proposes the IDEF3 standard, ~~but now it become~~which is now obsolete~~,~~ and not supported by many tools. ~~On the contrary~~However, the BPMN 2.0 standard (Object Management Group (OMG), 2014) is widely used and implemented in many businesses process modeling (BPM) and business process management system (BPMS) tools such as SAP, Bizagi, Business Studio, ELMA etc.

BPMN model graphically depicts business process workflow in intuitive manner as depicted on the Figure 2.9. Model can be executed by BPMS system following all build-in logic (Dumas, Reijers, La Rosa, & Mendling, 2015).

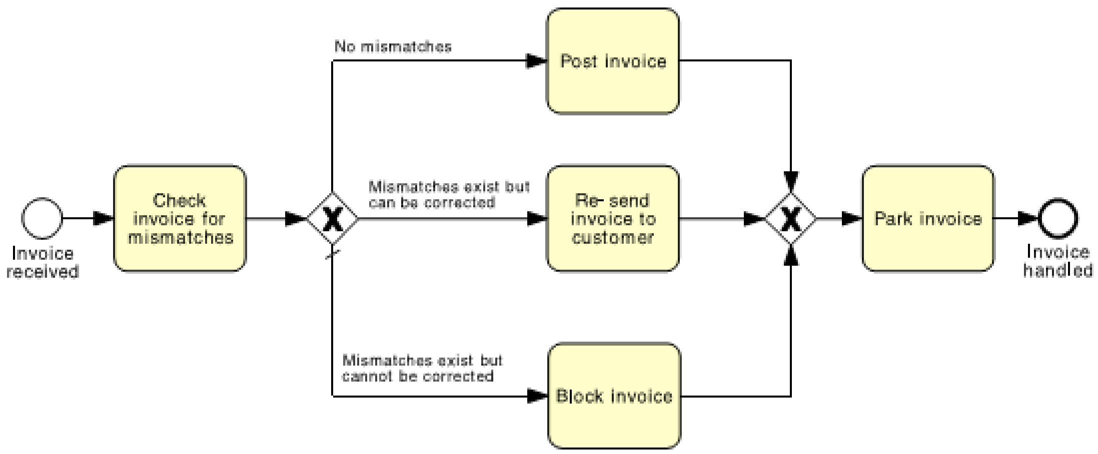


Figure 2.9 Simple BPMN process example (Dumas, Reijers, La Rosa, & Mendling, 2015)

The structure of process artifacts presented in BPMN format is shown on Figure 2.10. “We can see finite set of objects that can be represented both graphically and in XML formats. It consists of a list of elements, where some are optional (those with a dashed border) and others are mandatory (those with solid borders). The process element is mandatory and stores information about the process model. This consists of electronic data objects, events, tasks and flows” (Dumas, Reijers, La Rosa, & Mendling, 2015).

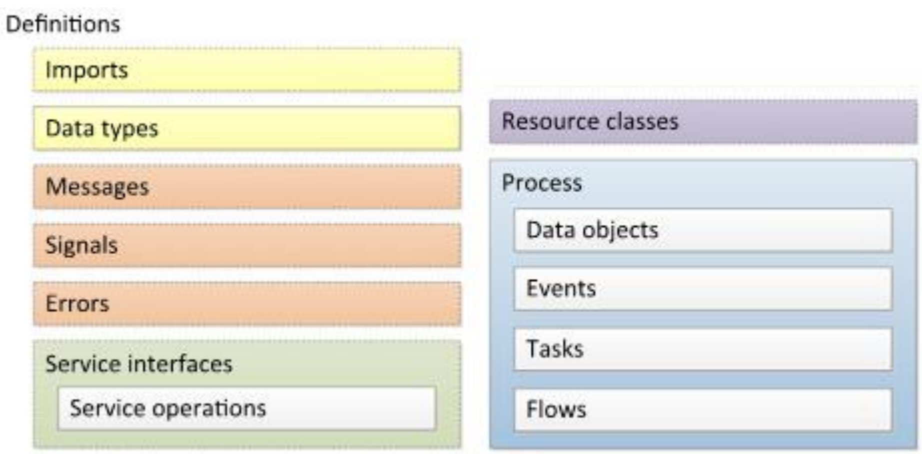


Figure 2.10 Structure of the BPMN format (Dumas, Reijers, La Rosa, & Mendling, 2015)

In the paper "Using the Zachman Framework to Achieve Enterprise Integration Based-on Business Process Driven Modelling" (Espadas, Romero, Concha, & Molina, 2008) authors showed the approach of applying BPMN 2.0 within the Zachman framework to achieve enterprise integration. Example of this approach is presented on Figure 2.11.

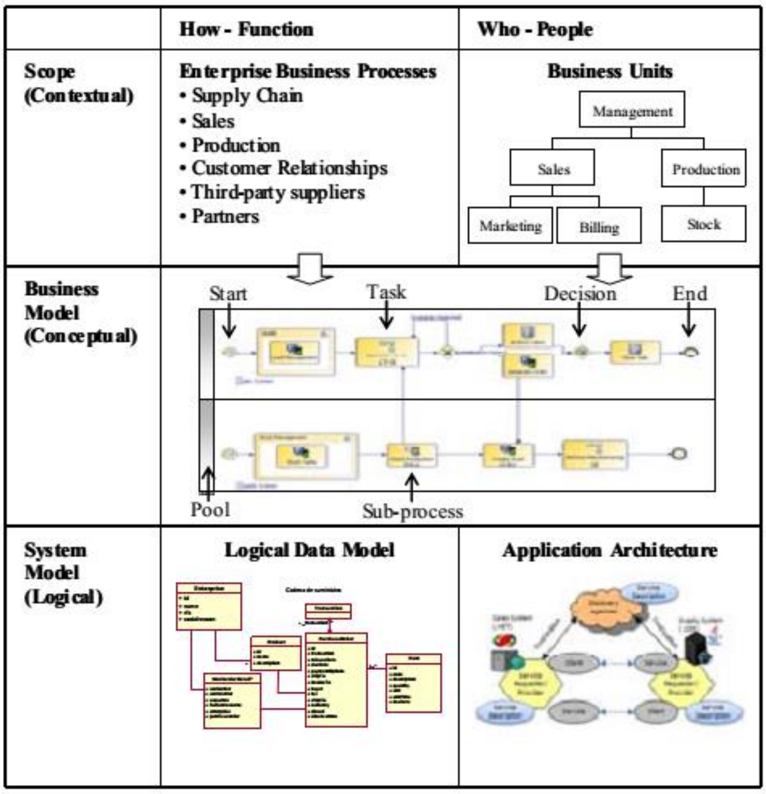


Figure 2.11 BPMN applyed to EA framework (Espadas, Romero, Concha, & Molina, 2008)

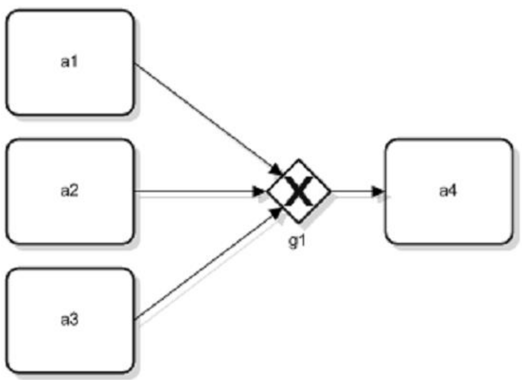
There are few formats for XML representation of BPMN standard (see Figure 2.12) proposed by different

pattern	standard						
	XPDL	UML	BPEL4WS	BPML	XLANG	WSFL	WSCI
1 (seq)	+	+	+	+	+	+	+
2 (par-spl)	+	+	+	+	+	+	+
3 (synch)	+	+	+	+	+	+	+
4 (ex-ch)	+	+	+	+	+	+	+
5 (simple-m)	+	+	+	+	+	+	+
6 (m-choice)	+	-	+	-	-	+	-
7 (sync-m)	+[6]	-	+	-	-	+	-
8 (multi-m)	-	-	-	+/-	-	-	+/-
9 (disc)	-	-	-	-	-	-	-
10 (arb-c)	+[7]	-	-	-	-	-	-
11 (impl-t)	+[8]	-	+	+	-	+	+
12 (mi-no-s)	+	-	+	+	+	+	+
13 (mi-dt)	+	+	+	+	+	+	+
14 (mi-rt)	-	+	-	-	-	-	-
15 (mi-no)	-	-	-	-	-	-	-
16 (def-c)	-	+	+	+	+	-	+
17 (int-par)	-	-	+/-	-	-	-	-
18 (milest)	-	-	-	-	-	-	-
19 (can-a)	-	+	+	+	+	+	+
20 (can-c)	-	+	+	+	+	+	+

Figure 2.12 A comparison of XPDL with other standards (van der Aalst, 2004).

institutions and pursuing different goals: XPDL UML BPEL4WS BPML XLANG WSFL WSCI. The XPDL format, proposed by The Workflow Management Coalition (WfMC, 2012) is one of the perspectives, especially intended for BPMN models exchange (WfMC, 2012) among different supporting BPMN tools. In Figure 2.12 we see, that XPDL covers a significant part of Petri-Nets (van der Aalst, 2011) workflow patterns (van der Aalst, 2004).

Thus, this standard is used for BPM tools exchange and process mining ProM (Kalenkova, de Leoni, & van der Aalst, 2014) and it can be used for gathering models for the XML representation of Enterprise DNA.




```
<Activities>
<Activity Id="3" Name="a1"/>
<Activity Id="4" Name="a2"/>
<Activity Id="5" Name="a3"/>
<Activity Id="6" Name="g1">
<Route GatewayType=”Exclusive”" MarkerVisible="TRUE"/>
</Activity>
<Activity Id="7" Name="a4">
</Activity>
</Activities>
<Transitions>
<Transition Id="8" Name="" From="3" To="6" FlowType="SequenceFlow"/>
<Transition Id="9" Name="" From="4" To="6" FlowType="SequenceFlow"/>
<Transition Id="10" Name="" From="5" To="6" FlowType="SequenceFlow"/>
<Transition Id="11" Name="" From="6" To="7" FlowType="SequenceFlow"/>
</Transitions>
```

Figure 2.13 Example of XPDL derived from BPMN (WfMC, 2012).

Thus, for XML representation of functions on the process level of Zachman Enterprise Architecture Framework, the XPDL format can be used. The example of XPDL format code, that was derived from BPMN merge gateway process workflow model is depicted on the Figure 2.13.

2.1.5.3 . Artifacts for network representation

The third perspective of Zachman Framework is Where (Network). It represents geographical and logistic structure of enterprise and its parts (Zachman, 2011) and (Hay, 2000). For Conceptual aspect (Noran, 2003) proposed Graph (see Figure 2.14) as a representation of company's logistic (supply chain) network. (Chatfield, Harrison, & Hayya, 2004) proposed The Supply Chain Modeling Language (SCML) as "a general, reusable, platform and methodology independent standard for describing a supply chain's structure and logic". ~~Based on~~According to the XML standart, Supply Chain Modeling Language (SCML) is "open, standard data format for supply chain modeling can make supply chain analysis, especially via simulation modeling, more robust, yet more accessible" (Chatfield, Harrison, & Hayya, 2004).

The five basic elements of SCML for describing a supply-chain are:

- Node
- Arc
- Component
- Action
- Policy.

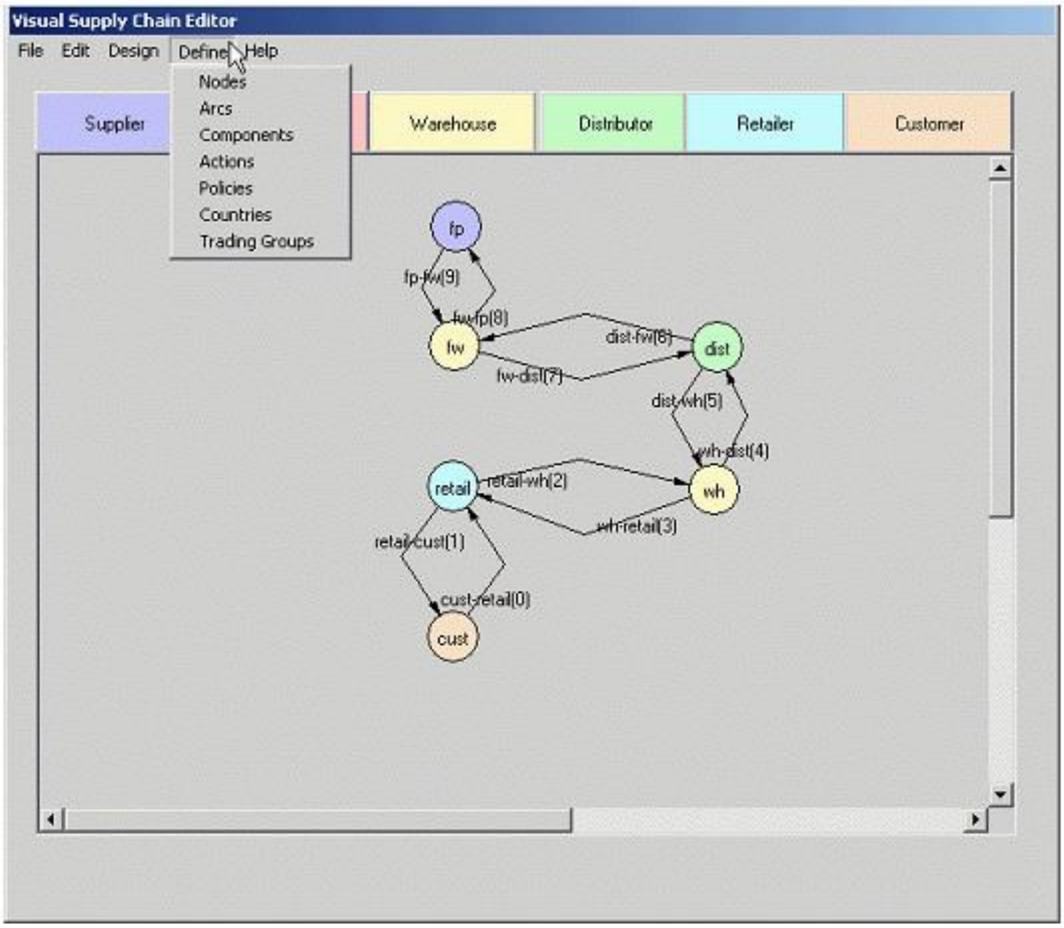


Figure 2.14 Supply chain topology network graph (Chatfield, Harrison, & Hayya, 2003).

Below is an example of document type definition of SCML (Chatfield, Harrison, & Hayya, 2004):

```
<!ELEMENT arc (arcContainer, arcCosts, arcTravelTime, arcPolicies, initialLevels)>
<!ATTLIST arc
name ID #REQUIRED
sourceNode IDREF #REQUIRED
destinationNode IDREF #REQUIRED
mode (land | rail | sea | air | telecomm |
other) "land"
distance CDATA #IMPLIED
unitOfDistance CDATA "miles"
capacity CDATA #REQUIRED
unitOfCapacity CDATA "unit"
containersOnHand CDATA #IMPLIED>
```

2.1.5.4 . Artifacts for people representation

The fourth perspective of Zachman Framework is Who (People). It represents company's structure and functions performers, which can be grouped in many ways with different degree of responsibility, centralization, thus, forming managerial hierarchy (Mintzberg, 1979). For Conceptual aspect (Noran, 2003) proposed Org Charts as representation, that is simple, straightforward and widely used by business entities. There are following main types of organizational structures (Legaard & Bindslev, 2006):

- Simple structure;
- Hierarchical system;
- Functional organization;
- Product organization;
- Matrix organization.

The Organizational Chart is one of possible Organizational Structure representations. Example of Org Chart of hierarchical organization is shown on Figure 2.15.

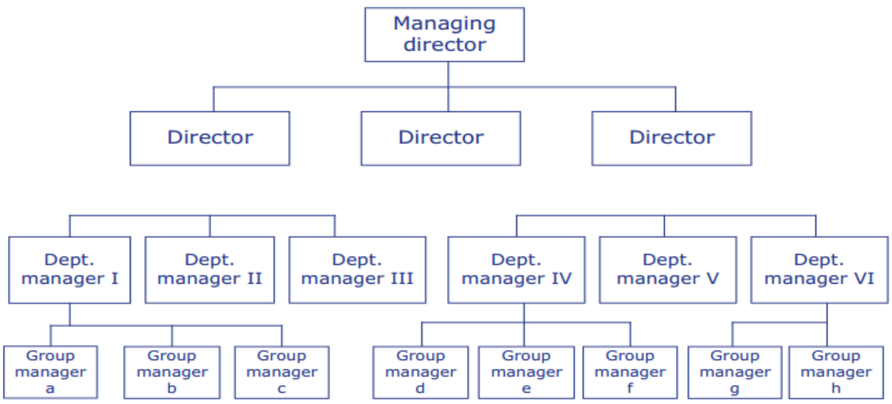


Figure 2.15 Organizational chart example (Legaard & Bindslev, 2006) .

The Organization Chart usually have hierarchical structure and can be represented in XML format (Bulajic & Filipovic, 2012), because both (Org Chart and XML) have a hierarchical structure. The example of organizational structure and its translation into XML format is depicted on Figure 2.16.

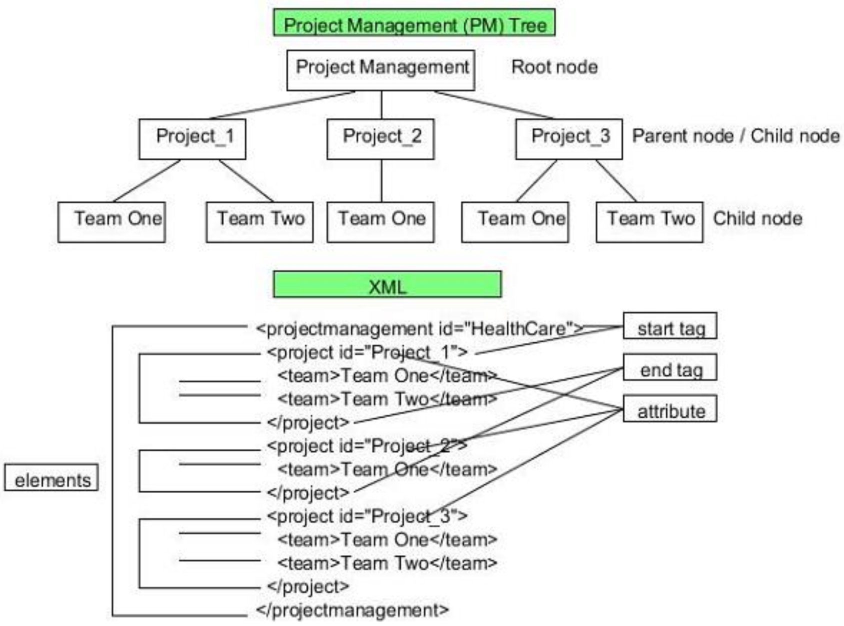


Figure 2.16 Example of XML format for Org Chart modelling (Bulajic & Filipovic, 2012)

2.1.5.5 . Artifacts for time representation

The fifth perspective of Zachman Framework is When (Time). It represents time of events significant for the operating of the company (Zachman & Sowa, 1992). It can be business processes start, pause and end events, projects timelines and other significant time events. The interconnected time events and project steps in Zachman framework can be represented as Gantt Charts (Noran, 2003). The example of classical Gantt Chart in project management tool is shown on Figure 2.17.

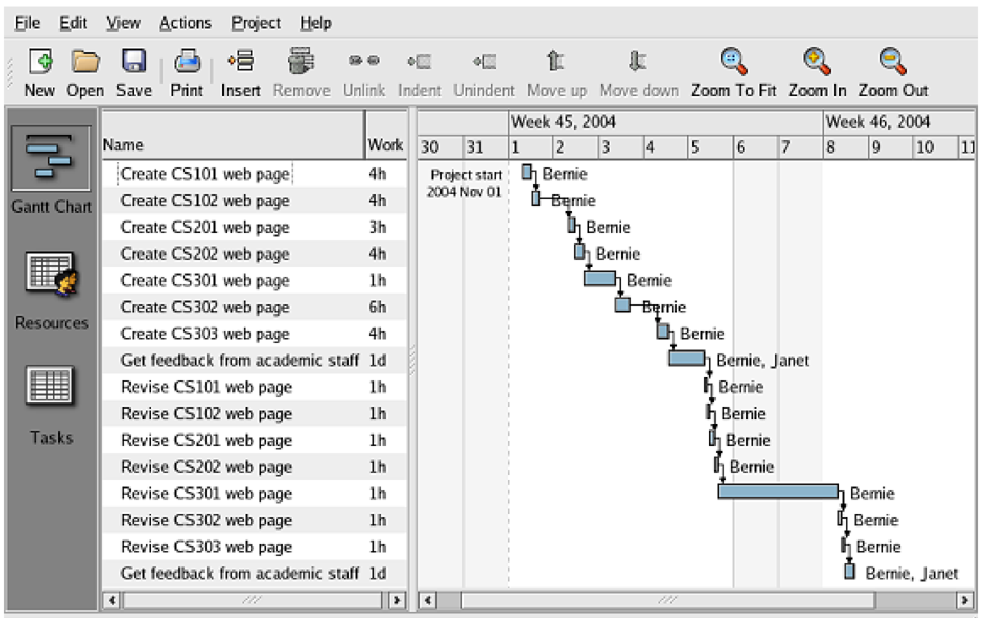


Figure 2.17 Gantt Chart example (Dale, Churcher, & Irwin, 2005).

(Dale, Churcher, & Irwin, 2005) showed XML Project Management Method (PMM) representation of Gantt Chart with PMXML (Project Management XML) format (see Figure 2.18).

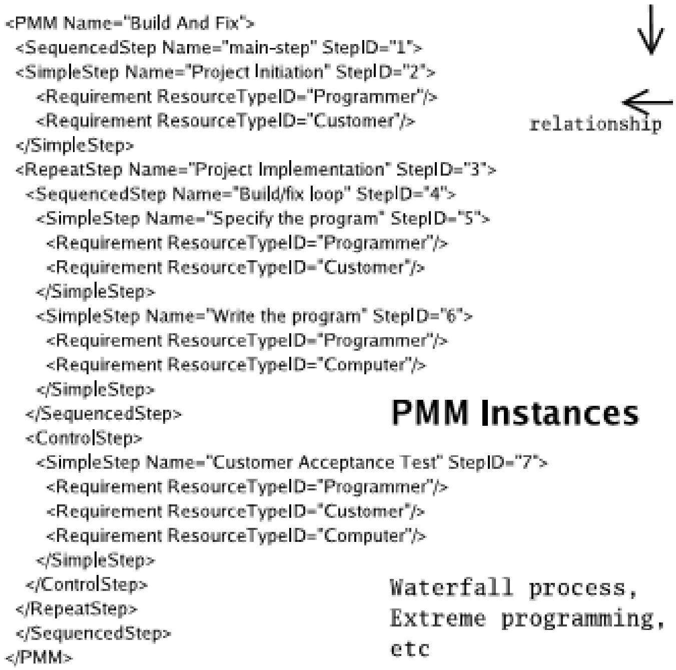


Figure 2.18 Gantt Chart XML representation (Dale, Churcher, & Irwin, 2005)

(Varela, Aparício, & Silva, 2005) also proposed XML representation of schedules and Gantt Charts, example of which is shown below:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE gantt SYSTEM "gantt.dtd">
<gantt>
  <problems>
    <problem id="0010" class="F3|n|Cmax">
        <results>
            <job id="J1" name="Job1"/>
              <machine id="M1" name="Machine1"/>
              <start>0</start>
              <finish>3</finish>
            </job>…
          </results>
    </problem>
  </problems>
</gantt>
```

2.1.5.6 . Artifacts for motivation representation

The last (sixth) perspective of Zachman Framework is Why (Motivation). It represents goals of the company (Zachman & Sowa, 1992). Business goals framework was proposed by (Kaplan & Norton, 1992) as Balanced Scorecard which became widely used by the business and consulting. Balanced Scorecard represent business goals connected and grouped by four perspectives: customer, finance, internal operations and development. Four perspectives of interrelated company's goals - Balanced Scorecards must be agreed with each other in order to achieve them. The original model of BSC, proposed by Kaplan and Norton is shown on Figure 2.19.

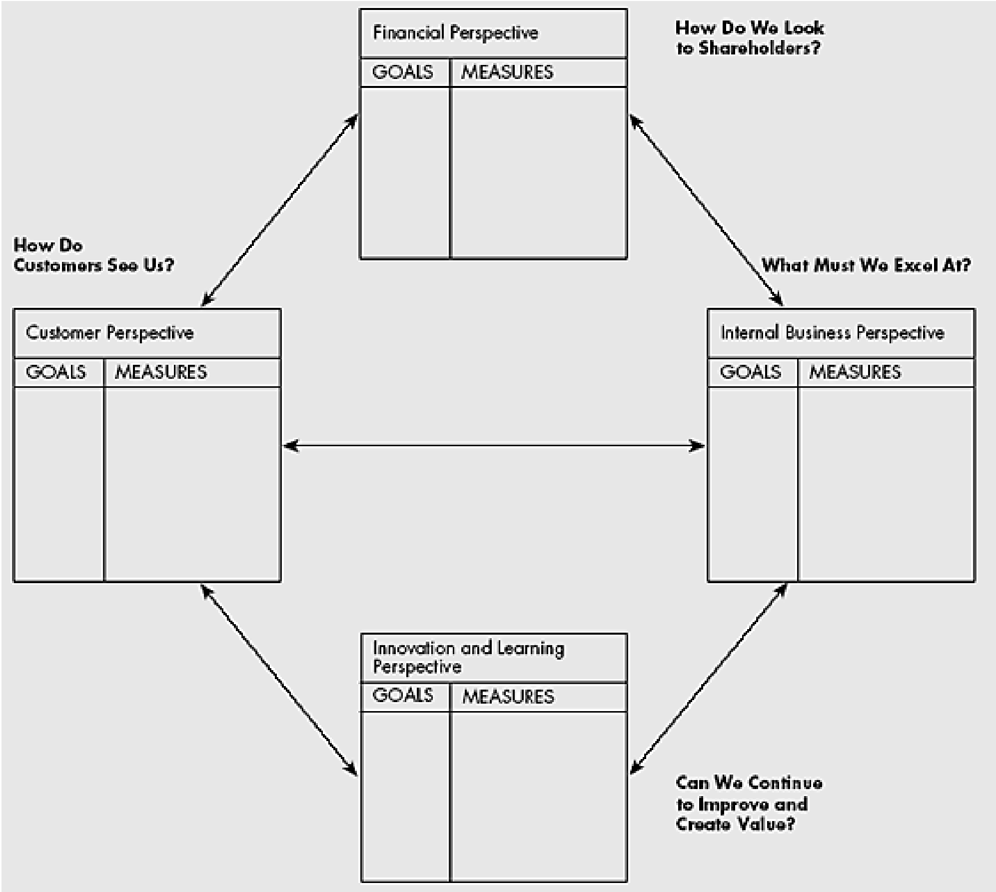


Figure 2.19 Balanced Scorecards performance measures (Kaplan & Norton, 1992).

(McGregor, 2003) showed that BSC XML draft standards can be used for XML representation of company's goals - Balanced Scorecards. The example of BSC XML implementation is depicted on Figure 2.20.

```
<BSCDoc language="EN">
<BSC id="SLBSC" name="Sell Loan Balanced Scorecard">
      <PERSPECTIVE id="SLC" name="Sell Loan Customer Perspective"
         perspectiveType="Customer">
            <Description>Customer Perspectives associated with Sell Loan
            Example</Description>
            <CUSTOMERTYPE id="C1" name ="Personal"/>
            <PRODUCTTYPE id="UPL" name="Unsecured Personal Loans"/>
            <Milestone id=CM0 name="Commence Sell Loan Process"/>
            <Milestone id=CM1 name="Request Loan"/>
            <Milestone id=CM2 name="Send Letter of Offer"/>
            <Milestone id=CM3 name="Receive Signed Letter of Offer"/>
            <Milestone id=CM4 name="Settle Loan"/>
            <OBJECTIVE id="CSL" name="Customer Service Levels" updateFrequency="" ownerRef="">
                  <MEASURE id="CM1M2" name="Send Letter of Offer" type="cycle time" unit="hours"
                  ownerRef="">
                        <Description>Cycle time target from completion of application to sending letter of
                        offer</Description>
                        <StartMilestone>CM1</StartMilestone>
                        <EndMilestone>CM2</EndMilestone>
                        <Target id="CM1M2" TargetValue="24"/>
                  </MEASURE>
                  <MEASURE id="CM2M2" name="Send Letter of Offer" type="cycle time" unit="hours"
                  ownerRef="">
                        <Description>Cycle time target to send letter offer</Description>
                        <StartMilestone>CM2</StartMilestone>
```

Figure 2.20 Excerpt of the BSC XML example (McGregor, 2003).

2.2. ARTIFICIAL LIFE

2.2.1. Concepts

~~The b~~Biology is a study of carbon-based life forms, while Artificial Life is a science of synthesis of artificial life forms~~,~~ that never existed before. ~~The b~~Biology as a science researches living organisms, ~~and decompose them starting from~~ biological organizational hierarchy, ~~then~~ specific species, their organs, tissues, cells and molecules. The dynamics of living organisms is non-linear, and its self-organization is hard to describe by means of analytical methods because it fundamentally depends on interaction of its different parts. While the biology studies organisms by their decomposed parts to give a whole picture the Artificial Life, on the contrary, is a synthetic approach to biology – it considers living things not separately, but as a whole, putting them all together (Langton, 1996).

There is a distinction between Artificial Intelligence (AI) and Artificial Life (ALife) modelling strategies as "there is important difference between the modeling strategies AI and ALife typically employ. Most traditional AI models are top-down-specified serial systems involving a complicated, centralized controller that makes decisions based on access to all aspects of global state. The controller's decisions have the potential to affect directly any aspect of the whole system" (Bedau, 2003b). On the contrary, ALife models populations of low-level autonomous "agents" that simultaneously interact with each other and every agent makes decisions based on its own situation. Thus, Artificial Life models copy the characteristics end development from examples in~~t the m~~nature (Bedau, 2003b).

Thus, ALife is a set of techniques and approaches ~~having in common the analogy with a natural life and thus,~~that simulating behavior and evolution of living species. (Langton, 1989). There are three main kind of Artificial Life:

Wet, Hard and Soft:

- **The "wet" artificial life** is intended "to create artificial cells out of biochemicals. Such artificial cells would be microscopic, autonomously self-organizing and self-replicating physical entities that assemble themselves out of non-living materials. Although artificial, they would repair themselves and adapt in an open-ended fashion, so for all intents and purposes they would be alive. The first artificial cells will probably just move through a fluid and process chemicals. To do even this flexibly and robustly, they must solve the functions of self-maintenance, autonomous control of chemical processing, autonomous control of mobility, and self-replication". (Bedau, 2003a)
- **"hard' life** produces hardware implementations of life-like systems, and 'wet' artificial life synthesizes living systems out of biochemical substances. Cognitive science and artificial life share some intellectual roots, and their subjects and methodologies are related. Now that artificial life has matured over the past decade or so, it is appropriate to review its achievements and speculate about its future connections with cognitive science. (Bedau, 2003a)
- **'Soft' artificial life** creates simulations or other purely digital constructions that exhibit life-like behavior, 'hard' artificial life produces hardware implementations of life-like systems (Bedau, 2003a). Soft ALIfe involves attempts to a) synthesize the process of evolution, b) in computers c) will be interested in whatever emerges in process, even if the results have no analogues in the 'natural' world (Langton, 1996).

Artificial Life employs such fundamental notions of biology as reproduction, evolution, natural and artificial selection, genetic connection, morphogenesis, ontogenesis and phylogenesis, genotype and phenotype, chromosome, organism (individual) and population (Tarasov, 2002). The most important artifacts are (Tarasov, 2002), (Langton, 1996):

- **Individual** – is a model of low-level agent that has specific properties and behavior;
- **Population** – is a set of interacting individuals;
- **Evolution** – is a process of adaptation of population and each individual to specific environment ~~conditions, that~~conditions that is measured by means of individual fitness. during the evolution individuals are replaced by their offspring;
- **Genotype** – it is an internal characteristic~~,~~ code which encodes ~~with~~ the linear sequence of nucleotide (Genes) with ~~the~~ instructions on how to build and organism, and is called DNA;
- **Phenotype** – is the behavioral property of existing organism that defines the adaptability to ~~the~~ a specific environment. Due to non-linear interactions of genes, the phenotype is a non-linear function of genotype.;
- **Chromosome** – the representation of individual Genotype (in the nature it is represented by the DNA)

For the purposes of this paper we consider the “Soft” Artificial Live that can be modelled by means of Evolutionary (Genetic) Algorithms.

2.2.2. A brief history of Artificial Life

The term “Artificial Life” and fundamental models of this research topic related to the work of the University of Cambridge professor John Horton Conway on cellular automata. He called it Game of Life and published in Scientific American magazine (Gardner, 1970) .

The phrase “Artificial Life~~”, that~~” that became the name of the science was proposed by Christopher Langton who was an unconventional researcher. He became interested by Conway's Game of Life and started to develop the idea that the computer could mimic living creatures. He intentionally applied the term as a title for the ‘‘interdisciplinary workshop on the synthesis and simulation of living systems’’ that he organized in September 1987, in Los Alamos, New Mexico (Langton, 1989).

Langton gave the definition of ALife as following: “Artificial life is the study of artificial systems that exhibit behavior characteristic of natural living systems. This includes computer simulations, biological and chemical experiments, and purely theoretical endeavors. Processes occurring on molecular, cellular, neural, social, and evolutionary scales are subject to investigation. The ultimate goal is to extract the logical form of living systems” (Langton, 1987).

In 1987, computer animator Craig Reynolds created flocking behavior in a computer program to animate groups of ~~boids~~birds. He used three simple rules to produce lifelike ~~modal~~ models to avoid obstacles ~~laying~~ on their path. This work emerged ~~many~~ as ~~alife~~ Alife research by movie producers who ~~was~~ were trying to find realistic animation of natural plants, creatures and inanimate nature for entertainment industry (Reynolds, 1987).

In 2006, Peter Turchin and Mikhail Burtsev were studying evolution of cooperative behavior by means of game theory and agent-based simulation. In their model, agents were endowed with a limited set of receptors, a set of elementary actions and a neural net in between. Behavioral strategies were not predetermined instead, the process of evolution constructed and reconstructed them from elementary actions. As a result, cooperative strategies could evolve even under such minimal restrictions and agents could perceive heritable external markers of other agent (Burtsev & Turchin, 2006).

2.2.3. Genetic algorithms

2.2.3.1 Concepts

Genetic algorithms (GA) are stochastic, heuristic optimization methods, first proposed by John Holland in 1975 (Holland, 1975). They are based on the idea of evolution through natural selection. In addition to finding the extremum more quickly, the positive properties of genetic algorithms include the finding of a "global" extremum. In tasks where the objective function has a significant number of local extremes, in contrast to the gradient method, genetic algorithms do not "get stuck" at the points of local extremum but allow us to find the "global" optimum.

Genetic algorithms work with a collection of individuals - a population where each individual represents a possible solution to this problem. It is estimated by the measure of its "fitness" according to how well the corresponding solution to the problem is. In nature, this is equivalent to assessing how effective an organism is in competing for resources. The most adapted individuals can "reproduce" the offspring by means of "crossing" with other individuals of the population. This leads to the appearing of new individuals, which combine some characteristics that they inherit from their parents. The least adapted individuals are less likely to reproduce their descendants, so that the properties they possessed will gradually disappear from the population in the process of evolution. Sometimes mutations occur, or spontaneous changes in genes (Holland, 1975).

Thus, from generation to generation, good characteristics spread among the population. Crossing the fittest individuals leads to the fact that the most promising parts of the search space are inherited. Ultimately, the population will converge to the optimal solution of the problem. The advantage of a GA is that it finds approximate optimal solutions in a relatively short time (Goldberg, 1989).

GA operates on the following terminology:

- **Chromosome** - the solution to the problem, the carrier of hereditary information. The set of chromosomes (the values of the parameters of the objective function) characterizes the individual. Chromosome consists of genes.
- **Genes** are coding elements of hereditary information (parameters of the objective function). As genes often acts bit-coding information.
- **An individual** is a set of chromosomes (a set of parameters for which the value of the objective function is sought).
- **Individual fitness** - the value of the objective function for a given set of parameters in relation to the desired value

GA performs on individuals the following actions

- Generation of **the initial chromosome population** - randomly selected values of the parameters of the objective function and for these values of the parameters is the value of the objective function.
- Breeding (**Selection**) - the choice of individuals with the best adaptability for reproduction (sorting by the value of the objective function). The better the individual's fitness, the higher its chances of the next generation crossing and inheriting its genes.

- **Crossover** - crossbreeding. The break point is randomly selected - the area between adjacent bits in the string. Both parent structures are broken into two segments at this point. Then, the respective segments of different parents are glued together, and two genotypes of descendants are obtained. There are variations with two point of cutting the chromosome.
- **Mutation** is a random change of genes. A randomly selected gene changes to another with some probability and bring new features and behaviour.

First, the GA function generates a certain number of possible solutions (individuals), and then calculates fitness for each one - proximity to the truth. These decisions give offspring (a crossover operation is performed). Better solutions have a greater chance of reproduction, and "weak" individuals gradually "die off." Thus, the process of evolution happens. At certain stages of this process, spontaneous changes in genes occur (mutations and inversions). Useful changes that lead to an increase in the individual's fitness give birth to their offspring, while “useless” changes die off. After crossing, mutations and inversions, the fitness of individuals of the new generation is again determined. The process is repeated until a solution is found or a sufficient approximation is obtained (Goldberg, 1989).

2.2.3.2 Genetic Algorithm work-flow

The classic generic Genetic Algorithm is shown on the Figure 2.21. There are plenty of variations of GA, mainly distinguished by the implementation of genetics operators: Initialization, Selection, Mutation and crossover. Moreover, the important difference in implementation of Genetic Algorithm for specific tasks is the representation of the individual.

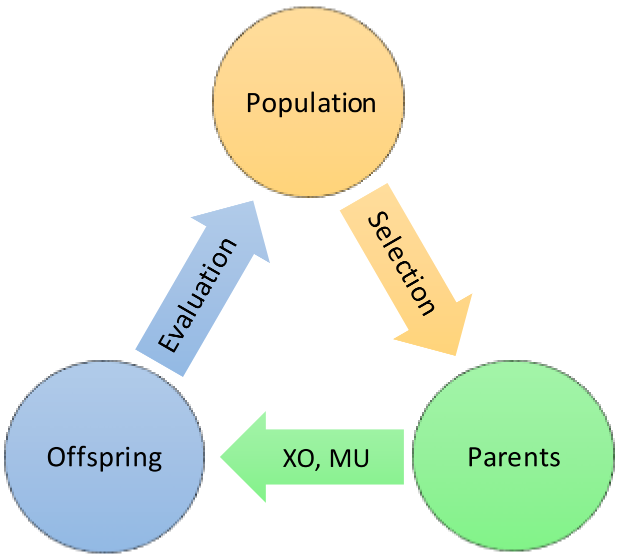


Figure 2.21 Generic Genetic Algorithm workflow.

The Genetic Algorithm have following stages (initialization step is not shown on the picture above), that are depicted on Figure 2.21:

- **Initialization** – setup of the population with the individuals, filling their chromosome with specific values, derived randomly or by means of particular algorithm;
- **Selection** – selecting individuals (For example couple of parents) for applying genetic operators.

- **Execution of Genetic Operators** such as crossover (XO) and mutation (MU). Implementation of genetic operators is specific for the particular task.
- **Evaluation** - is applying fitness function for each individual of the population to decide which of them have better fit for the environment in terms of the task. More fitted individuals form new population and cycle starts again.

After convergence of fitness function to the global optimum or iteration of specified number of epochs (generations) algorithm stops and the individual with the best fitness represent the solution of the problem.

2.2.3.3 Genetic Algorithm representation

For implementation of Genetic Algorithm, the DNA of the individual must be represented in digital and convenient way to make the solution of the problem possible. This representation must allow for the crossover and mutation of ~~to perform~~ genetic operators, ~~as crossover and mutation~~ and at the same time be reasonable for evaluation by means of fitness function. For the same task, different representations can be applied that will affect performance ~~and precise~~ of the algorithm within specific problem. There are several types of chromosome representation for GA:

- **Discrete** – usage of distinguishing numbers (usually binary representation – zeros and ones);
- **Real number** - representation in case if the genotype of the problem is within real number;
- **Ordinal** – used when the chromosome is permutation of numbers (Traveling salesman task) and must be consistent;
- **Tree-based** – used when individuals represented as a tree that can grow. Usually represent complex formulas.

Thus, the problem of best representation of the chromosome of the individual for the specific problem exists. For the purposes of this paper discrete representation considered (see example on Figure 2.22).

1	5	9	8	7	4	2	3	6	0

Figure 2.22 Discrete (integer) chromosome representation example.

3. METHODOLOGY

The aim of this paper is to create digital representation of Enterprise DNA which, strictly speaking, is the model for information system (artificial life simulation genetic algorithm), that, in turn, will simulate competitiveness of an Enterprises. For developing of new knowledge in Information Science the Design Science Research (DSR) discipline and framework was structured and developed (Hevner, 2004).

3.1. DESIGN SCIENCE RESEARCH

This framework defines steps for development of new Information Systems Science artifacts. In further work (Gregor & Hevner, 2013) summarized works in DSR discipline and structured researches by contribution type and research artifacts. From Table 3.1 we can figure out that this work is nascent design theory (Level 2) and as the result is the model (representation).

	Contribution Types	Example Artifacts
More abstract, complete, and mature knowledge	Level 3. Well-developed design theory about embedded phenomena	Design theories (mid-range and grand theories)
↕ ↕ ↕ ↕	Level 2. Nascent design theory—knowledge as operational principles/architecture	Constructs, methods, models, design principles, technological rules.
More specific, limited, and less mature knowledge	Level 1. Situated implementation of artifact	Instantiations (software products or implemented processes)

Table 3.1 Design science research contribution types (Gregor & Hevner, 2013)

Most of scientific paper aren't really "new." They are based on previous ideas, papers, artifacts etc. "Everything is made out of something else or builds on some previous idea. When is something really novel or a significant advance on prior work? A DSR project has the potential to make different types and levels of research contributions depending on its starting points in terms of problem maturity and solution maturity" (Gregor & Hevner, 2013). Thus, using DSR framework we can classify proposed idea as an **Invention** (see Figure 2.1),

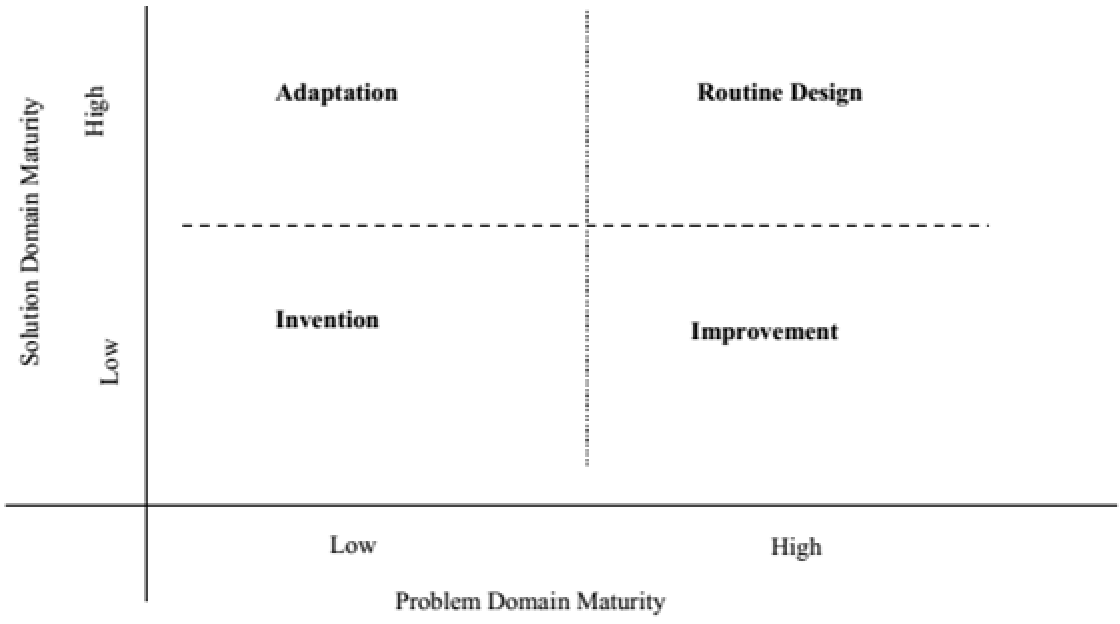


Figure 3.1 DSR Knowledge Contribution Framework (Vaishnavi, Kuechler, & Petter, 2017)

because solution domain (Enterprise's DNA digital representation) maturity and problem domain (Simulation of Enterprises competitiveness applying artificial life algorithms) maturity are low – author have not found other work on these subjects.

The Design Science Research methodology implies the following generic steps, showed in Table 3.2.

Section	Contents
1. Introduction	*Problem definition, problem significance/motivation, introduction to key concepts, research questions/objectives, scope of study, overview of methods and findings, theoretical and practical significance, structure of remainder of paper.* For DSR, the contents are similar, but the problem definition and research objectives should specify the ***goals*** that are required of the artifact to be developed.
2. Literature Review	*Prior work that is relevant to the study, including theories, empirical research studies and findings/reports from practice.* For DSR work, the prior literature surveyed should include any prior design theory/knowledge relating to the class of problems to be addressed, including artifacts that have already been developed to solve similar problems.
3. Method	*The research approach that was employed.* For DSR work, the specific DSR approach adopted should be explained with reference to existing authorities.
4. Artifact Description	A concise description of the artifact at the appropriate level of abstraction to make a new contribution to the knowledge base. This section (or sections) should occupy the major part of the paper. The format is likely to be variable but should include at least the description of the designed artifact and, perhaps, the design search process.
5. Evaluation	Evidence that the artifact is useful. The artifact is evaluated to demonstrate its worth with evidence addressing criteria such as validity, utility, quality, and efficacy.
6. Discussion	*Interpretation of the results: what the results mean and how they relate back to the objectives stated in the Introduction section. Can include: summary of what was learned, comparison with prior work, limitations, theoretical significance, practical significance,* and *areas requiring further work.* Research contributions are highlighted and the broad implications of the paper's results to research and practice are discussed.
7. Conclusions	*Concluding paragraphs that restate the important findings of the work.* Restates the main ideas in the contribution and why they are important.

Table 3.2 Publication schema for a DSR Study (Gregor & Hevner, 2013)

3.2. RESEARCH STRATEGY

The aim of this paper is to create digital representation of Enterprise according to Design Science Research methodology. For this goal following steps (see Table 3.3) were fulfilled and results obtained:

Step	Results
Introduction	• Statement of the problem and research objectives • Are there other paper on this subject?
Literature Review	Finding similar works and previous researches related to study explaining: • Typology of enterprises • Enterprise's architecture • Enterprise Architecture frameworks • Zachman framework • Artefacts for the representation of Enterprise aspects • Concepts of Artificial Life • History of Artificial Life • Genetic Algorithms
Method	Design Science Research method
Artifact Description	Digital representation of Enterprise's DNA using XML standard
Evaluation	• Interview with GA domain experts
Discussion	• Analysis of the created Enterprise DNA representation • Explaination how it can be used • Understand the pros and cons
Conclusions	Conclusions on the proposed Artefacts and further development of the research topic

Table 3.3 Research steps

4. RESULTS AND DISCUSSION

In this part of the research the Enterprise DNA model is synthesized and discussion regarding results conducted. The chromosome model is based on assumptions such as use of enterprise typology, Zachman Enterprise Architecture Framework as the foundation and further perspective of Artificial Life simulation. Digital representation of DNA phenotype artifacts as XML representation are proposed that resulted in model of genome. Interview with the topic experts performed for validation and discussion of results, then improvement of the model and possible future works proposed.

4.1. ASSUMPTIONS

For synthesis of the Enterprise's DNA for further simulation of company's competitiveness by means of Artificial Life following subjects were explored:

For finding Enterprise's DNA

- Typology of enterprises
- Enterprise's architecture
- Enterprise Architecture frameworks
- Zachman framework
- Artefacts for the representation of Enterprise aspects

For applying Artificial Life modelling

- Concepts of Artificial Life
- History of Artificial Life
- Genetic Algorithms

4.1.1. Typology of enterprises as Enterprise DNA Gene

Different typologies made by researchers and agencies can be used for adding to Enterprise's DNA Gene which will implement implicit aspects and properties of company, reflexing properties that cannot be figured out explicitly and which were found during other researches of Enterprise's taxonomies.

4.1.2. Zachman framework as a foundation for building Enterprise DNA Genes

Enterprise's architecture was considered as a main foundation for building the Enterprise DNA Genome. Particularly, Zachman Enterprise Architecture framework considered as a basement for consequent development of business architecture artifacts describing different aspects of the company, that, in turn, can be used as Company's Genes. The Owner's (Conceptual) view is assumed. Further, scientific papers which research XML representation of that artifacts were considered. Thus, we can build digital Enterprise's DNA representation, based in XML format. The Zachman framework have six aspects of the Enterprise, which describe most of company's behavioral and structural properties. These aspects of Zachman Framework with proposed artifacts and their XML representation are shown Table 4.1.

Aspect	Artifact	XML Representation
Data	ER Diagram	XML format for ER diagram representation
Function	BPMN	XPDL
Network	Network Graph	SCML
People	Organizational Structure	XML format
Time	Gantt chart	PMXML (Project Management XML) format
Motivation	Balanced Score Card	BSC XML

Table 4.1 Proposed artifacts and XML representation

4.2. MODEL OF THE DNA REPRESENTATION OF A GENERIC ENTERPRISES

4.2.1. Enterprise's DNA outline

Summarizing assumptions stated above the Enterprise DNA representation was synthesized. Proposed model consists of several sub genes based on Zachman framework artifacts, company reports data and service information. For the simplicity and applicability of the model, artifacts will be stored outside the individual~~,~~ to avoid redundancy, because genes can be duplicated among the simulation population. The outline of the proposed Enterprise's DNA representation is shown on the ~~Figure 4.1~~Figure 4.1 Outline of Enterprise's DNA representation~~Figure 4.1 Outline of Enterprise's DNA representation~~.

Enterprise ID	Enterprise typology	Data	Function	Network	People	Time	Motivation
Service info	Genes						

Figure 4.1 Outline of Enterprise's DNA representation

The Enterprise ID section contains identification information of the Company. Other sections are sub genes of the DNA. Enterprise typology sub genes show different taxonomies that can characterize company. Financial reports are classical financial reports that disclose Enterprise success in terms of profits, assets and other business parameters. The following artefacts: Data, Function, Network, People, Time, Motivation describe Enterprise aspects regarding Zachman Framework.

The main assumption for each Enterprise's DNA section is a fixed length and structure, which is needed for simplifying of crossover operations. Thus, Individual's Gene can be matched with the same Gene of its mate. The length itself depends on researcher's decisions about acquired reference models and artifacts used. For example, if for enumerating of Function artifacts we use APQS process framework (www.apqc.org, 2018) or SCOR reference model (Supply Chain Council, 2012) the length of sub gene "Function" must be set according utilized framework. Likewise, eTOM, ITIL and other process reference frameworks can be used.

4.2.2. Enterprise's DNA structure

The Enterprise's DNA consists of service information (Enterprise's ID) and Genes (sub DNA), which in turn contain all necessary for executing Artificial Life algorithm information. The proposed DNA structure is shown on the Figure 4.2.

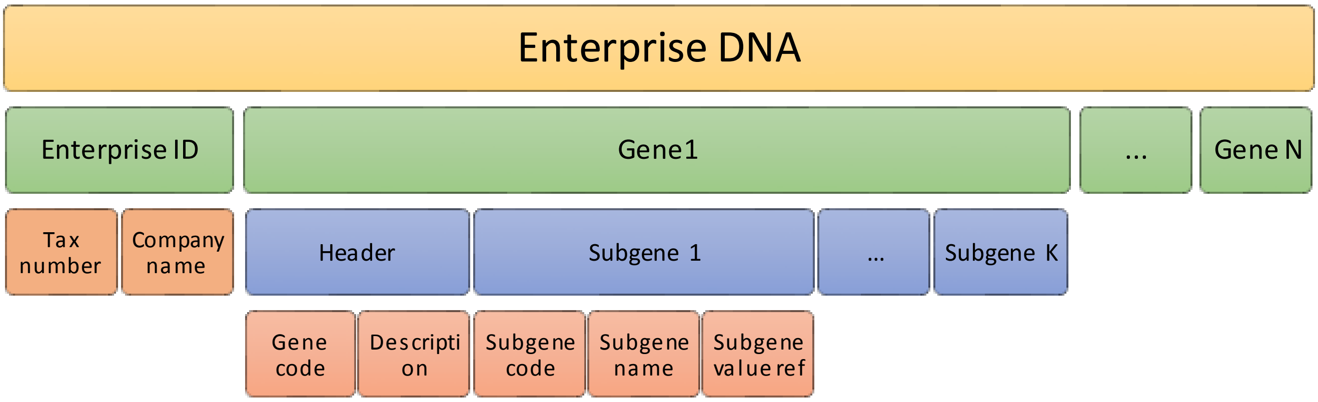


Figure 4.2 Proposed Enterprise's DNA structure

The XML representation of Enterprise DNA is textual and thus requires significant space in the computer memory. In the proposed model for the purpose of reducing the memory usage, the subgenes are represented by reference ID which point to gene\subgene specific dictionaries in which XML values of the artifacts are stored. Thus, for the genetic algorithm operations as initialization, crossover and mutation we used only one value instead of big XML text. For calculation of fitness function the XML value can be restored by "Subgene value ref" reference. The example of dictionary implementation is depicted Figure 4.3.

Subgene of typology Gene

SubGene code	SubGene name	Subgene value ref
K	Global Industry Classification Standard (GICS®)	**25503010**

Global Industry Classification Standard (GICS) dictonary

Sector		Industry Group		Industry		Sub-Industry	
25	Consumer	2550	Retailing	255030	Multiline Retail	**25503010**	Department Stores

Figure 4.3 Dictionary implementation.

Enterprise ID is service part of DNA (see Figure 4.4) which allotted for identification of specific Enterprise during simulation. It will help to produce insights after simulation by comparing initial and final chromosome of the company.

Enterprise ID	
Tax number	Company name
00000001	Company 1

Figure 4.4 Enterprise ID structure.

Enterprise typology represents a set of Enterprise's taxonomies used as a part of Enterprise's DNA. These typologies carry a lot of implicit properties of companies, that were resulted from researches performed by these taxonomies developers. There are many typologies and significant part of them can be used as the

Enterprise typology sub DNA**.** The structure of sub DNA, where K is the number of sub Genes in the Gene is shown on Figure 4.5.

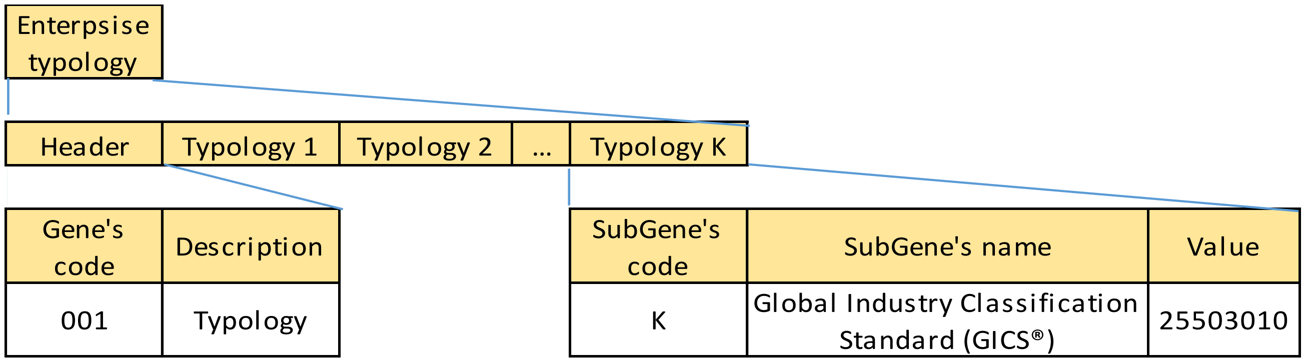


Figure 4.5 Enterprise typology sub DNA representation.

Data represents a set of Enterprise's databases used for business activities automation. Presence or absence of such systems affects company efficiency, for example, by authors experience, companies, that use CRM systems usually have higher sales. The database structure and the data included that represented in ER diagram affects the quality end completeness of computer systems used. A cCompany can have many different databases for sales, manufacturing, HR, Supply Chain, Budgeting, online Sales etc., so the Data Gene consist of several subgenes and header with identification information. The subgenes store subgene's header with their codes and names and reference to ER diagrams XML dictionary in which XML itself is stored. The structure of Data subDNA where K is number of possible automation systems used is show on the Figure 4.6.

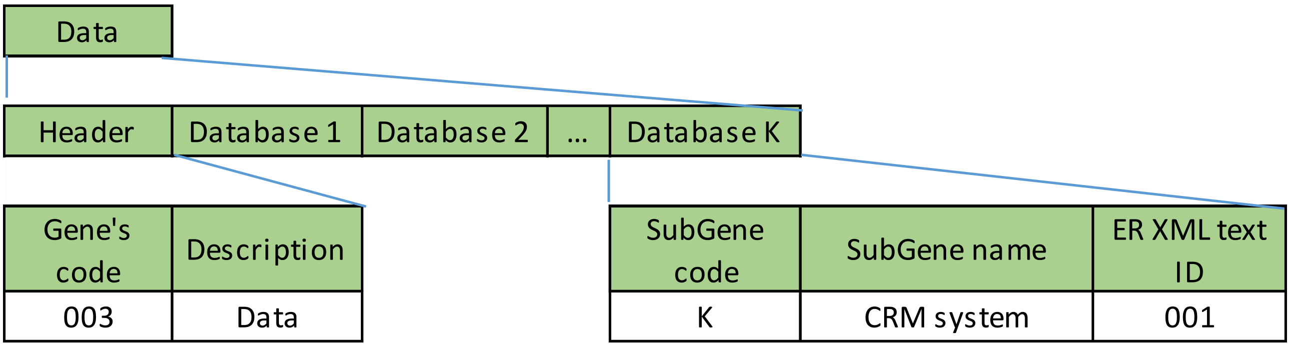


Figure 4.6 Data sub DNA representation.

The XML representation of Data gene can be retrieved from the dictionary with "ER XML test ID" field as it is shown in the Figure 4.7.

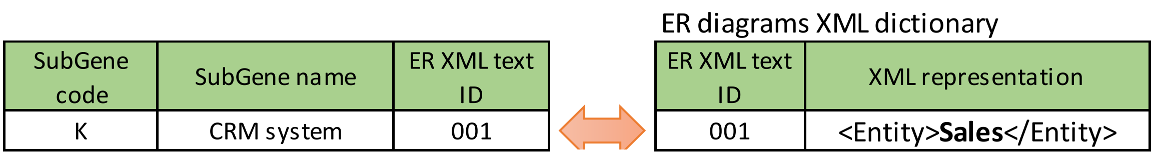


Figure 4.7 Data gene relation to ER diagrams XML dictionary.

Function represents a set of Enterprise's business processes used as a part of Enterprise's DNA. These processes define the functioning of a company and are defined depending of the it's maturity that affects productivity. The number of subgenes depends on process reference framework used, as was described above. The implementation of specific processes stored as XPDL representation in the special BPMN XPDL dictionary. The structure of sub DNA, where K is the number of sub Genes in the Gene is shown on Figure 4.8.

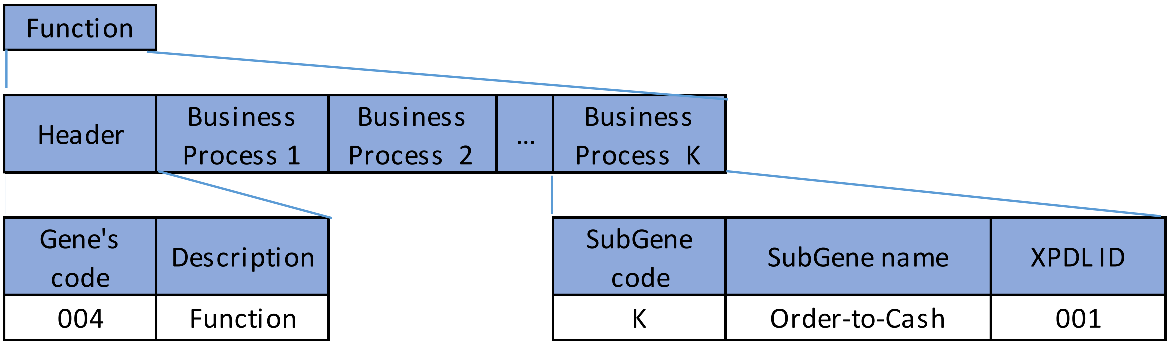


Figure 4.8 Function sub DNA representation.

The XML representation of Function gene can be retrieved from the dictionary with “XPDL ID” field as it is shown in the Figure 4.9.

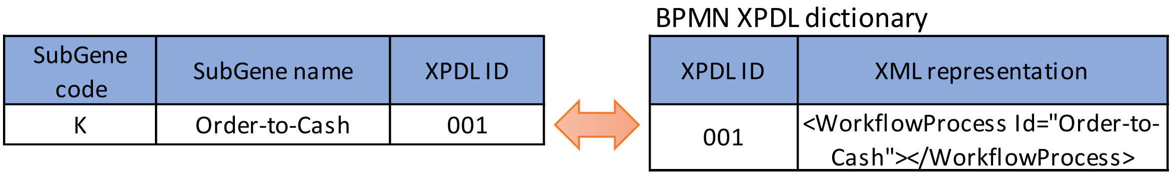


Figure 4.9 Function gene relation to BPMN XPDL dictionary.

Network represents a geographical logistic system structure used as a part of Enterprise’s DNA. This system is represented ~~with~~ by Supply chain topology graph and defines places where the company operates and relationships among them that defines logistic efficiency and geographic coverage. The Supply chain topology network graph represented with SCML language with is stored in special dictionary and its ID is used for DNA representation. The structure of Gene is depicted on Figure 4.10.

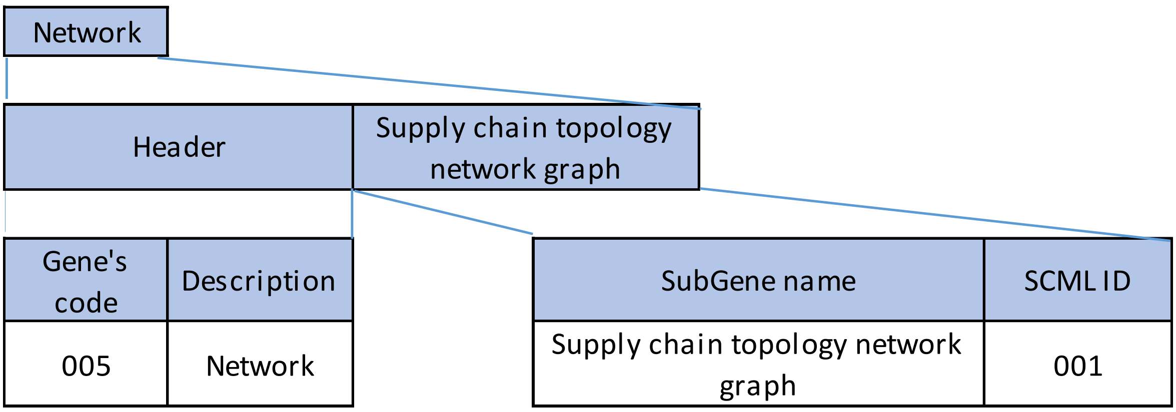


Figure 4.10 Network sub DNA representation.

The XML representation of Network gene can be retrieved from the dictionary with “SCML ID” field as it is shown in the Figure 4.11.

SubGene name	SCML ID
Supply chain topology network	001

Network Graph diagrams XML dictonary

SCML ID	XML representation
001	sourceNode IDREF #REQUIRED

Figure 4.11 Network gene relation to Network Graph diagrams XML dictionary.

People represents a human resources part of Enterprise's DNA. The organizational structure affects productivity of company thus becoming as an interesting artifact for the research. The type of organizational chart can also affect company' success. The organizational chart is represented as pure XML and stored in the special XML organizational chart diagrams XML dictionary. The structure of sub DNA is shown on Figure 4.12.

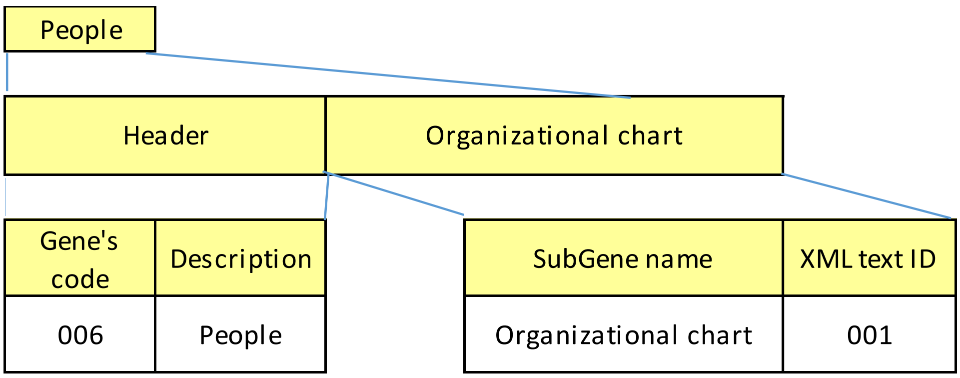


Figure 4.12 People sub DNA representation.

The XML representation of People gene can be retrieved from the dictionary with "XML text ID" field as it is shown in the Figure 4.13.

SubGene name	XML text ID
Organizational chart	001

XML organisational chart diagrams XML dictionary

XML text ID	XML representation
001	<?xml version="1.0" encoding="UTF-8"?> <Department id="Administration"> <id>2</id> <team>Team One</team>

Figure 4.13 People gene relation to XML Organizational Chart diagrams XML dictionary.

Time is a set of Enterprise's project milestone, strategy deadlines and business processes execution time artifacts represented as Gantt diagrams. These artifacts reveal the functioning of company in terms of time and time efficiency, that affects its success. The number of subgenes is defined by the researcher depending of on research project complexity, coverage and goals. The implementation of specific Gantt Diagrams stored as PMXML representation in the special PMXML Gantt Charts dictionary. The structure of sub DNA, where K is the number of sub Genes in the Gene is shown on Figure 4.14.

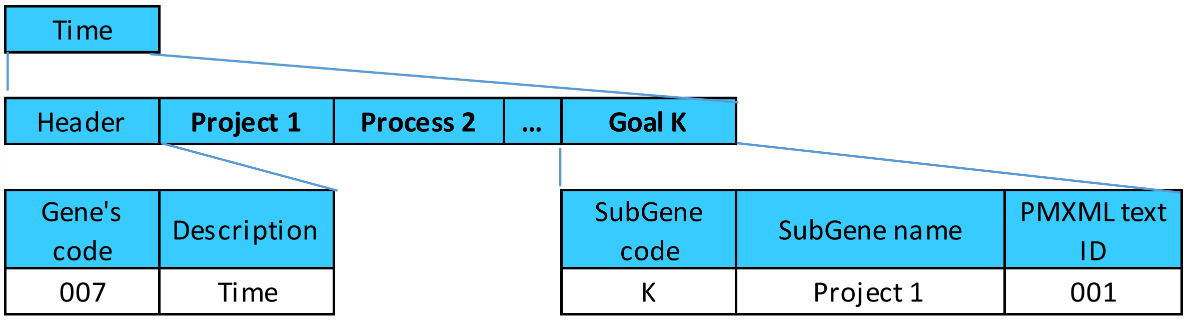


Figure 4.14 Time sub DNA representation.

The XML representation of Time gene can be retrieved from the dictionary with "PMXML ID" field as it is shown in the Figure 4.15.

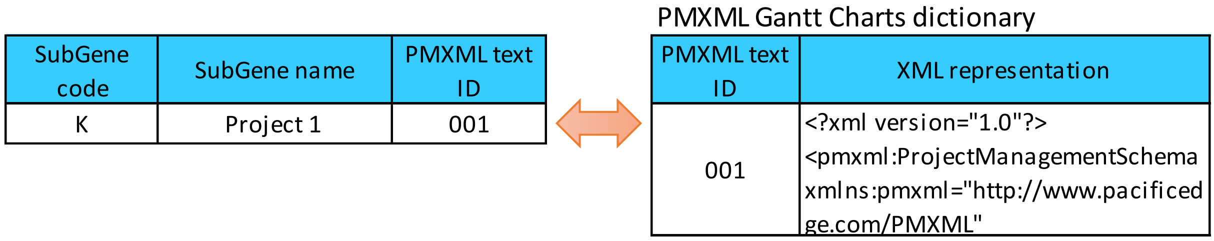


Figure 4.15 Time gene relation to PMXML Gantt Charts dictionary.

Motivation represents a set of Enterprise's goals and used as a part of Enterprise's DNA. Different goals lead to different actions and thus more or less success. Achieving stated plans and goals also important for company's success. The motivation of Enterprise defined as Balanced Scorecards indicators which are stored as BSCML representation in the special BSC XML dictionary. The structure of sub DNA is shown on Figure 4.16.

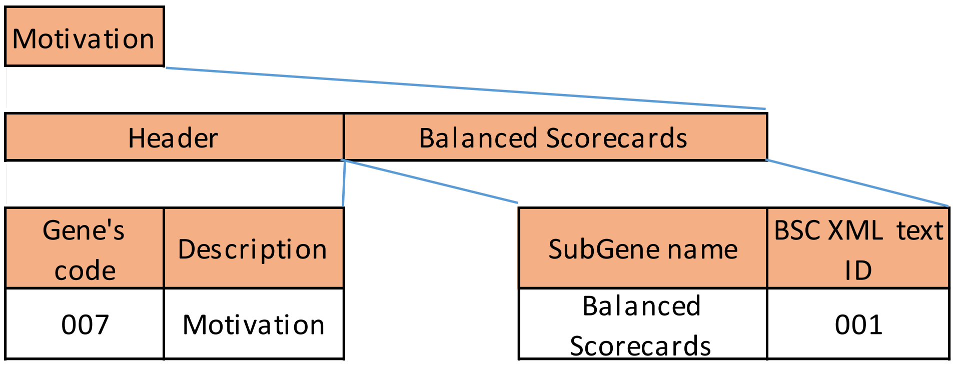


Figure 4.16 Motivation sub DNA representation.

The XML representation of Motivation gene can be retrieved from the dictionary with "BSC XML text ID" field as it is shown in the Figure 4.17

SubGene name	BSC XML text ID
Balanced Scorecards	001

BSC XML dictionary

BSC XML text ID	XML representation
001	<BSCDoc language="EN"> <BSC id="SalesBSC" name="Sales Balanced Scorecard"> <PERSPECTIVE id="SLC" name="Sales Customer Perspective"

Figure 4.17 Motivation gene relation to BSC XML dictionary.

4.2.3. DNA structure

The example of proposed DNA structure representation is shown on the Fig. 42. The number of sub genes in each gene is defined by researcher regarding subject domain and applied reference models and must be fixed. The Enterprise ID information is necessary for identification of a company after running the Artificial Life algorithm. Thus, we can unveil genes and subgenes that were changed, ~~and~~ which of them ~~are good that is make~~ lead to company success~~ful~~.

Enterprise ID		Gene Typology						Gene Data					
		Sub Gene 1					...	Sub Gene 1					...
Tax number	Company name	Gene's code	Description	SubGene code	SubGene name	Value ID		Gene's code	Description	SubGene code	SubGene name	ER XML text ID	
00000001	Company 1	001	Typology	K	Global Industry	25503010		003	Data	K	CRM system	001	

Gene Function						Gene Network				
Sub Gene 1					...	Sub Gene 1				...
Gene's code	Description	SubGene code	SubGene name	XPDL ID		Gene's code	Description	SubGene name	SCML ID	
004	Function	K	Order-to-Cash	001		005	Network	Supply chain topology network graph	001	

Gene People				Gene Time					
Sub Gene 1				Sub Gene 1					...
Gene's code	Description	SubGene name	XML text ID	Gene's code	Description	SubGene code	SubGene name	PMXML text ID	
006	People	Organizational chart	001	007	Time	K	Project 1	001	

Gene Motivation			
Sub Gene 1			
Gene's code	Description	SubGene name	BSC XML text ID
007	Motivation	Balanced Scorecard	001

Figure 4.18 Example of final DNA representation.

4.3. VALIDATION

Proposed model is theoretical because implementation of Artificial Life algorithm simulation on real data is out of the scope of this research, so the problem of validation arises. As a solution of this problem the interviews with experts in evolutionary algorithms from academia were organized. For the review of work: a usage of proposed Enterprise Architecture Framework, discovered artifacts, their XML representation and the final Enterprise DNA model were invited following experts:

- **Mauro Castelli,** Assistant Professor at NOVA Information Management School, Doctor in Computer Science - University of Milano Bicocca - Italy. He was the chair of the main European Conference on Genetic Programming (EuroGP) and he is in the editorial board of Genetic Programming and Evolvable Machines journal. He participated in the program committee of several international conferences and he is a reviewer for top-tier journal in the area of Artificial Intelligence. His current area of scientific activity is in the field of computer science. In particular, he is working in the following areas: Evolutionary Computation, Genetic Programming, Genetic Algorithms, Swarm Intelligence, Machine Learning, Soft Computing, Deep Learning and Neuroevolution.
- **Flávio L. Pinheiro**, Invited Assistant Professor of Data Science at NOVA Information Management School, Ph.D. in Physics – University of Minho. His research interests focus in understanding how the network structure of socio-economic systems impacts the strategic decisions of agents at different scales. At the macro-level, how does the structure of Economic and Innovation systems impact the success of agents is constrained when they have to take actions over the elements that make up such systems. At the individual scale, how is the evolution of ideas, diseases, opinions, behaviors and knowledge affected by the social network they are part of. His work has resulted in the publication of several manuscripts at top-tier peer review journals such as Physical Review Letters, New Journal of Physics, Journal of Theoretical Biology and PLoS Computational Biology and has been presented at international conferences in Complex Systems, Artificial Intelligence, Mathematical Biology, Network Science and Computational Social Sciences.

The interviews were performed for receiving of the feedback on usefulness and quality of proposed model, research strategy and the assumptions for the improvement of the model. The three following questions were asked:

1. Usefulness of having an Enterprise DNA model
2. Agreement with the adopted strategy to build the Enterprise DNA
3. Recommendations for improvement

4.4. DISCUSSION

The interview answers provided by experts ~~allow us to consider proposed~~confirm that the Enterprise DNA model i~~a~~s valuable and useful for the goal~~s~~ of company's competition simulation by means of Artificial Life algorithms. ~~For~~ At the moment this model is a good starting point - the way used for modelling of these enterprises by means of DNA is very interesting. Strong point of this abstraction is that enterprises and business models are evolving in time and let us ~~to~~ see development of the company in dynamics. The adopted strategy was discussed and approved and recommendations for improvement were given.

On the first question, ~~asked~~ to determine usefulness of having Enterprise's DNA model - all experts gave positive feedback. It is an interesting way of abstracts to see or study and try to understand how the companies or business models evolve through a biological view. They stated~~, that~~ the model is a nice option for achieving better understanding of what strength and weaknesses the company ~~have~~has. This would be particularly clear in the future after performing ~~of~~ a simulation, putting together different companies and ~~trying to understand, to see what happen,~~determining who will survive. This model is beneficial ~~in the sense that~~ because it is different with respect to existing models that are commonly used for modelling of enterprise. ~~because with~~ T~~t~~his model is able to somehow capture the interaction between companies~~,~~ the other competitors, ~~that~~ which is very important and very valuable.

On the second question about agreement with the adopted strategy to build the Enterprise DNA also were given positive feedbacks with some remarks. Experts agreed with this kind of modelling ~~and~~ but ~~made~~ remarked~~s~~ that ~~regarding~~ mimicing ~~of~~ real interaction among the companies ~~which~~ may require ~~to add to this~~addition of modelling strategy usage of external factors, not only the company itself. Despite the "nice design" of the model, many classifications used in Enterprise Typology gene can have very high correlation and it makes sense to use smaller set of subgenes – to reduce dimensionality. Usage, for example, APQC processes framework for Function gene with about 1000 types of processes as subgenes is tempting but complicated and would be hard to implement, so there is recommendation to reduce number of elements in gene. Another interesting suggestion is to make use as model for competitiveness simulation of only one aspect of the Zachman Framework. For example, compare companies' performance only by means of organization structure (People) or other.

The last, third question was asked about recommendations for improvement. Interviewed experts provided few interesting suggestions regarding model improvement and running the artificial life simulation, which are:

- To include in the model other factors that are not specific to companies but are external to them: like the political situation of the particular region in which company operates the economic situation, the worldwide economic situation – some kind of PEST analysis. PEST means political, economic, socio-cultural and technological. These additional genes will lead to improvement of the general performance of the model;
- To consider not only simulation of competition of different companies on the same market, but also behavior of similar companies in different markets, with various economic, geographic, political, etc. conditions. For example, we can model success of big company in specific country. Let's say McDonald's revenue per each restaurant in various countries, when they have the same processes, structure etc. of the business unit, but existing in different environment;
- To reduce the complexity of the model when implement it because it has too many variables, very big subgene set. Reduce the dimensions and number of variables, especially that can be correlated – similar typologies etc.;
- To consider necessity of Typology Gene because sometimes the classification is based on someone's subjectivity and not really represents the field.

Some interesting suggestions were discussed regarding future work and Artificial Life algorithm simulation implementation:

- To add real data part and understand whether they reflex real behavior or are redundant;
- Reverse engineering for fitness function design. The dependencies of factors (subgenes) and their evaluation with respect to fitness function could be hard and not straightforward. So, taking the temporary data of the company's fitness (For example profit) and DNA values can be used for design of fitness function from its value to specific formula or model.

- As the main application of this work is simulation of Alife and getting of insights there is a second approach: To look at temporal data of evolution of the companies and look how they evolve, and the market operates

Thus, considering the validation interview and answers of experts on questions asked we can conclude that proposed model is valuable and useful for researchers, consultants and managers by getting various insights on the company composition, evolution and strategy. The adopted strategy of building Enterprise's DNA by means of Zachman Enterprise Architecture Framework and its simulation with Artificial Life algorithm that mimic biological evolution makes sense and is different from existing approaches. Moreover, after adding external environment factors model can be used for analysis of their influence on business performance. Some other improvements of the model can be made.

5. CONCLUSIONS

The main goal of this work was to find the digital Enterprise's DNA representation, which will describe internal aspects of the company for better understanding of ~~which of them are~~what is important, how they were implemented at successful business units and how they affect the performance. DNA representation can be simulated in the future works by means of Artificial Life evolutionary algorithms, that can model interaction of companies on the specific market with respect to fitness function that will define the success or failure of competitors. Also, behavior of similar companies in different markets and environments can be simulated to find more attractive ones.

5.1. SYNTHESIS OF THE DEVELOPED WORK

The main purpose of this work was to find the Enterprise's DNA theoretical model based on Enterprise Architecture framework and its digital representation for further simulation by means of Artificial Life evolutionary algorithms. For that objective following steps were performed:

- Research of Typology of enterprises
- Research of Enterprise Architecture
- Research of Enterprise Architecture Frameworks
- Research of Zachman framework with respect to the goals of this paper
- Research of Artefacts for the representation of Enterprise aspects
- Research of concepts of Artificial Life and Genetic Algorithms
- Synthesis of the Enterprise's DNA digital model
- Validation of the model

The first step of the research was study of typology of enterprises, which was considered as possible part of Enterprise's DNA. Several taxonomies developed by different institutions were found which can provide some insight on the specific company. As a result, the gene (or we can call it sub DNA) "Enterprise Typology" was included in the model.

The second stage was the research of Enterprise Architecture science with respect to company modelling. In modern understanding of EA as architecture of business rather than IT systems it works for the goal of representation of properties of the Enterprise. It gives the outline, the skeleton, that can be detailed further for the specific implementation.

T~~On t~~he next step was the research of existing Enterprise Architecture Frameworks was performed~~, and the Zachman EA Framework was chosen~~. The reason was wide use, flexibility and simplicity. Evolved from Information Systems architecture to Business Architecture this framework became a workhorse of enterprise architects, business analysts, consultants and researchers.

The fourth step of the research was application of Zachman framework to the goals of this paper. For that purpose, Zachman Framework aspects and views were studied. As a result, as an Enterprise Model the Conceptual (business owner's) view was proposed. This level is not too common and not too detailed for the DNA representation and consider specific artifacts that can be useful for model development. Also, all six aspects of the Framework were considered and were chosen as valuable for the Enterprise's DNA model.

The fifth important step was the research of Artefacts for the representation of Enterprise aspects. There were papers with the proposal of some artifacts, but some of them were obsolete, thus some new proposals were made. The resulting artifacts list for the representation of the enterprise by means of Zachman Framework is:

- ER Diagram
- BPMN
- Network Graph
- Organizational Structure
- Gantt chart
- Balanced Score Card

These artifacts form the phenotype of the company and for the genotype – digital representation and fitness function evaluation the XML representation of each artifact was proposed based on existing previous research works. The XML standards proposed for each type of artifact (gene) are respectively:

- XML format for ER diagram representation
- XPDL
- SCML
- XML format
- PMXML (Project Management XML) format
- BSC XML

The sixth task of this research was studying the concept of Artificial Life and Genetic Algorithms with respect to goals of this paper. The Alife idea and evolutionary computation technic were explored and stated that the representation of GA is one of most important parts of its implementation which affects quality and performance of the simulation.

The core part of this work is the synthesis of the Enterprise's DNA digital model. Based on literature review of companies' taxonomies, EA concepts, Zachman Framework, proposed artifacts and their XML representation the enterprise genome structure was synthesized. This DNA consists of genes and subgenes, which with some identification information code the genome of the company. The gene consist of service information and IDs – references to the special dictionaries that store the XML representation of each subgene. Genetic operations such as crossover and mutation can be performed on these IDs (references) and evaluation (fitness function) calculation will retrieve XML values from external storages (dictionaries).

The last step of research is the validation of the model. For this purpose, two experts were invited. Regarding the asked questions the positive feedback was received along with some remarks and suggestions for improvement. The main outcome is that the proposed model is valuable and useful and can provide insights on the company performance and success. There are perspectives of applying of Artificial Life simulation for the proposed DNA to different companies on specific market (finding best genes) and the similar or to equal companies on different markets (finding best market). After all, some ideas for improvement of the model and its simulation were given by the experts.

Finally, two main contributions to the research topic were made: building an Enterprise's DNA for company behavior Alife simulation and applying of the Enterprise Architecture Framework for it. The results of the research will allow to model the enterprise by means of mimic of biological organisms and their evolution and interaction. During the complexity of the task, immaturity of the topic and lack of scientific works on it there are some limitations of developed work. Also, as the idea is undeveloped there is wide perspective for further works applying the proposed approach.

5.2. LIMITATIONS

Since the Enterprise DNA representation and simulation topic is undeveloped there were few limitations that discussed below:

- The proposed model is mostly based on Zachman EA framework, which does not consider intangible company's assets as a brand, a culture, an atmosphere that can influence competitiveness.
- This model does not consider tangible assets of the company such as buildings, equipment, row materials, money, etc. that can also define some success of the Enterprise. In general, they can be bought or rented if the company needs them, but for specific cases possession of them can bring some advantage for the company.
- Proposed model is theoretical one and for the future purposes the question of real data arises: where real data can be obtained? Is it possible to find all necessary data? These questions are out of scope of this work.
- The full implementation of simulation with defining the fitness function and evolutionary algorithm requires significant time and are out of scope of this work as this particular work was concentrated on DNA representation.

5.3. FUTURE WORKS

5.3.1. Development of DNA representation

The proposed model can be improved in the future by adding to business architecture of tangible and intangible assets genes that were not included in the baseline representation.

Proposed in this paper Enterprise's DNA representation can be extended by intangible factors that some businesses are trying to measure last few decades. These factors are goodwill, brand, corporate culture, company's atmosphere and reputation. All that factors can bring some value to the model.

The next suggestion for the improvement of the DNA representation model is adding to it company's assets, such as capital, buildings, equipment, communications, money and other assets. The value of such DNA extension is the objective of future research.

Possibly it makes sense to add to the model accounting and financial information. The advantage of this idea is that some financial and accounting data can implicitly describe and reflex some properties of the organization that were not represented in proposed model. On the other hand, most of accounts of the balance sheet represent assets, that can be modelled differently (see previous suggestion of future works). Moreover, the finance reports have different continuous numerical data that would be hard to convert into discrete representation. This idea requires further clarification and investigation.

One of possible future works is about including into genes of environmental factors that are external to the company. This idea was proposed by both experts during the interview, but from different point of views. Such information, that is usually result of PEST analysis that includes political, economic, socio-cultural and technological factors or geographical view can affect performance of the company on specific market. It means that companies' DNA in specific environment will result in different performance, so the internal factors will be multiplicated by outer conditions and thus the model can be more precise.

Another application of PEST and other outer factors in the model is search of the best marked for internally equal company. For example, geography affects number of visitors and average order for the restaurant or supermarket, though they have almost identical internal DNA. This approach, proposed by experts during interview is very interesting, because not only enterprises but also markets can be simulated and compared, and many managerial business decisions can be made with this kind of analysis.

5.3.2. Simulation be means of Artificial life evolutionary algorithms

Study of Artificial Life and application of Evolutionary (Genetic) Algorithms for Alife revealed opportunity to implement programmatic solution for modelling of competing population of Enterprises within specific environment. Genetic Algorithm will allow to model bunch of real-like companies, functioning on market and to show their success by means of fitness function, which can be based on Enterprise's DNA, modelled market conditions and other actions and success of other individuals.

The Artificial Life modelling consist of three main steps:

- Finding the representation of individual (the goal of this paper)
- Defining of fitness function
- Implementation of GA evolution using genetics operators (Implementation, Mutation, Crossover)

The purpose of this paper is defining digital representation for Alife simulation algorithm, that is the first step of Genetic Algorithms implementation. The synthesized model - Enterprise's DNA representation can be used in future works for further development of fitness function and evolutionary biologically inspired modelling environment that will mimic nature.

The next step that must be done for Artificial Life simulation of organizations competition is synthesizing of the fitness function. This is very complex task, because relationships between subgenes artifacts and the function value, that defines company's success is not straightforward and probably nonlinear. Thus, existing or missing subgene along with other subgenes represented can have multiplicative effect on overall fitness function.

The reverse engineering methodology for fitness function design, that was proposed during interview with Flávio Pinheiro looks like a very perspective approach, that can help to overcome complexity and uncertainty of dependences between genes and fitness function value. Consulting companies have bunch of business data from real companies along with number that characterize company's success and that data can be used for this kind of fitness function design.

The goal of this work is to create a foundation for answering real business questions by performing Artificial Life simulation of companies' behavior, interaction and following success of failure. This kind of simulation can be performed with respect to the company DNA or successful market or both, as was stated above. Alife simulation, that can be implemented by means of genetic Algorithms is programmatic implementation with some restrictions regarding to the specific market. This simulation can be implemented as a library for computing language, specific program or a out of box tool that can be used by non-programmers auditory for everyday research or business need.

BIBLIOGRAPHY

Bedau, M. (2003a). Artificial life: organization, adaptation and complexity from the bottom up. *Trends in Cognitive Sciences*, Vol.7 No.11.

Bedau, M. (2003b). Artificial Life. In L. Floridi (ed.), *The Blackwell Guide to the Philosophy of Computing and Information* (pp. pp. 505-512). Blackwell.

Bogomolova, M. (2016). *Enterprise Architecture: Tutorial.* Samara: PSUTI.

Bulajic, A., & Filipovic, N. (2012). Implementation of the Tree Structure in the XML and Relational Database. *Proceedings of Informing Science & IT Education Conference (InSITE).*

Burtsev, M., & Turchin, P. (2006). *Evolution of cooperative strategies from first principles.* Nature, 440 (7087): 1041–1044.

Chatfield, D., Harrison, T., & Hayya, J. (2003). *The supply chain modeling language (SCML).* Retrieved from https://www.researchgate.net/publication/228499797_The_supply_chain_modeling_language_SCML

Chatfield, D., Harrison, T., & Hayya, J. (2004). XML-based supply chain simulation modeling. *Proceedings of the 2004 Winter Simulation Conference.*

Chen, P. (1976). The Entity–Relationship Model – Toward A Unified View of Data. *ACM Transactions on Database Systems, New York: Association for Computing Machinery. 1 (1)*, 9–36.

Chen, P. (2002). Entity-relationship modeling: historical events, future trends, and lessons learned. In B. Manfred, & D. Ernst, *Software pioneers: contributions to software engineering* (pp. 296-310). Heidelberg: Springer-Verlag Berlin.

Companies Act. (2006). The Stationery OYce Limited. Retrieved from http://www.legislation.gov.uk/ukpga/2006/46/section/1

Dale, T., Churcher, N., & Irwin, W. (2005). A Framework for Linking Projects and Project Management Methods. In F. K.-S. Editors: Bomarius, *Product Focused Software Process Improvement.* th International Conference, PROFES 2005, Oulu, Finland, June 13-18, 2005, Proceedings. Retrieved from http://www.cosc.canterbury.ac.nz/research/RG/svg/index.html

Dumas, M., Reijers, H., La Rosa, M., & Mendling, J. (2015). *Fundamentals Of Business Process Management.* Springer-Verlag Berlin Heidelberg.

Espadas, J., Romero, D., Concha, D., & Molina, A. (2008). OTM 2008 Workshops, EI2N Conference Proceedings. *Using the Zachman Framework to Achieve Enterprise Integration Based-on Business Process Driven Modelling* (pp. 283–293). Monterrey, Mexico: Springer-Verlag Berlin Heidelberg.

European Communities. (2008). *NACE Rev. 2 – Statistical classification of economic activites in the European Community.* Luxembourg: European Commission.

Farmer, J. D., Lapedes, A. S., Packard, N., & Wendroff, B. (1986). *Evolution, Games, and Learning: Models for Adaptation in Machines and Nature.* Amsterdam: North Holland Physics Publishing. 22: vii–xii.

Franceschet M., Gubiani D., Montanari A., Piazza C. (2009). From Entity Relationship to XML Schema:. In H. E. In: Bellahsène Z., *Database and XML Technologies. XSym 2009. Lecture Notes in Computer Science, vol 5679.* Berlin, Heidelberg: Springer. Retrieved from Department of Mathematics and Computer Science, University of Udine: https://users.dimi.uniud.it/~massimo.franceschet/publications/xsym09.pdf

FTSE Russell. (2008). *Industry Classification Benchmark (Equity).* London: London Stock Exchange Group plc.

Gardner, M. (1970). *MATHEMATICAL GAMES The fantastic combinations of John Conway's new solitaire game "life".* Scientific American 223, 20-123.

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA.

Gregor, S., & Hevner, A. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly Vol. 37 No. 2, June 2013*, (pp. pp. 337-355).

Hay, D. (2000). *A Different Kind of Life Cycle: The Zachman Framework*. Retrieved from Essential Strategies, Inc: https://www.semanticscholar.org/paper/A-Different-Kind-of-Life-Cycle-%3A-The-Zachman-Hay/c74558e1988528489145b461c7f77e1abe9f5b7f

Hevner, A. (2004). Design Science in Information Systems Research. *MIS Quarterly Vol. 28 No. 1, pp. 75-105/March 2004.*

Holland, J. (1975). *Adaptation in Natural and Artificial Systems.* MIT Press.

Kalenkova, A., de Leoni, M., & van der Aalst, W. (2014). Discovering, Analyzing and Enhancing BPMN Models Using ProM. In B. W. L. Limonad, *BPM Demo Sessions 2014* (pp. 36-41). Netherlands - Eindhoven, Netherlands: PublisherCEUR-WS.org.

Kaplan, R., & Norton, D. (1992). The Balanced Scorecard - Measures That Drive Performance. *Harvard Business Review, Jan-Feb*, pp 71-79.

Langton, C. (1987). *Langton CG*. Retrieved from Announcement of artificial life: https://groups.google.com/group/news.announce.conferences/browse_thread/thread/776a2af47daf7a73/7ab70ed772c50437

Langton, C. (1989). *"Artificial Life", in Artificial Life, Langton (ed).* Addison-Wesley:Reading, MA.

Langton, C. (1996). Artificial Life. In M. A. Boden, *The Philosophy of Artificial Life* (pp. 39-94). Oxford University Press.

Lankhorst, M. (2017). *Enterprise Architecture at Work. Modelling, Communication and Analysis.* Springer-Verlag Berlin Heidelberg.

Legaard, J., & Bindslev, M. (2006). *Organizational theory.* Ventus Publishing ApS.

Mayr, E. (1969). *Principles of systematic zoology.* New-York: McGrow-Hil.

McGregor, C. (2003). Balanced Scorecard Driven Business Process Definition Using XML. *Proceedings of the 36th Annual Hawaii International Conference on System Sciences.*

McKelvey, B. (1975). Guidlines for the Empirical Classification of Organizations. *Administrative Science Quarterly*, Vol 20, No 4, pp 509-525.

McKelvey, B. (1978). *Organizational Systematics: Taxonomic Lessons from Biology, Management Science, Vol 24, No. 13, pp 1428-1440.*

Mintzberg, H. (1979). *The structuring of organizations.* Prentice-Hall.

Noran, O. (2003). An analysis of the Zachman framework for enterprise architecture from the GERAM perspective. *Annual Reviews in Control 27*, 163–183.

Object Management Group (OMG). (2014). *About the buisness process model and notation specification version 2.0.2*. Retrieved from https://www.omg.org/spec/BPMN/2.0.2/

Reynolds, C. (1987). *Flocks, herds, and schools: A distributed behavioral model.* Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'87).

Rich, F. (1992). The organizational taxonomy: definition and design. *The Academy of Management Review*, Vol 17., No. 4 758-781.

Standard & Poor. (2006). *Global Industry Classification Standart (GICS).* New-York: Standard & Poor.

Sung, J., & Kang, W. (2007). *Mapping Rules for ER to XML Using XML Schema.* Retrieved from https://aisel.aisnet.org/sais2007/38

Supply Chain Council. (2012). *Supply Chain Operations Reference Model Revision 11.0.* SCM Council.

Tarasov, V. (2002). *From multi-agent systems to intellectual organization: philosophym scicology, informatics.* Moscow: Editorial URSS.

United Nations. (2008). *The International Standard Industrial Classification of All Economic Activities (ISIC).* New York.

Vaishnavi, V., Kuechler, W., & Petter, S. (2017). *Design Science Research in Information Systems.* Retrieved from http://www.desrist.org/design-research-in-information-systems/

van der Aalst, W. (2004). Business Process Management Demystified: A Tutorial on Models, Systems and Standards for Workflow Management. In J. Desel, W. Reisig, & G. Rozenberg, *Lectures on Concurrency and Petri Nets: Advances in Petri Nets* (pp. 1-65). Berlin: Springer.

van der Aalst, W. (2011). *Process Mining: Discovery, Conformance and Enhancement of Business Processes.* Verlag Berlin Heidelberg: Springer.

Varela, L., Aparício, J., & Silva, C. (2005). *Production Scheduling Concepts Modeling through XML.* Retrieved from https://www.researchgate.net/publication/228579996_Production_Scheduling_Concepts_Modeling_through_XML

W3C. (1999a). *The Cambridge Communiqué.* Retrieved from W3C NOTE 7: https://www.w3.org/TR/schema-arch/

W3C. (1999b). *Resource Description Framework (RDF) Model and Syntax Specification.* Retrieved from www.w3.org: https://www.w3.org/TR/REC-rdf-syntax/

WfMC. (2012). *Workflow Management Coalition. Process Definition Interface - XML Process Definition Language.* WfMC.

www.apqc.org. (2018). Cross industry process classification framework Version 7.2.1. Houston, Texas 77024-7797, USA.

Zachman, J. (1987). *A framework for information systems architecture.* IBM systems journal, vol 26. no 3.

Zachman, J. (2011). *The Zachman Framework Evolution*. Retrieved from Zachman International, Inc.: https://www.zachman.com/resource/ea-articles/54-the-zachman-framework-evolution-by-john-p-zachman

Zachman, J., & Sowa, J. (1992). *Extending and formalizing the framework for information systems architecture.* IBM systems journal, vol 31, no 3.