INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

An Information Architecture for the Public Administration

Ricardo Mendes Castelão

Dissertação para obtenção do Grau de Mestre em

Engenharia Informática e Computadores

Júri

Presidente:	Professor Doutor Alberto Silva
Orientador:	Professor Doutor André Vasconcelos
Co-Orientador:	Professor Doutor José Tribolet
Acompanhante:	Dra. Maria João Marques (AMA)
Vogal:	Professor Doutor Pedro Sousa

Outubro de 2010

Para os meus pais

Agradecimentos

Ao chegar a hora de entregar esta dissertação, e quando as palavras nunca parecem suficientes, queria agradecer ao meu pai e à minha mãe pelo apoio incondicional ao longo de todos os anos da minha formação académica. De facto, não fora pelo constante sentido de ambição incutido pelo meu pai e pelo carinho da minha mãe, certamente não teria chegado onde estou hoje e por isso merecem a minha eterna gratidão.

Agradeço ao Professor André Vasconcelos, meu orientador, por todo o apoio que me deu durante a realização desta dissertação, demonstrando sempre disponibilidade, amabilidade e tempo para atender as minhas dúvidas, ao mesmo tempo que sempre me motivou a desenvolver um trabalho de excelência.

Agradeço ao Professor José Tribolet, meu co-orientador, pelas conversas informais que tivemos, que certamente nunca poderiam ter acontecido com outra pessoa. O seu conhecimento, visão e capacidade de inovação foram, e são, uma fonte de inspiração e um exemplo a seguir.

Agradeço ao Professor Pedro Sousa pelas reuniões que tivemos, de onde resultaram ideias chave para esta dissertação ter seguido este rumo. A sua perspectiva e conhecimento sobre estes temas foram uma preciosa e determinante ajuda para a realização deste trabalho.

Estendo o meu agradecimento a todos os docentes do Instituto Superior Técnico, sítio onde recebi cinco maravilhosos anos de formação académica e onde tive o privilégio de estudar sobre alçada dos melhores.

Agradeço à Dra. Maria João Marques, acompanhante por parte da AMA, e à Dra. Zaida Chora e Dra. Anabela Ferreira, por parte do ITIJ, pela disponibilidade para acolherem a execução deste trabalho e pelo tempo que dedicaram a ajudar-me a desenvolvê-lo o melhor possível.

Agradeço à minha madrinha Ana e ao Carlos, por toda a preocupação e apoio incondicional dado ao longo dos tempos e por todos os cafés às nove da manhã quando saía das minhas reuniões de orientação.

Agradeço aos meus tios José e Maria, Manuel e Isilda e Joaquim e Maria, por sempre terem acreditado em mim e me terem motivado a continuar.

Agradeço aos meus primos Rui, Paulo, Dulce, Isabel, Carlos, Sofia e Paula, não só pela disponibilidade que sempre demonstraram para me ajudar a concluir esta tese, mas pelo carinho que sempre tiveram ao cuidar de mim.

Agradeço a todos os meus amigos, em especial à Raquel pelo apoio ao concluir esta dissertação, ao João e ao Messias, por todas as vezes que me tiraram de casa para espairecer e pela sincera amizade ao longo de muitos anos, à Joana, à Inês, à Margarida, ao Tiago, ao Jorge, ao Pedro, ao João, ao Nuno e a todos os outros, cada um à sua maneira e cada um de forma única e especial.

A todos devo a concretização deste trabalho. Muito obrigado.

Abstract

In today's society, the role played by information is growing in significance as the everyday actions executed by people, and inside the organizations, rest on its manipulation.

Aware of this reality, enterprises invest large amounts of their budget in information technology in order to shape up to their needs, which if not properly planned and executed, can result in incompatibilities between different information systems on the same organization.

This work proposes a reference information architecture for the Portuguese public administration, and the methodology used to develop it.

The development of the architecture was achieved by analyzing multiple contexts and their specific information architectures, following the different steps within the proposed methodology, that allowed to gather the set of information entities and respective attributes transversal to the public administration.

As the Portuguese public administration presents as a vast context, this work is focused on the citizen's lifecycle perspective, it's main activities, and different intervenient in the execution of the main processes, specifically the act of being born and dying.

The output of the work is a revised version of the reference information architecture, modeled with the CEO framework, and a methodology for gathering a set of transversal information entities from the analysis of specific contexts, that we used for developing this work.

Keywords: Architecture, Citizen, Entity, Information, Methodology, Public Administration

Resumo

Na sociedade actual, a importância do papel desempenhado pela informação assume cada vez maior importância visto que as acções desempenhadas pelas pessoas, no contexto das organizações, residem na sua manipulação.

A par com esta realidade, as organizações investem grandes quantias do seu orçamento em tecnologias de informação de forma a conseguirem oferecer resposta às suas necessidades, o que se não for devidamente planeado e executado, pode resultar em incompatibilidades entre diferentes sistemas de informação na mesma organização.

Este trabalho propõem uma arquitectura informacional de referência para a administração pública Portuguesa e a metodologia usada para a desenvolver.

O desenvolvimento da arquitectura foi conseguido através da análise de múltiplos contextos e das suas arquitecturas informacionais específicas, seguindo os diferentes passos da metodologia proposta, que permitiram obter o conjunto de entidades informacionais e respectivos atributos transversais à administração pública.

Como a administração pública Portuguesa se apresenta como um contexto muito vasto, este trabalho está focado na perspectiva do ciclo de vida do cidadão, das suas principais actividades e dos intervenientes na execução dos principais processos, especificamente o nascimento e a morte.

O resultado deste trabalho é uma revisão da arquitectura informacional de referência, modelada com a framework CEO, e uma metodologia para o levantamento de um conjunto de entidades informacionais transversais a partir da análise de contextos específicos, que foi usada para o desenvolvimento deste trabalho.

Palavras-Chave: Arquitectura, Cidadão, Entidade, Informação, Metodologia, Administração Pública

Contents

List of Figures

List of Tables

Acronyms

AMA	Agência para a Modernização Administrativa
BPD	Business Process Diagram
BPMN	Business Process Modeling Notation
CEO	Centro Engenharia Organizacional
DEMO	Dynamic Essential Modeling of Organizations
DTD	Document Type Definition
EAP	Enterprise Architecture Planning Methodology
E-R	Entity-Relationship
ERD	Entity Relationship Diagram
ERM	Entity Relationship Model
HDF	HL7 Development Framework
HL7	Health Level Seven
IEEE	Institute of Electrical and Electronic Engineers
ISO	International Organization for Standardization
ITIJ	Instituto das Tecnologias de Informação na Justiça
OSI	Open Systems Interconnection
RIM	Reference Information Model
SGML	Standard Generalized Markup Language
TIC	Tecnologias de Informação e Comunicação
UML	Unified Modeling Notation
XML	Extensible Markup Language
XSD	XML Schema Definition

"When you know a thing, to hold that you know it;
and when you do not know a thing, to allow that you do not know it
- this is knowledge."

Confucius, *The Confucian Analects*

1 Introduction

1.1 Context

Nowadays, enterprises are becoming increasingly information based and the improvement in their information activities is a priority among their set of objectives [1]. The activities and processes, executed in the context of an organization, majorly depend on the use, creation, sharing and exchange of information, shifting the costs associated with these practices into high values unless there is some work developed to counteract this. In the light of these occurrences, information technologies are a important tool to achieve organizational goals, and at present constitute a considerable portion of organizational expenditure [2]. Justify the investments in information technologies, however, is not something that all organizations are capable of doing, and therefore, management researchers, focus on understanding the gap existing between the desired and the realized level of objectives in this area.

A key to achieve these objectives and improvements is developing an Information Architecture [1]. A great number of organizations have in place an *ad hoc* information architecture, which does not comply with what the organization will need in the future because it has an incoherent framework, unnecessary duplications, incompatibilities, missing elements and lack of standards.

Within the public sector, organizations are mounting pressure to upgrade their information systems and optimize the way they conduct business, leading to the fact that they are trying to emulate the practices of the private sector [2].

Given these problems and scenarios, it is clear that a globalized concern about information should be rapidly emerging in order for the enterprises to shape up with its current needs.

Through the next sections we will address these problems, in the context of the Portuguese public administration, as well as provide the first glance towards a methodology to achieve the desired solution.

1.2 Objectives

Due to the nature of this work, and the way the development of the solution was conducted, it is divided among two components that complement themselves: an academic-oriented component, and an enterprise-oriented component. This means that the work developed uses academic knowledge and best practices, to address a concrete problem giving to the researcher the possibility of checking the validity, applicability, and relevance of his work.

In the next sections we will formally define the problem, as well as the set of contributions that we are aiming at achieving with this work, and finally how it will be validated.

1.2.1 Clarification of the Problem

Considering all the issues identified previously, nowadays, different organisms of the Portuguese public administration have in place different information systems, which creates incompatibilities when a need to transfer data among different locations arises. This situation also creates problems when we need to find where a set of information is stored, or which entities have access and manipulate them. It is by developing an information architecture that we can perceive which are the relevant, and transversal, information entities to the public administration, as well as the relations among them.

After clarifying the context of the problem, we take place by defining the set of objectives and contributions associated with this work. In terms of objectives, this work will aim at answering to the main question Q1 and the set of sub-questions Q1.1 through Q1.5:

Q1: *What are the common Information Entities through the Portuguese public administration, within the context of the citizen's lifecycle?*

In order to do so, we identified different information entities and processes in various contexts of the Portuguese public administration and then elaborated a set of those who are transversal to the entire set of organisms. Within the process of answering to the problem, we also worked towards determining what are the main processes in the life cycle of a citizen, so that we can relate the processes to the information, and align both, as well as

Q1.1: *How can we elaborate a global information architecture from specific data models?*

None the less, the sequence of steps to be followed, in order to identify the common entities and develop the architecture, raises another question of

Q1.2: *What methodology should be used to develop the information architecture?*

Yet another issue that should be noted is the representation of the final work. Therefore, after gathering the set of information entities and other information that will allow us to develop the architecture,

Q1.3: *How will the information architecture be represented?*

Finally, since we are manipulating the whole information entities, there should be work developed towards defining the attributes included in those entities, which lays on the question of

Q1.4: *Which attributes should be included in each information entity identified?*

On top of the work developed, there is the issue of its validity and alignment with the current reality of the Portuguese public administration regarding the citizen's lifecycle. Another question should then be posed, namely

Q1.5: *How do we evaluate the alignment between the information architecture, and the reality of the Portuguese public administration in terms of the citizen's lifecycle?*

By doing so, we successfully address the problem underlying, and are able to develop *An Information Architecture for the Public Administration*.

1.2.2 Contributions

While developing an information architecture, we are aiming at discovering what are the relevant information entities in our context, and how they relate to certain processes. Particularly, we want to infer what are the information entities which existence and manipulation is relevant, for the execution of the major processes on a citizen cycle of life, *e.g.* Born, Die, and Have Children.

In order to do so, we established which are these main processes, analyzed different contexts, elaborating a list of the existent information entities, and finally described the relations among them, developing the information architecture. To accomplish these tasks we have applied different methodologies and frameworks that enabled us to use best practices to gather information, as well as assure the validity of the obtained results.

The academic contribution relies in the usage and investigation of the previously stated methodologies and frameworks. By doing so, we were able to devise a set of methods, enterprise languages and frameworks that have applicability in these kind of problems, and so, draw conclusions about what are the best enterprise languages and frameworks to use for developing an information architecture, and how should the process of determining the information entities be conducted. The conclusions of this work might serve as base to extend the information architecture to further departments, or as starting point to develop an information architecture on another country or institution. By formally describing the methodology used during the development of this work, we intend to provide a foundation for another kind of similar works developed around this subjects.

In the context of the Portuguese public administration, the concern about the information architecture is integrated into the project of the TIC Network (*Rede TIC*, in Portuguese) [3]. The TIC network is composed of thematic workgroups of representatives from the public central, local and islands and its main objective is to implement an inter-ministerial network for TIC governance in the public administration, by proposing a set of guidelines in the domain of the information interoperability, electronic identification and integration. The organization responsible for the execution of this project is *AMA – Agência para a Modernização Administrativa*[1]and that is where the work developed under this thesis took place.

[1] More about AMA in http://www.ama.pt/

As the development of the information architecture is already taking place, we behaved like external consultants from the academic context with active presence, and advise in terms of the state of the art in best practices concerning methodologies and frameworks for developing the information architecture. Also, we took into consideration the documentation given by the enterprise, the ease of access to certain contexts that it can provide, and used that information to develop our solution for the problem.

Hence, the main contribution is, by aligning the objective of our work with the academic and enterprise contributions, successfully develop an information architecture that can be validated by AMA and the TIC network, having as basis the work already developed by AMA.

For each of the questions raised on section 1.2.1, there is a contribution given by this work towards obtaining a definite answer. In order to summarize and wrap up this section, we enumerate our main contributions for each of the questions raised:

- **Q1 -** Our contribution in answering this question is the final result of our work. With the work developed, we were able to review the existent version of the architecture, and provide a new one, already incorporating the knowledge gained during all of our research, therefore containing the set of information entities that we concluded as transversal to the Portuguese public administration.
- **Q1.1 –** We targeted this question while researching the field of related work. During our investigation, we further researched on different frameworks and enterprise modeling languages that, not only allowed us to conclude our work, but also put us one step closer towards having a solid knowledge base that allows us to answer this question. As the data models may vary between different types of work, and also the specific contexts as well, each work should assess their particular problems and objectives. Concretely, during our work, we found the Spewak's EAP very useful in addressing this problematic, hence having integrate it in our section of related work. The knowledge contained on that section constitutes a working base for each future work to pick on, and then evolve within its context.
- **Q1.2 –** Despite we have researched the state of the art methodologies for addressing the problem underlying, and considering the nature and contingencies of our work, we chose to develop a methodology and apply it directly to the development of our work. As we were able to achieve the objectives intended, the methodology itself proved useful when applied to the problem, hence establishing an initial validation for it. Therefore, the devised methodology behaves as our answer to this question and as our main contribution to this subject.
- **Q1.3 –** During the development of this work, we had the need of using different tools and languages to represent different stages of it. However, based on our research, we chose to represent our proposed version of the architecture using the CEO framework. Our modeling of the architecture using this framework, demonstrates its potential for

representing information architectures, being our contribution towards answering this question.

- **Q1.4 –** Within our proposed methodology, there is a strong component that relies on the capacity of cross analyzing different contexts and establishing relations between the architectures/entities and respective attributes of those specific contexts, and the set of transversal ones. By doing these set of mappings and establishing these relations, we were able to identify which attributes should be subject of change within the entities already on the version of the architecture developed by AMA. As a matter of fact, our proposed methodology addresses this question directly on the steps it includes, providing a contribution on this field.
- **Q1.5 –** Our main contribution towards answering this question rests on our practical case and the tools used to analyze the results obtained. By investigating the two projects, modeling their main processes, and identifying the relation between them and the information entities under analysis, we were able to define a CRUD matrix and evaluate the alignment, while validating our work. Therefore, we were able to demonstrate how the CRUD matrixes can be applied to this type of work, enriching the academic knowledge base, demonstrating how we can validate a conceptual work against a real-life scenario.

The sum of the contributions identified for each question posed, represent the overall contribution of this work to the academic and enterprise fields.

1.2.3 Work Evaluation

The work developed under this thesis was performed in close proximity with AMA, as it is the responsible for the execution of the project of the information architecture for the Portuguese public administration. Another major player in this context is the TIC network, for it is the mechanism that supports the government decisions in matters of the public administration and modernization initiatives.

This means that the AMA's project for the information architecture, and consequently our work and contributions, will be later validated with AMA in conjunction with the TIC network, and after that submitted to a Minister's Council Resolution for approval, as it is the highest authority responsible for the TIC network. The validation by the TIC network consists on the formulation of a specific workgroup per each entity represented on the network, that will evaluate if their information requirements are present on the finished work resultant from our collaboration with AMA.

Despite the validation on behalf of AMA and TIC network, we persecuted and developed a practical case, that allowed us to instantiate part of our proposed methodology, while simultaneously taking the first step towards validating and testing it. The work developed around this subject is later presented in section 5.2.

By enrolling on the practical case, we were given the opportunity of increasing our set of contexts under analysis, and test/validate part of the proposed methodology by reviewing the impact of the chosen contexts against the achieved results. As we later conclude, analyzing the projects

under the scope of our practical case, allowed us to reunite data about their alignment with the version of the architecture proposed by AMA, which is a necessary early action towards completing the set of steps within our methodology.

1.3 Structure of the Document

In this document starting with section this section, we began by introducing the context of our work and its objectives, the problem underlying and the academic and enterprise contributions. The aim is to motivate the reader into the problematic and challenges that surrounded this work, transmitting what we wanted to achieve with the work described along the follow sections.

In section 2 we focus on presenting a set of relevant work developed in this area, part of which played an important role during the execution of this work, and other might serve as an input towards developing future work. By scrolling down this section, the reader can have a perception of what were the foundations towards developing our solution and get familiarized with the concepts later presented. On the next section, a theoretical overview over what were the methodologies and principles used during the research and analysis executing this work is presented, and the scientific and action-research methodologies are presented.

Leaving the themes approached during our initial research, in section 4 we present our proposition for a methodology towards achieving the results we were committed to. As it is one of our contributions, the definition of a methodology for addressing this problem is presented in a series of steps along the section, and a description about the type of work realized in each of them, as well as the evolution of the different knowledge gathered around the theme, is also described.

Shifting to section 5, the proposed methodology is put to practice, and the information obtained is presented. Therefore, the results from the instantiation of our methodology, our research, and our analysis, built up to our preposition for a revised version of the information architecture which is presented here.

Finally, through sections 6 to 8, we discuss our work and the results obtained, propose opportunities for improvement in terms of future work, and culminate with some considerations under the topic of conclusion.

2 Related Work

After having defined the objectives and contributions of this work, we skim the academic areas of interest that relate to the domain of the thesis. Here we present the main concepts revolving this subject as well as the current work being developed around it.

The chapter begins with an overview of some fundamental concepts to understand the environment in which the work is being developed. After that, in section 2.2, we present a set of enterprise modeling languages and frameworks that served as basis to conduct the work in order to achieve the intended objectives. We start by describing the Enterprise Architecture Planning, Zachman Framework, Business Process Modeling Notation and DEMO Methodology, and then

ArchiMate, CEO Framework, Unified Modeling Language, the Entity-Relationship Model and finally the Extensible Markup Language. Next, in section 2.3, we focus on describing the Health Level 7, their objectives, their work and their Reference Information Model. In section 2.4 the subject is the information architecture and it aggregates a set of information about the topic. We start by defining the subject and some related matters, such as the concept of information entity and data types, and as we walk through the section we also present the objectives of the information architecture as well as its benefits.

In the last section we provide an overview of the current initiatives that are taking place in Portugal, in terms of the public administration modernization, and relate to the subject of the thesis. Their importance has to be noted, as the development of the information architecture for the public administration modernization will relate, and be applied, to some of them.

2.1 Concepts

In this section we aim at proving a formal definition of some concepts that are on the basis of this work and will follow its execution closely. We start by describing the concepts of enterprise architecture and public administration, due to their relation to this work's title, and then we distinguish between specific and canonical data models, since a major part of this work is based on the analysis of specific domains, with the objective of achieving a transversal information architecture.

2.1.1 Enterprise Architecture

Being currently applied over a broader scope than just in technical or IT domains, an enterprise architecture is commonly referred as an architecture at the level of an entire organization.

In order to fully understand the concept of enterprise architecture, one must first seek the definitions of architecture and enterprise themselves. According to IEEE Standard 1471-2000, an architecture is the fundamental organization of a system embodied in its components, their relationships to each other and to the environment. In this context and based in The Open Group, an enterprise can be defined as any collection of organizations that have a common set of goals and/or a single bottom line. Therefore, and according to Mark Lankorst et al. [4], this leads us to the definition that an enterprise architecture is a coherent whole of principles, methods and models that are used in the design and realization of an enterprise's organizational structure, business processes, information systems and infrastructure.

An enterprise architecture captures the essentials of the business, IT and its evolution, while still allowing for flexibility and adaptability, providing an holistic view of the enterprise and contributing to achieve business success. Also, it benefits the organization by transmitting a much better understanding of its structure, products, operations, technology, the web relations tying all of them together, and connecting the organization to its surrounding environment.

2.1.2 Public Administration

Public administration can be viewed as the development, implementation, and study of branches of government policy, aiming at the pursuit of public good by enhancing the civil society and ensuring a fair and effective public service.

In a more concise definition, public administration is the typical *activity of organisms and individuals which, under supervision of the political power, realize in name of a collectivity, the task of satisfying regularly and continuously the collective needs for security, culture, and economical and social well-being, in the terms of the applicable legislation and under the control of the appropriate courts* [5, in Portuguese]. The definition of public administration, in the Portuguese context, is ambiguous as it can be perceived in different manners. To help defining it on a more precise way, Marcello Caetano [6, in Portuguese], defines public administration in a material sense, as the *group of decisions and operations in terms of which the state and other public entities search, among the general orientations driven by politics and directly or impelled by a stimulus, coordination and orientation of the private activities, to assure the regular satisfaction of the collective needs for security and well-being of the individuals, obtaining and employing rationally, with that purpose, the adequate resources*.

2.1.3 Specific Data Model

A specific data model is a model which defines a way of communicating within a particular type of domain [7].

Concretely, in our work, we considered a specific data model as the representation of a group of low level information entities for a specific context under analysis, *e.g.*, Segurança Social, of which a representation can be viewed on Appendix III, figure 31.

2.1.4 Canonical Data Model

A canonical data model is a model which serves as mediator between specific data models, establishing a common way of communication [7].

From our perspective, for the development of our work and specially during the stage of analysis of our proposed methodology, the set of XML provided by AMA and relating to the Portuguese interoperability platform, assumed the role of a canonical data model since they mediate the exchanges of information between a set of different specific contexts. if we target the information architecture proposed and assume the fact that it can be modeled into a set of low-level information entities specifically a, *e.g.*, database, it can also be viewed from the perspective of a canonical, reference data model.

2.2 Enterprise Modeling Languages and Frameworks

In order to make possible the development of the intended solution, research had to be made regarding the themes of interest that would contribute to the execution of our work.

Therefore, in this section, we start by presenting the Enterprise Architecture Planning which we used as basis for the development of our proposition of a solution's architecture, after that, we will approach the Zachman Framework for its capabilities of analyze and structure information of a certain context. Then we proceed to the Business Process Modeling Notation and the DEMO Methodology which gave us an insight of how to represent the processes that we defined and analyzed on the initial steps of the solution's architecture, after which we will describe ArchiMate and the CEO Framework whose characteristics fitted onto our necessities for representing the information architecture, and finally we will make a reference to the Unified Modeling Language and the Entity-Relationship Diagrams for their strategic importance when analyzing specific contexts, as well as to the Extensible Markup Language.

2.2.1 Zachman Framework

The Zachman Framework is a framework for describing an enterprise architecture, which provides a formal and structured way of viewing and defining an enterprise, and can be viewed from the perspective of a schema, an ontology, a meta-model and a basis for an architecture [10]. It consists of a two dimensional matrix (fig. 2), intended for classification, that should not be viewed as a methodology.

From the perspective of a schema, it is the intersection between two classifications. One is the fundamentals of communication represented over the form of the primitive interrogatives *What, How, When, Who, Where* and *Why*. The other is derived from reification, which is the instantiation of an abstract idea by transformation. It is the answers provided to the interrogatives that constitute the total set of descriptive representations relevant for describing an enterprise and allows the decomposition of complex ideas into simpler ones.

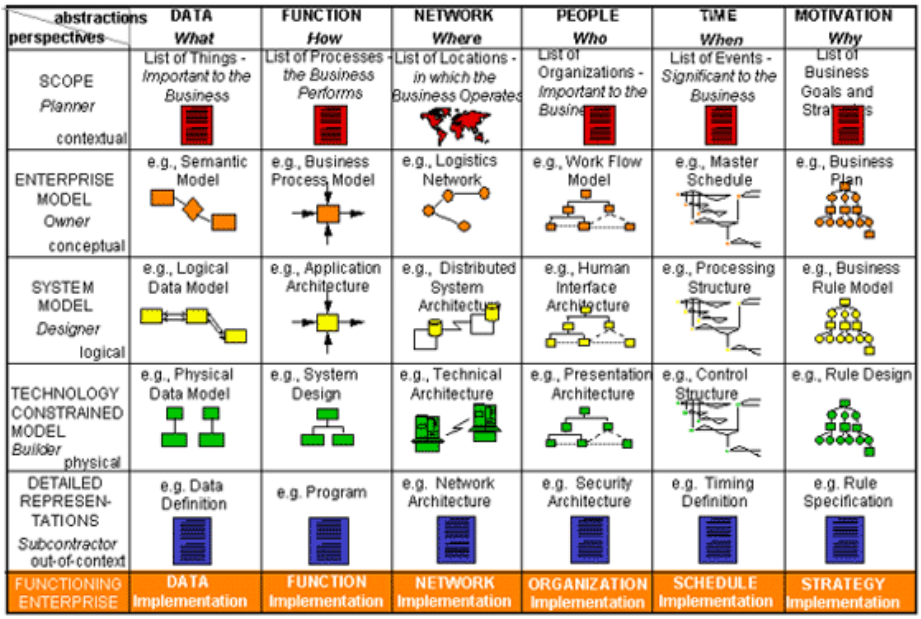


Figure 1: Zachman Framework

From the perspective of an ontology, it is a theory of existence of a structured set of essential components of an object. It is not a methodology, as a methodology is a process, and implies transformation, whether an ontology is a structure and provides definition.

From the perspective of a meta-model and opposing to a methodology, it does not imply that *e.g.* we make an architecture or simply build a primitive model. Also, it does not specify how the architecture is done or what should be done first. Those are the roles of a methodology which are not the concern.

According to the author, John Zachman [10], and due to the fact that we live in the information age in which enterprises are constantly changing and increasing in complexity, this framework has some "*profound significance in putting definition around Enterprise Architecture"*, therefore constituting the basis for architecture.

In the perspective of contributions to our work, each column of the Zachman Framework, represents information categories that, each of the intervenient present in the framework's perspectives, fills according to their responses to the questions *What*, *How*, *Where*, *Who*, *When*, and *Why*. On the same line of thought, the most important columns, from our perspective, are the *Data* that represents the structural relations between objects of the organization, the *Function* that gives a functional description of the organization and her business process model, and the *People*, that describes the aspects related to the internal structure of the organization. From the set of chosen columns, the first is where the dictionary and the diagram of entities is represented, the second is where the CRUD Matrix subject fits through the existence of a *Processes x Entities* matrix, and at last, the third column is where we can find the matrix of the *Processes x Organization*, respectively.

2.2.2 Enterprise Architecture Planning

Developed by S. Spewak [8] the Enterprise Architecture Planning (EAP) uses and complements the Zachman Framework, focusing on the two top levels of it. Spewak considers that these top two levels are directly connected to what the enterprise is, instead of the bottom four which are more concerned with the design, development, and implementation phases of the enterprise architecture planning.

The EAP methodology aims at defining the enterprise architecture for the top two perspectives of the business model and objectives, *i.e.*, define and plan the top two levels of the Zachman framework [9].

As in the Zachman framework, the EAP is strongly data-driven and business-driven, because information is the basis of business operation, therefore being of major importance to it [8]. With this, the EAP defines that the information model and information used by business, in order to run the business, must be defined first, derived from the business context and strategy, which ultimately results in the fact that the enterprise architecture definition should follow a strict sequence of steps, comprehending at first the data architecture, then the application architecture defined from it, and at last the definition of the technology architecture from the application architecture. The purpose of all this actions is to ensure the alignment between the different architectures and the business, essential

factor for obtaining strategic alignment. The Spewak's EAP presents as a four-step process illustrated on figure 1 [8]:

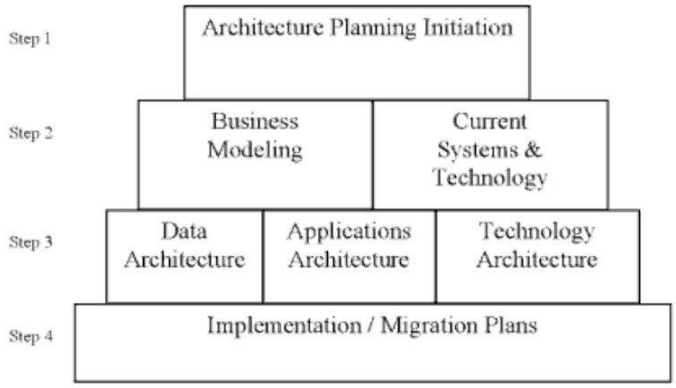


Figure 2: Spewak EAP Methodology

- **Step 1 -** is the initial planning of the architecture, focusing on the scope, objectives, roles, responsibilities and methodology to be used, considering the project plan as well as the resource requirements.
- **Step 2 -** is the modeling of the *As Is* of the information system environment, involving the modeling of the business, technologies and systems as they are today. Business modeling defines the preliminary functional business model, which takes into consideration the business activities, processes, and information used.
- **Step 3 -** defines the architecture vision for the future, *i.e. To Be*, using as ground the business vision and strategy which include, respectively, the target data architecture, application architecture, and technology architecture. The data architecture is intended to define the business entities and the entity-relationship diagrams, which are required to support the target business model and activities. Application architecture defines the conceptual applications needed to support the business functions and information processing, in order to make the business model effective. The technology architecture defines the distribution of data, and technology platforms, required to assure the existence of a technology infrastructure to run the business operational model support applications.
- **Step 4 -** defines the new architecture implementation and migration plans, including the applications to be implemented, the migration plan, costs and benefits, and recommendations.

In terms the contributions to our work, the EAP methodology served as basis for us to develop our proposed solution's architecture. However, we could not apply it directly, as we are not going to fulfill the entire set of required steps since they are not comprehended in the scope of our work. None the less, the first steps of the methodology gave us a coherent mean of achieving our goals, hence, with some adjustments considering our problem, it served as a starting point for defining our strategy towards the solution. Therefore, our solution's architecture presented in section 4 takes into consideration this methodology.

2.2.3 Business Process Modeling Notation

The Business Process Modeling Notation (BPMN) is a standard modeling language developed by the Business Process Management Initiative since May, 2004.

The main goal of BPMN is to provide a fairly easy to read and understand notation, that all business users can understand [11]. It is aimed to all the people involved in business, from the business analysts who create the first drafts of the processes, passing by the technical developers responsible for implement the technology, to the business people who will manage and monitor those processes. Also, the BPMN creates a standardized bridge that links the business processes design to process implementation.

BPMN defines a Business Process Diagram (BPD), which is based on a flowcharting technique, designed to create graphical models of business processes (fig. 3) [12].

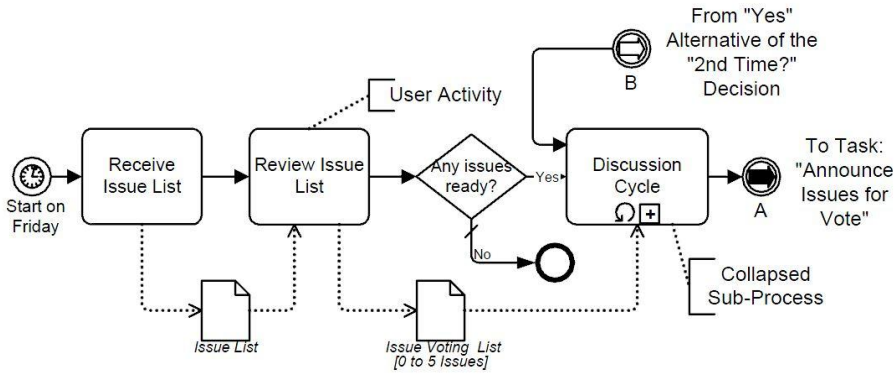


Figure 3: Example of a Business Process Diagram

A BPD is composed of a set of graphical elements subdivided into four categories:

- **Flow Objects -** set of core elements that help simplifying the graphic identification of shapes. Are divided into *Event*, *Activity* and *Gateway*.

- **Connecting Objects -** provides the basic structure of a business process by connecting the different existent flow objects. Are divided into *Sequence Flow*, *Message Flow* and *Association.*

- **Swimlanes -** based on some modeling methodologies, allows activities to be organized into separate visual categories illustrating different functional capabilities or responsibilities. Are divided into *Pool* and *Lane.*

- **Artifacts -** allows extensions over the basic notation and modeling of specific contexts. They can be in any number and created by modelers. Are divided into *Data Object*, *Group* and *Annotation.*

Since it has the capacity to cover many types of modeling, BPMN is used to communicate a wide variety of information aimed at different audiences. Due to the nature of its specification, BPMN is generally used for creating BPD of *Collaborative (Public) B2B Processes* and *Internal (Private) Business Processes* [11].

From our perspective, the BPMN is the language in which we represented the processes that we encountered during the development of the solution. Since we attempted to find what are the transversal information entities that take part in the major processes of a citizen's life, we needed to represent the processes that we analyzed, and for its wide use and versatility, BPMN proved useful on this task.

2.2.4 Dynamic Essential Modeling of Organizations Methodology

The Dynamic Essential Modeling of Organizations (DEMO) is a methodology for modeling business processes, while understanding the relationship between them and the information systems [13]. It is used for modeling an organization from the responsibility and communication perspective, by brokering the *world* into major business processes, then each of them into *actors* and *interaction*, and finally each of them further down into *action rules* and *coordination acts* (fig. 4) [14].

Business Processes	
Actors	**Interaction**
Action Rules	**Coordination Acts**

Figure 4: DEMO framework

An *actor* is a basic unit of responsibility, meaning that actors are units of responsibility and can be fulfilled by *subjects*, *i.e.*, people. Not only a subject can perform the role of more than one actor, as well as different subjects can represent the role of the same actor. *Interaction* is the communication that is performed between actors to allow them to perform tasks. *Coordination acts* present as the basic acts that allow actors to coordinate their work, and therefore interact with other actors, in order to accomplish their tasks. Finally, *action rules* are rules that define how an actor responds to a certain interaction, and what kind of response does it trigger.

According to DEMO, an organization is made up of subjects performing acts, which are divided into two categories:

- **Coordination Acts (C-acts) -** Acts referring to commitments between subjects.
- **Production Acts (P-acts) -** Acts that produce things, not necessarily material, in the world.

Coordination acts arouse when actors communicate with each other while performing acts. A communicative act is, by definition, the communication between two subjects about a certain preposition. Typically, one subject initiates the act, assuming the role of performer, and informs the

other subject, the addressee, about the intention and preposition of the communication. The proposition is the fact about which the two communicate, and the timeframe concerning that fact. Concerning the proposition, the intention determines what the performer wants to communicate. When the communicative act finishes, depending on the act, a communicative fact is created. Therefore, a coordination act generates a coordination fact, and a production act generates a production fact.

The responsibility to perform a one, and only one, type of P-fact is what was early designated as an actor. However, an actor does not simply perform its designated P-act an undetermined number of times. An actor will only perform a P-act if another actor requests the first actor to do so, meaning that an actor will only perform the P-act as a response for a series of C-facts. The cycle of an actor's job is represented in the figure 5.

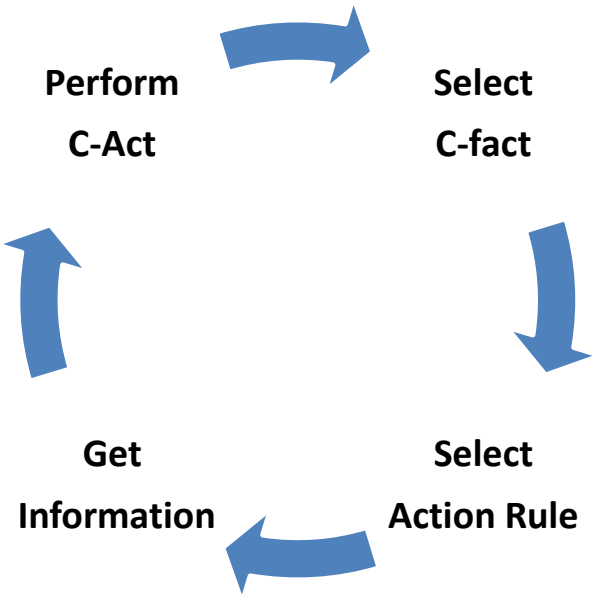


Figure 5: Actor's job cycle

Once a certain performer completes a C-act, a C-fact is created, and then the action rule will determine what are the acts, as well as from and which can be chosen, can the actor choose. A sequence of C-acts, C-facts, P-acts, and P-fact constitute a transaction, and the composition of connected transactions define a business process. The entire DEMO consists on a large set of actors interconnected by transactions.

Finally, the DEMO methodology comprehends various diagrams and tables that can be used to visualize the model of an organization (fig. 6) [15]:

Name
Actor-Transaction Diagram
Actor-Transaction Table
Process Phase Diagram
Process Step Diagram
Object Property Table
Object Fact Diagram
Actor-Bank Diagram
Actor-Bank Table
Action Rule Specification

Figure 6: DEMO diagrams and tables

DEMO is deeply connected to the subject of enterprise ontology, which would be the greatest application for our work. Enterprise ontology is focused on the essence of the operation of an organization, meaning that it is fully independent of the (current) realization and implementation of the organization. The theory that underlies the notion of enterprise ontology as presented by Dietz [46] is called the PSI-theory. Dietz uses this theory to construct a methodology providing an ontological model of an organization, i.e. a model that is coherent, comprehensive, consistent, and concise, and that only shows the essence of the operation of an organization model. This methodology is DEMO.

Compared to its implementation model, the ontological model of an enterprise offers a reduction of complexity of over 90% [46]. This reduction of complexity makes an organization for a manager intellectual manageable and transparent. It also shows the coherence between all fields within the enterprise, like business processes, workflow, organization structure, etc.

DEMO has been widely accepted in both scientific research and practical appliance. It has been used as a base for formalizing enterprise architecture and governance [47, 48, 49] and for formalizing the splitting and allying of enterprises [50]. Further research has extended DEMO by constructing an ontological model of an information system supporting the enterprise model (the BCI-3D method) [51]. More recently enterprise ontology has been used to construct a formal framework for service specification [52]. An extensive ten year study executed with 28 projects concluded that DEMO is a good method for the fast (re)design of organizations [53].

From our perspective, the DEMO Methodology offers a different approach of BPMN when representing the business processes, giving the possibility of defining different viewpoints, illustrating different relations. Despite the fact that our goal is not to emphasize the topic of the business processes over the topic of the information architecture, but hence promote the alignment between the two parts of our work, the DEMO Methodology would be an interesting approach of added value to the accomplished solution, therefore justifying our interest on this subject.

2.2.5 Archimate

ArchiMate is an enterprise modeling language, developed in Netherlands by cooperation between several Dutch partners, with the purpose of providing an integrated view of the different domains inside an organization [16].

The need for a language capable of doing this, appeared when the complexity of the modern organizations started to force the architects to find a way of expressing architectures as clearly as possible not only for themselves, but for the different types of stakeholders as well. In order to solve this problem, ArchiMate describes a taxonomy for mapping the different architectural components and comprehends a set of viewpoints aimed for different stakeholders. Although ArchiMate does not aim to be a language that replaces the need for specific domain languages such as *e.g.* UML, it can describe enterprise architectures in a precise way across the different domains and to different kinds of stakeholders, a feature that other languages in the same area do lack [17]. It provides the means for integration by allowing the production of models that represent the high-level structures not only within domains, but also between domains.

One of the ArchiMate's objectives is to aid *Change* management by presenting a *number of techniques that help architects and stakeholders to compare alternative designs [...] and to be able to study the impact of a change to the design* [18]. Another one is to facilitate the integration between the business, application and technology layers (fig. 7) [19].

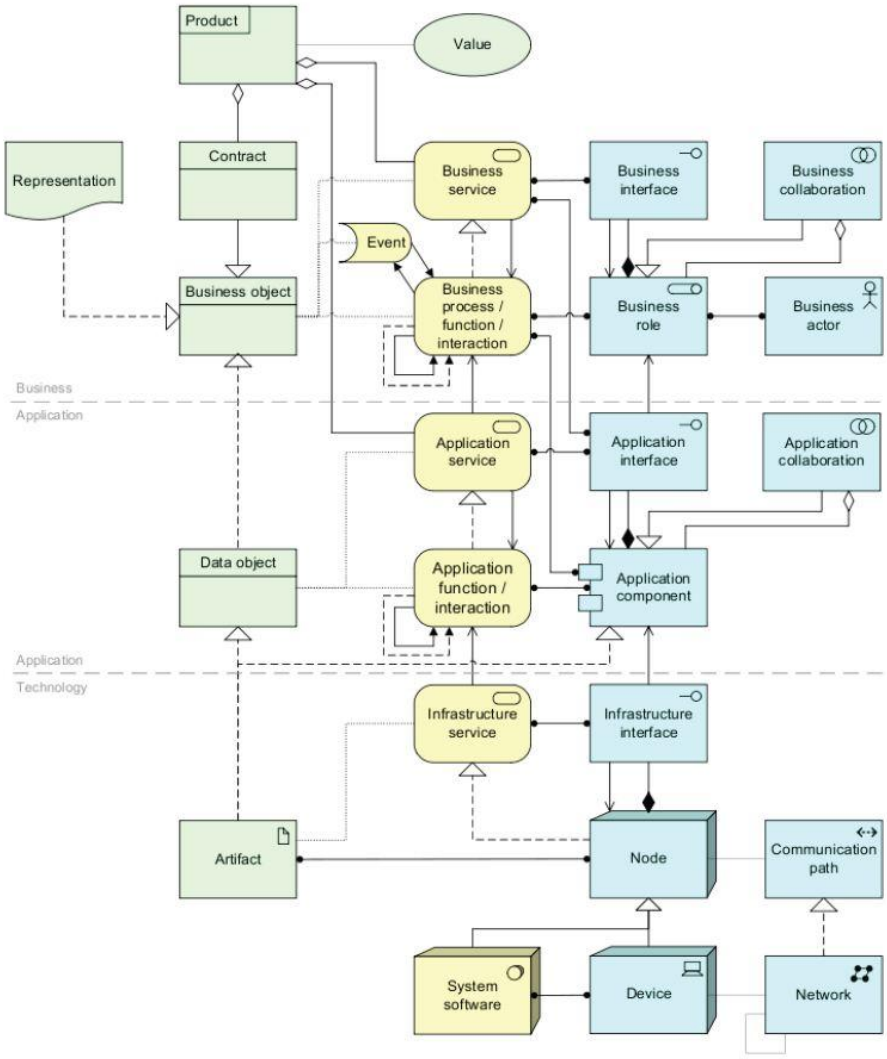


Figure 7: ArchiMate meta-model

As visible on the ArchiMate meta-model, the purposed concepts are divided among three different layers, and segmented in three different types particularly: the passive structure, behavior, and active structure, respectively represented on figure 7 in green, yellow, and blue. The three layers consist in:

- **Business layer -** addresses products and services offered by the organization to its external customers and environment. Business processes, executed by business actors or roles, provide support to them;

- **Application layer -** provides support for the business layer using application services which are realized by the technology layer;

- **Technology layer -** set of infrastructural services, realized by software, hardware infrastructure and networking services, that support the upper layer and are needed to run applications;

In relation to our work, ArchiMate does not comprehend the explicit concept of information entity, instead it has the information domain mapped onto the passive structure of the business layer (fig. 8) [20].

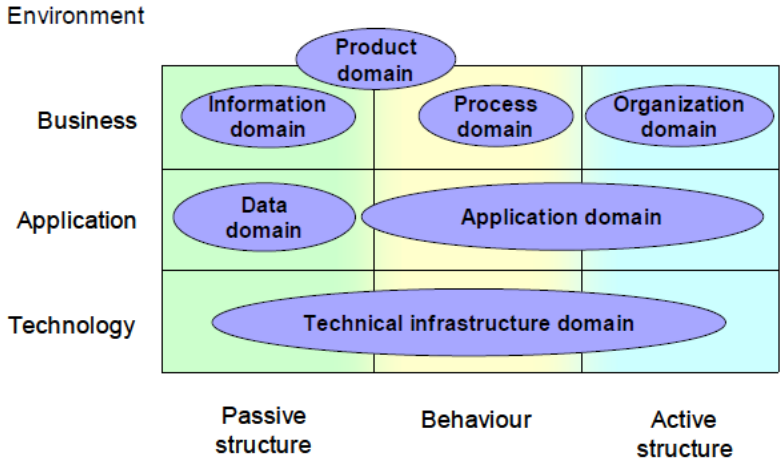


Figure 8: ArchiMate domains

From our perspective this means that we would have to make a comprehensive modeling of the different layers of the ArchiMate and extend the framework to include the concept of information entity. Despite the fact that our goal is not to model these layers, ArchiMate is still a powerful framework with known benefits, that could be adapted to our work objectives.

2.2.6 CEO Framework

The CEO framework was purposed by the Organizational Engineering Center (CEO, in Portuguese) as a standard UML profile (fig. 9) for describing, linking and tracing organizational concepts at multiple levels, allowing a common language to be used not only in business, but in software domain as well [21]. These levels are subdivided into strategy and goals, business processes and information systems. By using UML, it is used a common representation language in both business and software domain.

Although a business process does not have a universal accepted definition [21], it can be described as a set of activities that, in conjunction, transform the inputs, creating outputs which have value to a customer. Also, a business process has goals and is triggered by one or more events either internal or external to the organization, so is by using goals that one can assess, quantitatively or qualitatively a business process.

Since it is by analyzing and modeling, goals and processes, that an organization can trace its strategy down to the business processes, they reveal themselves to be of great importance, to which is not always given the proper attention. In order to represent the goal model, the CEO framework uses a Balance Scorecard approach [22] and extends the concept of goal, problem and contradiction. By doing so, a larger amount of business and strategic knowledge is captured.

As the alignment is one of the concerns, the business processes are represented as a hierarchic decomposition of the different activities that compose them, and provide added value in the achieving of a goal. Using this decomposition, a separation is made between the core and the supporting processes, as well as between goals and processes, allowing strategy and processes to be correlated. As the same approach is applied to the information system level, the system components that support the processes can also be represented.

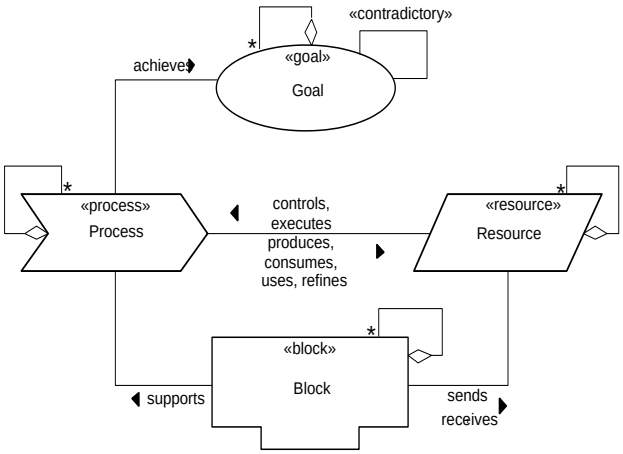


Figure 9: UML Meta-model of the CEO Framework

The framework aims at formally describing the business goals, processes and information systems, as well as the dependencies between them. In order to do so, it contemplates three separate levels (fig. 10) corresponding, each one, to the previously referred concepts and garnishes them with the means of representing the notions about the considered layer.

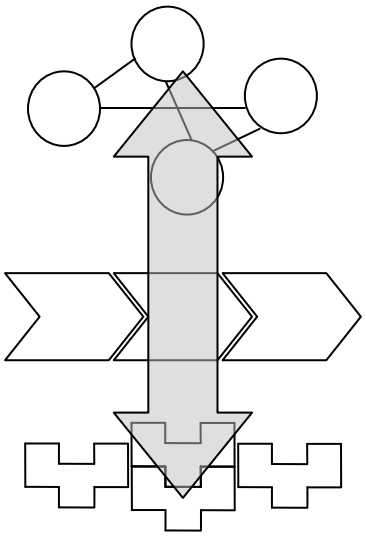


Figure 10: Goal / Processes / System framework

In the first level, the objective is to describe the business strategy through the goals set, which must be achieved through one or more business processes. The second level describes the business processes and its existence is relevant in order to satisfy one or more goals. Also, business processes interact with resources, in order to realize work, and may or may not be supported by information systems. Finally, the information systems layer aims at modeling the system components supporting the business. Although there is a separation among layers and the concerns are distinct, the dependencies and relations between them are also considered.

In terms of the relation with this work, the CEO framework comprehends the concept of information architecture, as well as the explicit concept of information entity, under the level of the information system architecture (fig. 11), which are two major advantages considering the objectives of the work [23].

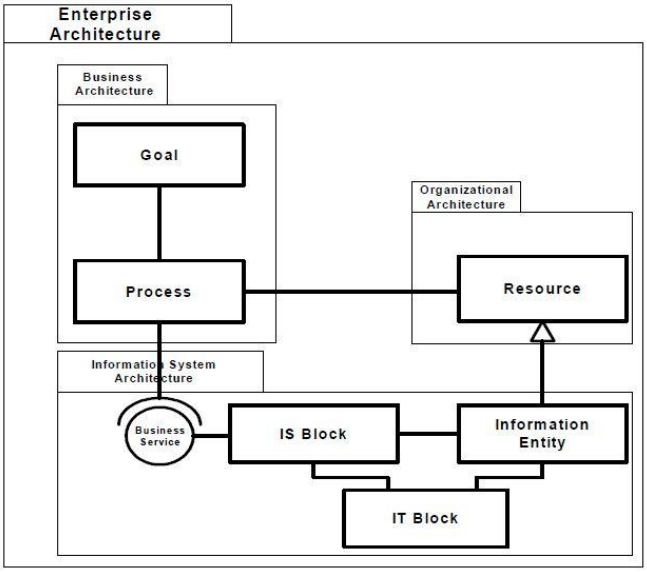


Figure 11: Simplified Meta-model of the FCEO for Information Systems Architecture

This framework is, therefore, of high value to this thesis as it constitutes a well-suited form of representation for our solution, providing us with the means needed for describing the output of our work. After analysis, the CEO Framework became the effective means for representing our work.

2.2.7 Unified Modeling Language

The Unified Modeling Language (UML), a relatively open standard controlled by the Object Management Group, is a family of graphical notations, defined over a single meta-model, which helps designing and describing software systems, particularly those which use an object-oriented approach [24]. It is a language for specifying, visualizing, constructing and documenting the artifacts of a software system and, all together, it represents a set of best engineering practices in its area of expertise.

The UML definition consists in *UML Semantics*, *UML Notation Guide*, *UML Extension for the Object Process for Software Engineering* and *UML Extension for Business Modeling*.

From Martin Fowler and Steve Mellor perspective [24], the UML is used in three different modes: sketch, blueprint and programming language. The most common mode is for sketching, by using it to clarify and communicate some aspects of a system. In this case, there are two different approaches: it can be used in a forward-engineering perspective, in which the UML is drawn prior to developing the system and helps it implementation, or in a reverse-engineering perspective, in which it is drawn from the interpretation of the system to help people understand it. When it is used for blueprinting, the difference is in the completeness with which, the UML, is drawn. Commonly, this method is used to develop a detailed schematic for a system that is going to be coded by a programmer (fig. 12). The schematic should be complete and precise enough, at the point that all important decisions are clarified so that the work of the programmer will be simplified if he follows it. However, in many places, UML is not enough and other non-UML diagrams are proven very useful [24].

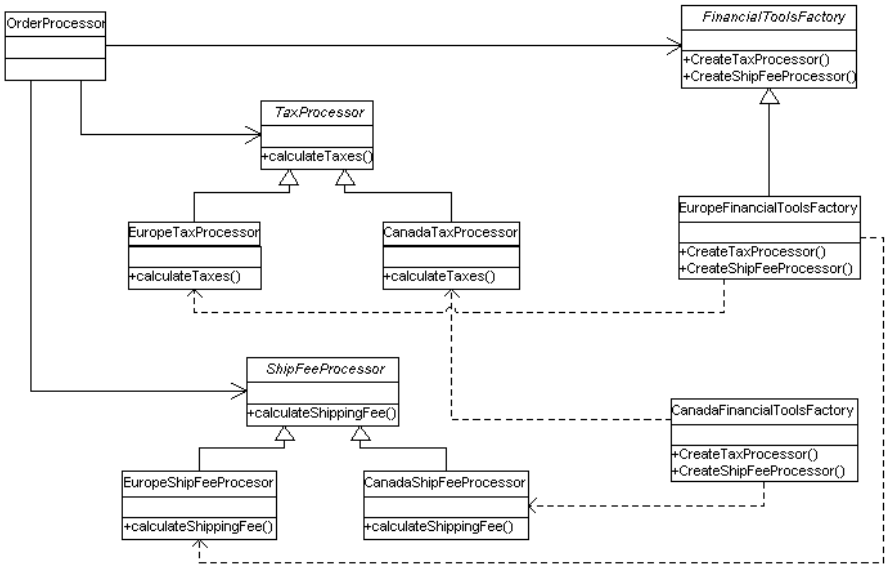


Figure 12: Example of a UML diagram

From our perspective, since we analyzed different contexts and looked into data layers and models, having a deeper understanding of a specific language for that purpose, like UML, helped us understanding and better analyzing them, hence the importance to our work.

2.2.8 Entity-Relationship Model

The Entity Relationship Model (ERM) is a data model which incorporates semantic information about the real world, and is better known as a diagrammatic technique tool for designing databases, the

Entity-Relationship Diagrams (ERD) (fig. 13). It is also used as a basis for unification of different views of data, particularly the network model, the relational model, and the entity set model [25].

The ERM model adopts the view that the world consists of entities and relationships, and by doing so, it reunites most of the advantages of the three data models mentioned above. Also, it is based on set and relation theories and can achieve an high degree of data independence.

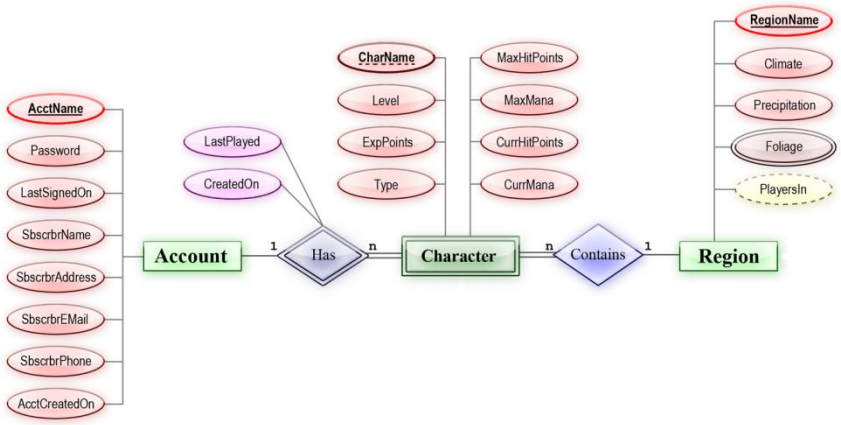


Figure 13: Example of a Entity-Relationship Diagram

The ERD consist of a number of conventions among which are the fact that entities are represented as rectangles, relationship sets as diamonds and attributes as ovals. If an entity set participates in a relationship set, a line is drawn between them. A line is also drawn between the attributes and the entities or relationship sets to which they belong. Should also be taken into consideration the cardinality constraints expressed with the numbers, the double lines and the underlines.

As there is more than one diagramming convention technique, *e.g.* Chen, IDEF1X, and Bachman, the notation can differ between diagrams.

From our perspective, the usage of ERD allows for the representation of information entities in a low-level architecture, much closer to a technological layer, ready to be mapped into a, *e.g.*, database. As we will later see, the first version of the information architecture purposed by AMA is represented in a entity-relationship model, and therefore the knowledge on this subject will allow us to better analyze the work, using it as basis for the development of our solution.

2.2.9 Extensible Markup Language

Extensible Markup Language (XML) is a simple, flexible text format derived from SGML, originally designed to meet the challenges of large-scale electronic publishing. Currently, XML is also playing an increasingly important role in the exchange of a wide variety of data on the web and elsewhere [42].

XML was developed by an XML Working Group (originally known as the *SGML Editorial Review Board*), formed under the auspices of the World Wide Web Consortium (W3C) in 1996, with the follow design goals:

- XML shall be straightforwardly usable over the Internet.
- XML shall support a wide variety of applications.
- XML shall be compatible with SGML.
- It shall be easy to write programs which process XML documents.
- The number of optional features in XML is to be kept to the absolute minimum.
- XML documents should be human-legible and reasonably clear.
- The XML design should be prepared quickly.
- The design of XML shall be formal and concise.
- XML documents shall be easy to create.
- Terseness in XML markup is of minimal importance.

The XML specification defines an XML document as a text which is well-formed, *i.e.*, that satisfies a list of syntax rules provided in the specification. In addition to being well-formed, an XML document may be valid. This means that it contains a reference to a Document Type Definition (DTD), that its elements and attributes are declared in that DTD, and follow the grammatical rules for them that the DTD specifies.

XML documents are made up of storage units called entities, which contain either parsed or unparsed data. Parsed data is made up of characters, some of which form the character data in the document, and some of which form markup. Markup encodes a description of the document's storage layout and logical structure. XML provides a mechanism to impose constraints on the storage layout and logical structure. A software module called an XML processor is used to read XML documents and provide access to their content and structure. It is assumed that an XML processor is doing its work on behalf of another module, called the application. XML processors are classified as validating or non-validating depending on whether or not they check XML documents for validity. A processor which discovers a validity error must be able to report it, but may continue normal processing. A DTD is an example of a schema or grammar.

Since the initial publication of XML 1.0, there has been substantial work in the area of schema languages for XML. Such schema languages typically constrain the set of elements that may be used in a document, which attributes may be applied to them, the order in which they may appear, and the allowable parent/child relationships.

The XML prove useful to our work during the analysis phase, by serving as a tool to map attributes and entities across different domains.

2.2.10 Critical Analysis

In this section we have presented the related work relating to the problem. During the approach to the different frameworks and languages, we have expanded the scope of our research, and therefore finished with a set of tools at different levels. None the less, each of them has an application to one,

or more steps of the architecture later proposed for trying to address and solve the problem underlying.

The Enterprise Architecture Planning will serve as ground for developing of our solution's architecture, as the results we want to achieve are comprehended within the steps of the methodology, which lead us to planning our development having it as basis. This way, we are using a known methodology with proven results, which will assure the validity of our proposed methodology.

The Zachman Framework could have been applied at an early stage of the development of the solution, when there was a need to analyze, and classify the different contexts, but ultimately it was not used focused our work and didn't feel necessary to analyze a universe that we would not contempt.

When analyzing each of the contexts, we needed to represent the processes that were being identified in a certain standard or language, and that is where we used the BPMN. As future work would be interesting making the same analysis using the DEMO methodology. While still in this step of the development of the solution, in certain contexts, we needed to analyze the work in a more technological perspective, revealing the need of using a language more appropriate for this task. In this step, UML and E-R should be proved useful, *e.g.* the first revision of AMA's work is modeled in E-R, mainly the E-R. The XML proved useful when representing the data collected and establishing relations between the different attributes and entities.

After gathering all the information we needed, to develop the proposed information architecture, we had to represent the information entities as well as the relations among each other, as well as the result of our work. Here, the CEO framework was used, but we could have also used Archimate to develop this task. However, Archimate would have required a more integrated and vertical approach, and facing the time available and the quantity of contexts to be analyzed it would not have been reasonable under the threat of not finishing with valid results in the expected time.

As should be expected, the correct identification of the relevant information entities, is not a trivial procedure, and it must not came as a surprise that we needed to analyze the contexts at different levels, *i.e.*, business, application, and technology, proving that the work should be supported on languages and frameworks that are design to assess each of these levels. A qualitative analysis on the possible contributions to this work, of several of the analyzed languages and frameworks, is represented on Appendix X.

2.3 Health Level Seven

The Health Level Seven (HL7) was founded in 1987 with the vision of creating the best and most widely used standards in healthcare [43]. It presents itself as a non-profit organization, ANSI-accredited standards developing organization dedicated to providing a model and related standards for the exchange, integration, sharing, and retrieval of electronic health information, used to support clinical practice as well as the management, delivery and evaluation of health services.

Its mission rests on providing standards for interoperability that improve care delivery, optimize workflow, reduce ambiguity, and enhance knowledge transfer among all stakeholders, *e.g.*, healthcare providers, government agencies, and patients. The name *Level Seven* refers to the

seventh level of the International Organization for Standardization (ISO) seven layer communications model for Open Systems Interconnection (OSI), *i.e.*, the application level.

Hospitals and other healthcare provider organizations typically have many different computer systems used for everything from billing records to patient tracking. Ideally, these systems should communicate with each other when they receive new information. HL7 specifies a number of flexible standards, guidelines, and methodologies by which various healthcare systems can communicate with each other. Such guidelines exist as a set of rules that allow information to be shared and processed in a uniform and consistent manner. These data standards are meant to allow healthcare organizations to easily share clinical information. Theoretically, this ability to exchange information should help to minimize the tendency for medical care to be geographically isolated and highly variable, while maximizing the capability for integration.

HL7 also develops conceptual standards (*e.g.*, HL7 Reference Information Model), document standards, application standards, and messaging standards (*e.g.*, HL7 v2.x and v3.0). The last version of HL7 is version 3.0 [44], and it is based on a formal methodology, the HDF.

The HL7 version 3 Development Framework (HDF) [45] is a continuously evolving process that seeks to develop specifications that facilitate interoperability between healthcare systems. The HL7 RIM, vocabulary specifications, and model-driven process of analysis and design combine to make HL7 version 3 one methodology for development of consensus-based standards for healthcare information system interoperability. The HDF documents the processes, tools, actors, rules, and artifacts relevant to the development of all HL7 standard specifications..

HL7 specifications draw upon codes and vocabularies from a variety of sources, and the v3 vocabulary work aims at assuring that the systems implementing HL7 specifications have an unambiguous understanding of the code sources and code value domains they are using.

2.3.1 Reference Information Model

The Reference Information Model (RIM) is the cornerstone of the HL7 Version 3 development process. The RIM is an object model created as part of the version 3 methodology that can be defined as a large, pictorial representation of the HL7 clinical data, indentifying the life cycle that a message or groups of related messages will carry. It is a shared model between all domains and, as such, is the model from which all domains create their messages.

The RIM is an ANSI approved standard, and it is illustrated in Appendix II.

2.3.2 Critical Analysis

In this section we have described the work of an organization, whose mission revolves around the production and definition of standards for the healthcare industry. The standards and models available for this type of industry gave us another perspective for this work, as they are not directly related to politics, rules, or technology, but to the concept of the living being. This type of approach proves itself very useful when attempting to achieve the goal of our work, since it relates to the lifecycle of a citizen and therefore ontological questions arise when trying to define concepts directly related to a human being and ideally independent from politics and technology, *e.g.*, life, death.

The HL7's RIM, specifically the part of the model containing the type *entities* that constitute the standard, form a possible information architecture for the healthcare industry. Hence, we have considered this architecture as yet another scenario and integrated it in our analysis, with the purpose of enriching our work with the ontological perspective given by this non-politic, health-related, human being based architecture. As the information should be considered a first class citizen, and therefore be independent from the current political structure and used technologies, it makes sense to compare, and include in the result of our work this kind of vision. Due to its characteristics, the HL7's work proved a valuable ally in the subject.

2.4 Information Architecture

There is no clear distinction between data and information. The differences can only be seen depending on the context, since both are materialized into bytes and both have syntax and semantic. Information can be perceived as data which is useful to someone at a specific time, being simply data outside that time frame. The utility itself is relative and depends on the context and purpose to which the data is being applied. The role of the information system in this matter is well defined, and resides on processing the data so that its utilization can be maximized.

An information architecture can be viewed as a structured set of multidimensional interrelated elements that support all information processes. It is also an effective framework for acquiring, organizing, and prioritizing a wide range of technological knowledge, facilitating the ability to effectively and appropriately apply it. It provides the framework for planning and implementing a rich, standards-based, digital information infrastructure with well-integrated services and activities [1].

The structuring of the information entities necessary to pursue the business processes of an organization is also a role associated with the information architectures. Additionally, it serves the purpose of defining which information entities are necessary and how they relate to each other.

2.4.1 Information Entity

An information entity can be perceived as any concept *e.g.* people, place or object, which has meaning to the business context and about which is relevant for the organization to store information about.

An information entity is characterized for having a name, a unique identifier, a description and its relation to processes, other entities and information systems [26].

2.4.2 CRUD Matrix

A *CRUD* Matrix represents the type of actions that a business process has over a certain information entity [27]. *CRUD* stands for *Create*, *Read*, *Update* and *Delete*:

- **Create** Implies at least the creation of the identifier of the entity.
- **Read** Implies that the business process accesses the entity.

- **Update** Implies a change on the state of the identifier associated with the entity.

- **Delete** Implies at least the invalidation of the entity's identifier, and it can no longer be manipulated.

In order to assure the alignment between the processes architecture and the information architecture, all processes should create/update, and read at least one entity, which implies that all the entities should be created, read, and deleted, assuring the completeness.

With the increasing size of the matrix, it becomes harder to administrate. To address this issue, the different lines and columns are combined accordingly to access patterns, which does not affect the architecture:

1. Similar lines and columns should be united.
2. The processes should be aggregated by the entities they create.
3. The processes that update entities should be aggregated with the processes who create them.
4. Proceed to the identification of the systems suggested by the aggregations.

After the proper manipulations and depending on the granularity of the elements that constitute the matrix, it is possible to identify different systems like domains, solutions or applications:

- **High-level Processes x High-level Entities** generate *Domains*.
- **Processes x Entities** generate *Solutions*.
- **Activities x Entities** generate *Applications*.

2.4.3 Data Types Taxonomy

The different data types determine the characteristics of the data access and of the information systems who manage them. The data can be classified in three different dimensions according to Immon [28]: Historical *vs.* Projected, Primitive *vs.* Derived and Public *vs.* Private.

- **Historical -** Record facts, previously occurred, with precise and correct values.
- **Projected -** Are approximations or previsions, of facts, that are yet to occur.
- **Primitive -** Depend on a single occurrence or fact inside the organization.
- **Derived -** Depend on multiple facts or occurrences inside the organization.
- **Public -** Their integrity is maintained by the organization and they can be the result of a single fact or record inside the organization. Also, they are relevant to more than one individual inside the organization.
- **Private -** They represent the immediate needs of each individual, which is also their owner. Also, they are relevant to one unique individual.

2.4.4 Objectives of Information Architectures

Boyton et al. views the importance of the information architecture as a key framework to increase the capacity of an organization to adapt to technology [29]. Cook [30] points that a prime purpose of an

information architecture, as a framework, is to develop standards. Vasconcelos et al. clarifies that the major purpose of an information architecture is to identify and define the major data types that support the business development and support the information architecture [31].

2.4.5 Benefits of Information Architectures

Enterprises are becoming increasingly information based, creating an urge to improve their information activities so that they can assure continuous competitiveness. Also, organizations suffer from a series of problems that affect the way they conduct business. Among them is the non-existence of an organization or process for creating and sustaining an information architecture, the lack of standards, and the fact that changes in technology occur faster than the ability of an organization, to bring new services and products to their workspace. Due to these facts, developing an enterprise information architecture is a main concern, as it helps to address this concerns.

The benefits that arose from an information architecture relate to strategic advantages, increase effectiveness, and reduce costs, particularly [1][32]:

- Easier information sharing and exchange.
- Improved security and privacy.
- More effective response to customer requirements through easier and faster building of information services.
- Increasingly effective matrix organization structure because of the use of common information services, resources, and tools.
- Promote data sharing.
- Easier incorporation of outside vendors within chains of needed capabilities and better integration with the rest of the departments, academic community, and industry.
- Reduced cost by eliminating data redundancies.
- Reduce time spent on future data collection projects.

2.4.6 Work Developed by AMA

As described earlier, this work was developed in collaboration with AMA, the agency responsible for the execution of the project of the information architecture for the Portuguese public administration. Despite the fact that we integrated the project in a relatively early stage, where there isn't yet an initial version approved, there is already some work developed that deserves a reference in this document, as it constitutes the state of the art in the subject in Portugal and will provide us a initial basis to work on (fig. 14) [33].

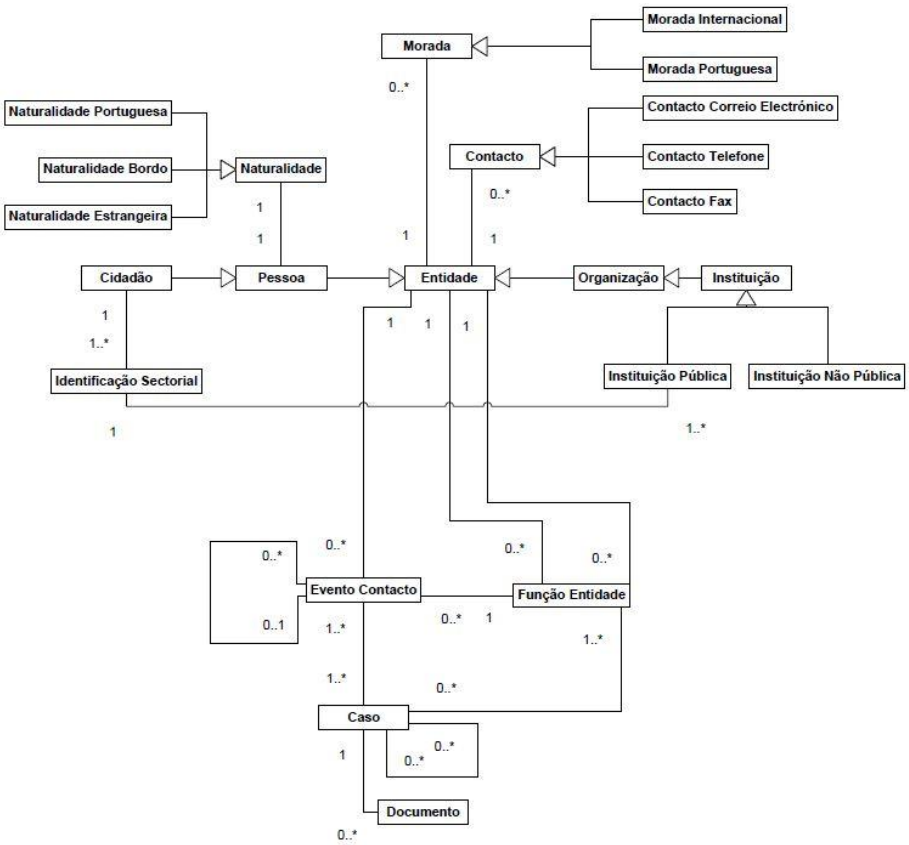


Figure 14: AMA's proposed E-R initial version of the relations between entities, *in Portuguese*

This information architecture gives effort to the areas of identification, relationship, and reception, reflecting the information entities needed for the identification and execution of contacts, and issues registration. The architecture aims at providing a basis for each organization to develop its own architecture, giving the necessary support for the execution of processes related to new requests, and facilitating the use of common code and language through the different organizations [33].

The presented architecture gave us a formal basis to work on, providing us with a perspective into the problem, and allowed us to use it as starting point towards the expansion onto other contexts, with the objective of accomplishing our work's goals and contributions already defined.

2.4.7 Critical Analysis

In this chapter we have made an overview of some of the relevant factors that contribute to the development of an information architecture. Due to the nature of this work, it is important to first understand what is an information architecture, then its objectives and advantages, and finally the concept of information entity. We have also made reference to the CRUD matrix which can be used later, on the development of the solution, helping us to maintain the alignment between the information entities and the business processes by analyzing the relation among each other. Since our primary goal is to proceed to the identification of these entities, the CRUD matrix proved to be a

useful tool in assisting with this task. Finally, by using the data types taxonomy a greater level of detail was accomplished when identifying the information entities.

2.5 Public Administration Modernization

In order to be able to develop the intended architecture, a initial phase of the proposed solution's architecture consists on the analysis of several specific contexts, from where should be inferred and defined a set of information entities associated with each of them. This was the foundation that allowed us to perceive what are the transversal information entities, and to do so, we had to understood how each of the contexts related to problem.

Considering our problem, in the context of the Portuguese public administration modernization, there are several initiatives currently being developed or scheduled related to the creation of the TIC network [34]. Since we want to determine the relation between the information entities and the major processes in the life of a citizen, we had to chose the contexts that could provide us the documentation and the means to do so. Being this work developed in collaboration with AMA, we had access to projects currently under its orientation, and this appeared as a major advantage, for projects like the citizen card and the citizen portal that already integrate a fair set of information regarding the citizen, are under the supervision of AMA. Another advantage was the ease of access that we had to other organisms of the public administration, as the project of the TIC network and the interoperability platform serve as a mean of integration for them, and are also under the supervision of AMA.

In order to further understand what each of this projects and initiatives are, a brief description is provided:

- **TIC Network** (*Rede Interministerial para as Tecnologias da Informação e Comunicação*, in Portuguese) This network is composed by a group of technology information and communication agents, enrolled in TIC projects, with the objective of defining TIC and interoperability regulations to be used on the entire public administration. The definition of these regulations is based on the transversal debate of the themes by the public administration intervenient on the subject, allowing for the knowledge exchange and creation of synergies between the public administration, the society, and the academic context. The figure 15 illustrates the organizational model of the TIC network, in Portuguese, and the existent groups [35].

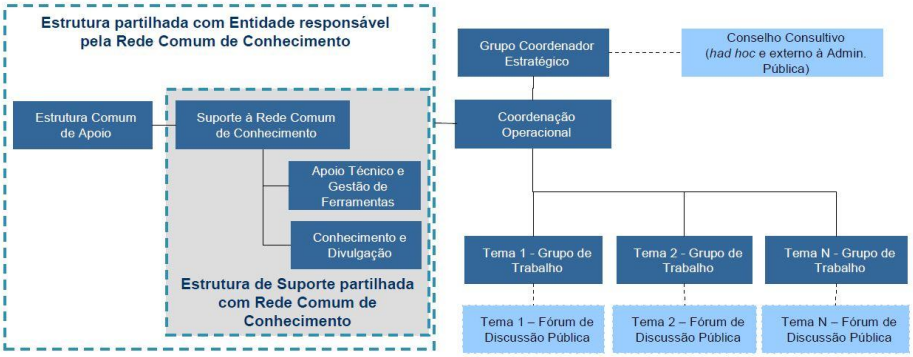


Figure 15: Modelo Organizacional da Rede Interministerial TIC, *in Portuguese*

- **Interoperability Platform for the Public Administration** (*Plataforma de Interoperabilidade para a Administração Pública*, in Portuguese) The project aims at creating a single platform, using open standards, that will be implemented in the public organizations. It will support the exchange of information in order to provide a mean of integration for heterogeneous systems, and therefore allowing multi channel electronic services among different organisms.
- **Citizen Card** (*Cartão de Cidadão*, in Portuguese) Being one the most visible results on the use of the interoperability platform for the public administration, the Citizen Card aggregates, on a single document of citizenship, the identification documents of several public organisms.
- **Citizen Portal** (*Portal do Cidadão*, in Portuguese) Its main goal is to simplify the relationship between the citizens and the public organisms, being a privileged channel to the services provided by the public administration. In 2009, the portal was making available to the citizen 670 services, provided by 125 organisms and public entities.

Although not directly related to the Interoperability Networks and TIC, there are other initiatives which could be relevant to this work in a future version. Therefore, a brief description is also presented.

- **Project Nascer Cidadão** The project *Nascer Cidadão*, in Portuguese, allows a newborn to be registered immediately after the birth. The act of registration takes room directly in the place of birth, being it an hospital or a maternity, and is performed by a member of the *registo civil* that periodically travels to the health care units. Among the objectives of the project is the reduction of the percentage of newborns whose father or mother are unknown, the effective declaration of all born children, and the simplification of the birth register since the parents no longer have to go to the *registo civil*.
- **Project SICO** This project plays an important role, not only in augmenting the rate of modernization within the scope of the public administration, but also on the increasing of the service quality delivered on the conservatories from the citizen's point of view. Project

SICO allows for the emission of the death certificates, in digital format, and their propagation to other organisms, while it helps materializing the concept of *'interoperability'* within the public administration.

- **IMTT** (*Instituto da Mobilidade e dos Transportes Terrestres*, in Portuguese) Its main objective is the coordination of inland transport, and presents as an independent entity with administrative and financial autonomy, as well as jurisdiction, over the national territory.
- **Land Register** (*Registo Predial*, in Portuguese) The land register aims at maintaining a record of the legal situation of the buildings, particularly, the composition of a certain building, to whom does it belong, and what kind of charges apply.

Contrary to the previous initiatives, which are direct responsibility of AMA, these are not. However, it was possible to access them through the TIC network, and by doing so collect the necessary information.

2.5.1 Critical Analysis

The initiatives described in this section constituted a major part of the contexts described as to be analyzed in the proposed *Solution's Architecture*. It is therefore important to understand what is the context in which these initiatives are integrated, and what kind of contribution can they make in order to help addressing the problem.

As one can see, the diversity of the contexts allows for the investigation of a wide set of major activities present in the life of an individual, giving us the possibility of inferring different information entities, over different initiatives, and by doing so over different contexts. The TIC network was a major player in this step of the development as not only it integrates a vast quantity of information about the citizen, but also allowed us to access other organisms, *e.g.* ITIJ, to obtain the information we needed.

3 Methodology for Work Evaluation

In this chapter we will aim at describing what kind of approach, method and methodology was used for developing the work and evaluate the obtained results.

Accordingly to Salomon [36], **scientific investigation** is defined as a "*work developed methodologically, when facing a problem, and for which one seeks a adequate solution of scientific nature".* The same author defines that **scientific work** is characterized by a "*written scientific investigation and processing of the methodologically approached questions*".

Hence, when developing a solution, the approach and methods used are what assures the credibility of the results obtained. Therefore, when developing a certain type of work, the validity of the obtained results must be supported through the utilization of a scientific method that, not only supports the work, but also allows for its repeatability [37].

3.1 Approach

The existent approaches are divided into two major categories [38]: **Quantitative** approaches and **Qualitative** approaches. Since the objective of this work is not the pursuit of the theory revolving work evaluation methodologies, we will only make an overview of the chosen methods.

In this work we followed a qualitative approach. This kind of approach has its roots on the social sciences and allows investigators to study the social and cultural phenomena. The main characteristics are the fact that the role of the investigator is fundamental in order to collect data, multiple data sources are used, and the investigator has a high proximity with the studied phenomena [39]. Considering the type of work being executed, this approach fitted as we were inserted into the real context revolving the main problem and performing multiple analyses over different sources of information *e.g.* different departments and services of the public administration.

In terms of methods, we used the **inductive** scientific research method and the **action research** methodology. In the inductive method, we start from particular observations and try to develop a generic solution. Once again, the justification for the usage of this method relates to the fact that we analyzed the different information entities, existent over different departments, and then identified which one of them were common to the public administration. The inductive method follows these steps:

1. Identification of the problem and formulation of the question.
2. Observations in the context of the problem.
3. Development of a possible solution to the problem.
4. Testing and validation of the developed solution.
5. Interpretation of the results and conclusion.

3.2 Action Research Methodology

Finally, after establishing the method, we describe the methodology that was used during the execution of this work. We adopted an **action research** methodology in which the objective is to respond to the immediate concerns (action), as well as learn with the process and increase the knowledge base of the scientific community. In this kind of methodology it is expected from the investigator to be enrolled in the target organization, instead of acting only as an observer [40], and the relation between the theory and the practice resembles a cycle in which the theory and the practice inform and depend on each other [41]. The figure 16 illustrates a possible sequence of steps on an action research based methodology.

Figure 16: Example of an action research methodology steps

The solution's architecture described in the next chapter was planned considering this scientific methodology, which is why there is a relation between the different phases of the solution's architecture, and the different steps on the action research methodology. The figure 17 demonstrates the relations between each step of the action research (on top), and the solution's architecture (on bottom).

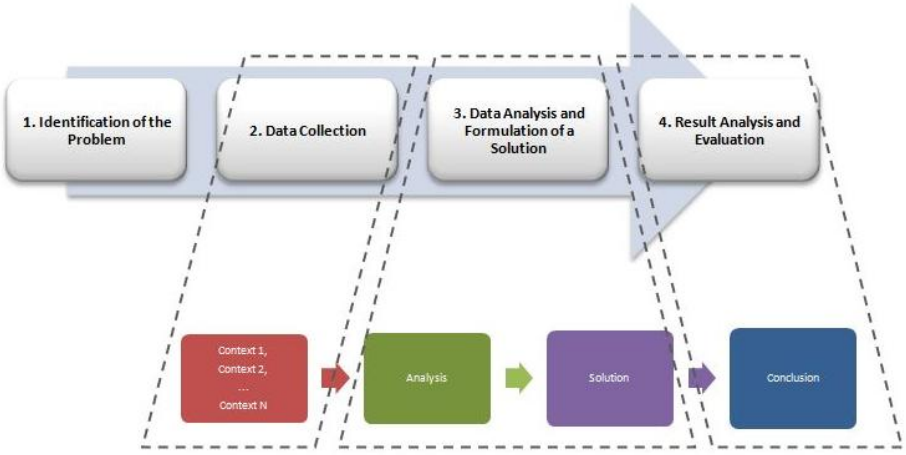


Figure 17: Relation between the solution's architecture and action research

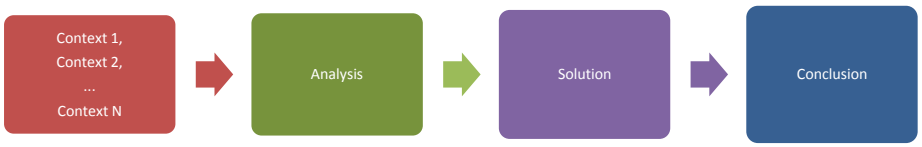
The first step of the action research does not have a direct correspondence to the solution's architecture, because it was performed in a early stage, and is already documented in this work. Therefore, the work produced after the execution of this step is described in the section 1 of this document, and corresponds to the definition and clarification of the problem, objectives, and contributions. The next step of the action research corresponds to the first step of the solution's architecture, which is where we analyzed the different selected contexts and proceed to the data collection as expected. The third step of the action research is related to the execution of the second and third steps of the solution's architecture, particularly when analyzing the information resulting from the previous context analysis and designing the architecture. The last step of the action research corresponds to the last step of the solution's architecture, where we evaluated the results obtained and draw conclusions.

The chronogram included on the Appendix XI provides a higher detail view of the different activities that will be conducted during the development of the solution. Each of the steps of the solution's architecture is directly mapped onto the chronogram, and is followed by a set of activities that were executed within that phase of the work.

4 Proposed Methodology for Developing the Information Architecture

In this chapter we present the different set of steps included in our proposed methodology for developing the architecture, which we followed during the execution of this work, whereas in the previous chapter we presented the scientific methodology used while developing this work. It is therefore imperative for the development of the solution to be aligned with the methodology in order to enable its application, and the collection of valid scientific results.

Facing the problem identified on chapter 2, considering the existing modeling languages, frameworks, methodologies - particularly the Spewak's Enterprise Architecture Planning, and the evaluation methodology to be applied, we devised a coherent sequence of steps/phases that allowed for the obtaining of a solution, and are an adaptation to our problem from Spewak's EAP methodology. Figure 18 contains a global perspective over the four main steps that constitute the proposed methodology, as well as a activity diagram developed in Enterprise Architect ©, that provides further detail on the activities realized in each of the defined steps.



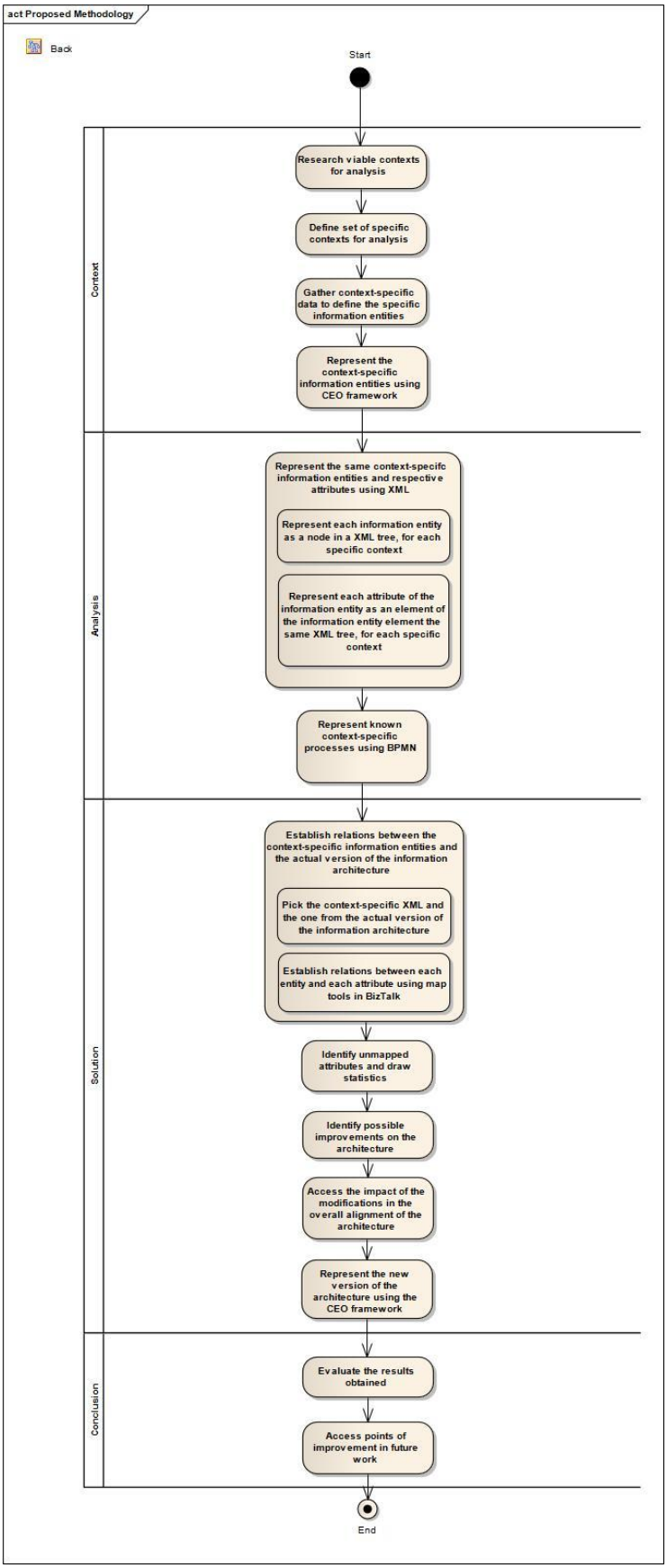


Figure 18: Proposed Methodology

The development of the solution started with a step pre-dating the one's represented on figure 18, aligned with the first step of the EAP, from which resulted this particular work. As defined by the EAP, and described through this work, we proceeded to definition of the scope, objectives, contributions, and project plan, in order to achieve the desired solution.

During the next sections we provide further insight on each of the steps of the proposed methodology, as well as on the work realized in them.

4.1 *Context*

Maintaining the alignment with the EAP methodology, shifting to its second step, in *Context* we aimed at defining which areas of the citizen's lifecycle we would want to target and proceeded to the analysis of the different set of contexts relevant for our work. This included not only different departments of the public administration, which represent an important role during the life of a citizen and with whom he interacts, but also two projects directly related to the processes of being born and dying. The contexts and projects that served as basis for this work are enumerated as follows:

- Administração Fiscal
- Identificação Civil
- Saúde
- Segurança Social
- Cartão de Cidadão
- Information Architecture developed by AMA
- HL7
- Project *Nascer Cidadão*
- Project *SICO*

For each of these contexts/projects we proceeded to its representation using the CEO Framework from a set of gathered data, *e.g.*, specific data models, project documentation, and interviews with enrolled people, and aimed at representing the context-specific information entities while focusing on the interaction with the citizen (fig.19). Here, we were able to gather different sets of information entities that allowed us to create a vision of the *As Is*.

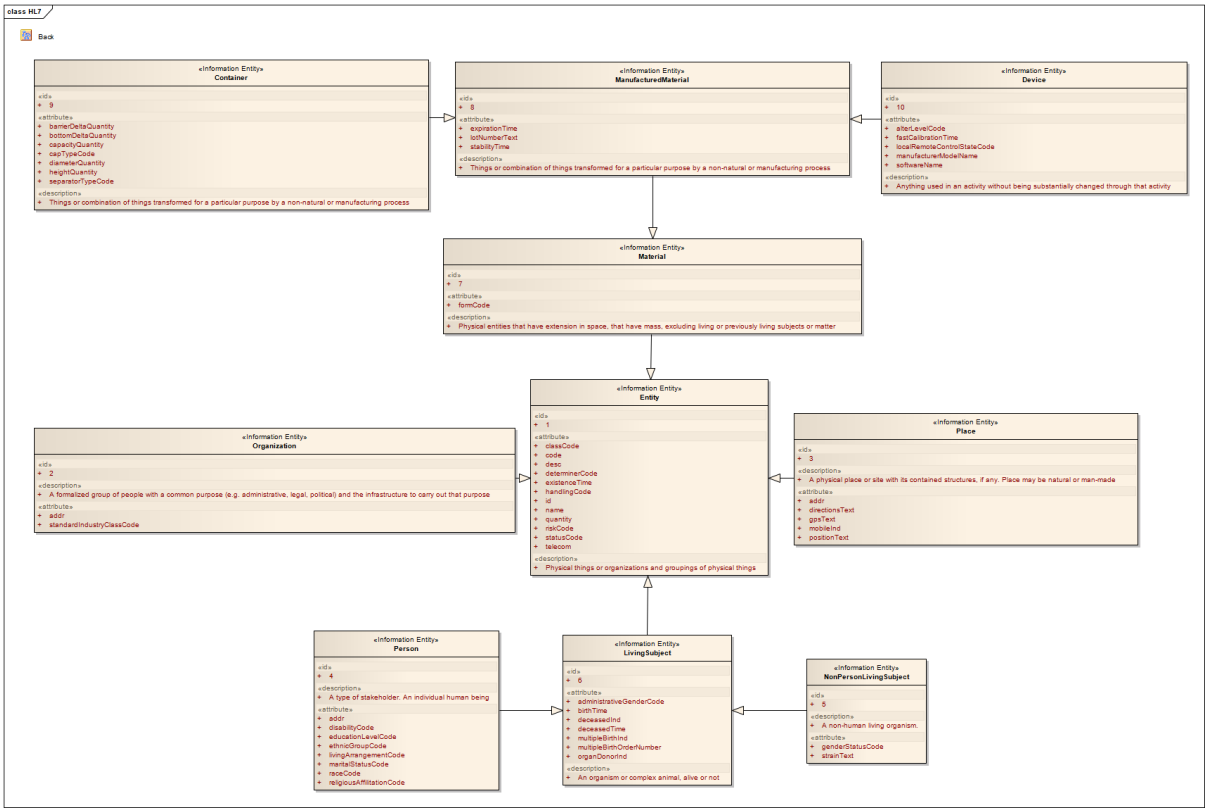


Figure 19: Example of a representation using the CEO Framework

(HL7 illustrated)

In order to assess and define the set of viable contexts for analysis, we started by doing a conceptual work of understanding what were the main activities in the life of a citizen. Then, by doing some brainstorming with people from AMA, ITIJ, and the academic environment, as well as some research on what were the departments of the public administration enrolled in the provision of the services related to that activities, we defined the set of contexts already enumerated. In order to model their specific information entities and attributes, we mainly used information provided by AMA. In the case of project *Nascer Cidadão* and *Project SICO*, we decided to form a practical case and went one step ahead of our methodology, representing their main processes with ITIJ's support. Finally, HL7 was represented based on our research developed for the field of related work.

In compliance with our methodology, we represented each of the context-specific information entities using the CEO framework, with the exception of the ones included in our practical case which we managed later.

4.2 *Analysis*

Next, in *Analysis*, still aligned with the second step of the EAP methodology, we reunited and cross-reference the information gathered from the previous step and, using XML, proceeded to the representation of the entities gathered and their attributes, in a structured manner. Also, after further

gathering information about the two projects under analysis, we proceeded to their representation using BPMN. This ended the definition of the *As Is*, and start the processes of definition of the *To Be*.

Each of the specific contexts was represented using a XML tree, and for each context-specific information entity, a new node was included in that three. The same procedure was followed to deal with the information entities' attributes, this time by creating elements for each corresponding node of the XML tree. As we did not have enough information to elaborate a similar representation for our subjects of the practical case, a different approach was followed, choosing to represent their main activities using BPMN and, with the information available, later associate them to the information entities instantiated during their execution. This was a disruptive action in terms of the devised methodology, but despite having considered the contexts on the practical as yet another set of contexts under analysis, our aim was to use them to validate part of the work, meaning that their treatment had to be different from the rest.

At this point we had documentation about the contexts, a representation of each of them using the CEO Framework, and a structured representation using XML. None the less, we had already developed work around the subject of traceability, and identified relations inside the contexts analyzed. This was a necessary step towards understanding the relations between the entire universe under investigation.

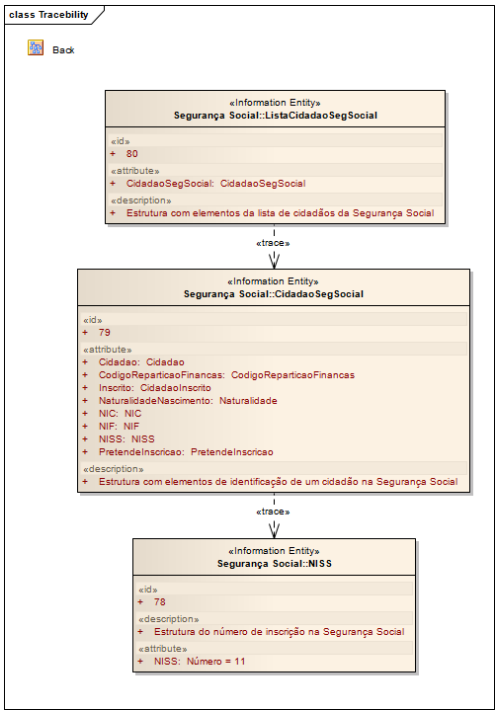


Figure 20: Example of indentified relations inside a context

(Segurança Social illustrated)

As can be seen on figure 20, the entity *ListaCidadaoSegSocial* has an attribute which is itself described as another entity, in this case *CidadaoSegSocial*, which once again has an attribute which is traced back to the entity *NISS*.

With the work developed until here, we were in position to start crossing information and analyzing the relation between the contexts, the existent version of the information architecture, and the two processes modeled from the projects analyzed. The result of this last actions was later the main trigger that initialized the implementation of the solution, supported on data.

4.3 *Solution*

In *Solution*, aligned with the third step of the EAP, we elaborated the information architecture based on the work developed here and until here. In order to do so, we took all the knowledge gathered from the previous phases and analyzed the relations between the different topics investigated.

While the results are presented on the next section, we started by picking up each context's XSD and match it against the one correspondent to the actual revision of the information architecture. To do so, we introduced the use of the mapping tools available in BizTalk, and mapped as much attributes as possible between the two sides, using the application and the XML provided by AMA, which originate on the *Portuguese Interoperability Platform* . If we recall the previous step of our methodology, we had started with a base representation of the information entities using the CEO framework, developed an equivalent representation using XML, mapping the contexts as trees, the entities as nodes, and the attributes as elements, and now, having each context represented in XML, the version of the architecture proposed by AMA, and BizTalk, we used the map tools available in it, to establish the relations between the entities and attributes of each of the specific contexts, and the ones in the current version of the architecture.

To maintain the consistency with the *Portuguese Interoperability Platform*, and to give another insight into the topic of mappings using roles, and the modeling work performed during the step of *Analysis*, we established the same relations as in the documentation provided, grouping the entities by roles. This can be perceived as a drill-up on the work realized with the XSDs and BizTalk.

Taken that we were able to model two processes that represented precisely the beginning and ending of a citizen's lifecycle, and with limited information and resources, we tried to identify the relation between the activities in the processes and the information entities in the context. This was based on the information provided by ITIJ during the interviews, and therefore the results are subject of discussion as they do not incorporate a complete vision of the information manipulated within the two processes analyzed.

Following this work, another analysis phase took place and at the end of it we were able to identify what were the most recurrent attributes to not have a mapping with the current version of the architecture. In order to do so, we picked on the results of the mappings established, and identified which attributes did not have a correspondence with the architecture proposed by AMA. We devised a set of statistics, were we identified how many times a certain attribute appeared as unmapped and on how many contexts, the global number of attributes and entities analyzed, and the ratios between the attributes analyzed and the ones categorized as unmapped. The formulas used are presented later

on section 5.4.3. This and other data made possible to come up with a set of opportunities for improving the alignment between the reference architecture and the different contexts, namely what were the top attributes categorized as unmapped and their dispersion over the contexts analyzed. By considering these two key factors, we chose a set of attributes viable for inclusion on our version of the architecture. After having defined the set of attributes to include in the new version of the architecture, we applied traceability on the attributes, used the XML to define which information entities were affected by their inclusion, and developed a new representation of the architecture using the framework CEO. based on the same rules of equivalence used to make the representation from FCEO to XML.

With the information from ITIJ, and the representation of the processes using BPMN, we developed a CRUD matrix for each process, and represented the manipulations on the information circulating within each process. As the practical case was intended for validation, after analyzing the CRUD matrixes, and effectively gather de information managed within each process, we checked if each piece of information could be matched against an attribute in the XML of the AMA's proposed architecture. By considering each piece of information manipulated as an attribute, we developed a similar process as for each of the specific contexts previously analyzed. This concluded our stage of validation, and by the end of it we had a proposed information architecture, considering the new attributes from the instantiation of our proposed methodology, and the validations done with our practical case.

Hence, the most noticeable output of this step and of this work took shape, and a new version of the existent architecture is proposed.

4.4 *Conclusion*

Finally, in *Conclusion*, we evaluated the work done and made considerations about the future work to be developed. As the definition of the migration and implementation plans towards the solution is out of the scope of our work, the fourth step of the EAP was not performed. Still we looked into it, and considered any relevant information that might exist, which could benefit/complement this work. Also, we assessed the impacted of our modifications, and evaluated the alignment improvement between our proposed version of the information architecture, and the one proposed by AMA.

5 Proposed Information Architecture

Through this chapter we provide the results obtained with the concretization and execution of the different phases of our proposed methodology.

We start by presenting our representation of the information entities that we identified within the different contexts analyzed. From the fact that we assumed an pro-active role in our investigation and realized field work in order to get information specially from two of our contexts, *e.g.,* interviews, we united the two projects analyzed into the section 5.2, and provided further information on them.

Although having already a description of what they encompass, we investigated further (without support of previous or related work), and modeled their main processes.

On the next section designated *Mappings*, we present the result of our work with BizTalk and XML in order to establish the relations between the attributes existent in the contexts we analyzed and AMA's developed information architecture.

Finally, in section 5.4, we present our proposal for a revised version of the information architecture developed by AMA, containing modifications that during the execution of our methodology we concluded that could improve the current version.

5.1 Information Architecture As-Is

Although having already provided an overview of the methodology used in the development of this work, throughout this chapter and the ones that follow, we will drill-in into each of the phases established, and further describe the work done.

The work under this chapter mainly concerns the *Context* phase of our methodology, as the majority of the work realized under the *Analysis* phase was divided between chapters 5.2 and 5.3.

At the beginning of this phase, our main concern was how could we restrict the scope of the work we were attempting to realize, so that it could became viable in the time window available for its realization. Attending to this matter, we made an option of focusing on the perspective of the citizen's lifecycle, leaving other universes of analysis out of our scope, *.e.g.,* organizations. While focusing on the citizen, and following the devised methodology, we needed to identify what were the main processes that we wanted to investigate further. This became a very delicate work, as we had to consider several implications that could arouse from our options. After discussing the matter with AMA's representatives, and getting to know the existence of different projects like *Project Nascer Cidadão*, *Project SICO* and *Framework de Serviços Comuns*, it was clear that an adequate choice would be focusing on the birth and death of a citizen. This choice was made taking into consideration several key factors like the alignment with our and AMA's objectives, the projects available for us to research on that supported this two activities, and also the support documentation available at our disposal. After some initial analysis of what the two projects consisted of, and with what entities did they interact, we focused on where to gather information about those contexts. This was where the project of the *Framework de Serviços Comuns*, appeared. This framework, as already described, supports the communication between different departments of the public administration, and AMA had available useful documentation that allowed us to identify exactly what was the information being transmitted. From this multiple sources we gathered the information we needed and constituted our universe of contexts for analysis.

As we discussed earlier, we also wanted to do some work less connected to the processes and technology already existent, and more related to an conceptual and ontological perspective. As we could no focus our work solely on that subject, we researched and concluded that another source of information and best practices in class would be the HL7. In order to provide integrate potential advantages from this context, we also decided to include it in our pool of contexts, and the final set obtained was as follows:

- *Administração Fiscal*, in Portuguese
- *Identificação Civil*, in Portuguese
- *Saúde*, in Portuguese
- *Segurança Social*, in Portuguese
- *Cartão de Cidadão,* in Portuguese
- Information Architecture developed by AMA
- HL7

For each of the contexts, based on documentation provided by AMA and using the Framework CEO, we modeled each of the information entities. The meta-model used in Enterprise Architect for the Framework CEO is presented on Appendix I. The results obtained are presented in Appendix III, grouped by context. Depending on the available information, each entity is identified by a unique id, and contains the attributes related to it, as well as a brief description. Appendix VI contains a full list of the information entities modeled, their correspondent id, and a color code indentifying the context they belong to.

5.2 Practical Case: Project *Nascer Cidadão* and Project *SICO*

This section describes the work developed around the two contexts that proved to be more of a practical case than yet another context analysis. In order to develop work around these projects, in collaboration with AMA and ITIJ[2] we have performed field work aligned with our research methodology, conducting interviews to key representatives, and on field analysis. With the results from this analysis, we were able to identify within the key activities of the projects at study, the information needed and in circulation between the different intervenient.

After establishing the specific objectives for this part of the work, we proceeded to the collection of the information needed, in order to understand which entities were involved within the scope of these projects and in what measure were the projects supported on the same information architecture as the one proposed by AMA. The modeling was achieved through the use of BPMN and the results are illustrated on figures 21 and 22. The figures illustrate, not only the main activities of each process, but also the information entities instantiated during them. The viewpoint provided only has de information entities effectively instantiated, in order to facilitate the reading. Entities like, *e.g.*, *Pessoa* and *Cidadao* are manipulated during the entire two processes by means of readings and update. However, these entities are not represented in association with all activities, since the information about their manipulation is represented on the CRUD matrixes, which are explained next.

Since one of our objectives was to collect data about the type of information used along the execution of the processes, we tried to get hold of documentation and examples of web services being used in production. However, due to the nature of the information, we were not able to get complete and precise data on this subject, and had to rely on the information provided during the interviews and meetings realized. Although without formal documentation support, the information we

[2] More on ITIJ in http://www.itij.mj.pt/sections/home

received allowed us to produce a CRUD matrix for each of the processes in order to infer the relation between them and the manipulated information within them. The result of this work is illustrated on tables 1 and 2.

Having determined which was the information managed inside these two processes, and in the scope of these two projects, an analysis needed to be made in order to perceive the relation with the type of information managed and the attributes of the entities contained in the existent version of the information architecture. Table 3 provides the correlation between these two contexts.

Being on the line of the work described in section 5.3, we chose to present the analysis directly related to these component of our work, as in our understanding it should be faced in a distinct manner, as it was a practical case. Never the less, the work realized is aligned with our methodology and with the rest of our work, being a complement of the analysis realized in section 5.3.

The development of this practical case followed our proposed methodology, but in a somewhat disruptive manner when considering the path that our analysis of the specific contexts followed. Assessing the major differences, they relate to the fact that on the practical case we modeled the main processes and did not represent the specific information entities using first the CEO framework and then XML. Never the less, the ideal behind our proposed methodology was still followed, and in the *Context* step, we research on each project and conducted a first interview towards understanding their behavior. in *Analysis* we took further insight, performed another set of interviews, gathered more information and were able to identify each activity within the processes, and the information manipulated within them, even with limited knowledge. In *Solution*, we iterated on their already developed representation using BPMN, established the mappings between the information manipulated and the version of the architecture proposed by AMA, and proceeded to the validation of the existence of the attributes against the ones in the architecture proposed by AMA.

With the conclusion drawn from these validations, we were able to fulfill the role intended for the practical case, within the scope of our work.

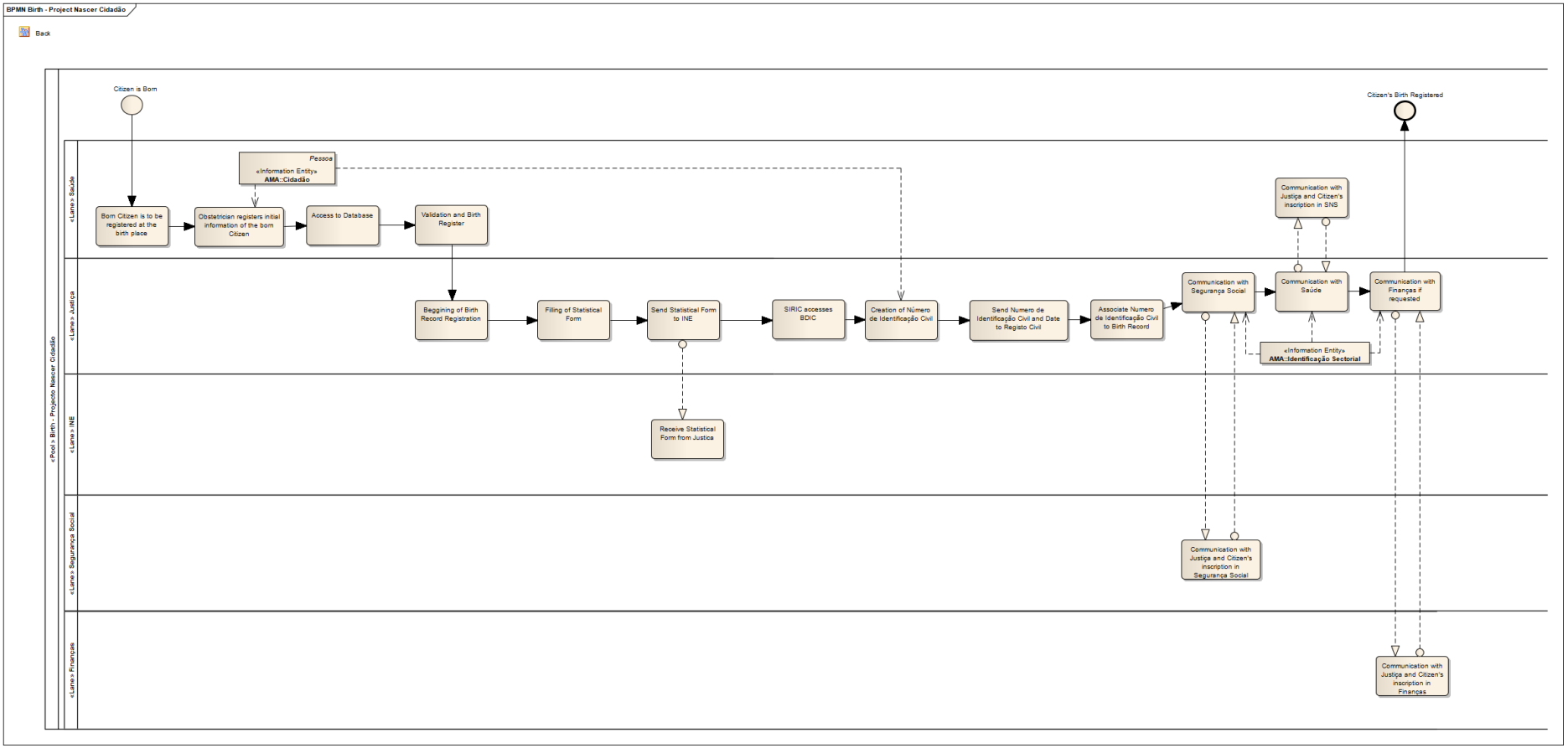


Figure 21: Representation of *Projecto Nascer Cidadão* main activities

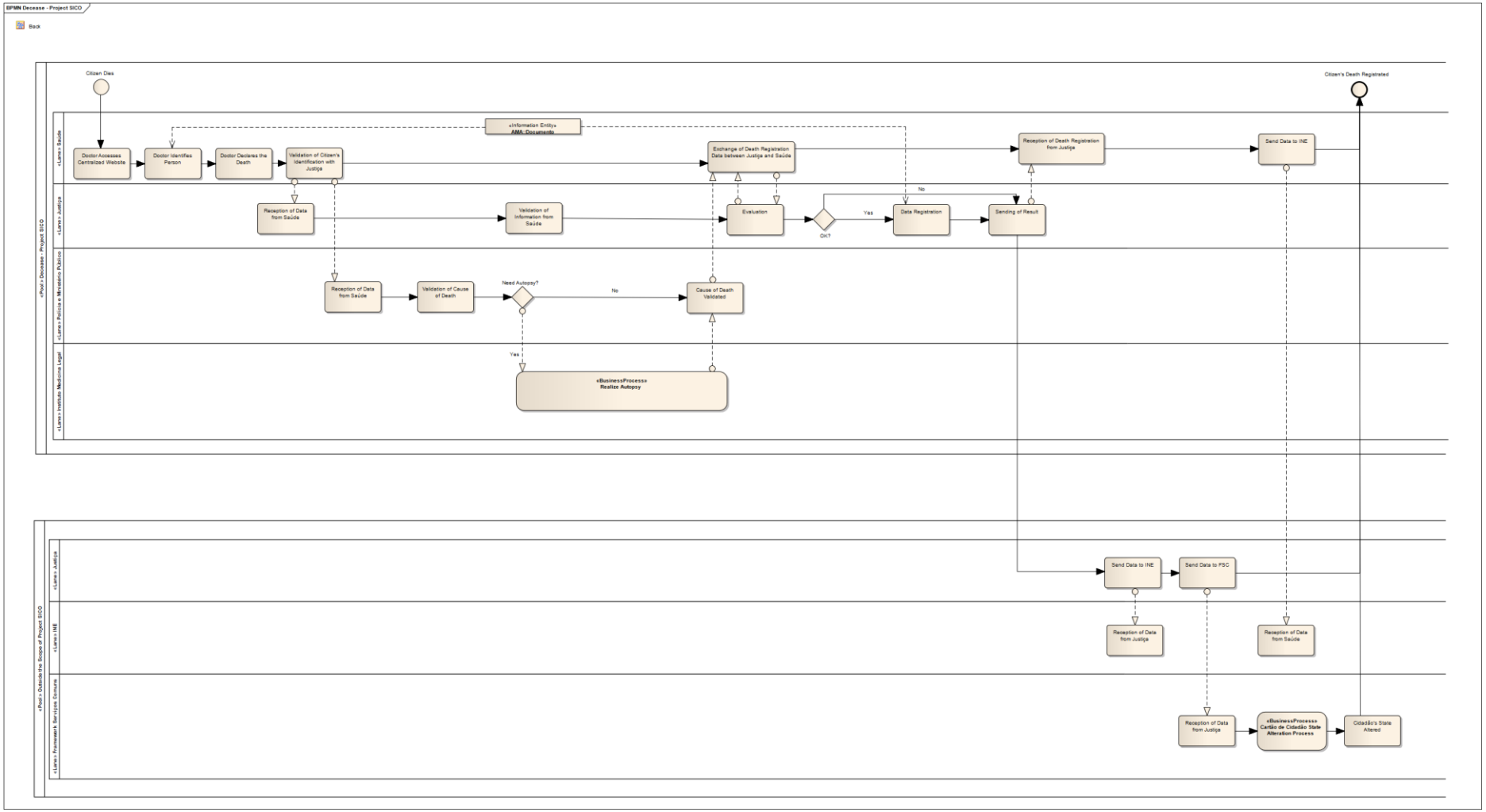


Figure 22: Representation of *Projecto SICO* main activities

Activities / Attributes	Obstetrician registers initial information of the born Citizen	Access to Database / Validation and Birth Register	Filling of Statistical Form	Receive Statistical Form from Saúde	SIRIC accesses BDIC	Creation of Número de Identificação Civil	Associate Numero de Identificação Civil to Birth Record	Communication with Justiça and Citizen's inscription in Segurança Social	Communication with Justiça and Citizen's inscription in SNS	Communication with Justiça and Citizen's inscription in Finanças
Born Citizen's Date and Time of Birth	CRU	R	R	R	R	RU	RU	RU	RU	RU
Mother's Name	CRU	R			R	RU		RU	RU	RU
Born Citizen's Gender	CRU	R	R	R	R	RU		RU	RU	RU
Born Citizen's Name					R	CRU		RU	RU	RU
Father's Name					R	CRU		RU	RU	RU
Mother's Place of Birth			R	R	R	CRU				
Mother's Address			R	R						
Mother's Marital Status			R	R	R					
Father's Marital Status			R	R	R					
Grandmother's Name										
Grandfather's Name										
Born Citizen's NIC						CRU	RU	RU	RU	RU

Born Citizen's Nationality			R	R	R	CRU		RU	RU	RU
Born Citizen's Place of Birth			R	R	R	CRU				
Born Citizen's Marital Status			R	R	R	CRU				
Birth Record Number						CRU	RU			
Birth Record Year						CRU	RU			
Birth Record Conservatory						CRU	RU			
Born Citizen NISS								CRU		
Mother's NISS								RU		
Born Citizen NSNS									CRU	
Mother's NSNS									RU	
Born Citizen NIF										CRU
Mother's NIF										RU

Table 1 : CRUD matrix for *Project Nascer Cidadão*'s activities

Activities / Attributes	Doctor Identifies Person	Validation of Citizen's Identification with Justiça	Validation of Information from Saúde	Exchange of Death Registration Data between Justiça and Saúde	Evaluation	Data Registration	Reception of Death Registration from Justiça	Send Data to INE	Send Data to FSC	Cartão de Cidadão's State Altered
Medical Report	CRU	RU	RU	RU	RU	R				
Death Declaration						CRU	RU	R	R	
Citizen's Name	R	R	R	R	R	R				RU
Citizen's Mother Name	R	R	R	R	R	R				RU
Citizen's Father Name	R	R	R	R	R	R				RU
Citizen's Place of Birth	R	R	R	R	R	R		R		RU
Citizen's Birth Date	R	R	R	R	R	R		R		RU
Citizen's Nationality	R	R	R	R	R	R		R		RU
Citizen's Marital Status	R	R	R	R	R	R		R		RU
Citizen's Gender	R	R	R	R	R	R		R		RU
NIC	R	R	R	R	R	R			R	RU
NIF	R				R	R				RU
NISS	R				R	R				RU
NSNS	R				R	R				RU

Table 2 : CRUD matrix for *Project SICO*'s activities

Information Managed within analyzed Projects	Attributes from AMA's Proposed Architecture
Born Citizen's Date and Time of Birth / Citizen's Birth Date	DataNascimento
Mother's Name / Citizen's Mother Name	NomeProprio, Apelido
Born Citizen's Gender / Citizen's Gender	Sexo
Born Citizen's Name / Citizen's Name	NomeProprio, Apelido
Father's Name / Citizen's Father Name	NomeProprio, Apelido
Mother's Place of Birth	Naturalidade
Mother's Address	Morada
Mother's Marital Status	EstadoCivil
Father's Marital Status	EstadoCivil
Grandmother's Name	NomeProprio, Apelido
Grandfather's Name	NomeProprio, Apelido
Born Citizen's NIC / NIC	NI
Born Citizen's Nationality / Citizen's Nationality	Nacionalidade
Born Citizen's Place of Birth / Citizen's Place of Birth	Naturalidade
Born Citizen's Marital Status / Citizen's Marital Status	EstadoCIvil
Birth Record Number	Documento.Dados
Birth Record Year	Documento.Dados
Birth Record Conservatory	Documento.Dados
Born Citizen NISS / NISS	NI
Mother's NISS	NI
Born Citizen NSNS / NSNS	NI
Mother's NSNS	NI
Born Citizen NIF / NIF	NI
Mother's NIF	NI
Medical Report	Documento
Death Declaration	Documento

Table 3: Relation between information in projects and AMA's architecture attributes

The results produced under this step of our work, suggest that the proposed information architecture is aligned with the two projects studied. By having a correlation between the entire set of information elements gathered, and the attributes present in the proposed information architecture, we can perceive that the last one is transversal to the universe in the projects analyzed.

Let us remind that it was not possible to exhaustively collect the information needed to ensure the completeness of the practical case analysis, due to confidentiality issues, but the sample we have analyzed supports the work present in the current version of the proposed information architecture.

5.3 Mappings

At this point of our work, while we already had modeled the different context-specific information entities using the Framework CEO and Enterprise Architect, this method of representation made it hard to execute the work we intended in terms of mappings.

To transform the information we already had into a format that allowed us to easier establish the relations between AMA's proposed architecture and each of the contexts under analysis, we chose to represent each of the contexts using XML. However, instead of just represent the entities by themselves, as we had information about the interactions existent between each of the contexts, using *Plataforma de Serviços Comuns* of the public administration, we chose to represent each of the contexts using roles. Therefore, for each of the contexts under analysis we identified the roles existent, elaborated a XSD containing that information, and represented the attributes used on each role. Appendix IV contains an example of a XML, a XSD, and the internal mapping diagram representation of BizTalk for one of the analyzed contexts. In concrete terms, figure 23 illustrates the obtained result with the definition of the XSDs and the mapping of the attributes between each of the specific-context roles and the AMA's proposed information architecture.

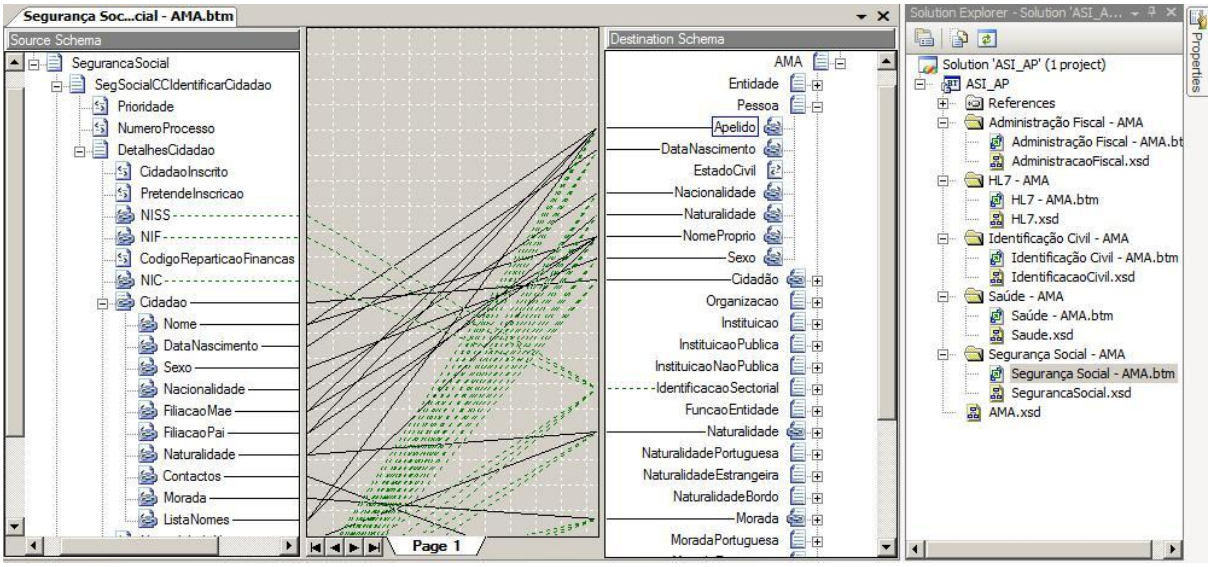


Figure 23: Representation of mappings in BizTalk

After having realized this work, our knowledge about what were the most recurrent attributes without mapping was very present, and we were in position to perceive what were the set of attributes that should be added to the current version of the information architecture developed by AMA, in order to increase its alignment with the specific contexts, that it intends to represent in a transversal manner.

Following this work around mappings, we picked up on our representation of the contexts *As-Is* and using the roles and relations established in BizTalk, complemented our representation with the relations between each entity of the specific context and the entities in AMA's proposed architecture.

The Appendix V contains the results of the interaction between each context analyzed and AMA's proposed architecture representation using roles, and the Appendix VII summarizes the results of our analysis after completing the mappings of each of the contexts, to the architecture proposed by AMA. The global appreciation of the work results achieved during this phase is illustrated in table 4.

Statistics from Mappings

Overall Unmapped Attributes
256

Overall Attributes
546

≈ Overall % Unmapped Attributes
46,88644689

≈ Overall % Mapped Attributes
53,11355311

Table 4: Result of mapping analysis in BizTalk

Drilling into the statistics presented, we were also able to identify what was the group of unmapped attributes that repeated more often (fig. 24).

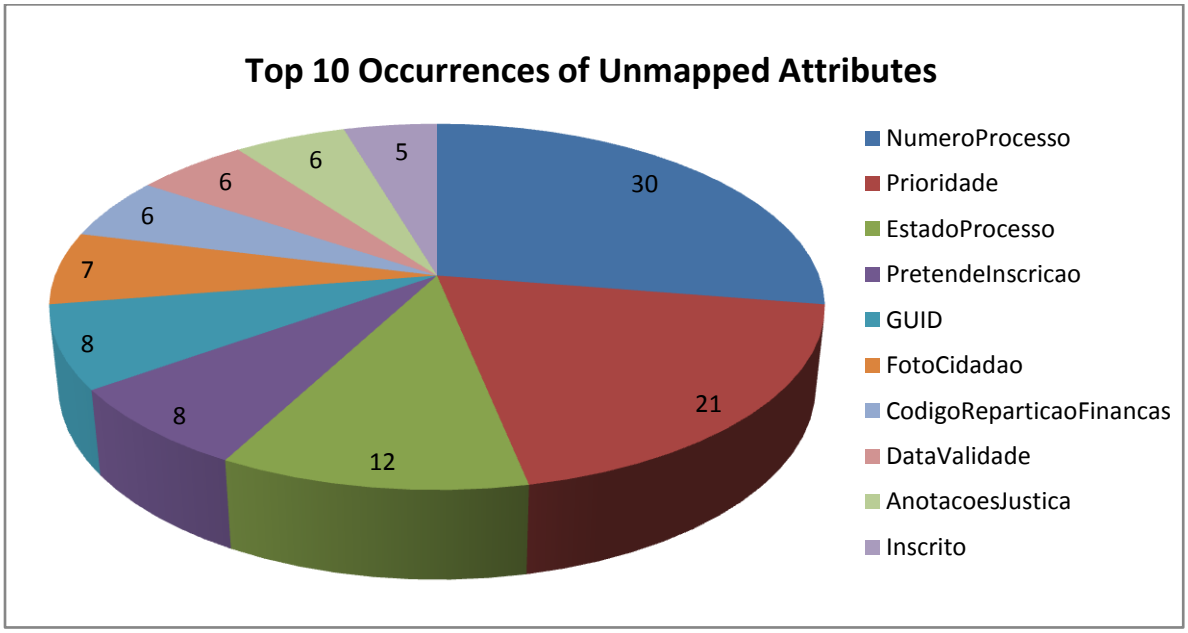


Figure 24: Top 10 Unmapped Attributes

Although being important to find what are the attributes that appear most commonly through the different contexts, since they will be our base for improvement, the fact is that this data just by itself cannot form a solid base for arguing that a specific attribute is transversal to the public administration, and therefore should be added to the existent version of the architecture. To mitigate that problem, we also analyzed the attributes from the perspective of their existence among the different contexts. The table 5 demonstrates the result achieved with the analysis on the dispersion of the top ten unmapped attributes among the contexts analyzed.

Unmapped Attributes	# Occurrences	Contexts						# Contexts Present
		Administração Fiscal	Cartão de Cidadão	HL7	Identificação Civil	Saúde	Segurança Social	
NumeroProcesso	30	X			X	X	X	4
Prioridade	21	X			X	X	X	4
EstadoProcesso	12	X			X	X	X	4
PretendeInscricao	8	X				X	X	3
GUID	8				X			1
FotoCidadao	7				X			1
CodigoReparticaoFinancas	6	X					X	2
DataValidade	6		X		X			2
AnotacoesJustica	6				X			1
Inscrito	5	X				X	X	3

Table 5: Presence of unmapped attributes in analyzed contexts

Just by focusing on the ranking of the unmapped attributes, we might be taken into consider the inclusion of the, *e.g.*, top six attributes, however, if we also take into consideration their dispersion between the different contexts, we perceive that the attribute *GUID* and *FotoCidadao* should not qualify as they only have presence in one of six contexts. In practical terms, what we can infer is that most likely this attributes are specific of a certain context, and therefore not transversal to the public administration, reason why they should not be included into the revised version of the architecture.

In next section we unite the data gathered from this analysis, the best practices from HL7, and the results of our analysis of the two projects related to this work, and propose a revised version of the actual architecture, with the inclusion of new attributes that increase its capacity for being transversal to the public administration.

5.4 Information Architecture To-Be

This section contains the major output of this entire work, *i.e.*, the proposed information architecture. The intermediate results from each step of our methodology are therefore combined, making possible the achievement of the proposition of the new architecture.

We start by identifying the set of alterations proposed to the current version of the architecture, then we present the revised version of the architecture taking into consideration our proposals, and finally we make some considerations about them.

5.4.1 Proposed Changes

In order to assess the list of modifications proposed to the actual version of the information architecture, we must first consider the methodology our work was based on, the type of work developed, the intermediate results achieved, and the context of the work.

Hence, and since this work is focused on the citizen's perspective, the first consideration that should be made is that the changes will specially target the entities *Pessoa* and *Cidadão*. However not restricted to them, due to the type of work performed, it was unlikely to have obtained relevant data to justify modifications to entities like *Organização* or *Instituição*. Another relevant fact is that, since we do not have performed specific ontological work, we will most likely not be proposing major structural changes, *i.e.*, we will maintain the entities already defined, and the coherence of the relations between them, not arguing if, *e.g.*, the entity *Morada* should be an attribute of the entity *Pessoa*, instead of being an information entity. What we are looking is to proposed changes in terms of attributes that should be added, removed, or moved, in order to maximize the capability of the proposed architecture to be transversal.

Considering the work realized until now and the methodology proposed, there are three key factors that we should take into consideration:

1. What are the highest ranked attributes with presence in the contexts analyzed and not in the information architecture?
2. What was the information present in the projects analyzed that did not have mapping to the attributes in the information architecture?
3. What is the contribution of the HL7?

Evaluating the results achieved from the data relevant to the first point, and considering the results from table 5, there are ten attributes noticed to be recurrently identified as unmapped. The reason why we chose these ten relates to the fact that, from there onwards, the hypothetical candidates begun to show very small number of unmapped occurrences, as show on Appendix VIII. These ten attributes represent the best compromise between the number of contexts analyzed and the total number of attributes existent. The small number o unmapped occurrences of other attributes, suggest that they are most likely specific to a context, rather than transversal. However, as we did not achieve the required grade of completeness, to have quantitative metrics to support this arguing, focused on our scope is reasonably to consider the set of these ten attributes. However, we cannot go point blank adding the ten attributes to the existent version of the architecture. If we analyze table 5, we perceive that there are a subset of attributes that, although having a substantial occurrence count, do so in only one or two contexts. This once again suggests that they are specific to that context and should not be considered for this selection. Since our number of presences in a context varies from one to four, for the reasons previously stated, we will pick on the top bottom half, *i.e.*, attributes with high unmapped occurrences and with presence in three or four contexts.

By doing so, we achieve the set of attributes that should be added to the architecture, based on the work developed on the analysis of specific contexts:

- NumeroProcesso, Prioridade, EstadoProcesso, PretendeInscricao and Inscrito

Considering the results displayed in section 5.2, from the analysis of two projects specifically related to the concept of citizen, the data gathered indicates that there is a full alignment between the information circulating in the processes of the two projects, and the reference information architecture. From here, we can perceive that the attributes already under the entities related to these projects should not be removed. Unfortunately, with the data gathered it is not possible to make other considerations, and propose changes to the current version of the architecture. However, we should not forget that, in a work at this level, there is always a need for some degree of validation of the work being developed, with real cases. Since we are using the actual version of the architecture as foundation, what we have done is guarantee that our new version will also have this grade of alignment.

Finally, considering the special characteristics of the HL7, we will consider its contributions on two different perspectives. For once, similar to what we have been doing, we will consider importing some attributes from HL7 to our architecture, and on the other hand we will use it to help us reviewing and categorizing the distribution of the attributes, on the current version of the architecture, and at last to answer so conceptual and ontological questions that were risen during the entire execution of this work.

Starting by the ontological questions, next we will formulate a series of questions, answering with the support of HL7:

- How can the current version of the architecture support the information about a person/citizen with a disability, which is relevant for contexts like *Segurança Social*?
    - We can introduce the attribute *disability code*, associated with the entity *person.*
- How can the current version of the architecture support the fact that a person is dead?
    - As a person is different from a citizen, the state alteration in the documents that identify a citizen should not ontologically identify the fact that a person is dead, therefore, a *deceased indicator* should be added to the entity *person.*
- In this last case, a citizen/person cannot simply disappear from the system, because its current status and time of dead is relevant for other entities. How can the architecture information support the time at which a citizen/person died?
    - There is no attribute defined to represent the time of death, then, an attribute like *deceased Time* should be added.
- In case of a woman being pregnant with more the one fetus, and a medical or other type of intervention is needed, how can the architecture support the individual identification of one fetus?
    - Since the fetus is not yet considered a citizen, at first glance, the architecture does not provide support for such situation. Based on HL7, an attribute *multiple birth order* should be added to the entity *person.*

Since we have already defined the affected entity by the attributes from the HL7, we must now also associate the five identified attributes, from the analysis of specific contexts, to information entities. Considering the five attributes identified, we are targeting two different entities. As the first

three relate to a process initiated by a person or citizen, they should be integrated into the entity *EventoContacto*. However, this entity already possesses an attribute capable of representing the state of a process, which we were not able to guarantee that represents the same information as we want to express with *EstadoProcesso*. Regarding the objectives of our work, we will alter the current name of the attribute from *Estado* to *EstadoProcesso*, to reflect our modifications, but if turns out that the two attributes do not represent the same information, an alternative representation should be created. The other two attributes should be directly added. With the purpose of maintaining the alignment and consistency with the work already developed by AMA, we will represent the attributes from HL7 in Portuguese, using an appropriate translation. The remaining attributes should be added to the entity *Cidadao*. From an ontological perspective, they represent the act of a citizen requesting the registration in one system, or the fact that it is already registered.

These decisions are the result of all the work and research done until this point, and the choices made take into consideration the description that AMA has provided for each information entity in its proposed architecture. The results of this analysis and the record of modifications to be made, are represented in Table 6.

Attribute	Type of Modification	Destination Information Entity
Numero Processo	Add attribute	Evento Contacto
Prioridade	Add attribute	Evento Contacto
Estado Processo	Add attribute	Evento Contacto
Pretende Inscricao	Add attribute	Cidadao
Inscrito	Add attribute	Cidadao
Codigo Deficiencia	Add attribute	Pessoa
Indicador Obito	Add attribute	Pessoa
Data Falecimento	Add attribute	Pessoa
Indicador Ordem Gravidez Multipla	Add attribute	Pessoa

Table 6 : Record of modifications proposed to the architecture

Next, we present the revised version of the architecture, including the proposed modifications.

5.4.2 Proposed Information Architecture

This section presents the final result of our work. The revised version of the architecture, with the set of changes proposed is represented on figure 25, using the framework CEO.

The Appendix IX provides another viewpoint of the proposed architecture, where the entities affected with our proposals are highlighted in green.

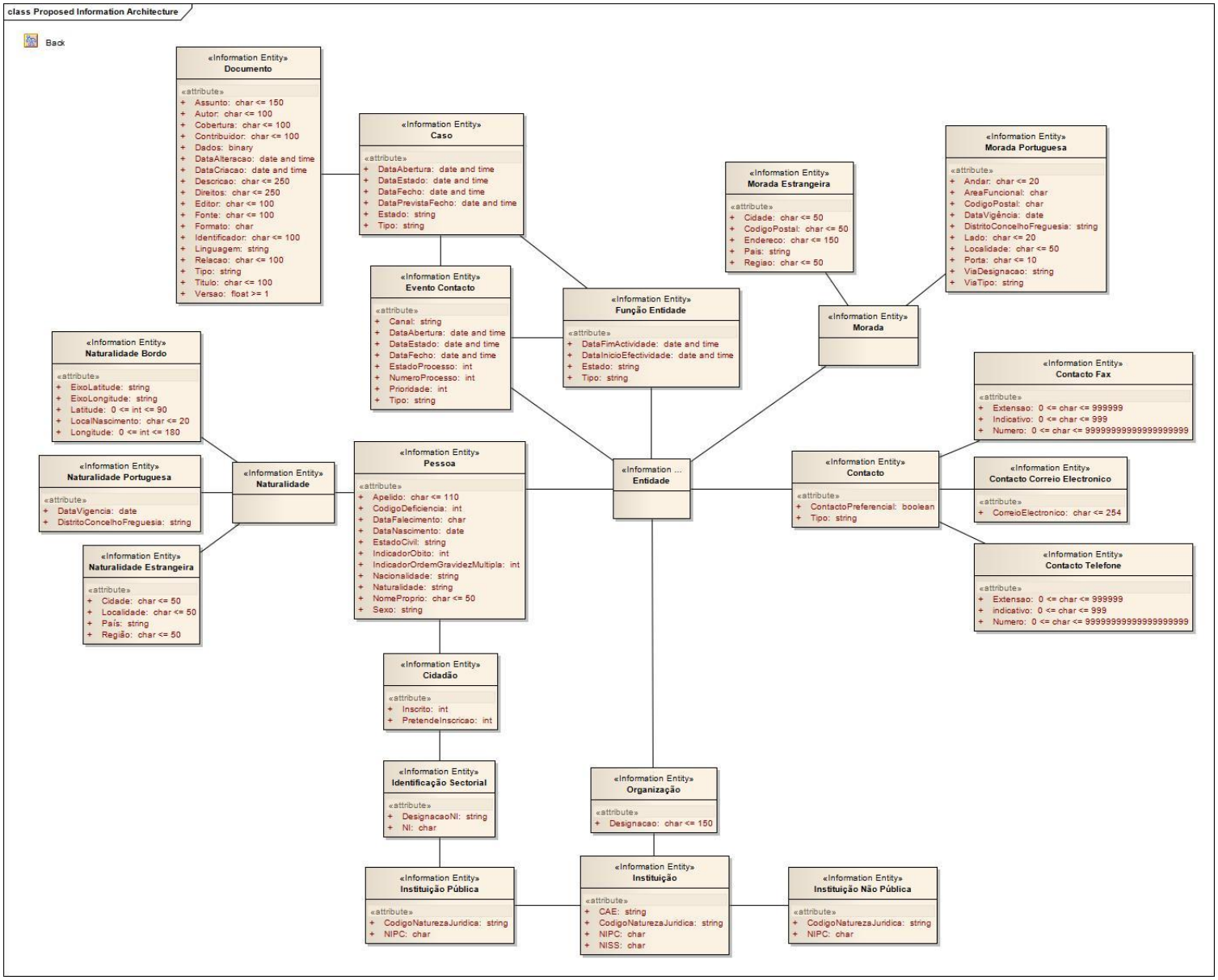


Figure 25: Representation of the proposed information architecture

5.4.3 Impact of the Changes Proposed

In this section we evaluate the impact on the alignment of our version of the information architecture against the one currently available.

We will use as base for our analysis the statistics produced during our step of context analysis, and compare them to a set of new ones produced using as basis our proposed information architecture. Although this can provide us a quantitative metric, a qualitative set of considerations should also be presented. In fact, we cannot correctly assess the impact that the modifications based on HL7 will have, as we do not have data that supports that analysis. None the less, there are some considerations that should be made, and therefore are presented ahead.

Referring to the results displayed in table 4 and figure 24, consider that the percentage of unmapped attributes is given by using the formula:

$$\frac{\#\ \text{Overall Unmapped Attributes}}{\#\ \text{Overall Attributes}} \times 100$$

We can then recalculate the data using our version of the architecture as basis, by updating the values implicated in the formula. Since our architecture contains modifications towards mitigate the existence of the top unmapped attributes, in order to obtain the updated number of *overall unmapped attributes* we can subtract the occurrences related to the attributes now added to our version of the architecture, from the value obtained before. Hence, considering the values on table 4 and table 5, the updated value is given by:

$$\begin{aligned} \#\ Overall\ Unmapped\ Attributes \\ &= Previous\ \#\ Overall\ Unmapped\ Attributes \\ &- Unmapped\ Ocurrences\ of\ (NumeroProcesso + Prioridade + EstadoProcesso \\ &+ PretendeInscricao + Inscrito) \end{aligned}$$

$$\#\ Overall\ Unmapped\ Attributes = 256 - (30 + 21 + 12 + 8 + 5) = 180$$

As the value for # *Overall Attributes* has not been altered, we can then apply the formula and obtain the new percentage of unmapped attributes:

$$\%\ Unmapped\ Attributes = \frac{182}{546} \times 100$$

$$\%\ Unmapped\ Attributes \sim 33{,}3\ \%$$

By comparing the previous statistics with the current ones, we obtain the results presented in table 7.

	Statistics from AMA's Proposed Architecture	Statistics from our Proposed Architecture	Alignment Improvement
# Overall Unmapped Attributes	256	**182**	**13,6 %**
# Overall Attributes	546	546	
Overall % Unmapped Attributes	46,9	**33,3**	
Overall % Mapped Attributes	53,1	**66,7**	

Table 7 : Statistics comparison between previous and revised architecture

Attending to the results achieved, we are looking at an alignment improvement of 13,6% over the AMA's version of the architecture. This set of data proves that our proposal increases the percentage of mapped attributes, providing an higher rate of alignment between the information architecture and the contexts and processes analyzed.

Hence, the answer to question Q1 is the set of information entities and respective attributes presented earlier. For being capable of effectively providing an answer to the main question underlying this work, and having produced results that prove an higher rate of alignment of the information architecture with the public administration context, we consider that our main work objective was achieved with success.

Although not being able to effectively measure the grade of contribution from the work developed around the ontological field with HL7, we believe that the fact that the added attributes represent a solution for a number of raised questions, its contributions are noticeable and have a structuring role in the continual improvement of our proposed architecture. Despite not having developed ontological work with an higher degree on impact, it is still our understanding that the contribution given by this context, was successful on including a new perspective into the work developed until now. For that, we consider that we were able to achieve the specific objectives defined for this subject.

6 Global Critical Analysis

In this section, now that we have described the steps taken that lead to the accomplishment of our objectives and the results obtained, we adopt a critic sense and look back into the work developed in order to make an appreciation of what has been done and could have been done differently.

From a global perspective it is our opinion that considering the objectives of the work, the means at our disposal, and the time frame available, we positively contributed to improve the existent version of the information architecture for the public administration, while also performing work that will enrich the academic universe, and can serve as basis for future research on similar subjects.

During our phase of investigation towards finding related work in the field, despite having found useful sources, we noticed that there was a gap about work on these matters, in Portuguese organizations. It is therefore with pleasure that we see the work we develop became part of the academic community, with the possibility of serving as future reference for another researchers.

On the following sub-sections, we divide our global critical analysis between what we consider to be our main and accessory contributions, and the limitations of the proposed architecture, as well as the rest of the work developed.

6.1 Main Contributions

Concretely focusing on the work developed, the research performed in terms of related work, and the fact that we had the support of AMA and its work, was a major key factor into achieving the results we had. By already having a draft version of a proposition for an information architecture, we were able to confine the scope of our work to improve part of it, making it more viable that attempting to develop an information architecture from ground zero. It surely had been impossible to get to know all the public departments implicated on the design of the architecture, and produce a coherent and viable solution within the time available. With the methodology we devised and implemented, we were able to perform a full cycle of the investigation methodology, from the data collection, to the solution, passing by the analysis and implementation. Each of the steps allowed us to perceive and strict the relations between the different contexts and entities, allowing us to always have a back work developed that supported the new step we were trying to execute, *e.g.*, describe the relations between the entities from AMA's proposed architecture and a specific context using roles, while already having the specific contexts modeled.

The results obtained reflect the changes that, based on our research and analysis, improve the alignment of the reference architecture with each of the contexts analyzed. Given the type of work performed, the granularity of the changes and improvements proposed is aligned with our objectives.

Considering the questions rose in section 1.2.1, the work developed addressed them in a transversal manner, and the answers are not confined to a specific section. Instead, it was by developing the answer to Q1.1 that we were able to answer Q1 as well. The answer to Q1.1 is mainly answered in section 4, where we propose a methodology to realize this type of work, and the validation of the proposal was achieved by realizing this work, proving that it was indeed possible to find a transversal model from the analysis of specific contexts. The answer to Q1 was achieved through the development of this entire work, and the concrete result presented in section 5.4.

From our perspective there are two main contributions provided by the work developed, namely the information architecture and the proposed methodology. In terms of the proposed information architecture, it is the result of the entire set of activities performed during this work, and it is the direct answer to Q1. Overall, not only it behaves like a main contribution to the academic field, as it also has implications on the accessory contributions in the enterprise field. None the less, it also has implications in the answers to question Q1.1 through Q1.5, specifically by representing the results achieved with the manipulation of the information entities attributes. The methodology proposed behaves like the other main contribution, particularly to the academic field. The methodology, despite

being implicated in the answers/contributions to questions Q1.1 through Q1.5, specifically targets the Q1.2.

6.2 Accessory Contributions

With the work we developed we could improve the actual vision that this version of the architecture provides over the concept of citizen and person, and increase its alignment with the reality in the public administration.

The validation of the results was an iterative process, since for each new context analyzed we were verifying the alignment with the proposed architecture. Not only did this provided us with a reasonable appreciation of how well had we achieved our objectives, but AMA will also validate this work as stated in section 1.2.4.

An accessory contribution of the results of this work is the impact on AMA. As we were integrated in the project of the development of the architecture, and the project is not yet finalized, the input of our proposals will be considered and validated by AMA, providing another perspective over the work developed until now. Therefore, our accessory contributions are focused on the enterprise field. None the less, the entire set of activities performed during the execution on this work, were able to help answering the questions Q1 and Q1.1 through Q1.5. The steps within our methodology that comprehend the manipulation of the entities' attributes, the establishing of relations, and the execution of the practical case, helped answering the remaining raised questions and provided contributions towards that end.

If we focus our main contributions on the proposed architecture and methodology, the representation chosen for the architecture, the work developed during the *context* and *analysis* stages, the mappings, and the practical case and CRUD matrixes, qualify as accessory contributions to the field targeted by this work.

6.3 Limitations of the Proposed Architecture

Had we followed the path were we had analyzed the problem from an ontological perspective, the impact of the changes suggested would have been certainly greater, as we would have obtained results that would justify the creation of new entities or the alteration of the relations between them.

However, pursuing the idea of realizing this work strictly focused on the ontological perspective would have triggered a series of problems. On a first note, for representing the results achieved from this type of work we would have used DEMO and implying that we would have needed to make an extensive research around that subject, and learned enough to be able to use it accordingly. From our entire section of related work, DEMO was the only one about which we had no previous knowledge, suggesting that the learning curve would not have benefited us, and we still had the risk of not being able to apply it directly to our work. None the less, this type of approach would have required us to develop physical work in all the contexts analyzed, since it would be an entire new type of work, that would have triggered a set of bureaucratic questions, namely access issues, the need of establishing written agreements, the availability from the part of the organizations, and

another rather large issue that relied on our capability of precisely understand the way that each department, from each context, works and perceives the concept of citizen. By balancing the time and risk of overcoming all these issues, we came to an agreement of not strictly focus on this perspective, under the premises of compromising the completion of the work in due time.

However, a note should be made about the adaptability of this architecture. As stated before, information architectures should be as independent as possible, meaning that they are the least probable item on the subject of enterprise architectures to go under major alterations across the time. Being aware of that, we used the best practices from HL7 because it already provides this sort of vision over this subject. As we were focused on the citizen, we mitigated that problem by taking into consideration the HL7 as another context for analysis. Never the less, as we describe in section 7, for further improve the architecture and extend it to other perspectives, the conceptual and ontological perspectives should become a major priority, and play important roles in each of the new iterations.

7 Future Work

In order to provide a clear vision for what the future can and should be, we must make some considerations first. Taking into account the scope of this work, *i.e.*, public administration, as well as the time and effort it would require to develop a solution that would cover it entirely, one of the first measures we had to take was to restrict the amount of data to be analyzed and processed. The restriction was applied not only to the set of contexts analyzed but also, in a higher level to the perspective in which the work was focused. For instance, while this work is focused on the citizen, there are other perspectives that should be analyzed, *e.g.*, companies. Based on the methodology we presented, with direct relation to the percentage of the entire universe of the public administration analyzed, the more of this universe is analyzed, the likely it is to obtain an architecture closer to a final and stable version. As stated earlier, this work is focused on the perspective of the citizen and even so it is not complete, as the contexts analyzed and the two projects chosen for the practical case, provide mainly insight in the act of being born and dying. Relating to our future work, and enrolling other departments in the process, a necessary step towards the completion of the citizen's perspective analysis is also analyze contexts related to, *e.g.*, the act of getting married, buying a car, and buying a house. However, shifting towards another perspective, the architecture should be complemented with other work. The work that we have done for the concept of a citizen should be extended to the analysis of the lifecycle of, *e.g.*, a company, a building, a document, and a vehicle. For each of this perspectives a conceptual work should be done, towards defining what are the most important activities in their lifecycles, to what departments of the public administration are they related, who/what else intervenes and its relevant for their lifecycle, and which is their transversal relation to the rest of the information entities present in the public administration.

As important as this work, closer to the as-is, proves itself to be towards obtaining a final version of the architecture, another step should be taken while developing future work. As we have seen before, the information should be considered a first class citizen and in the context of the theme enterprise architectures, it should also be the more stable and less likely to change. As we have also

seen, this comes from the fact that the information architecture should ideally be independent from the technology and installed processes. Since previously we suggested that more field work should be performed, we can rapidly see that, as much universe as we cover, we could not approach an ideal solution without developing conceptual and ontological work around the subject. Referring to our related work, it is not by mere chance that, even not using them, we included Archimate and DEMO. An innovative and important approach to this work would be to have a dedicated group conducting more business-layer oriented work, identifying processes, and working on the development of a ontological view over the different concepts. After that, already having concluded the field and as-is work, another group could enroll in the process of integrating the entire set of results obtained, and achieve a solution which includes both sides. In this case, the focus would not be only on the part of information entities, but we would have the capacity of representing the public administration using an integrated, from business to technology, framework. Here, the CEO framework could still be used or another approach could be followed that could culminate in the use of Archimate. Nevertheless, identifying these opportunities for future work is important and should be considered.

After having related and contextualized the opportunities for future work surrounding the theme, they can be listed as follows:

- Analyze other key actions in the lifecycle of a citizen, continuously upgrading the work already developed
- Identify the set of other key concepts to include in the architecture besides citizen, *e.g.*, company, building
- Analyze these key concepts in terms of their lifecycle, and gather the set of main actions that execute in their context, during their lifecycle
- Identify the transversal information entities being manipulated during the concretization of each one of the actions, and cross-match them with the existent version of the architecture
- Develop an ontological and conceptual perspective around this entire theme, using DEMO
- Consider the entire scope of the CEO framework or Archimate, and after having accomplished all the previous tasks, aim for an integrated representation of the information architecture, inside the greater context of the enterprise architecture

8 Conclusion

The Portuguese public administration is being target of a modernization initiative which, among other aspects, contemplates the development of an information architecture. This action aims at addressing the incompatibilities existent between different information systems, on different departments, and their inefficiency in terms of dealing with information. Despite some initiatives already taking place, *e.g.*, the interoperability platform which starts aiming towards a solution for this problem, there is a need for developing an information architecture, so that can be perceived which are the entities and departments enrolled in the manipulation of the information by the major existent processes, and therefore improve the work efficiency as well as the experience perceived by a citizen when triggering one of them.

With this work, we began by clarifying the problem and establish the motivation towards the solution. Then, we researched on the theme of the information architecture, and the state of the art languages and frameworks that allowed us to develop a coherent work. We researched on the scientific investigation methodologies that we used to support our work on scientific basis, and therefore guarantee its validity and continuity. The action research methodology has proven to be useful since an early stage of this work, and the solution's architecture is also aligned with its steps. After having a sense of perspective into the context of the work and the approaches to be taken, we devised a methodology for working towards a solution, with support on activities and objectives defined for each step, as well as on the Spewak's Enterprise Architecture Planning methodology which fits our necessities for the current work.

By using the knowledge gathered under this phase, and by applying it into the further development of the work, we were able to analyze and process the different data gathered and ultimately provide a revised version of the existent architecture, with improvements based on the entire work we developed.

Finally, and after achieving our objectives, we adopted a critical perspective and made considerations about the work developed and eventual opportunities of improvement in future work around the subject.

The subject and context of this work, as well as the people with whom we worked, were a sincere motivation towards its realization and completion. None the less, the fact that we integrated a project of significant dimension, already on the move, with multi-disciplinary teams and organizations involved, and for which there is a real need, put us in a position where we had the means and level of participation of the different intervenient that allowed us to perform the planned activities and with relative flexibility. Sometimes the elevated dimension of the project also proved to be a problem, especially while shifting between organizations. As we know, with sensitive materials there is always a pipeline of approvals that need to be granted and all that consumes a valuable resource: time.

We always had the notion that this was an ambitious work, and that the time range available would not suffice to approach it from all the perspectives we intended. We do not hide the fact that we would have preferred to develop more ontological and conceptual work, as suggested in the previous section, but the fact is that probably we would not have the means and support needed to develop something concrete and ready for presentation, while in this case we were capable of developing relevant work for both the academic community and AMA, for who we hope that our contributions prove useful.

Finally, and on that note, we would like to leave the challenge open to anyone who continues working on this theme, to attempt introducing this ontological perspective on it. We are positive that its contribution to the related work on the area would be enormous, as well as a very differentiating factor and solid base to develop similar future work.

9 References

1. Watson, Richard W.: An Enterprise Information Architecture: A Case Study for Decentralized Organizations. IEEE, New York (2000)
2. Nandi, Madhavi L., Gopal, K. Nayak: Information Systems Management in Public Sector Organizations. In: International Conference on Information Technology. IEEE, New York (2008)
3. Diario da República - 2 de Outubro de 2009, Série.1, N.192 (2009)
4. Lankhorst, M., et al.: Enterprise Architecture at Work, pp. 1—10. Springer, Heidelberg, ISBN 3540243712 (2005)
5. Amaral, Diogo Freitas do: Curso de Direito Administrativo (Vol. I), pp. 49. Almedina, Coimbra (1986)
6. Caetano, Marcelo: Manual de Direito Administrativo (Vol. I), pp. 5. Almedina, Coimbra (1991)
7. Olivé, A., Pastor, Joan A.: Advanced Information Systems Engineering, 9th International Conference, CAiSE'97, Barcelona, Catalonia, Spain, June 16-20, Proceedings (Lecture Notes in Computer Science). Springer, Heidelberg, ISBN 3-540-63107-0, (1997)
8. Spewak, S. H., Hill, S. C.: Enterprise Architecture Planning: developing a blueprint for data, applications, and technology. John Wiley & Sons, New York, ISBN 0471599859 (1992)
9. Chew, K., Gottschalk, P.: Information Technology Strategy and Management: Best Practices. Information Science Reference, ISBN 1599048027, (2008)
10. Zachman, J.: John Zachman´s Concise Definition of the Zachman Framework. Zachman International, USA (2008)
11. White, Stephan A.: Introduction to BPMN. IBM Corporation (2004)
12. Business Process Model and Notation (BPMN), Object Management Group, Version 1.2 (2009)
13. Mallens, P., Dietz, J., Hommes, B.: The Value of Business Process Modeling with DEMO Prior to Information Systems Modeling with UML. Faculty of Information Technology and Systems, Department of Information Systems and Software Engineering, University of Technology, The Netherlands (2009)
14. Dietz, J. L. G.: Introductie tot DEMO (1996)
15. Reijswoud, V. E., Dietz, J. L. G.: DEMO Modeling Handbook (1999)
16. Lankhorst, M., et al.: Enterprise Architecture at Work, pp. 86—113. Springer, Heidelberg, ISBN 3540243712 (2005)
17. Lankhorst, M., et al.: Enterprise Architecture at Work, pp. 84—85. Springer, Heidelberg, ISBN 3540243712 (2005)
18. Lankhorst, M., et al.: Enterprise Architecture at Work, pp. 191—198. Springer, Heidelberg, ISBN 3540243712 (2005)

19. Lankhorst, M., et al.: Enterprise Architecture at Work, Appendix A. Springer, Heidelberg, ISBN 3540243712 (2005)
20. Lankhorst, M., Drunen, Hans v.: Enterprise Architecture Development and Modeling, Combining TOGAF and ArchiMate. Via Nova Architectura (2007)
21. Vasconcelos, A., Caetano, A., Neves, J., Sinogas, P., Mendes, R., Tribolet, J.: A Framework for Modeling Strategy, Business Processes and Information Systems. IEEE, New York (2001)
22. Balanced Scorecard Institute, http://www.balancedscorecard.org/
23. Vasconcelos, A.: Arquitectura dos Sistemas de Informação: Representação e Avaliação. PhD Thesis, pp. 99—100, Instituto Superior Técnico (2007)
24. Fowler, Martin: UML Distilled (A Brief Guide To The Standard Object Modeling Language). Addison-Wesley, ISBN 0321193687 (2003)
25. Chen, Peter P.: The Entity Relationship Model - Towards a Unified View of Data. In: ACM Transactions on Database Systems, Vol. 1, No. 1, pp. 9—36. ACM (1976)
26. Vasconcelos, A.: Arquitectura dos Sistemas de Informação: Representação e Avaliação. PhD Thesis, pp. 124—125, Instituto Superior Técnico (2007)
27. Sousa, P., Caetano, A., Vasconcelos, A.: Aula 10-12 - Arquitectura de Sistemas de Informação, Arquitectura Processos e Ferramentas de Sistemas de Informação. Support Material, Instituto Superior Técnico (2009)
28. Immon, William H.: Data Architecture: The Information Paradigm. QED Information Sciences (1993)
29. Boyton, Andrew C., R. Zmud, G.Jacobs: The Influence of IT Management Practice on IT Use in Large Organizations, pp. 299—319. MIS Quarterly, V.18, N.3 (1994)
30. Cook, M.: Building Enterprise Information Architectures. Prentice Hall, New York (1996)
31. Vasconcelos, A., Silva, Miguel M. da, Fernandes, A., Tribolet, J.: An Information System Architectural Framework for Enterprise Application Integration. IEEE, New York (2004)
32. Linnell, J., Tucker, J., Linger, T., Smith, S., Alcorn, W.: Information Architecture For An Electronics Recycling National Coordinating Entity. IEEE, New York (2005)
33. AMA, I.P.: Arquitectura Informacional - Entidades e Relações, v0.9, Internal Documentation (2009)
34. AMA, I.P.: Projecto Modelos Interministeriais - Relatório de Funcionamento, Normas e Procedimentos, v1.0, Internal Documentation (2009)
35. AMA, I.P.: Projecto Modelos Interministeriais - Relatório de Governance e Organizacional para Grupos de Interoperabilidade e Grupos TIC, v1.0, Internal Documentation (2009)
36. Salomon, G.: Distributed Cognitions: Psychological and Educational Considerations (Learning in Doing – Social, Cognitive, and Computational Perspectives). Cambridge University Press (1996)
37. Vasconcelos, A.: Arquitectura dos Sistemas de Informação: Representação e Avaliação. PhD Thesis, Instituto Superior Técnico (2007)
38. Myers, M. D.: Qualitative Research in Information Systems, pp. 241—242. MIS Quarterly, V.21, N.2 (1997)

39. Bryman, A.: Research Methods and Organization Studies. Unwin Hyman (1989)
40. Baskerville, R., Wood-Harper, A. T.: Diversity in information systems action research methods. In: European Journal of Information Systems, V.7, pp. 90—107 (1098)
41. Checkland, P. B.: From framework through experience to learning: the Essential nature of action research, Nissen, H., Klein, H. K., Hirschheim, R. A. (eds.): Information Systems Research: Contemporary Approaches and Emergent Traditions, pp. 397—403. North-Holland, Amsterdam (1991)
42. Harold, Elliotte R., Means, W. Scott: XML in a Nutshell, O'Reilly Media, ISBN: 978-0-596-00764-5 (2004)
43. HL7 International: About HL7, http://www.hl7.org/about/index.cfm?ref=nav (acedido a 07-10-2010)
44. HL7 International: HL7 Standards, http://www.hl7.org/implement/standards/index.cfm?ref=nav (acedido a 07-10-2010)
45. Coller, L., Daoust, N.: HDF - HL7 Development Framework, http://hdf.wikispaces.com/Proj (acedido a 07-10-2010)
46. J. L. G. Dietz: Enterprise Ontology. Springer-Verlag, Berlin Heidelberg, 2006
47. J. A. P. Hoogervorst: Enterprise Architecture: Enabling Integration, Agility and Change. International Journal of Cooperative Information Systems, 13(3):213-233, 2004.
48. J. A. P. Hoogervorst: Enterprise Governance & Architectuur - Corporate, IT en enterprise governance in samenhangend perspectief. Sdu Uitgevers bv, Den Haag, 2007.
49. J. Dietz and J. Hoogervorst: Enterprise Ontology and Enterprise Architecture - how to let them evolve into effective complementary notions. GEAO Journal of Enterprise Architecture, 1, 2007.
50. M. Op 't Land: Applying Architecture and Ontology to the Splitting and Allying of Enterprises. PhD thesis, Delft University of Technology, Schildmos 13, 3994 LS Houten, Netherlands, June 2008.
51. A. Albani and J. Dietz: The benefit of enterprise ontology in identifying business components. In WCC '06: Proceedings of the IFIP World Computer Congress, Santiago de Chile, Chile, 2006.
52. G. Hardjosumarto: An Enterprise Ontology based Approach to Service Specification. Master's thesis, Delft University of Technology, Delft, the Netherlands, October 2008.
53. J. Mulder. *Rapid Enterprise Design*. PhD thesis, Delf University of Technology, 2006.

10 Appendixes

10.1 Appendix I – Metamodel used for CEO Framework

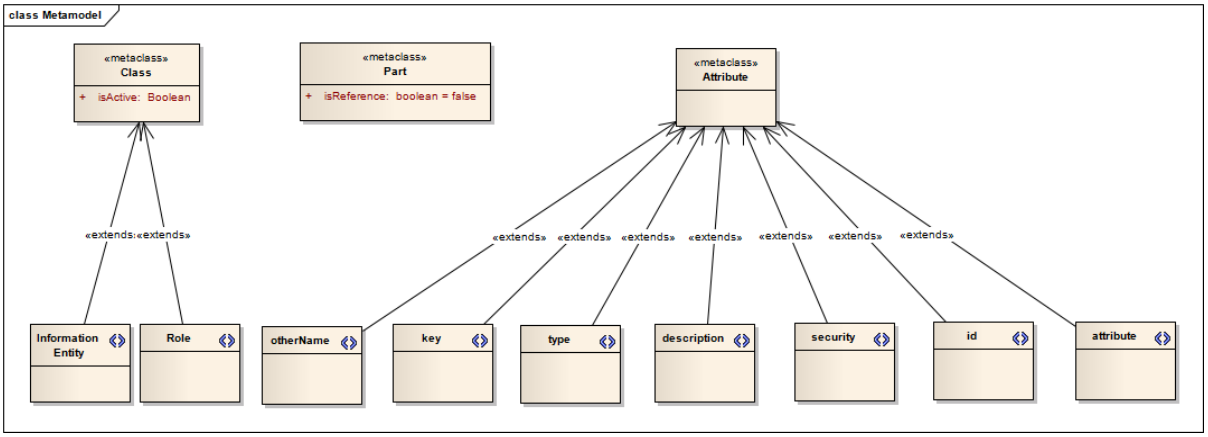


Figure 26: Representation of CEO's Framework metamodel

10.2 Appendix II – HL7 Reference Information Model

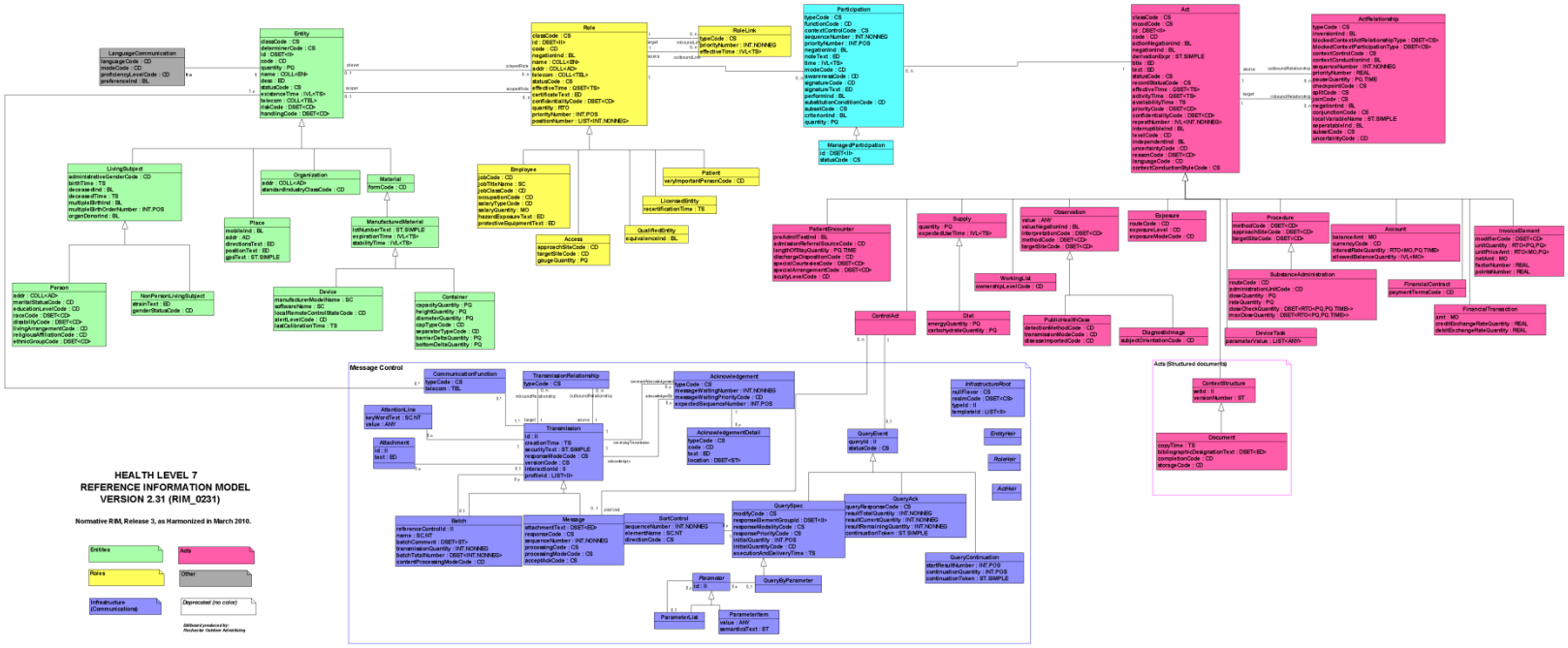


Figure 27: Representation of HL7 Reference Information Model

10.3 Appendix III – Representation of Information Entities on Analyzed Contexts

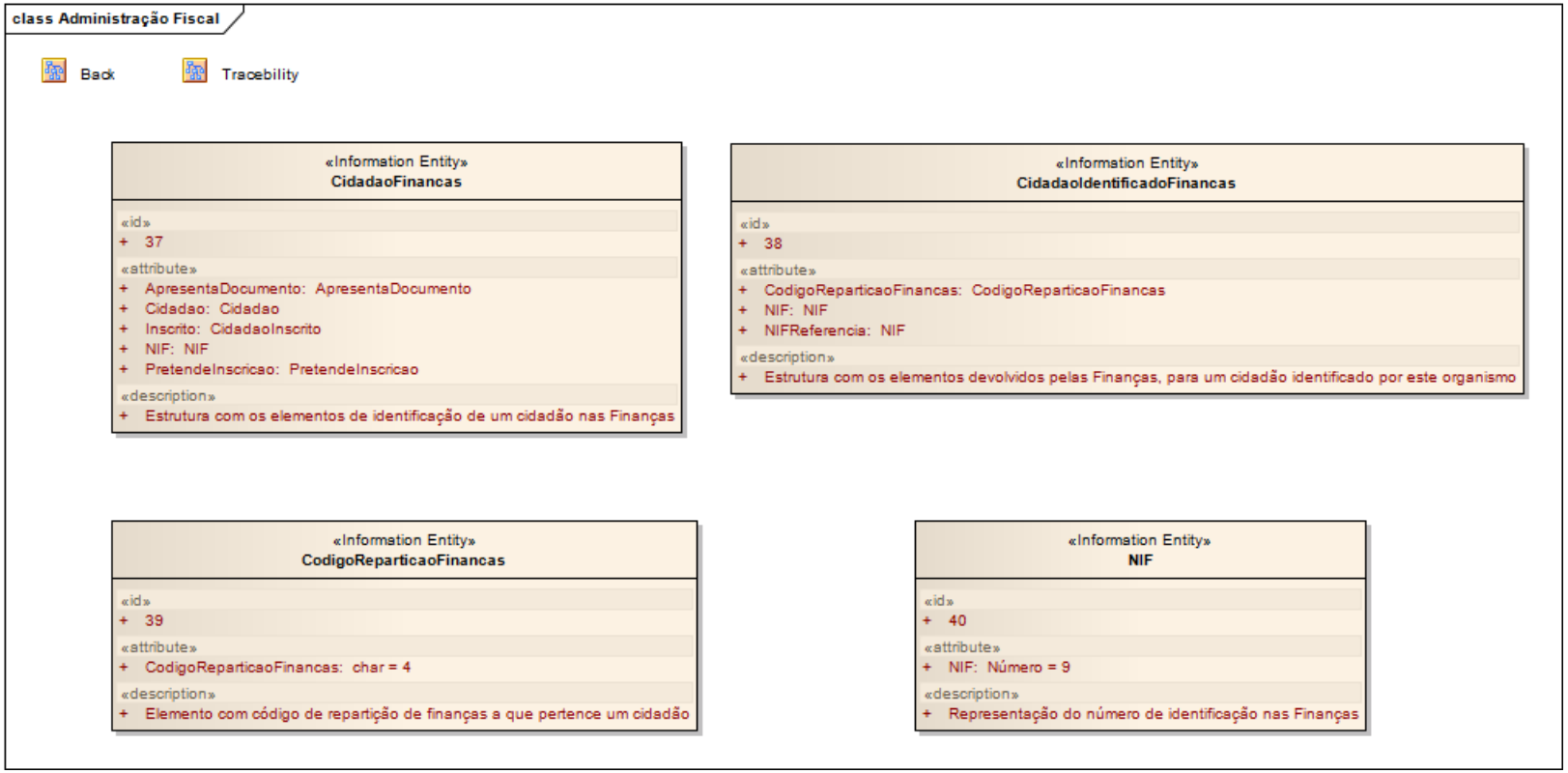


Figure 28: Representation of information entities in the context of *Administração Fiscal*

Figure 29: Representation of information entities in the context of *Identificação Civil*

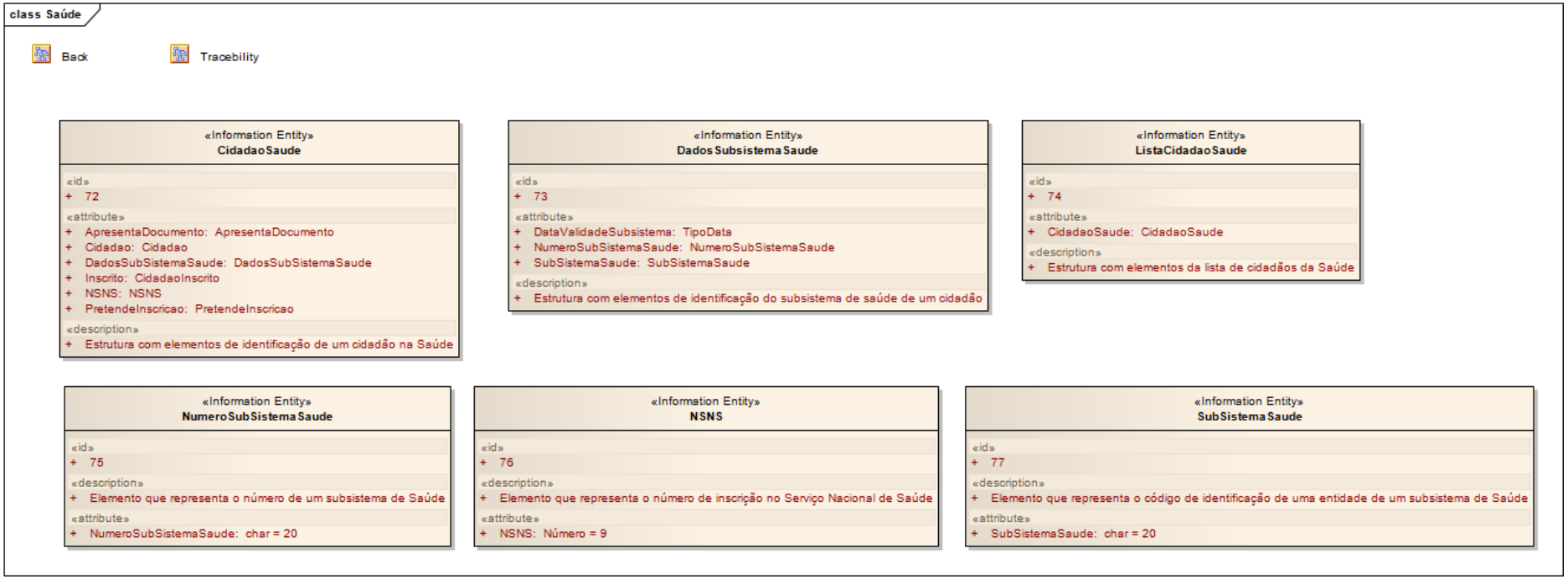


Figure 30: Representation of information entities in the context of *Saúde*

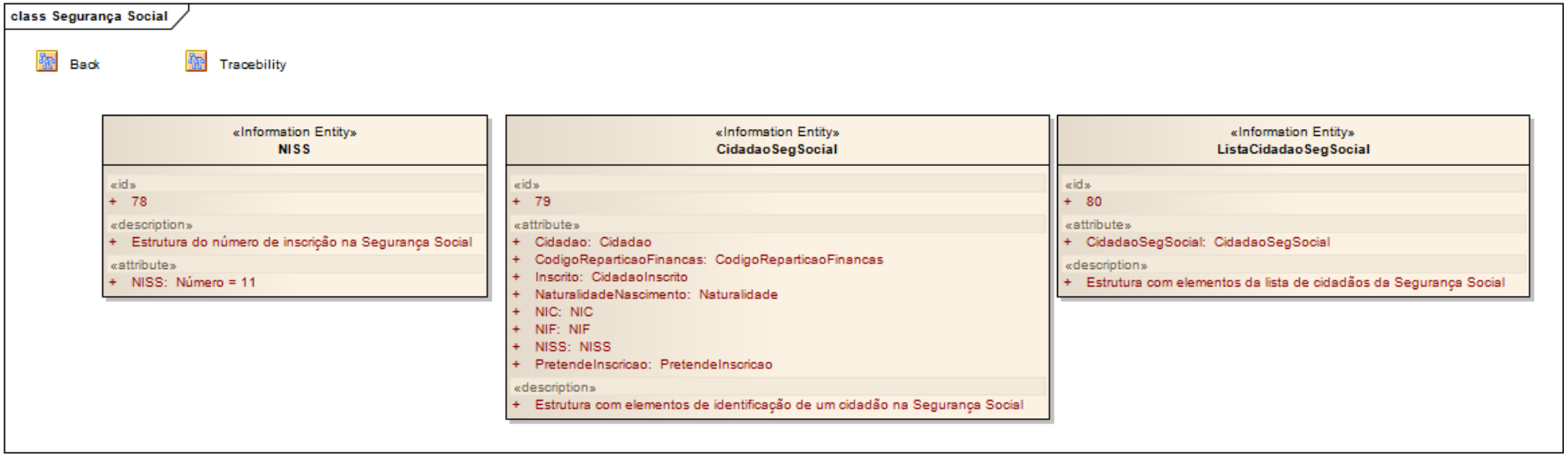


Figure 31: Representation of information entities in the context of *Segurança Social*

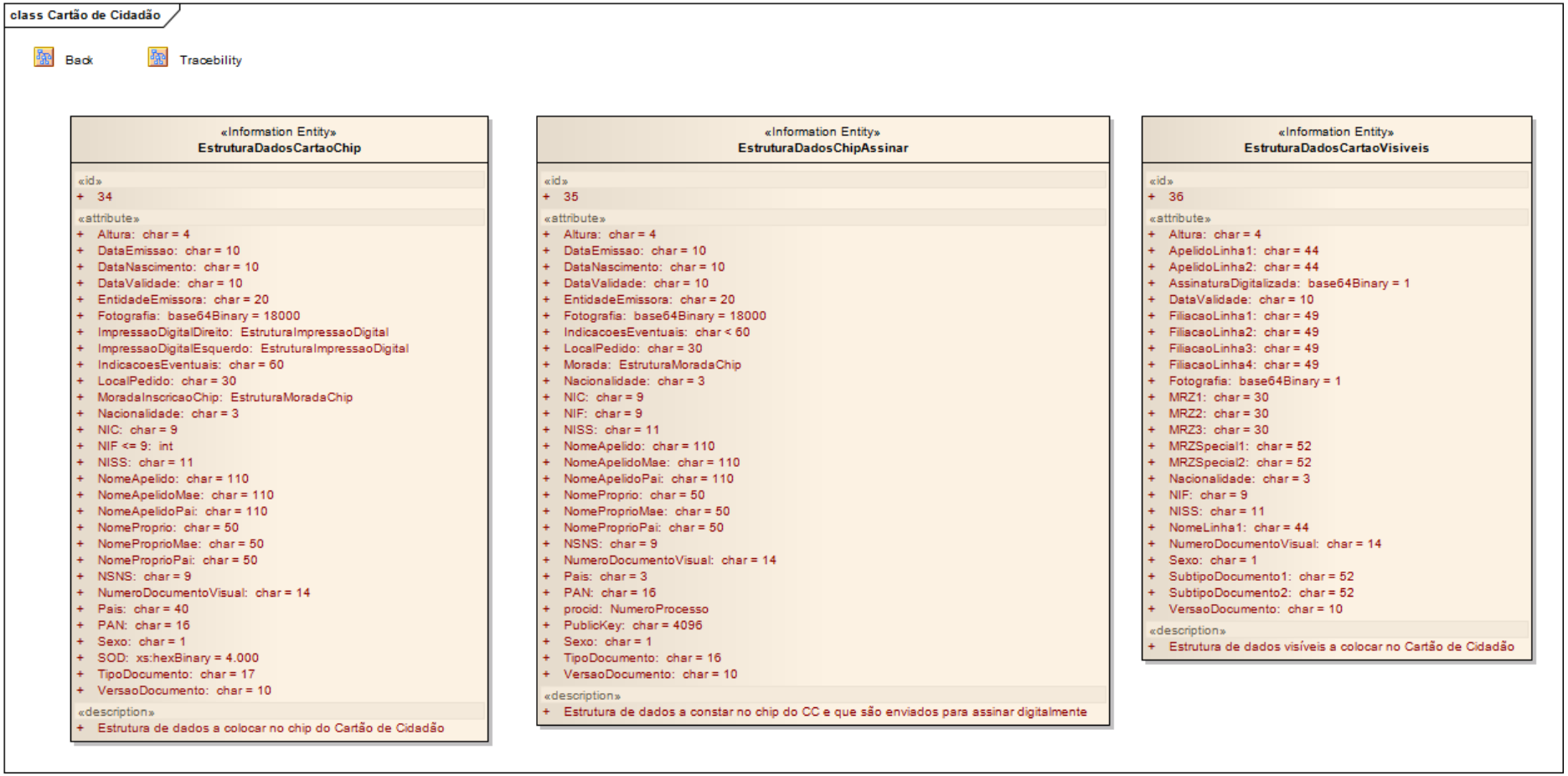


Figure 32: Representation of information entities in the context of *Cartão de Cidadão*

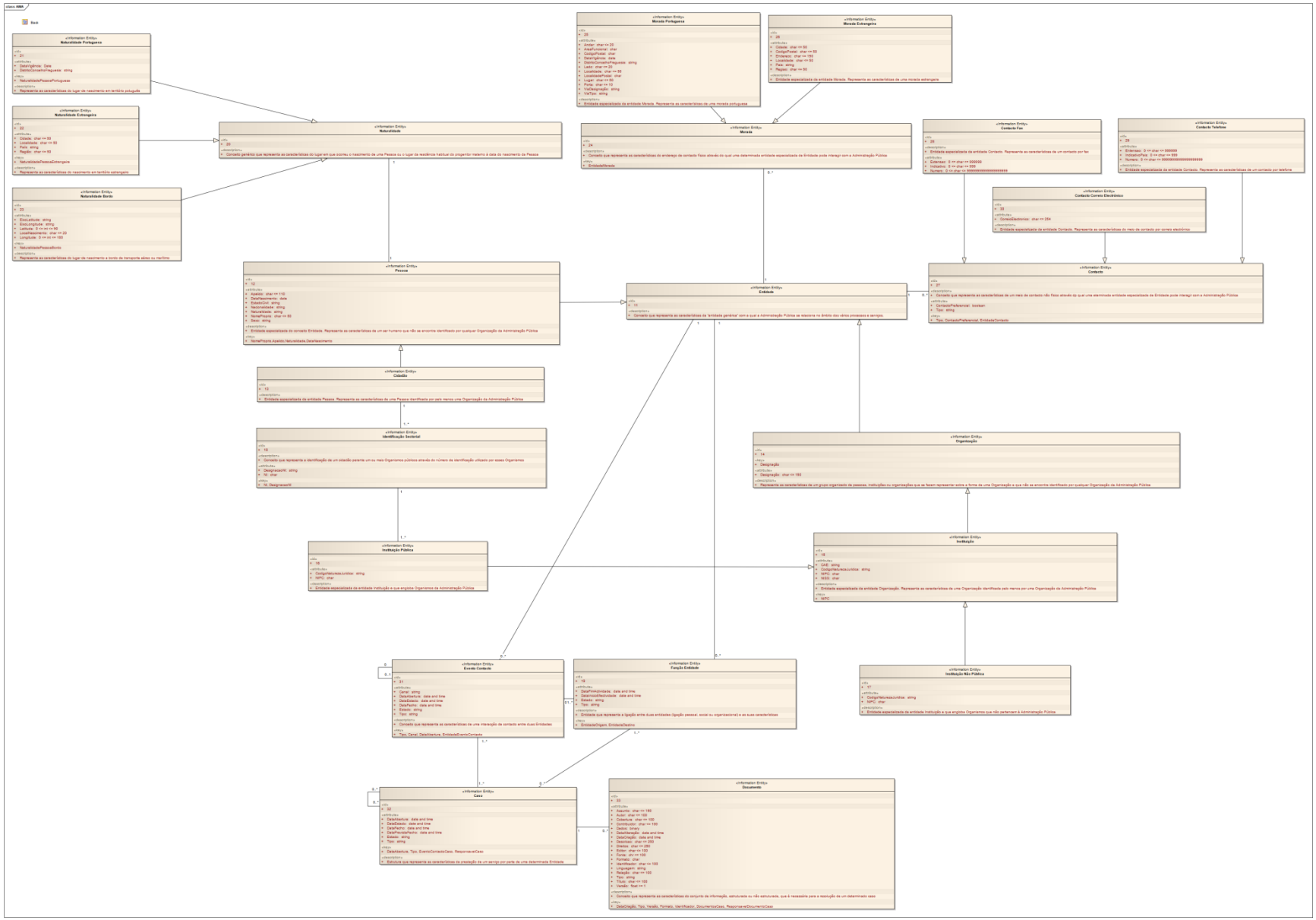


Figure 33: Representation of AMA's developed information architecture

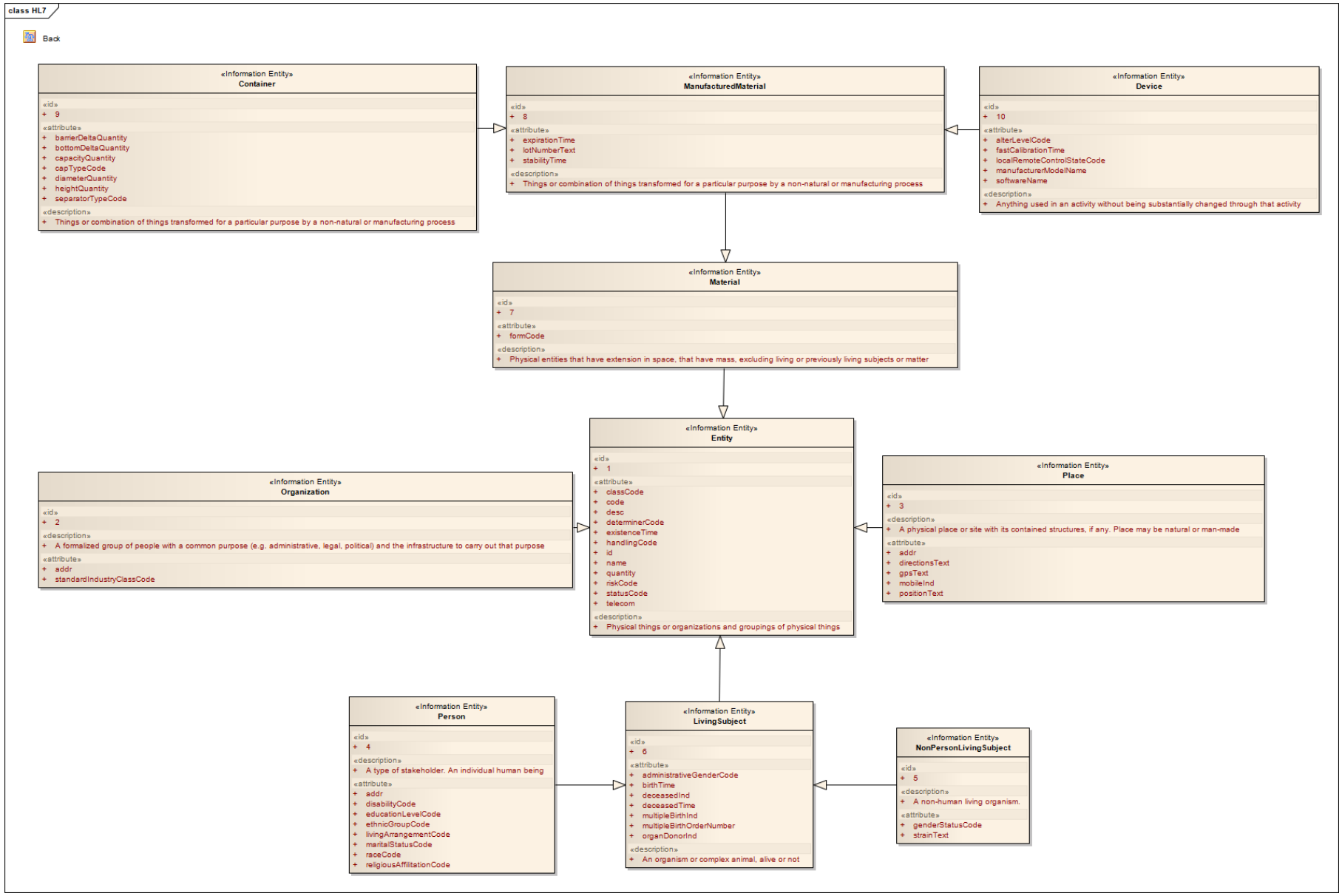


Figure 34: Representation of information entities in the context of *HL7*

10.4 Appendix IV – Examples of Representations of Analyzed Contexts using XML

The appendix contains an example of the usage given to XML on the course of this work.

First we provide the XSD used to represent the information architecture developed by AMA, followed by an instantiation of the very same XSD. As the XML served for the purpose of enabling the elaboration of mappings using BizTalk, we did not aim at extensively and completely model the specific details of each context. This means that although the entities and attributes are accurate, according to the data we gathered, aspects like the type of each attribute may not entirely correspond to the reality. We used XML for its structured, clear representation capabilities and not with the intention of providing a precise representation, closer to an hypothetical technology layer, as that was not an objective of our work.

Finally, we provide a view over the internal representation, that also uses XML, of the way a BizTalk Map is stored.

10.4.1 XSD of the Information Architecture by AMA

```
<?xml version="1.0" encoding="utf-16"?>
<xs:schema xmlns:b="http://schemas.microsoft.com/BizTalk/2003"
xmlns="http://ASI_AP.AMA" targetNamespace="http://ASI_AP.AMA"
xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="AMA">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="Entidade">
          <xs:complexType />
        </xs:element>
        <xs:element name="Pessoa">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="Apelido" type="xs:string" />
              <xs:element name="DataNascimento" type="xs:string" />
              <xs:element name="EstadoCivil" type="xs:string" />
              <xs:element name="Nacionalidade" type="xs:string" />
              <xs:element name="Naturalidade" type="xs:string" />
              <xs:element name="NomeProprio" type="xs:string" />
              <xs:element name="Sexo" type="xs:string" />
            </xs:sequence>
          </xs:complexType>
        </xs:element>
        <xs:element name="Cidadão">
          <xs:complexType />
        </xs:element>
        <xs:element name="Organizacao">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="Designacao" type="xs:string" />
            </xs:sequence>
          </xs:complexType>
        </xs:element>
        <xs:element name="Instituicao">
          <xs:complexType>
            <xs:sequence>
```

```
            <xs:element name="CAE" type="xs:string" />
            <xs:element name="CodigoNaturezaJuridica" type="xs:string" />
            <xs:element name="NIPC" type="xs:string" />
            <xs:element name="NISS" type="xs:string" />
          </xs:sequence>
        </xs:complexType>
      </xs:element>
      <xs:element name="InstituicaoPublica">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="CodigoNaturezaJuridica" type="xs:string" />
            <xs:element name="NIPC" type="xs:string" />
          </xs:sequence>
        </xs:complexType>
      </xs:element>
      <xs:element name="InstituicaoNaoPublica">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="CodigoNaturezaJuridica" type="xs:string" />
            <xs:element name="NIPC" type="xs:string" />
          </xs:sequence>
        </xs:complexType>
      </xs:element>
      <xs:element name="IdentificacaoSectorial">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="DesignacaoNI" type="xs:string" />
            <xs:element name="NI" type="xs:string" />
          </xs:sequence>
        </xs:complexType>
      </xs:element>
      <xs:element name="FuncaoEntidade">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="DataFimActividade" type="xs:string" />
            <xs:element name="DataInicioEfectividade" type="xs:string" />
            <xs:element name="Estado" type="xs:string" />
            <xs:element name="Tipo" type="xs:string" />
          </xs:sequence>
        </xs:complexType>
      </xs:element>
      <xs:element name="Naturalidade">
        <xs:complexType />
      </xs:element>
      <xs:element name="NaturalidadePortuguesa">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="DataVigencia" type="xs:string" />
            <xs:element name="DistritoConcelhoFreguesia" type="xs:string"
/>
          </xs:sequence>
        </xs:complexType>
      </xs:element>
      <xs:element name="NaturalidadeEstrangeira">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="Cidade" type="xs:string" />
            <xs:element name="Localidade" type="xs:string" />
            <xs:element name="Pais" type="xs:string" />
            <xs:element name="Regiao" type="xs:string" />
          </xs:sequence>
```

```
          </xs:complexType>
        </xs:element>
        <xs:element name="NaturalidadeBordo">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="EixoLatitude" type="xs:string" />
              <xs:element name="EixoLongitude" type="xs:string" />
              <xs:element name="Latitude" type="xs:string" />
              <xs:element name="LocalNascimento" type="xs:string" />
              <xs:element name="Longitude" type="xs:string" />
            </xs:sequence>
          </xs:complexType>
        </xs:element>
        <xs:element name="Morada">
          <xs:complexType />
        </xs:element>
        <xs:element name="MoradaPortuguesa">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="Andar" type="xs:string" />
              <xs:element name="AreaFuncional" type="xs:string" />
              <xs:element name="CodigoPostal" type="xs:string" />
              <xs:element name="DataVigencia" type="xs:string" />
              <xs:element name="DistritoConcelhoFreguesia" type="xs:string"
/>
              <xs:element name="Lado" type="xs:string" />
              <xs:element name="Localidade" type="xs:string" />
              <xs:element name="LocalidadePostal" type="xs:string" />
              <xs:element name="Lugar" type="xs:string" />
              <xs:element name="Porta" type="xs:string" />
              <xs:element name="ViaDesignacao" type="xs:string" />
              <xs:element name="ViaTipo" type="xs:string" />
            </xs:sequence>
          </xs:complexType>
        </xs:element>
        <xs:element name="MoradaEstrangeira">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="Cidade" type="xs:string" />
              <xs:element name="CodigoPostal" type="xs:string" />
              <xs:element name="Endereco" type="xs:string" />
              <xs:element name="Localidade" type="xs:string" />
              <xs:element name="Pais" type="xs:string" />
              <xs:element name="Regiao" type="xs:string" />
            </xs:sequence>
          </xs:complexType>
        </xs:element>
        <xs:element name="Contacto">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="ContactoPreferencial" type="xs:string" />
              <xs:element name="Tipo" type="xs:string" />
            </xs:sequence>
          </xs:complexType>
        </xs:element>
        <xs:element name="ContactoFax">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="Extensao" type="xs:string" />
              <xs:element name="Indicativo" type="xs:string" />
              <xs:element name="Numero" type="xs:string" />
```

```
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="ContactoTelefone">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="Extensao" type="xs:string" />
      <xs:element name="IndicativoPais" type="xs:string" />
      <xs:element name="Numero" type="xs:string" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="ContactoCorreioElectronico">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="CorreioElectronico" type="xs:string" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="EventoContacto">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="Canal" type="xs:string" />
      <xs:element name="DataAbertura" type="xs:string" />
      <xs:element name="DataEstado" type="xs:string" />
      <xs:element name="DataFecho" type="xs:string" />
      <xs:element name="Estado" type="xs:string" />
      <xs:element name="Tipo" type="xs:string" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="Caso">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="DataAbertura" type="xs:string" />
      <xs:element name="DataEstado" type="xs:string" />
      <xs:element name="DataFecho" type="xs:string" />
      <xs:element name="DataPrevistaFecho" type="xs:string" />
      <xs:element name="Estado" type="xs:string" />
      <xs:element name="Tipo" type="xs:string" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="Documento">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="Assunto" type="xs:string" />
      <xs:element name="Autor" type="xs:string" />
      <xs:element name="Cobertura" type="xs:string" />
      <xs:element name="Contribuidor" type="xs:string" />
      <xs:element name="Dados" type="xs:string" />
      <xs:element name="DataAlteracao" type="xs:string" />
      <xs:element name="DataCriacao" type="xs:string" />
      <xs:element name="Descricao" type="xs:string" />
      <xs:element name="Direitos" type="xs:string" />
      <xs:element name="Editor" type="xs:string" />
      <xs:element name="Fonte" type="xs:string" />
      <xs:element name="Formato" type="xs:string" />
      <xs:element name="Identificador" type="xs:string" />
      <xs:element name="Linguagem" type="xs:string" />
      <xs:element name="Relacao" type="xs:string" />
```

```
            <xs:element name="Tipo" type="xs:string" />
            <xs:element name="Titulo" type="xs:string" />
            <xs:element name="Versao" type="xs:string" />
          </xs:sequence>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:element>
 </xs:schema>
```

10.4.2 Instantiation of the Information Architecture's by AMA XSD

```
<ns0:AMA xmlns:ns0="http://ASI_AP.AMA">
  <Entidade />
  <Pessoa>
    <Apelido>Apelido_0</Apelido>
    <DataNascimento>DataNascimento_0</DataNascimento>
    <EstadoCivil>EstadoCivil_0</EstadoCivil>
    <Nacionalidade>Nacionalidade_0</Nacionalidade>
    <Naturalidade>Naturalidade_0</Naturalidade>
    <NomeProprio>NomeProprio_0</NomeProprio>
    <Sexo>Sexo_0</Sexo>
  </Pessoa>
  <Cidadão />
  <Organizacao>
    <Designacao>Designacao_0</Designacao>
  </Organizacao>
  <Instituicao>
    <CAE>CAE_0</CAE>

<CodigoNaturezaJuridica>CodigoNaturezaJuridica_0</CodigoNaturezaJuridica>
    <NIPC>NIPC_0</NIPC>
    <NISS>NISS_0</NISS>
  </Instituicao>
  <InstituicaoPublica>

<CodigoNaturezaJuridica>CodigoNaturezaJuridica_0</CodigoNaturezaJuridica>
    <NIPC>NIPC_0</NIPC>
  </InstituicaoPublica>
  <InstituicaoNaoPublica>

<CodigoNaturezaJuridica>CodigoNaturezaJuridica_0</CodigoNaturezaJuridica>
    <NIPC>NIPC_0</NIPC>
  </InstituicaoNaoPublica>
  <IdentificacaoSectorial>
    <DesignacaoNI>DesignacaoNI_0</DesignacaoNI>
    <NI>NI_0</NI>
  </IdentificacaoSectorial>
  <FuncaoEntidade>
    <DataFimActividade>DataFimActividade_0</DataFimActividade>

<DataInicioEfectividade>DataInicioEfectividade_0</DataInicioEfectividade>
    <Estado>Estado_0</Estado>
    <Tipo>Tipo_0</Tipo>
  </FuncaoEntidade>
  <Naturalidade />
  <NaturalidadePortuguesa>
    <DataVigencia>DataVigencia_0</DataVigencia>
```

```
<DistritoConcelhoFreguesia>DistritoConcelhoFreguesia_0</DistritoConcelhoFre
guesia>
  </NaturalidadePortuguesa>
  <NaturalidadeEstrangeira>
    <Cidade>Cidade_0</Cidade>
    <Localidade>Localidade_0</Localidade>
    <Pais>Pais_0</Pais>
    <Regiao>Regiao_0</Regiao>
  </NaturalidadeEstrangeira>
  <NaturalidadeBordo>
    <EixoLatitude>EixoLatitude_0</EixoLatitude>
    <EixoLongitude>EixoLongitude_0</EixoLongitude>
    <Latitude>Latitude_0</Latitude>
    <LocalNascimento>LocalNascimento_0</LocalNascimento>
    <Longitude>Longitude_0</Longitude>
  </NaturalidadeBordo>
  <Morada />
  <MoradaPortuguesa>
    <Andar>Andar_0</Andar>
    <AreaFuncional>AreaFuncional_0</AreaFuncional>
    <CodigoPostal>CodigoPostal_0</CodigoPostal>
    <DataVigencia>DataVigencia_0</DataVigencia>

<DistritoConcelhoFreguesia>DistritoConcelhoFreguesia_0</DistritoConcelhoFre
guesia>
    <Lado>Lado_0</Lado>
    <Localidade>Localidade_0</Localidade>
    <LocalidadePostal>LocalidadePostal_0</LocalidadePostal>
    <Lugar>Lugar_0</Lugar>
    <Porta>Porta_0</Porta>
    <ViaDesignacao>ViaDesignacao_0</ViaDesignacao>
    <ViaTipo>ViaTipo_0</ViaTipo>
  </MoradaPortuguesa>
  <MoradaEstrangeira>
    <Cidade>Cidade_0</Cidade>
    <CodigoPostal>CodigoPostal_0</CodigoPostal>
    <Endereco>Endereco_0</Endereco>
    <Localidade>Localidade_0</Localidade>
    <Pais>Pais_0</Pais>
    <Regiao>Regiao_0</Regiao>
  </MoradaEstrangeira>
  <Contacto>
    <ContactoPreferencial>ContactoPreferencial_0</ContactoPreferencial>
    <Tipo>Tipo_0</Tipo>
  </Contacto>
  <ContactoFax>
    <Extensao>Extensao_0</Extensao>
    <Indicativo>Indicativo_0</Indicativo>
    <Numero>Numero_0</Numero>
  </ContactoFax>
  <ContactoTelefone>
    <Extensao>Extensao_0</Extensao>
    <IndicativoPais>IndicativoPais_0</IndicativoPais>
    <Numero>Numero_0</Numero>
  </ContactoTelefone>
  <ContactoCorreioElectronico>
    <CorreioElectronico>CorreioElectronico_0</CorreioElectronico>
  </ContactoCorreioElectronico>
  <EventoContacto>
    <Canal>Canal_0</Canal>
```

```
    <DataAbertura>DataAbertura_0</DataAbertura>
    <DataEstado>DataEstado_0</DataEstado>
    <DataFecho>DataFecho_0</DataFecho>
    <Estado>Estado_0</Estado>
    <Tipo>Tipo_0</Tipo>
  </EventoContacto>
  <Caso>
    <DataAbertura>DataAbertura_0</DataAbertura>
    <DataEstado>DataEstado_0</DataEstado>
    <DataFecho>DataFecho_0</DataFecho>
    <DataPrevistaFecho>DataPrevistaFecho_0</DataPrevistaFecho>
    <Estado>Estado_0</Estado>
    <Tipo>Tipo_0</Tipo>
  </Caso>
  <Documento>
    <Assunto>Assunto_0</Assunto>
    <Autor>Autor_0</Autor>
    <Cobertura>Cobertura_0</Cobertura>
    <Contribuidor>Contribuidor_0</Contribuidor>
    <Dados>Dados_0</Dados>
    <DataAlteracao>DataAlteracao_0</DataAlteracao>
    <DataCriacao>DataCriacao_0</DataCriacao>
    <Descricao>Descricao_0</Descricao>
    <Direitos>Direitos_0</Direitos>
    <Editor>Editor_0</Editor>
    <Fonte>Fonte_0</Fonte>
    <Formato>Formato_0</Formato>
    <Identificador>Identificador_0</Identificador>
    <Linguagem>Linguagem_0</Linguagem>
    <Relacao>Relacao_0</Relacao>
    <Tipo>Tipo_0</Tipo>
    <Titulo>Titulo_0</Titulo>
    <Versao>Versao_0</Versao>
  </Documento>
   </ns0:AMA>
```


10.4.3 Internal representation of Mappings between HL7 and AMA


```
<?xml version="1.0" encoding="utf-16"?>
<!-- Generated using BizTalk Mapper on Sun, Jul 11 2010 06:49:51 PM -->
<mapsource Name="BizTalk Map" BizTalkServerMapperTool_Version="2.0"
Version="2" XRange="100" YRange="420" OmitXmlDeclaration="Yes"
TreatElementsAsRecords="No" OptimizeValueMapping="Yes"
GenerateDefaultFixedNodes="Yes" PreserveSequenceOrder="No" CopyPIs="No"
method="xml" xmlVersion="1.0" IgnoreNamespacesForLinks="Yes">
  <SrcTree>
    <Reference Location=".\HL7.xsd" />
  </SrcTree>
  <TrgTree>
    <Reference Location="..\AMA.xsd" />
  </TrgTree>
  <ScriptTypePrecedence>
    <CSharp Enabled="Yes" />
    <ExternalAssembly Enabled="Yes" />
    <VbNet Enabled="Yes" />
    <JScript Enabled="Yes" />
    <XsltCallTemplate Enabled="Yes" />
    <Xslt Enabled="Yes" />
  </ScriptTypePrecedence>
  <TreeValues>
```

```
    <TestValues />
    <ConstantValues />
  </TreeValues>
  <Pages>
    <Page Name="Page 1">
      <Links>
        <Link LinkID="1" LinkFrom="/*[local-
name()='&lt;Schema&gt;']/*[local-name()='HL7']/*[local-
name()='LivingSubject']/*[local-name()='administrativeGenderCode']"
LinkTo="/*[local-name()='&lt;Schema&gt;']/*[local-name()='AMA']/*[local-
name()='Pessoa']/*[local-name()='Sexo']" Label="" />
        <Link LinkID="2" LinkFrom="/*[local-
name()='&lt;Schema&gt;']/*[local-name()='HL7']/*[local-
name()='LivingSubject']/*[local-name()='birthTime']" LinkTo="/*[local-
name()='&lt;Schema&gt;']/*[local-name()='AMA']/*[local-
name()='Pessoa']/*[local-name()='DataNascimento']" Label="" />
        <Link LinkID="3" LinkFrom="/*[local-
name()='&lt;Schema&gt;']/*[local-name()='HL7']/*[local-
name()='Person']/*[local-name()='maritalStatusCode']" LinkTo="/*[local-
name()='&lt;Schema&gt;']/*[local-name()='AMA']/*[local-
name()='Pessoa']/*[local-name()='EstadoCivil']" Label="" />
        <Link LinkID="4" LinkFrom="/*[local-
name()='&lt;Schema&gt;']/*[local-name()='HL7']/*[local-
name()='Person']/*[local-name()='addr']" LinkTo="/*[local-
name()='&lt;Schema&gt;']/*[local-name()='AMA']/*[local-name()='Morada']"
Label="" />
      </Links>
      <Functoids />
    </Page>
  </Pages>
   </mapsource>
```

10.5 Appendix V – Entities in Analyzed Contexts and AMA's proposed Architecture

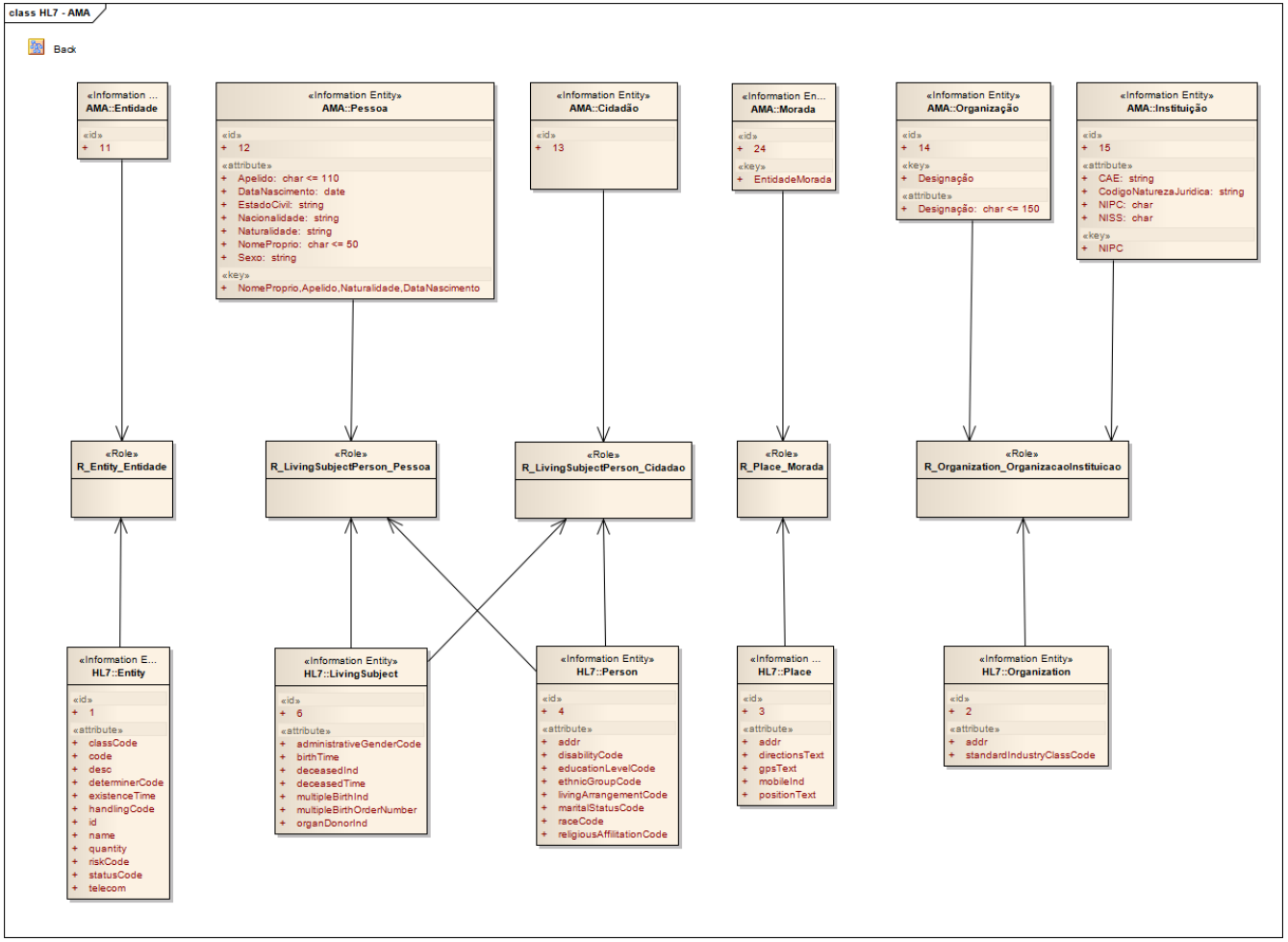


Figure 35: Relation between HL7 and AMA's proposed architecture using roles

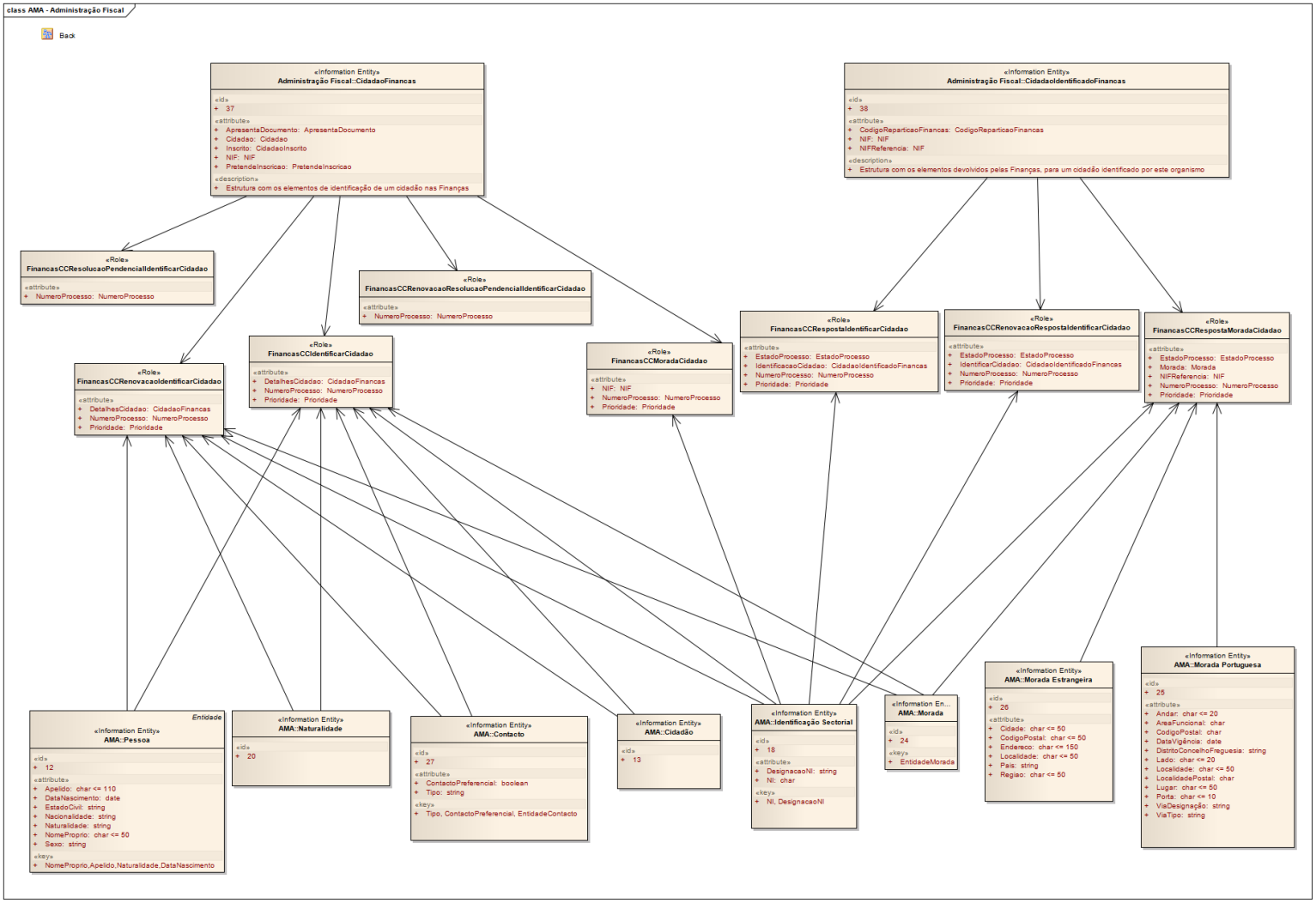


Figure 36: Relation between *Administração Fiscal* and AMA's proposed architecture - roles

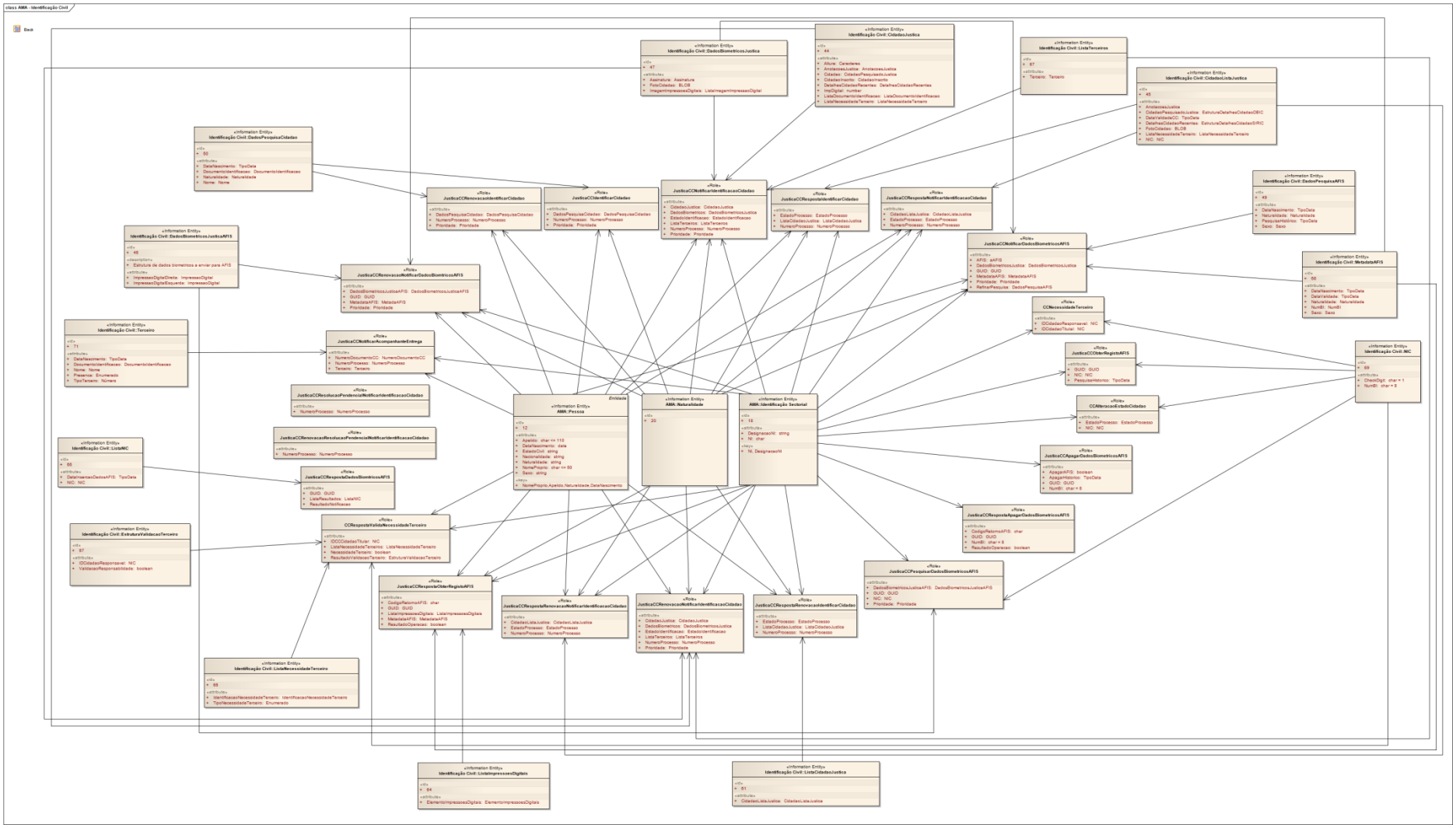


Figure 37: Relation between *Identificação Civil* and AMA's proposed architecture using roles

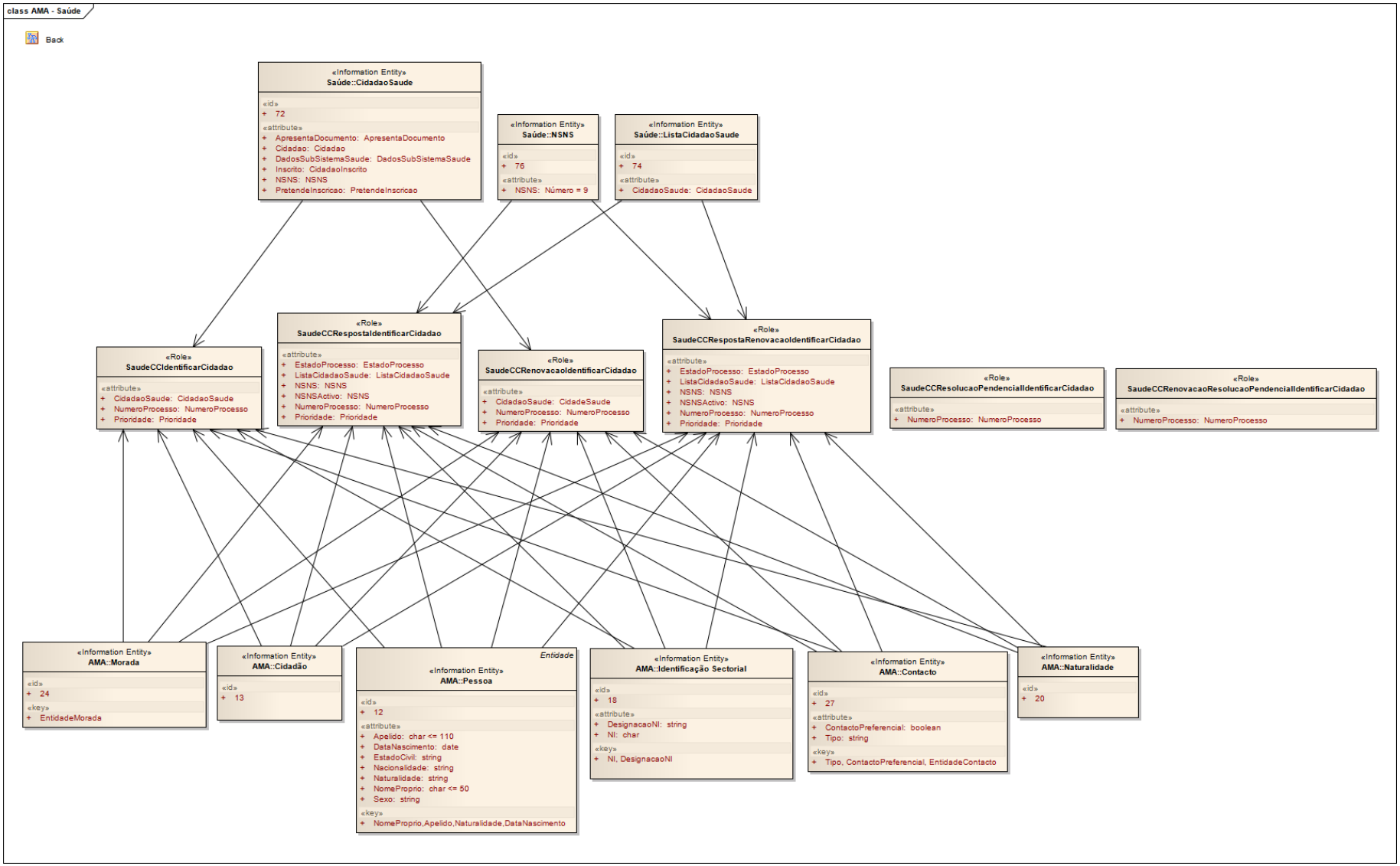


Figure 38: Relation between *Saúde* and AMA's proposed architecture using roles

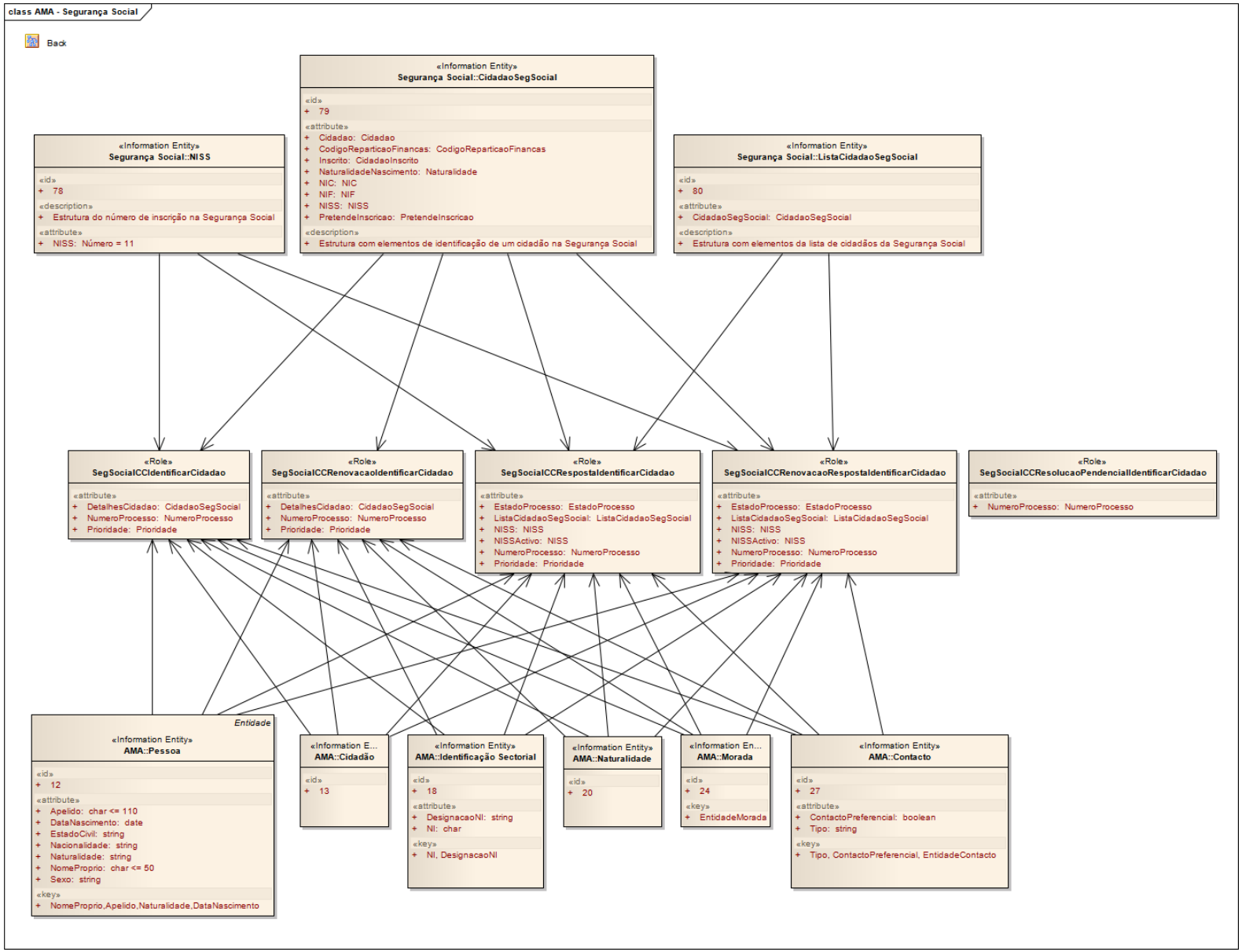


Figure 39: Relation between *Segurança Social* and AMA's proposed architecture using roles

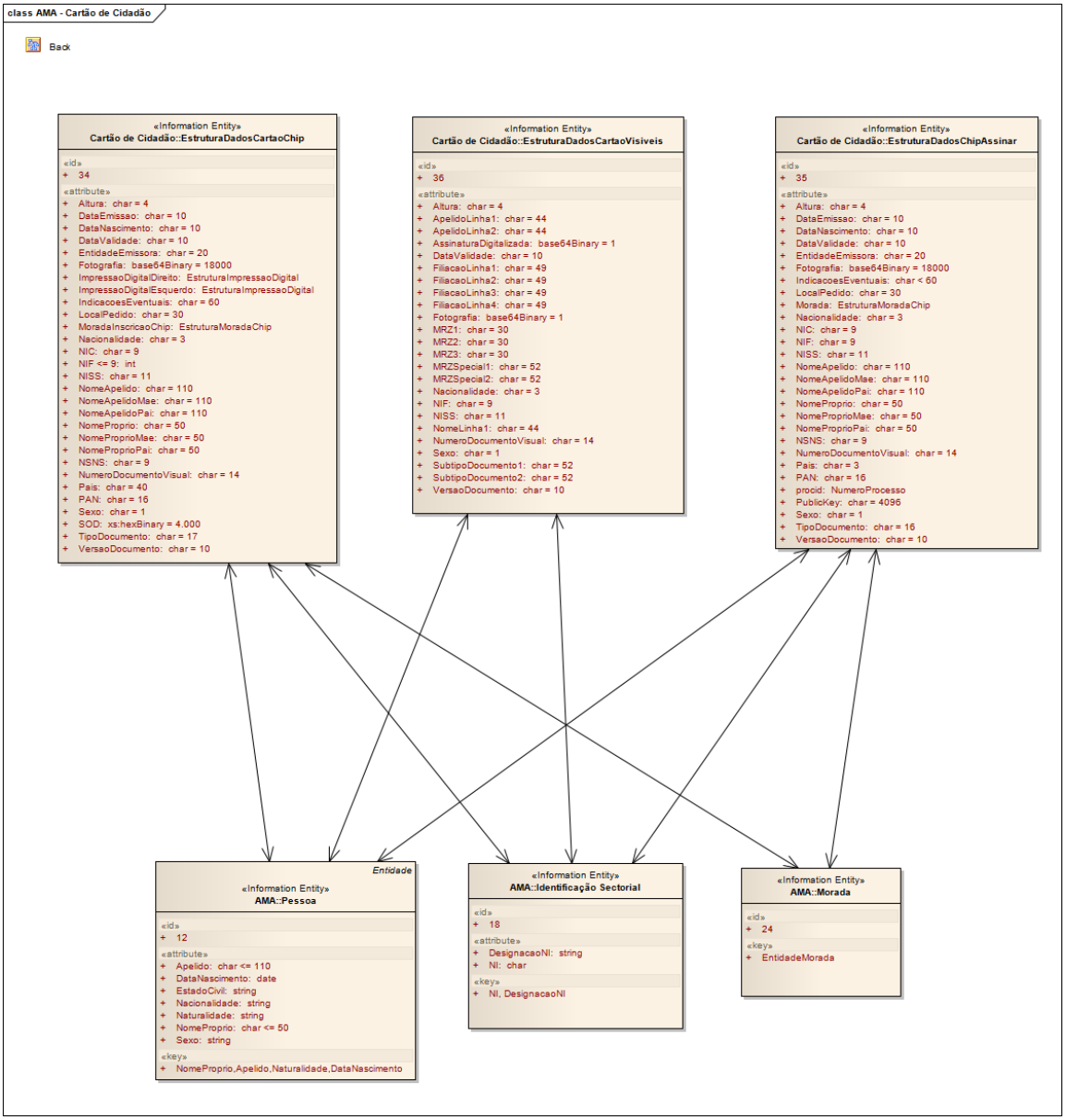


Figure 40: Relation between *Cartão de Cidadão* and AMA's proposed architecture using roles

10.6 Appendix VI – List of Entities in Analyzed Contexts

Id	Name
1	Entity
2	Organization
3	Place
4	Person
5	NonPersonLivingSubject
6	LivingSubject
7	Material
8	ManufacturedMaterial
9	Container
10	Device
11	Entidade
12	Pessoa
13	Cidadão
14	Organização
15	Instituição
16	Instituição Pública
17	Instituição Não Pública
18	Identificação Sectorial
19	Função Entidade
20	Naturalidade
21	Naturalidade Portuguesa
22	Naturalidade Estrangeira
23	Naturalidade Bordo
24	Morada
25	Morada Portuguesa
26	Morada Estrangeira
27	Contacto
28	Contacto Fax
29	Contacto Telefone
30	Contacto Correio Electrónico
31	Evento Contacto
32	Caso
33	Documento
34	EstruturaDadosCartaoChip
35	EstruturaDadosChipAssinar
36	EstruturaDadosCartaoVisiveis
37	CidadaoFinancas
38	CidadaoIdentificadoFinancas
39	CodigoReparticaoFinancas
40	NIF
41	Assinatura
42	AnotacoesJustica
43	AssentoNascimento
44	CidadaoJustica
45	CidadaoListaJustica
46	CidadaoPesquisadoJustica
47	DadosBiometricosJustica
48	DadosBiometricosJusticaAFIS
49	DadosPesquisaAFIS
50	DadosPesquisaCidadao
51	DetalheIdentificacaoTitular
52	DetalhesCidadaoRecentes
53	DocumentoIdentificacao
54	ElementoImpressoesDigitais
55	EstruturaDetalhesCidadaoBDIC
56	EstruturaDetalhesCidadaoSIRIC
57	EstruturaValidacaoTerceiro
58	IdentificacaoNecessidadeTerceiro
59	ImagemImpressaoDigital
60	ImpressaoDigital
61	ListaCidadaoJustica
62	ListaDocumentoIdentificacao
63	ListaImagemImpressaoDigital
64	ListaImpressoesDigitais
65	ListaNecessidadeTerceiro
66	ListaNIC
67	ListaTerceiros
68	MetadataAFIS
69	NIC
70	OutroDocumento
71	Terceiro
72	CidadaoSaude
73	DadosSubsistemaSaude
74	ListaCidadaoSaude
75	NumeroSubSistemaSaude
76	NSNS
77	SubSistemaSaude
78	NISS
79	CidadaoSegSocial
80	ListaCidadaoSegSocial
81	EstruturaImpressaoDigital
82	EstruturaMoradaChip
83	Cidadao
84	CidadaoInscrito

Context	Color Code
Information Architecture - HL7 :	
Information Architecture - AMA :	
Information Architecture - Cartão de Cidadão :	
Information Architecture - Administração Fiscal :	
Information Architecture - Identificação Civil :	
Information Architecture - Saúde :	
Information Architecture - Segurança Social :	
Other:	

Table 8 : List of entities in analyzed contexts

10.7 Appendix VII – Results of Mappings Analysis

Context	Role	Unmapped Attributes	# Unmapped Attributes	# Attributes	% Unmapped Attributes
Administração Fiscal	FinancasCCIdentificarCidadao	Prioridade NumeroProcesso Inscrito PretendeInscricao ApresentaDocumento	5	16	31,25
	FinancasCCResolucaoPendencialIdentificarCidadao	NumeroProcesso	1	1	100
	FinancasCCRespostaIdentificarCidadao	NumeroProcesso Prioridade EstadoProcesso CodigoReparticaoFinancas	4	6	66,66666667
	FinancasCCRenovacaoIdentificarCidadao	Prioridade NumeroProcesso Inscrito PretendeInscricao ApresentaDocumento	5	16	31,25
	FinancasCCRenovacaoResolucaoPendencialIdentificarCidadao	NumeroProcesso	1	1	100
	FinancasCCRenovacaoRespostaIdentificarCidadao	NumeroProcesso Prioridade EstadoProcesso CodigoReparticaoFinancas	4	6	66,66666667
	FinancasCCMoradaCidadao	Prioridade NumeroProcesso	2	3	66,66666667
	FinancasCCRespostaMoradaCidadao	NumeroProcesso	6	22	27,27272727

		Prioridade			
		EstadoProcesso			
		Versao			
		Estado			
Cartão de Cidadão	EstruturaDadosCartaoChip	Altura			
		DataEmissao			
		DataValidade			
		EntidadeEmissora			
		Fotografia			
		ImpressaoDigitalDireito			
		ImpressaoDigitalEsquerdo			
		IndicacoesEventuais			
		LocalPedido			
		NumeroDocumentoVisual			
		Pais			
		PAN			
		SOD			
		TipoDocumento			
		VersaoDocumento	15	29	51,72413793
	EstruturaDadosChipAssinar	Altura			
		DataEmissao			
		DataValidade			
		EntidadeEmissora			
		Fotografia			
		IndicacoesEventuais			
		LocalPedido			
		NumeroDocumentoVisual			
		Pais	14	28	50

		PAN			
		Procid			
		PublicKey			
		TipoDocumento			
		VersaoDocumento			
	EstruturaDadosCartaoVisiveis	Altura			
		AssinaturaDigitalizada			
		DataValidade			
		FiliacaoLinha1			
		FiliacaoLinha2			
		FiliacaoLinha3			
		FiliacaoLinha4			
		Fotografia			
		MRZ1			
		MRZ2			
		MRZ3			
		MRZSpecial1			
		MRZSpecial2			
		NomeDocumentoVisual			
		SubtipoDocumento1			
		SubtipoDocumento2			
		VersaoDocumento	17	24	70,83333333
HL7	Entity	classCode			
		code			
		desc			
		determinerCode			
		existenceTime			
		handlingCode	12	12	100

		id			
		name			
		quantity			
		riskCode			
		statusCode			
		telecom			
	LivingSubject	deceasedInd			
		deceasedTime			
		multipleBirthInd			
		multipleBirthOrderNumber			
		organDonorInd	5	7	71,42857143
	Person	disabilityCode			
		educationLevelCode			
		ethnicGroupCode			
		livingArrangementCode			
		raceCode			
		religiousAffiliationCode	6	8	75
	NonPersonLivingSubject	genderStatusCode			
		strainText	2	2	100
	Organization	addr			
		standardIndustryClassCode	2	2	100
	Place	addr			
		directionsText			
		gpsText			
		mobileInd			
		positionText	5	5	100
	Material	formCode	1	1	100

	ManufacturedMaterial	expirationDate idNumberText stabilityTime	3	3	100
	Device	alterLevelCode firstCalibrationTime localRemoteControlStateCode manufacuredModelName softwareName	5	5	100
	Container	barrierDeltaQuantity bottomDeltaQuantity capacityQuantity capTypeCode diameterQuantity heightQuantity separatorTypeCode	7	7	100
Identificação Civil	JusticaCCIdentificarCidadao	Prioridade NumeroProcesso DocumentoIdentificacao	3	6	50
	JusticaCCNotificarIdentificacaoCidadao	Prioridade NumeroProcesso CidadaoInscrito Altura ImpDigital AnotacoesJustica EstadoIdentificacao FotoCidadao Assinatura	11	31	35,48387097

		ImagemImpressoesDigitais			
	JusticaCCResolucaoPendencialNotificarIdentificacaoCidadao	NumeroProcesso	1	1	100
	JusticaCCRespostaIdentificarCidadao	NumeroProcesso EstadoProcesso FotoCidadao DataValidadeCC AnotacoesJustica	5	16	31,25
	JusticaCCRespostaNotificarIdentificacaoCidadao	NumeroProcesso EstadoProcesso FotoCidadao DataValidadeCC AnotacoesJustica	5	16	31,25
	JusticaCCRenovacaoIdentificarCidadao	Prioridade NumeroProcesso DocumentoIdentificacao	3	6	50
	JusticaCCRenovacaoNotificarIdentificacaoCidadao	Prioridade NumeroProcesso CidadaoInscrito Altura ImpDigital AnotacoesJustica EstadoIdentificacao FotoCidadao Assinatura ImagemImpressoesDigitais	11	31	35,48387097
	JusticaCCRenovacaoResolucaoPendencialNotificarIdentificacaoCidadao	NumeroProcesso	1	1	100
	JusticaCCRespostaRenovacaoIdentificarCidadao	NumeroProcesso	5	16	31,25

		EstadoProcesso			
		FotoCidadao			
		DataValidadeCC			
		AnotacoesJustica			
	JusticaCCRespostaRenovacaoNotificarIdentificacaoCidadao	NumeroProcesso			
		EstadoProcesso			
		FotoCidadao			
		DataValidadeCC			
		AnotacoesJustica	5	16	31,25
	CCAlteracaoEstadoCidadao	EstadoProcesso	1	2	50
	JusticaCCNotificarDadosBiometricosAFIS	Prioridade			
		GUID			
		FotoCidadao			
		Assinatura			
		ImagemImpressoesDigitais			
		AFIS			
		DataValidade			
		PesquisaHistorico	8	15	53,33333333
	JusticaCCRenovacaoNotificarDadosBiometricosAFIS	Prioridade			
		GUID			
		DadosBiometricosJusticaAFIS			
		DataValidade	4	8	50
	JusticaCCRespostaDadosBiometricosAFIS	GUID			
		ListaResultados			
		ResultadoNotificacao	3	3	100
	JusticaCCApagarDadosBiometricosAFIS	GUID			
		ApagarAFIS	3	4	75

		ApagarHistorico			
	JusticaCCRespostaApagarDadosBiometricosAFIS	GUID ResultadoOperacao CodigoRetornoAFIS	3	4	75
	JusticaCCPesquisarDadosBiometricosAFIS	Prioridade GUID DadosBiometricosJusticaAFIS	3	4	75
	JusticaCCObterRegistoAFIS	GUID PesquisaHistorico	2	3	66,66666667
	JusticaCCRespostaObterRegistoAFIS	GUID DataValidade ListaImpressoesDigitais ResultadoOperacao CodigoRetornoAFIS	5	9	55,55555556
	JusticaCCNotificarAcompanhanteEntrega	NumeroProcesso NumeroDocumentoCC Presenca TipoTerceiro	4	7	57,14285714
	CCNecessidadeTerceiro	-	0	2	0
	CCRespostaValidaNecessidadeTerceiro	NecessidadeTerceiro	1	4	25
Saúde	SaudeCCIdentificarCidadao	Prioridade NumeroProcesso Inscrito PretendeInscricao ApresentaDocumento SubSistemaSaude NumeroSubSistemaSaude	8	20	40

		DataValidadeSubsistema			
	SaudeCCResolucaoPendencialIdentificarCidadao	NumeroProcesso	1	1	100
	SaudeCCRespostaIdentificarCidadao	Prioridade NumeroProcesso EstadoProcesso	3	6	50
	SaudeCCRenovacaoIdentificarCidadao	Prioridade NumeroProcesso Inscrito PretendeInscricao ApresentaDocumento SubSistemaSaude NumeroSubSistemaSaude DataValidadeSubsistema	8	20	40
	SaudeCCRenovacaoResolucaoPendencialIdentificarCidadao	NumeroProcesso	1	1	100
	SaudeCCRespostaRenovacaoIdentificarCidadao	Prioridade NumeroProcesso EstadoProcesso	3	6	50
Segurança Social	SegSocialCCIdentificarCidadao	Prioridade NumeroProcesso CidadaoInscrito PretendeInscricao CodigoReparticaoFinancas	5	19	26,31578947
	SegSocialCCResolucaoPendencialIdentificarCidadao	NumeroProcesso	1	1	100
	SegSocialCCRespostaIdentificarCidadao	Prioridade NumeroProcesso EstadoProcesso CidadaoInscrito	6	22	27,27272727

		PretendeInscricao			
		CodigoReparticaoFinancas			
	SegSocialCCRenovacaoIdentificarCidadao	Prioridade			
		NumeroProcesso			
		CidadaoInscrito			
		PretendeInscricao			
		CodigoReparticaoFinancas	5	19	26,31578947
	SegSocialCCRenovacaoRespostaIdentificarCidadao	Prioridade			
		NumeroProcesso			
		EstadoProcesso			
		Inscrito			
		PretendeInscricao			
		CodigoReparticaoFinancas	6	22	27,27272727

Table 9 : Results of mappings analysis

10.8 Appendix VIII – Results of Attribute's Mapping Analysis

Unmapped Attributes	# Occurrences
NumeroProcesso	30
Prioridade	21
EstadoProcesso	12
PretendeInscricao	8
GUID	8
FotoCidadao	7
CodigoReparticaoFinancas	6
DataValidade	6
AnotacoesJustica	6
Inscrito	5
Altura	4
CidadaoInscrito	4
ApresentaDocumento	4
DataValidadeCC	4
Fotografia	3
VersaoDocumento	3
Assinatura	3
ImagemImpressoesDigitais	3
DataEmissao	2
EntidadeEmissora	2
IndicacoesEventuais	2
LocalPedido	2
NumeroDocumentoVisual	2
Pais	2
PAN	2
TipoDocumento	2
addr	2
DocumentoIdentificacao	2
ImpDigital	2
EstadoIdentificacao	2
PesquisaHistorico	2
DadosBiometricosJusticaAFIS	2
ResultadoOperacao	2
CodigoRetornoAFIS	2
SubSistemaSaude	2
NumeroSubSistemaSaude	2
DataValidadeSubsistema	2
Versao	1
Estado	1
ImpressaoDigitalDireito	1
ImpressaoDigitalEsquerdo	1
SOD	1
Procid	1
PublicKey	1
AssinaturaDigitalizada	1
FiliacaoLinha1	1
FiliacaoLinha2	1
FiliacaoLinha3	1
FiliacaoLinha4	1
MRZ1	1
MRZ2	1
MRZ3	1
MRZSpecial1	1
MRZSpecial2	1
NomeDocumentoVisual	1
SubtipoDocumento1	1
SubtipoDocumento2	1
classCode	1
code	1
desc	1
determinerCode	1
existenceTime	1
handlingCode	1
id	1
name	1
quantity	1
riskCode	1
statusCode	1
telecom	1
deceasedInd	1
deceasedTime	1
multipleBirthInd	1
multipleBirthOrderNumber	1
organDonorInd	1
disabilityCode	1
educationLevelCode	1
ethnicGroupCode	1
livingArrangementCode	1
raceCode	1
religiousAffiliationCode	1
genderStatusCode	1
strainText	1
standardIndustryClassCode	1
directionsText	1
gpsText	1

mobileInd	1
positionText	1
formCode	1
expirationDate	1
idNumberText	1
stabilityTime	1
alterLevelCode	1
firstCalibrationTime	1
localRemoteControlStateCode	1
manufacuredModelName	1
softwareName	1
barrierDeltaQuantity	1
bottomDeltaQuantity	1
capacityQuantity	1
capTypeCode	1

diameterQuantity	1
heightQuantity	1
separatorTypeCode	1
AFIS	1
ListaResultados	1
ResultadoNotificacao	1
ApagarAFIS	1
ApagarHistorico	1
ListaImpressoesDigitais	1
NumeroDocumentoCC	1
Presenca	1
TipoTerceiro	1
NecessidadeTerceiro	1

Table 10 : Results of attribute's mapping analysis

10.9 Appendix IX – Proposed Architecture with Highlighted Changes

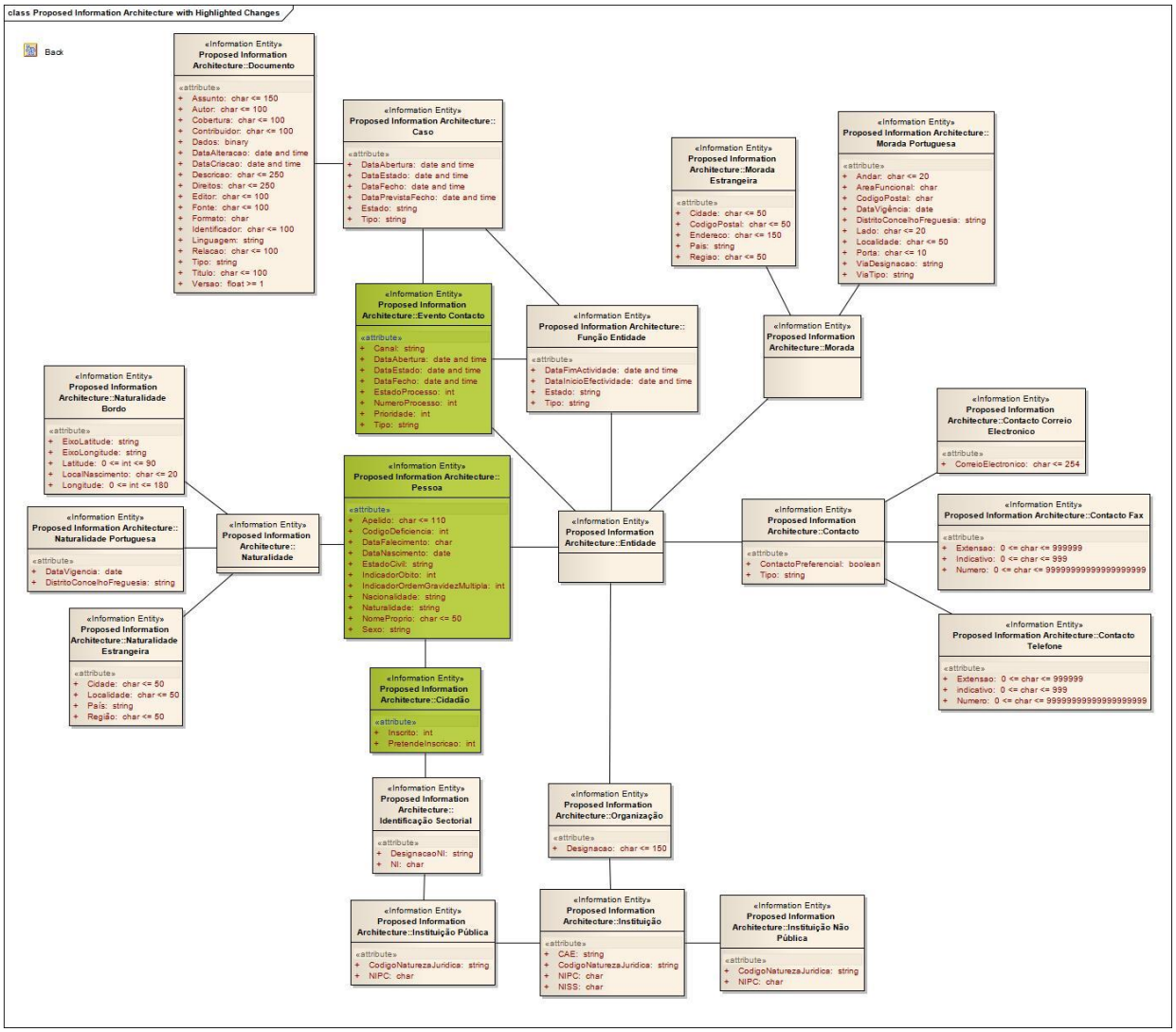


Figure 41: Proposed architecture with highlighted changes

10.10 Appendix X – Qualitative Analysis of Languages and Frameworks

	Representation of the Architecture	Representation of the Information Entities	Definition of the Methodology for developing the Architecture	Representation of Business Processes	Representation of Data Models	Establishing Relations between Entities	Establishing Relations between Attributes	Evaluate the Alignment between Entities and Processes
Spewak's EAP			X					
BPMN				X				X
Archimate	X			X	X	X		X
CEO Framework	X	X		X	X	X		X
UML					X			
E-R	X	X			X			
XML							X	
CRUD Matrix								X

Table 11: Qualitative analysis of languages and frameworks' contributions to this work

10.11 Appendix XI – Activities Chronogram

	Task Name	Duration	Start	Finish
1	**Masters Degree Project**	**84 days**	**Tue 15-09-09**	**Thu 07-01-10**
2	**Initial Research**	**34 days**	**Tue 15-09-09**	**Fri 30-10-09**
3	Articles	23 days	Tue 15-09-09	Thu 15-10-09
4	Conferences	12 days	Wed 30-09-09	Thu 15-10-09
5	Practical Case Documentation	23 days	Wed 30-09-09	Fri 30-10-09
6	**Problem**	**34 days**	**Thu 15-10-09**	**Mon 30-11-09**
7	Identification of related problems	12 days	Thu 15-10-09	Fri 30-10-09
8	References to the related problems	12 days	Mon 02-11-09	Mon 16-11-09
9	Characterization of the main problem	23 days	Fri 30-10-09	Mon 30-11-09
10	**Solution's Methodology**	**23 days**	**Fri 30-10-09**	**Mon 30-11-09**
11	Purposed Contributions	23 days	Fri 30-10-09	Mon 30-11-09
12	Description of the Purposed Solution	12 days	Sun 15-11-09	Mon 30-11-09
13	Description of the Solution's Methodology	12 days	Sun 15-11-09	Mon 30-11-09
14	**State of the Art**	**48 days**	**Tue 03-11-09**	**Wed 06-01-10**
15	Structure	10 days	Tue 03-11-09	Sun 15-11-09
16	Initial Version	12 days	Mon 16-11-09	Tue 01-12-09
17	Revised Version	26 days	Wed 02-12-09	Wed 06-01-10
18				
19	Write Article	61 days	Thu 15-10-09	Wed 06-01-10
20	Article Submission	1 day	Thu 07-01-10	Thu 07-01-10
21				
22	**Masters Degree Dissertation**	**203 days?**	**Fri 08-01-10**	**Mon 18-10-10**
23	**Contexts**	**127 days?**	**Fri 08-01-10**	**Fri 02-07-10**
24	Identify Possible Contexts	32 days?	Fri 08-01-10	Fri 19-02-10
25	Chose Relevant Set of Contexts	15 days?	Mon 22-02-10	Fri 12-03-10
26	Define Set of Business Processes	20 days?	Mon 22-02-10	Fri 19-03-10
27	Research each Context	70 days?	Mon 22-03-10	Fri 25-06-10
28	Structure and Organize Collected Information	75 days?	Mon 22-03-10	Fri 02-07-10
29	**Analysis**	**134 days?**	**Mon 29-03-10**	**Thu 30-09-10**
30	Initial Analysis of Gathered Information	25 days?	Mon 29-03-10	Fri 30-04-10
31	Context-Specific Business Processes	58 days?	Mon 12-04-10	Wed 30-06-10
32	Context-Specific Information Entities	58 days?	Mon 12-04-10	Wed 30-06-10
33	Common Information Entities	72 days?	Mon 24-05-10	Tue 31-08-10
34	Definition of the Set of Transversal Information Entities	22 days?	Wed 01-09-10	Thu 30-09-10
35	**Solution**	**191 days?**	**Fri 08-01-10**	**Thu 30-09-10**
36	Identification of the Relations Between Entities	68 days?	Fri 08-01-10	Mon 12-04-10
37	Definition of the Set of Desired Viewpoints	58 days?	Fri 08-01-10	Mon 29-03-10
38	Initial Version of the Solution	41 days?	Mon 24-05-10	Mon 19-07-10
39	Revised Version of the Solution	49 days?	Mon 26-07-10	Thu 30-09-10
40	**Conclusion**	**9 days?**	**Mon 20-09-10**	**Thu 30-09-10**
41	Evaluation of Results and Future Work	9 days?	Mon 20-09-10	Thu 30-09-10
42				
43	Write Dissertation	158 days?	Tue 23-02-10	Thu 30-09-10
44	Dissertation Submission	1 day?	Mon 18-10-10	Mon 18-10-10

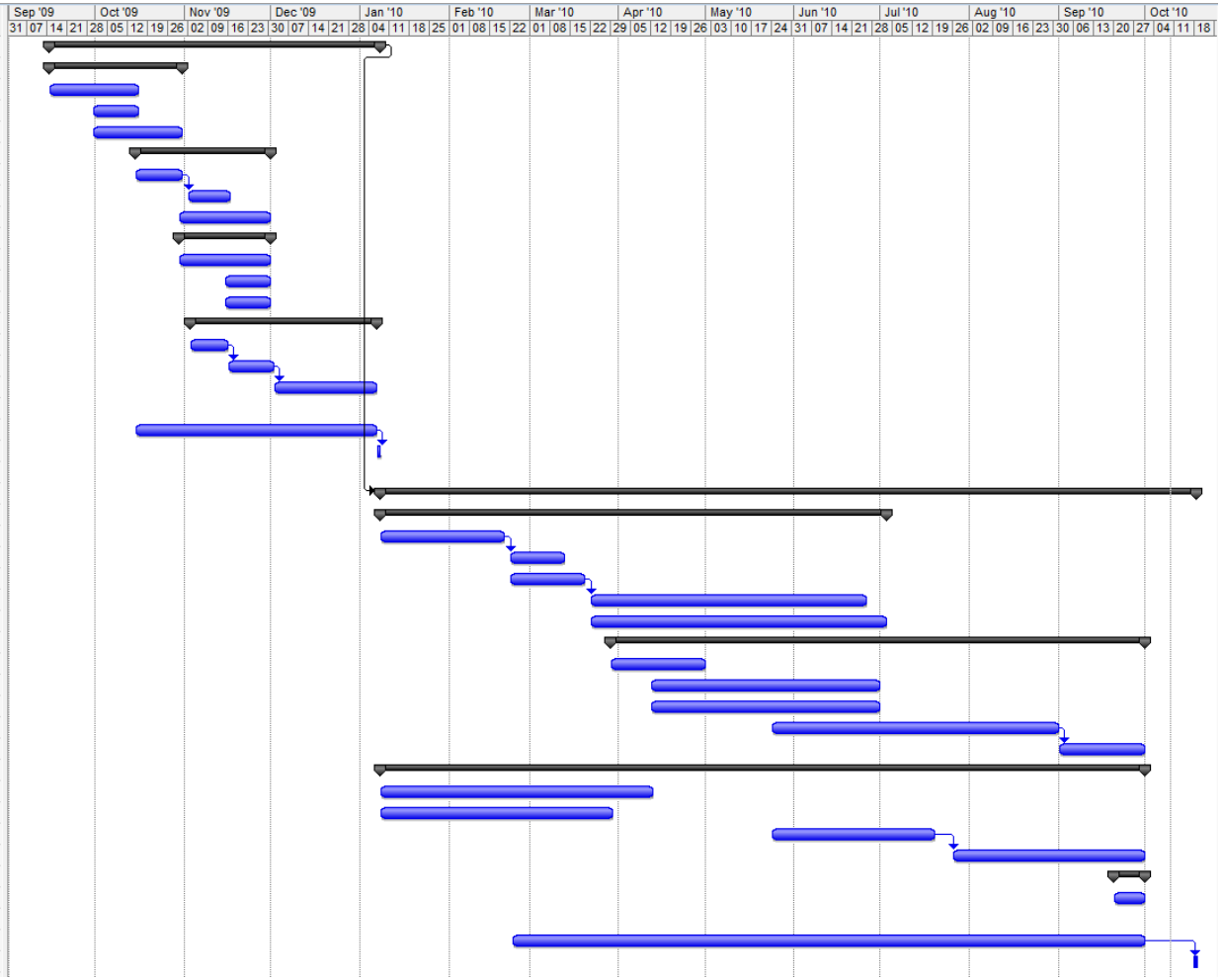


Figure 42: Activities chronogram for the work developed