Data Quality Management in e-Health Integration Platforms: the Case of Uruguay

Silvana Pidre, Laura González, Rodrigo Mendoza,
Mauricio Piñatares, Nicolás Granja, Flavia Serra, Raúl Ruggia
Universidad de la República, Facultad de Ingeniería,
Montevideo, Uruguay, 11300,
{*spidre,lauragon,fserra,ruggia*}@fing.edu.uy

Abstract

Healthcare organizations increasingly need to integrate their software systems with each other in order to exchange clinical data and carry out healthcare processes in coordinated ways. Integration platforms provide mechanisms, based on advanced middleware technologies like the Enterprise Service Bus (ESB), in order to facilitate such integration. In addition, data quality in the healthcare area is critical to take timely and accurate decisions for the health and well-being of patients. This paper proposes a solution for data quality management in the context of the Uruguayan e-health integration platform. The proposal comprises a data quality model, that includes quality characteristics identified from different types of sources, and extensions to an ESB-based integration platform which allows assessing and enforcing data quality requirements in inter-organizational data exchanges. The proposal was completely prototyped which allowed validating the technical feasibility of the proposed solution.

Keywords: data quality, integration platforms, e-health, enterprise service bus.

1 Introduction

The integration of software systems operating in different health organizations is increasingly necessary not only to exchange data but also to implement coordinated health care processes [1]. Such integration is essential, for instance, in order to access the electronic health record of a patient which may be distributed in different health organizations (e.g. laboratories, health service providers, ministry).

Integration platforms [2] constitute an approach to address the issues of inter-organizational systems integration. These platforms are specialized infrastructures providing an intermediate processing layer among organizations in order to facilitate integration issues (e.g. connectivity, heterogeneous data models and formats) [2]. Integration platforms are based on advanced middleware technologies, such as Enterprise Service Bus (ESB) [3]. In Government [4] and Health [5] areas, these platforms can be provided by government entities as part of broader interoperability frameworks. The overall goal is to facilitate and promote the interconnection between organizations providing services that generate economies of scale [4].

In Uruguay, the Salud.uy¹ program aims to strengthen the Nationwide Integrated Healthcare System (Sistema Nacional Integrado de Salud, SNIS) by promoting the conformation of a healthcare network through the use of Information and Communication Technologies (ICT) with the goal of providing citizens with a better access to quality health services. The program has designed, implemented and made available a National e-Health Platform [5] which provides support to the SNIS by allowing health organizations to provide their services in an integrated, complementary and user-centered way. In particular, the platform allows exchanging clinical data among health providers in order to support the National Electronic Health Record (Historia Clínica Electrónica Nacional, HCEN).

On the other hand, the quality of health data is a must for enabling timely and accurate decisions concerning patients' welfare and health. The relevance of this aspect is due to its impact on the improvement of patient care, on the definition of health policies for the population as well as on the management of funds for the maintenance of health services [6]. In order to improve the quality of the collected data and the information generated based on them it is necessary to implement quality control measures [6]. One of

¹<https://hcen.salud.uy/>

the first steps to achieve this is to identify and classify quality characteristics, which are usually organized hierarchically and documented in quality models.

While health organizations should implement these controls internally, there are a number of reasons motivating the implementation of such controls within an integration platform, in particular, when organizations exchange data. First, organizations participating in a health system at a national level are generally autonomous and diverse, so that data quality criteria may vary from one to another. Secondly, some quality requirements may refer to aspects of the entire health system so that their control may require data which are available only on the platform (e.g. previously exchanged data). Finally, it may be convenient for an integration platform to implement data quality controls so that organizations can delegate part of these controls to it, thus generating economies of scale.

Although integration platforms, particularly those based on ESBs, provide mechanisms that can be used to control the quality of the data exchanged through them, they do not provide native solutions to address this issue [7] neither in a general way nor for the health area.

This paper proposes a solution for data quality management in the context of the Uruguayan e-health integration platform. The proposal comprises a data quality model, that includes quality characteristics identified from different types of sources (e.g. international standards, Uruguayan regulations, experiences in other countries), and extensions to an ESB-based integration platform which, leveraging native ESB mechanisms (e.g. data transformation, routing), allows assessing and enforcing data quality requirements in inter-organizational data exchanges. The proposal was completely implemented with the Switchyard product which allowed validating the technical feasibility of the proposed solution.

This paper is a substantially extended and thoroughly revised version of [8]. Additional material includes:

- analysis of additional sources with the goal of identifying new data quality characteristics, in particular, experiences from other countries were included in Section 3.2.4
- an enhanced data quality model with new quality factors and metrics mainly identified from the additional information sources
- a detailed description of the architecture and design of the prototype

The rest of the paper is organized as follows. Section 2 presents a conceptual framework relevant to the proposal while Section 3 presents a data quality model applicable to the context of the Uruguayan health platform. In Section 4 extensions to an ESB-based integration platform to manage data quality are proposed. Section 5 describes a case study and the results of non-functional tests. Finally, Section 6 analyses related work and Section 7 presents conclusions and future work.

2 Background

This section presents a conceptual framework relevant for this work.

2.1 Data Quality in Health

Health information includes structured data collected from individual patients and summarized information about patients' experience with the health service provider [9]. They constitute a formal representation of a clinical event or patient's characteristic [6]. This data has special features given its sensitivity as well as the importance of having accurate, timely and accessible information for providing adequate health care services to the population.

The quality of this information is crucial for decision making and for defining public policies, as well as for the elaboration of statistics, epidemiological surveillance, service planning and distribution of resources. For health providers, good quality information enables the medical care coordination among the involved services and improves their quality, management and efficiency as well as the budget management for the continuity of the services [6].

A number of factors impact on the quality of the data used to generate health information, notably not using standards and the lack of a common vocabulary which would enable data legibility and sharing as well as data interpretation by all actors. Mistakes in the design of data collection forms also impact on data quality (e.g. paper or electronic) as well as the lack of controls in the entry of data and lack of trained personal, among others.

2.2 Data Quality Models

Identifying and classifying quality characteristics (e.g. syntactic correctness) is one of the main data quality management activities. These characteristics are usually organized hierarchically and documented through quality models.

The organization of data quality models vary according to different aspects. For instance, the work in [10] proposes a quality meta-model defining an organization based on dimensions, factors, metrics and methods. Dimensions capture a quality facet at a high level (e.g. accuracy) while factors represent a particular aspect of a dimension (e.g. syntactic correctness). Metrics are used to measure a quality factor (e.g. date format) and methods are processes implementing a certain metric (e.g. a program that checks whether a date has a certain format).

2.3 Enterprise Service Bus

An Enterprise Service Bus (ESB) is a standards-based integration platform which combines Web Services, messaging, data transformation and intelligent routing in order to reliably implement the interaction between software components with transactional integrity [3].

ESBs provide an intermediate layer with reusable integration capabilities in order to enable the interaction between clients and services in a SOA. ESBs receive message-based requests on which they perform mediation operations to overcome client-server heterogeneities.

Although actual functionalities provided by ESB may vary according to the vendor, they all provide similar capabilities regarding connectivity, message transformation, intelligent routing, mechanisms for implementing mediation flows, asynchronous messaging, monitoring and administration, among others.

Message transformation allows, for example, the interaction between applications using different formats or data models [3]. Most ESBs allow specifying and executing transformations with the XSLT² standard.

Intelligent routing is the capacity by which the ESB determines, at runtime, the destination of a message according to different factors. Some types of intelligent routing are content-based routing and itinerary-based routing [3]. In particular, content-based routing determines the destination of the message based on its content (e.g. a value in the message header, a value in the business data) [3].

Mediation flows allow specifying a sequence of mediation operations (e.g. transformation, routing) to be executed in messages. These flows are designed for simple and short-term processes, mainly addressing issues of integration and communication [11]. Several ESB products allow specifying mediation flows based on the Enterprise Integration Patterns (EIP) [12].

2.4 Standards in Health

Standardization in the health area has been motivated by the need of sharing clinical information with heterogeneous representations among different organizations. Standards are intended to homogenize the information as well as the interpretations about it in such a way that it can be understood and interpreted without errors.

Two important organizations that are creating and disseminating standards are *Health Level Seven* (HL7) *International* [13] and *International Health Terminology Standards Development Organization* (IHTSDO) [14].

In particular, HL7 developed the standards HL7 [13] and *Clinical Document Architecture* (CDA) [15]. HL7 is a messaging standard for the electronic exchange of clinical information, while CDA provides a model for the exchange of clinical documents. The last one enables documents to be electronically processed as well as easily found and used.

In turn, IHTSDO develops a dictionary providing codes, terms, synonyms and definitions used in clinical documentation and reports: *Systematized Nomenclature of Medicine Clinical Terms* (SNOMED CT) [14]. SNOMED is a collection of clinical terms systematically organized and electronically processable.

Finally, the standard of International Classification of Diseases (ICD) [16] was proposed by the World Health Organization for coding causes of morbidity and mortality and in order to be able to compare values at the National and international level. The ICD has become the international standard for health management and epidemiological studies. The current version of ICD is 11.

3 Health Data Quality Model

This section describes the Uruguayan e-Health Platform and the design of the proposed Data Quality Model for this context. In particular, the design of the model comprised three phases:

²<https://www.w3.org/TR/xslt>

- identification of data quality characteristics
- organization of the characteristics in dimensions and factors
- definition of data quality metrics for the proposed quality factors

It should be noted that the model is meant to be extensible, so its quality characteristics are not thought to cover all aspects of data quality within the context. On the contrary, the model includes quality characteristics which were considered the most relevant and representative for the scope of this work.

3.1 Uruguayan e-Health Integration Platform

The development of the National Electronic Health Record (HCEN) in Uruguay was proposed as one of the objectives of the Digital Agenda 2011 - 2015 [17]. In this context, the Salud.uy Program emerged as an agreement between the e-Government Agency in Uruguay (Agencia para el Gobierno Electrónico y la Sociedad de la Información y Conocimiento, AGESIC) [18], Presidency of the Republic, the Ministry of Economy and Finance, and the Ministry of Public Health. Salud.uy aims to propose solutions for a fluid and systematic integration of health services, in particular, for the exchange of clinical information, among participating institutions of the National Integrated Health System (SNIS). Salud.uy focuses on the development of the HCEN and on the creation of a National Bank of Health Registries [5].

In order to support the development of the HCEN, AGESIC and Salud.uy are working on the construction of an e-Health Integration Platform that leverages the Interoperability Platform (PDI) [4] launched by AGESIC and addresses specific problems of the Health area. This platform aims to allow the exchange of clinical information among the different actors in an efficient and effective manner as well as to provide specific services for the area. Figure 1 presents in a simplified way the Uruguayan e-Health Integration Platform.

Figure 1: Uruguayan e-Health Integration Platform

Health organizations exchange health information through the platform leveraging the web services technology. In particular, the exchange of information is performed through XML messages following the SOAP, WS-Security, HL7 and CDA standards, among others.

This exchange of messages must comply with requirements, particularly in the area of data quality, which may arise from laws, standards and guidelines established by Salud.uy and AGESIC, among others. These requirements can apply to the entire message (e.g. the message must be digitally signed), to a certain element of the message (e.g. the value of the element must be a SNOMED term), to a set of message elements (e.g. the value of the city element must correspond to a city of the country specified in the country element) or to a set of messages (e.g. sent by different organizations or the same organization at different times).

3.2 Identifying Data Quality Characteristics

This section presents the way in which data quality characteristics were identified for the context presented in Section 3.1. The identification process comprised the review and analysis of documentation coming from four types of sources:

- Health organizations
- Uruguayan regulations
- Works in the data quality area

- Experiences in other countries

Figure 2 presents the different types of sources that were reviewed and analyzed.

Figure 2: Identifying Data Quality Characteristics

Data quality characteristics were identified for each one of the sources. Some of them (e.g. Uruguayan regulations) specify data quality requirements (e.g. using specific values for a given data field) which lead to data quality characteristics (e.g. syntactic correctness). In other sources (e.g. works in the data quality area) data quality characteristics were directly identified.

Through a selection process, taken as input all these characteristics, a consolidated set of data quality characteristics for the health area was obtained. The selection of characteristics was based on the recurrent use of the characteristics throughout the different sources, the feasibility of its measurement within data exchanges in integration platforms and the relevance of the characteristic for the Uruguayan e-Health Platform.

3.2.1 Health Organizations

In order to identify data quality requirements, information sources from the following health organizations were analyzed: Health Level Seven (HL7), International Health Terminology Standards Development Organization (IHTSDO) and World Health Organization (WHO). The analysis of information sources coming from HL7 and IHTSDO leads to data quality requirements which specify, for example, that messages exchanges have to be compliant with HL7 and CDA as well as use ICD (International Classification of Diseases) codification and SNOMED terminology. These requirements lead to characteristics in the area of syntactic accuracy and compliance. Information sources coming from WHO [6] directly identify data quality characteristics such as accuracy, timeliness and completeness.

3.2.2 Uruguayan Regulations

The analysis of Uruguayan regulations comprised regulations coming from AGESIC (i.e. the e-Government agency in Uruguay), current Uruguayan laws and decrees, the Salud.uy program [19] (i.e. the e-Health initiative in Uruguay) and the Ministry of Public Health (Ministerio de Salud Pública, MSP) [20]. Some data quality requirements identified from these sources are: i) the use of Data Reference Models to exchange information regarding citizens and addresses [21], ii) the compliance with national laws and decrees (e.g. Act N° 18.600 [22] which acknowledges legal effects and validity of e-documents and e-signature), iii) the use of ICD-10 (i.e. version 10 of ICD) for specifying information regarding mortality and morbidity causes and, iv) compliance with the established minimal data sets³ for registering information of specific health events (e.g. patient discharge). From these requirements, data quality characteristics in the areas of syntactic accuracy, completeness and security were identified.

³<https://hcen.salud.uy/web/pub/conjunto-minimo-de-datos>

3.2.3 Works in the Data Quality Area

The analysis of works in the data quality area included academic works [23] [24] as well as industry standards such as ISO/IEC 25012:2008 [25]. The analysis of academic works led to the conclusion that there is not agreement in the set of relevant data quality characteristics to consider in the health area. This may be due to the fact that authors follow different methodologies to identify characteristics. In addition, throughout the sources authors use the same name for referring to characteristics with different meaning and use different names for referring to the same characteristic. Some of the characteristics identified from this analysis are: relational integrity in the consistency area and timeliness in the freshness area. In addition, some of the characteristics identified from the ISO/IEC 25012:2008 standard are compliance and syntactic correctness.

3.2.4 Experiences in Other Countries

The analysis of experiences in other countries aims to identify data quality characteristics considered in nationwide initiatives in the health area such as the one in Uruguay. The analysis focuses on sources from nationwide organizations with government support.

Table 1 presents the organizations that were considered in the analysis.

Table 1: Experience in Other Countries

Country	Acronym	Name	Description
Canada	CIHI	Canadian Institute for Health Information (https://www.cihi.ca/en)	Independent and not-for-profit organization that provides essential information on Canada's health systems and the health of Canadians.
Irlanda	HIQA	Health Information and Quality Authority (https://www.hiqa.ie)	Independent authority, reporting to the Minister for Health and the Minister for Children and Youth Affairs, established to drive high-quality and safe care for people using health and social care services in Ireland.
Australia	WA Health	Western Australian Health (http://ww2.health.wa.gov.au)	Western Australia's public health system. Its public health programs ensure the delivery of a safe, high quality and sustainable health system for all the population.
Australia	-	Australian Digital Health Agency (https://www.digitalhealth.gov.au)	This Agency is responsible for all national digital health services and systems. Its focus is on putting data and technology safely to work for patients, consumers and the healthcare professionals.
Australia	ACT Health	Australian Capital Territory Health (http://www.health.act.gov.au)	It is in charge of planning the health services of the Australian Capital Territory, financed with public funds, with the aim of providing safe, effective and patient-centered care.

The analysis of these experiences showed that each country takes as a base the experience of another country with the same characteristics (e.g. at the level of territory, political division, language, health system) to build its quality framework and to identify quality characteristics. Some of the characteristics identified from the review of these experience are: relational integrity, integrity over time, organizational integrity, inter-organizational integrity and coverage.

3.3 Organizing Data Quality Characteristics

After obtaining the relevant data quality characteristics, they were organized in a Data Quality Model with an structure based on the concepts presented in Section 2.2.

The Data Quality Model comprises the Dimensions and Factors defined in Table 2 and Table 3, respectively. The description of dimensions and factors results from the analysis of the definitions in the different sources, taking as a result the one that was considered most appropriate for the context of this work.

3.4 Defining Metrics

The last step in the design of the model consisted in defining metrics for the identified quality factors indicating for each of them: how it is measured (i.e. semantics), its unit (e.g. Boolean) and the measurable object to which it applies (i.e. granularity).

Since the focus of this work is evaluating the quality of the data included in the messages that are exchanged through the platform, four measurable objects were defined: the message that is exchanged, an element within the message, a set of elements and a set of messages. The granularity of each metric is then specified based on these objects.

Table 2: Data Quality Dimensions

Dimension	Description - Questions that answers
Completeness	Does the data exchanged match the actual set of entity data?
Freshness	Is the information up to date with respect to a established update frequency?
Security	Does the exchanged data is protected against illegal access and manipulation?
Interoperability	Does the data exchanged comply with the defined interoperability standards and guidelines?
Consistency	Can data integrity be assessed by comparing different sources and over time?
Accuracy	How close is the value of the exchanged data to the actual value? Does the data exchanged comply with the rules and definitions defined for the type of data?

Table 3: Data Quality Factors

Dimension	Factor	Description	Sources
Completeness	Density	The information of the entity must be complete.	[21] [19] [6] [26] [25]
	Coverage	The data must include the set of data describing a specific entity.	[27] [28]
	Minimal Data	The minimum data defined for the health event must be present.	[21] [19] [6] [26]
Freshness	Timeliness	The data must be timely for its purpose.	[24] [6] [26] [27] [29] [30] [28]
	Currency	The data must be current.	[26] [29] [6] [25]
Security	Data Safety	The data must not be altered.	[22] [23] [25] [31]
Interoperability	Compliance	Data must comply with the specified standards.	[14] [15] [13] [16] [6] [26] [27] [29] [30] [28] [25]
Consistency	Relational Integrity	Data must be consistent with each other.	[26] [27] [29] [30] [28] [24] [6] [25]
	Time Integrity	Data must be consistent over time.	[26] [27] [29] [30] [28]
	Organizational Integrity	Data sent by an organization must be consistent in a period of time.	[26] [27]
	Inter-Organizational Integrity	Data sent by different organizations must be consistent in a period of time.	[26] [27]
Accuracy	Syntactic Correctness	Data must represent valid values within a domain and established rules.	[16] [20] [14] [21] [19] [23] [6] [26] [27] [29] [30] [28] [25]
	Semantic Correctness	Data must represent the reality.	[24] [26] [27] [29] [30] [28] [6] [25]
	Precision	Data must be represented with the established precision.	[26] [6]

Table 4 presents the set of metrics that were identified for this work. This set aims to be descriptive regarding the metrics that the model could include.

Table 4: Data Quality Metrics

Dimension	Factor	Metric	Semantic	Unit	Granularity
Completeness	Density	Nulls()	Is data equal to the null value?	Boolean	Elemento
		EmptyInteger()	Is data equal to 0?	Boolean	Element
		EmptyText()	Is data equal to ""?	Boolean	Element
	Coverage	PopulationCoverage()	Percentage of people of the Health System included in messages	Percentage	Set of messages
	Minimal Data	EmergencyMinimalData()	For a message describing an Emergency event, are all the minimum data included?	Boolean	Element
Freshness	Timeliness	DataAge()	Difference between the date of the data in the message and the current date. It is the same metric used for the Currency factor.	Day	Element
	Currency	DataAge()	Difference between the date of the data in the message and the current date. It is the same metric used for the Currency factor.	Day	Element
Security	Data Safety	SignedMessage()	Is the message digitally signed?	Boolean	Message
		SignedElement()	Is a specific element in the message digitally signed?	Boolean	Element
Interoperability	Compliance	IsCIE10Code()	Does the element represent a valid ICD10 code?	Boolean	Element
		IsSNOMEDTerm()	Does the element represent a valid SNOMED term?	Boolean	Element
		HL7Element()	Does the message contain valid HL7 elements?	Boolean	Message
		CDAStructure()	Has the clinical document included in a message the CDA structure?	Boolean	Message
		CDALevel()	Return the CDA level of a clinical document included in a message (1, 2 or 3)	Integer	Message
Consistency	Relational Integrity	DepartmentCityIntegrity()	Does the city included in the message correspond to a city of the specified department?	Boolean	Set of Elements (Department, City)
		DiseaseSexIntegrity()	Is the disease indicated in the message consistent with the specified sex?	Boolean	Set of Elements (Disease, Sex)
	Time Integrity	EqualValueOverTime()	Is the data value the same in messages sent at different times?	Boolean	Set of Messages
		LowerValueOverTime()	Is the data value lower in messages sent at different times?	Boolean	Set of Messages
		GreaterValueOverTime()	Is the data value greater in messages sent at different times?	Boolean	Set of Messages
	Organizational Integrity	OrganizationlEqualValue()	Is data the same in different messages sent by an organization within a time period?	Boolean	Set of Messages
	Inter-Organizational	InterOrganizationlEqualValue()	Is data the same in different messages sent by different organizations within a time period?	Boolean	Set of Messages
Accuracy	Syntactic Correctness	DateFormat1()	Does the data comply with the format yyyy-mm-dd?	Boolean	Element
		SexDomain()	Is data within the domain 1, 2 and 9 whose values indicate Male, Female and Not Applicable, respectively?	Boolean	Element
		CountryISOAlpha3()	Is the data within the domain defined by ISO 3166-1?	Boolean	Element
	Semantic Correctness	DateBeforeToday()	Is the date less than or equal to today's?	Boolean	Elemento
		DateAfterToday()	Is the date equal to or greater than today's date?	Boolean	Element
		DistanceFromToday()	Distance between a date and today	Day	Element
	Precision	DecimalPrecision()	Number of decimals in the data	Integer	Element

4 Data Quality aware ESB-based Integration Platform

This section describes the proposed Data Quality aware ESB-based integration platform that allows evaluating and enforcing data quality requirements in messages exchanges between health organizations.

4.1 General Description

The general idea of the proposal is to extend the capabilities offered by an ESB-based Integration Platform in order to be able to monitor, control and ensure the quality of data in the messages that organizations exchange through it. Figure 3 presents how the Data Quality solution is included in an ESB-based e-Health Integration Platform.

Figure 3: Data Quality aware ESB-based e-Health Integration Platform [32]

The main components of the data quality solution are:

- *Gateway*: Component in charge of centralizing the access of clients to the services provided by organizations through the platform, as well as sending them the responses.
- *Backoffice*: Web application which allows managing the organizations that are integrated to the platform and the web services they publish. It also allows managing Data Quality Models in the platform as well as configuring Actions to be carried out when it is detected that a data quality requirement is not met.
- *Services Registry and Configuration*: Component that stores configuration data and metadata of the services that are published in the platform.
- *Quality Models*: Quality Models have evaluation methods which are in charge of evaluating if a quality requirement is met or not within a certain set of elements of a message. These elements are part of request and response messages exchanged when invoking operations of the services exposed in the platform.
- *Actions*: Different types of actions can be taken if the application of methods results in a non-compliance of data quality requirements. These actions range from notifying an administrator (e.g. SMS, email, etc), filtering the message by responding to the client with the corresponding error, or executing mechanisms to try to ensure quality, such as transforming messages.

Figure 4: Ensuring Data Quality within an ESB-based Integration Platform [32]

Figure 4 presents an example in which the general operation of the platform is described. In this example, a Uruguayan government organization (i.e. Dirección General de Registro de Estado Civil, DGREC) invokes the operation `getCertificate`, provided by the Ministry of Public Health (MSP), which returns a death certificate. There is also a Quality Model where a Measurement Method is specified that determines whether the format of the `deathDate` element, which is an input parameter of the aforementioned operation, has the format `yyyy-mm-dd`. Given that in this case the message does not comply with this control, the Actions associated with the Method are executed: an email is sent notifying the non-compliance with this quality control and an XSLT transformation is executed to the message in order to transform the format of the date to the expected one.

The complete flow that is carry out in the example is described as follows: DGREC sends a request to the platform in order to invoke the MSP service (1). The Gateway performs validations and logs the request message. In this case, depending on the operation and its elements, it is detected that MET1 method must be evaluated, so the message is redirected to the corresponding Method (2). The Method is then executed and, given that it does not comply with the format requirement, the message is redirected to the first of the Actions associated with it, which in this case corresponds to the sending of an email notification (3). Then the request is sent to the component responsible for executing the next Action, where an XSLT transformation (4) is carried out. The final message is sent to the service that the original client wants to consume (5). Once the MSP finishes processing the request, it returns the response to the platform (6), and since there are no configured methods for the elements of the response message, it is sent to DGREC by means of the Gateway(7).

4.2 Conceptual Model

Figure 5 presents the conceptual model of the proposed solution.

In particular, organizations publish web services that may have several operations. Each operation has a set of input and output elements. Each element can be associated with evaluation methods that return true or false according to criteria that are specified in the definition of the method. If the method returns a false value it means that the element does not meet some of the data quality requirements.

The system allows the specification of evaluation methods in two ways: defining rules or using custom methods.

The specification of methods using rules allows a user without programming knowledge to implement methods by means of the combination of comparison rules, set functions, logical operations and / or regular expressions. For example, a user can define a set of rules, regarding the elements of a message, which specify the logical relationship that they must fulfill in order to return a true value. As an example, two rules are presented in the Table 5: the first specifies that the value of the `sex` parameter must be contained in the list "1, 2, 9" (Male, Female, Not applicable) and the second specifies that the Document Parameter must have a value with a length of less than 9.

Table 6 presents a quality method for the metric with identifier 2 which returns true when the value of the `Sex` element is contained in the list "1, 2, 9" and the length of the value of the element Document is

Figure 5: Conceptual Model [32]

greater than or equal to 9. Finally, for each of the relationships between the set of elements and method, a serie of actions is defined, which will be executed in case the specified method returns a false value (e.g. send an e-mail).

Table 5: Examples of Rules

Rule ID	Operation	Parameter	Value	Method
R1	Content	Sex	1, 2, 9	1
R2	LengthLessThan	Document	9	1

Table 6: Example of Quality Method

Relationship between rukes	Metric	Action
R1 and (not R2)	2	3

In turn, the specification of custom methods implies that an advanced user should implement a software component with the logic for evaluating a quality requirement and should create a custom method in the system referencing this software component.

4.3 Logical Architecture

This section presents and describes the logical architecture of the platform. In particular, Figure 6 presents the main components of the solution as well as the interaction between them and the native ESB mechanisms they leverage (e.g. validate, transform, dynamic routing).

The main components of the solution are described as follows:

- *Gateway*: The Gateway is the component in charge of encapsulating the access to the platform and it is used to centralize the validation and transformation of messages to the required format within the ESB-based integration platform. All messages processed by the platform pass through the Gateway which initiates their processing. In particular, the Gateway carries out validation procedures in order to verify that messages comply with some requirements (e.g. it is validated that messages are compliant with the SOAP 1.2 standard and that WS-Addressing headers specify the service and operation to be executed as well as the organization that sends the message).
- *Core*: The Core component receives messages, which were already validated and transformed, from the Gateway (request messages) as well as from the Dispatcher (response messages). These messages are sent to the Router which after processing them returns them back to the Core. According to the response sent by the router, the Core evaluates if it should continue processing the message (i.e. data quality requirements were met or some action could be applied to meet them) or if it has to cancel it

Figure 6: Logical Architecture [32]

(i.e. data quality requirements were not met). If the processing of the message has to continue, the Core routes it to the Dispatcher (if it is a request) or to the Gateway (if it is a response) to send the response to the client that invoked the service. If the message has to be canceled, the Error Handler generates a response informing the reasons of the cancellation and sends it to the client through the Gateway.

- *Router*: This component is responsible for routing the message according to its content. If there are Pending Methods to execute on the message, it routes it to the Method Router. On the other hand, if Actions must be executed, it routes it to the Action Router. In case there are no Methods or Actions to apply, the message is returned to the Core.
- *Method Router*: Messages which require the application of Methods are routed to the Method Router which invokes the quality validation component configured for that Method (Quality Validator) and updates the message to continue with the rest of the Actions or Methods to be executed.
- *Action Router*: When a message fails the validation with a given Method, the message is routed to the Action Router where the Actions that must be executed are loaded. Then, they are executed sequentially on the received message.
- *Quality Validator*: It executes the Methods that were defined in the Quality Model and it is in charge of updating the internal structure of messages according to the results of those evaluations.
- *Actions*: It executes the different types of Actions supported by the platform (e.g. cancel the invocation, send notifications or alerts, transform a message). The platform is designed in such a way that it is possible to include new types of Actions as needed.
- *Dispatcher*: It obtains from the internal structure of the message the destination in order to send the message to the Service Invoker. When the Dispatcher receives the response from the Service Invoker it routes it to the Core.
- *Service Invoker*: It performs invocations to external services according to the message and destination indicated by the Dispatcher. After receiving the response, this component returns it to the Dispatcher.
- *Error Handler*: It generates response messages when an error must be returned, either because the invocation is incorrect, or because the message does not comply with the specified quality requirements.
- *Logger*: It logs data regarding the exchanged messages as well as the results of the evaluations performed by the Methods associated with those messages. Those data may be use for creating reports and graphs.

4.4 Implementation Details

This section presents implementation details of the prototype which was developed leveraging Java EE 7 and using the JBoss EAP 6.4 platform. A demonstration of the main functionalities provided by the prototype is available on line⁴.

Figure 7 presents the main products and technologies used to implement the main components of the solution.

Figure 7: Prototype Technologies and Products [32]

Apache Maven was used to define the modular structure of the project, compile and manage the dependencies to libraries. PostgreSQL was used to store all the structures managed by the Backoffice, such as Quality Models, organizations, web services and logs.

Hibernate was used for object-relational mapping in order to facilitate the access to the data handled by the platform.

Apache Camel is an open source Java framework focused on facilitating the integration of applications on different platforms. It provides implementation of the most used Enterprise Integration Patterns. In particular, this framework was used to perform the routing of messages in the platform.

Switchyard has the typical features of an ESB and it is integrated with Apache Camel in order to provide routing functionalities and message transformation, among others.

Java Server Faces (JSF) is a Java standard for the development of server-side web user interfaces⁵. Primefaces is an open source library, which has a set of rich components for JSF. JSF and Primefaces were used for the implementation of the Backoffice.

5 Case Study and Tests

This section presents a case study and the results of non-functional tests.

5.1 Case Study

The case study is based on the Electronic Death Certificate Consultation service⁶ provided by the MSP and describes how messages received by the platform are modified, in order to ensure compliance with established quality requirements.

One of the operations of the service returns a Certificate and receives as input parameters: certificateNumber, documentType, documentCountry and documentNumber. In the context of the case study, it is considered that a quality requirement has been defined for the documentCountry element based on the CountryISOAlpha3 metric of the Syntactic Correctness factor defined in the model presented in Section 3. The quality requirement establishes that the value of that metric must be true for that element (i.e. the value of the documentCountry element must be in the list of countries defined in ISO 3166-1). It is also considered that an action to be executed, in case that the quality requirement is not met, is defined: an

⁴<https://www.fing.edu.uy/inco/grupos/lins/demos/dq-esb.mp4>

⁵<http://www.oracle.com/technetwork/java/javase/javaserverfaces-139869.html>

⁶<https://www.agesic.gub.uy/innovaportal/v/5333/9/agesic/consulta-certificado-de-defuncion-electronico.html>

XSLT transformation that changes invalid values of the country document element for valid values according to ISO 3166-1.

Figure 8 presents the body of the SOAP message that the platform receives for invoking the aforementioned operation.

```
<soapenv:Body>
  <con:getCertificate>
    <certificateNumber>1000</certificateNumber>
    <documentType>68909</documentType>
    <documentCountry>URUGUAY</documentCountry>
    <documentNumber>44072453</documentNumber>
  </con:getCertificate>
</soapenv:Body>
```

Figure 8: SOAP Message - *Request*

Given that the message does not meet the quality requirement, the transformation shown in Figure 9 is executed so that the value of the element meets the requirement established before the service is invoked.

```
<xsl:stylesheet version="1.0" .....>
  <xsl:template match="node()|@*">
    <xsl:copy>
      <xsl:apply-templates select="node()|@*" />
    </xsl:copy>
  </xsl:template>
  <xsl:template match="documentCountry/text()[.='URUGUAY']">
    URY </xsl:template>
</xsl:stylesheet>
```

Figure 9: XSLT Transformation

Figure 10 presents the body of the SOAP message after the transformation. This way, the MSP service receives a valid request message according to the established requirements.

```
<soapenv:Body>
  <con:getCertificate>
    <certificateNumber>1000</certificateNumber>
    <documentType>68909</documentType>
    <documentCountry>URY</documentCountry>
    <documentNumber>44072453</documentNumber>
  </con:getCertificate>
</soapenv:Body>
```

Figure 10: SOAP Message - *Request*

Finally, Figure 11 presents the body of the SOAP message that a client would receive if the quality requirement is not met and a cancel action is configured instead of a transformation action.

5.2 Non-functional Tests

Tests were performed in order to evaluate the overhead that the proposed solution introduces in the response time of the services. In particular, five types of invocations were executed to a service: i) direct invocation (i.e. without passing through the platform), ii) invocation through the platform but without applying quality assessment methods, iii) invocation through the platform applying quality assessment methods with data that met the quality requirements, iv) invocation through the platform applying evaluation methods with data in the request that did not meet the quality requirements, and v) invocation through the platform applying evaluation methods with data in the response that did not meet the quality requirements.

Five clients were simulated executing requests to the web service concurrently in intervals of 1000ms during 60 seconds. The processing times were registered for each type of invocation in order to compare them. The average time and the overhead generated by the solution in each case is presented in Table 7.

The results of the tests shows that the platform by itself introduces an approximate overhead of 118ms (invocation type 2). This is due to the fact that once a message is received by the platform, a fixed processing

```

<soap:Fault>
  <soap:faultcode>400</soap:faultcode>
  <soap:faultstring>
    The quality requirement is not met.
  </soap:faultstring>
  <soap:detail><error>
    1152 – CountryISOAlpha3: documentCountry does not comply with the required format.
  </error></soap:detail>
</soap:Fault>

```

Figure 11: SOAP Error Message - *Response*.

Table 7: Overhead with the Different Types of Invocations

ID	Type of Invocation	Average (ms)	Overhead (ms)
1	Direct Invocation	7	-
2	Without Applying Assessment Methods	125	118
3	With Methods (request is not compliant)	132	125
4	With Methods (data are compliant)	256	249
5	With Methods (response is not compliant)	268	261

time is required in order to perform validation operations regarding the format and content of the message. In addition, there is a difference between the types of invocation without methods (i.e. invocation type 2) and the one with quality problems in the request (i.e. invocation type 3) with respect to those that do not have quality problems (i.e. invocation type 4) or that have quality problems in the response (i.e. invocation type 5). This may be explained given that for this test in invocations of type 2 Methods are not evaluated and in invocations of type 3 only seven methods are evaluated (i.e. the one applied to a message request that the service consumes). On the contrary, in invocation types 4 and 5, 34 Methods (that apply to the elements of the request or the response) are evaluated and there is also the additional time for invoking the target service.

On the other hand, Figure 12 shows a great similarity in terms of response times of messages not canceled and canceled in the response. This allows verifying that canceling a message almost does not imply overhead if the same amount of Methods are evaluated.

Figure 12: Overhead [32]

When running the tests, two clear flows can be distinguished by which a message can pass in the solution:

monitoring and validation. The monitoring flow does not add any logic to the message that arrives at the solution, rather than logging a record and routing it to the destination. On the other hand, validation requires the execution of various calculations and evaluations, which determine which Methods and Actions should be invoked.

The tests shows that the overhead increases with a very low exponent and that a large number of Methods / Actions (more than 30 for a message) is needed so that the response time exceeds the linear range. The overhead is generated mostly by having to execute Methods and Actions, rather than by the amount of them, by involving the components that are responsible for performing these tasks in the flow of the message. It is worth mentioning that part of the overhead could be significantly decreased (up to 20 %) by using fixed Switchyard components, and not invoking them dynamically according to the context or content of the message. In any case, implementing a dynamic invocation of them gives the flexibility to introduce new components in a more agile way than if these invocations could not be configurable.

The tests performed allows concluding that the proposed solution does not introduce an overhead that could originate important limitations in its use.

6 Related Work

This section presents other work related to our proposal.

Firstly, some works address the problem of data quality in the context of an ESB [7] [33] and propose solutions to tackle these issues by using such middleware infrastructure. For instance, in [7] authors consider that data quality management is particularly challenging when there is a large number of data sources. To address this problem authors claim that data quality process must move from the user level to the integration server. Unlike our proposal, [7] does not address data quality requirements in inter-organizational interactions and it mainly focuses on extending an ESB to provide solutions that allow the selection of data sources according to quality criteria based on requests for user data. This work neither address data quality issues specific to the healthcare domain.

In turn, ensuring data quality in the healthcare domain is a recurring research topic [34] [35] [36] [37], in particular, when it is necessary to interconnect heterogeneous systems in order to share health information. In [36] authors analyze the quality of the data in HL7 messages and they conclude that the HL7 standard is not always correctly used, because custom structures used to transport data generate inconsistencies when integrating such data. This work also proposes mechanisms to monitor data quality in this kind of messages. In [35], the authors consider data from a pervasive Health environment where data sources include wireless patient monitoring sensors. According to the authors, in a pervasive environment data can be seen from many perspectives that include quality dimensions (e.g. completeness, freshness, credibility) and its context. Although these works, like others in the area [6] [38], address data quality issues in a healthcare domain, none of them propose a quality model and neither present proposals to monitor and enforce data quality requirements in exchanges of inter-organizational information, in particular, using an ESB.

In addition, there are several works that propose using middleware platforms to address data quality issues. In [37] the authors present an application context similar to our proposal (i.e. organizations that exchange health data), in which an intelligent health platform is used for the communication between organizations. The platform, which addresses data quality issues, is based on a Service-Oriented Architecture (SOA) and it is implemented with different components including an ESB. In [39] users' quality requirements and potential conflicts among them are analyzed. The authors argue that such conflicts must be resolved automatically and for this, they propose a middleware based on configurable components. In turn, in the context of Internet of Things (IoT) [40] [41], data quality is considered a fundamental issue. For example, in [40] authors introduce a framework within a distributed middleware platform to address issues of privacy and data quality. In [41] the authors consider that a reconfigurable distributed system middleware can be applied in heterogeneous, unreliable, dynamic and large-scale systems environments. In particular, they propose a policy-driven middleware to address legal requirements that include data quality issues. Meanwhile, the works [42] [43] address the management of data quality in the exchange of master data using the new ISO 8000 standard. These papers propose the use of SOA to implement the control of such standard, especially parts 100 to 140, which specifically deal the exchange of master data.

Unlike our proposal, these works do not propose a data quality model for the healthcare domain and the capabilities of an ESB are not used to monitor and ensure compliance with data quality requirements in the exchange of information.

7 Conclusions and Future Work

This paper presents an ESB-based service integration platform specialized in data quality control in the health area. This platform evaluates and enforces data quality requirements in cross-organizational processes.

Specifically, the proposed platform consists of an extended ESB with monitoring, control and quality assurance functions, which are carried out through mediation operations (e.g. transformations) on the messages exchanged. The underlying Data Quality Model, which includes Dimensions, Factors and Metrics for the evaluation of quality, is based on international standards and frameworks as well as on Uruguayan health regulations. This paper presents a first version of this model including the aspects considered the most relevant to the scope of the work.

The main contributions of this work consist of the proposal of a specialized platform for data quality control which is built on a combination of service integration techniques, an ESB middleware and a Data Quality Model instantiated to the health area. In addition, this work presents an application example and experimental tests, particularly on the overhead generated when using the integration platform, which contribute to studying the technical feasibility of the proposed approach.

The overall goal is to specify these types of mechanisms to be applied in a specialized integration platform for the Health area like the one in Uruguay.

Future work consists of further developing the Data Quality Model for e-health. The main goals are to incorporate features of the new ISO 8000 standard for data quality as well as to improve aspects of implementation considering the tests performed.

It is important to highlight that the proposed approach is not restricted to the Data Quality Model, which may be extended within the specific health area and to other ones.

More generally, this work constitutes a step forward towards the development of compliance management mechanisms in the context of inter-organizational integration platforms.

References

- [1] G. J. Kuperman, “Health-information exchange: why are we doing it, and what are we doing?” *Journal of the American Medical Informatics Association*, vol. 18, no. 5, p. 678, 2011.
- [2] L. González and R. Ruggia, “A reference architecture for integration platforms supporting cross-organizational collaboration,” in *Proceedings of the 17th International Conference on Information Integration and Web-based Applications & Services*, ser. iiWAS ’15. New York, NY, USA: ACM, 2015, pp. 92:1–92:4.
- [3] D. Chappell, *Enterprise Service Bus: Theory in Practice*. O’Reilly Media, Jun. 2004.
- [4] L. González, R. Ruggia, J. Abin, G. Llambías, R. Sosa, B. Rienzi, D. Bello, and F. Álvarez, “A Service-Oriented Integration Platform to Support a Joined-Up E-Government Approach: The Uruguayan Experience,” in *Advancing Democracy, Government and Governance*, ser. Lecture Notes in Computer Science, A. Kő, C. Leitner, H. Leitold, and A. Prosser, Eds. Vienna, Austria: Springer Berlin Heidelberg, Jan. 2012, pp. 140–154.
- [5] J. Abin, H. Nemeth, and I. Friedmann, “Systems architecture for a nationwide healthcare system,” in *MedInfo*, 2015, pp. 12–16.
- [6] W. H. Organization *et al.*, *Improving data quality: a guide for developing countries*. Manila: WHO Regional Office for the Western Pacific, 2003.
- [7] K. Rasta, T. H. Nguyen, and A. Prinz, “A framework for data quality handling in enterprise service bus,” in *Third International Conference on Innovative Computing Technology (INTECH 2013)*, 2013, pp. 491–497.
- [8] Silvana Pidre, Laura González, Rodrigo Mendoza, Mauricio Piñatares, Nicolas Granja, Flavia Serra, and Raúl Ruggia, “A Data Quality aware Enterprise Service Bus for e-Health Integration Platforms,” in *XLIII Conferencia Latinoamericana en Informática (CLEI 2017)*, Córdoba, Argentina, 2017.
- [9] N. Davis and M. LaCour, *Introduction to Health Information Technology*, 2001.
- [10] L. Etcheverry, V. Peralta, and M. Bouzeghoub, “Qbox-Foundation: a Metadata Platform for Quality Measurement,” in *Proceedings of the 4eme Atelier Qualité des données et des Connaissances*, Nice, France, 2008.

- [11] M.-T. Schmidt, B. Hutchison, P. Lambros, and R. Phippen, "The enterprise service bus: Making service-oriented architecture real," *IBM Syst. J.*, vol. 44, no. 4, pp. 781–797, Oct. 2005.
- [12] G. Hohpe and B. Woolf, *Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions*. Addison-Wesley Professional, Oct. 2003.
- [13] Health Level Seven International. (2017, Apr.) Health Level Seven. [Online]. Available: <http://www.hl7.org/>
- [14] International Health Terminology Standards Development Organisation. (2017, Apr.) Systematized Nomenclature of Medicine Clinical Terms. [Online]. Available: <http://www.snomed.org/>
- [15] Health Level Seven International. (2017, Apr.) Clinical Document Architecture. [Online]. Available: http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7
- [16] World Health Organization. (2017, Apr.) International Classification of Diseases. [Online]. Available: <http://www.who.int/classifications/icd/en/>
- [17] AGESIC. (2017, Apr.) Agencia Digital Uruguay 2011-2015. [Online]. Available: http://www.agesic.gub.uy/innovaportal/v/1443/1/agesic/mapa_de_ruta:_\discretionary{-}{-}{agenda_digital_uruguay_2011-2015.html?menuderecho=11
- [18] AGESIC. (2017, Apr.). [Online]. Available: <https://www.agesic.gub.uy/>
- [19] AGESIC. (2017, Apr.) Programa Salud.uy. [Online]. Available: http://www.agesic.gub.uy/innovaportal/v/4422/19/agesic/que_es.html
- [20] Ministerio de Salud del Uruguay. (2017, Apr.). [Online]. Available: <http://www.msp.gub.uy/>
- [21] AGESIC. (2017, Apr.) Modelo de Referencia de Personas y Direcciones. [Online]. Available: <https://www.agesic.gub.uy/innovaportal/v/204/1/agesic/buenas-practicas.html?currentpage=1>
- [22] Poder Legislativo del Uruguay. (2017, Apr.) Ley n° 18.600 Documento Electrónico y Firma Electrónica. [Online]. Available: <https://legislativo.parlamento.gub.uy/temporales/leytemp3217058.htm>
- [23] I. Ibnualim and S. H. Supangkat, "Design of health social media to improve the quality of patient's recovery," in *2012 International Conference on Cloud Computing and Social Networking (ICCCSN)*. IEEE, 2012, pp. 1 – 4.
- [24] M. Kahn, M. Raebel, J. Glanz, K. Riedlinger, and J. Steiner, "A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research," 2012.
- [25] ISO/IEC 25012:2008. (2017, Jul.). [Online]. Available: <https://www.iso.org/home.html>
- [26] Canadian Institute for Health Information. (2017, Apr.). [Online]. Available: <https://www.cihi.ca/en>
- [27] Health Information and Quality Authority. (2017, Apr.). [Online]. Available: <https://www.hiqa.ie>
- [28] Australian Digital Health Agency. (2017, Apr.). [Online]. Available: <https://www.digitalhealth.gov.au>
- [29] WA Health, Government of Western Australia. (2017, Apr.). [Online]. Available: <http://www.health.wa.gov.au>
- [30] ACT Health, Government of Australian Capital Territory. (2017, Apr.). [Online]. Available: <http://www.health.act.gov.au>
- [31] I. Rafique, P. Lew, M. Q. Abbasi, and Z. Li, "Information quality evaluation framework: extending iso 25012 data quality model," *World Academy of Science, Engineering and Technology*, vol. 65, pp. 523–528, 2012.
- [32] Rodrigo Mendoza, Mauricio Piñatares, and Nicolás Granja, "Calidad de Datos en Enterprise Service Bus," May 2016. [Online]. Available: <https://www.fing.edu.uy/inco/grupos/lins/informes/pg-dqesb.pdf>
- [33] P. C. Brebner, "Real-world performance modelling of enterprise service oriented architectures: delivering business value with complexity and constraints," *SIGSOFT Softw. Eng. Notes*, vol. 36, no. 5, pp. 85–96, Sep. 2011. [Online]. Available: <http://doi.acm.org/10.1145/1958746.1958762>

- [34] K. Sartipi, M. H. Yarmand, and D. G. Down, “Mined-knowledge and decision support services in electronic health,” in *Systems Development in SOA Environments, 2007. SDSOA '07: ICSE Workshops 2007. International Workshop on*, 2007, pp. 10–10.
- [35] J. O'Donoghue and J. Herbert, “Data management within mhealth environments: Patient sensors, mobile devices, and databases,” *J. Data and Information Quality*, vol. 4, no. 1, pp. 5:1–5:20, Oct. 2012. [Online]. Available: <http://doi.acm.org/10.1145/2378016.2378021>
- [36] R. J. T. Ferreira, M. E. C. D. Correia, F. N. R. Gonçalves, and R. J. C. Correia, “Data quality in hl7 messages – a real case analysis,” in *2015 IEEE 28th International Symposium on Computer-Based Medical Systems*, 2015, pp. 197–200.
- [37] M. Fan, J. Sun, B. Zhou, and M. Chen, “The smart health initiative in china: The case of wuhan, hubei province,” *J. Med. Syst.*, vol. 40, no. 3, pp. 1–17, Mar. 2016. [Online]. Available: <http://dx.doi.org/10.1007/s10916-015-0416-y>
- [38] N. Botts, O. Bouhaddou, J. Bennett, E. Pan, C. Byrne, L. Mercincavage, L. Olinger, E. Hunolt, and T. Cullen, “Data Quality and Interoperability Challenges for eHealth Exchange Participants: Observations from the Department of Veterans Affairs’ Virtual Lifetime Electronic Record Health Pilot Phase,” *AMIA Annual Symposium Proceedings*, vol. 2014, pp. 307–314, Nov. 2014.
- [39] P. J. del Cid, N. Matthys, C. Huygens, S. Michiels, W. Joosen, and D. Hughes, “Sensor middleware to support diverse data qualities,” in *2011 Eighth International Conference on Information Technology: New Generations*, 2011, pp. 673–676.
- [40] S. Sicari, A. Rizzardi, D. Miorandi, C. Cappiello, and A. Coen-Porisini, “Security policy enforcement for networked smart objects,” *Computer Networks*, vol. 108, pp. 133 – 147, 2016.
- [41] J. Singh, T. F. J.-M. Pasquier, J. Bacon, R. Diaconu, J. Powles, and D. Eysers, “Big ideas paper:policy-driven middleware for a legally-compliant internet of things,” in *Proceedings 17th ACM/IFIP/Usenix Middleware*, ACM. ACM, December 2016.
- [42] I. Caballero, I. Bermejo, L. Parody, M. T. Gómez-López, R. M. Gasca, and M. Piattini, “Sla4dq-i8k: Acuerdos a nivel de servicio para calidad de datos en intercambios de datos maestros regulados por iso 8000-1x0,” in *X Jornadas de Ciencia e Ingeniería de Servicios (JCIS 2014)*, 2014, pp. 157–166.
- [43] “Towards a service architecture for master data exchange based on {ISO} 8000 with support to process large datasets,” *Computer Standards & Interfaces*, pp. –, 2016.