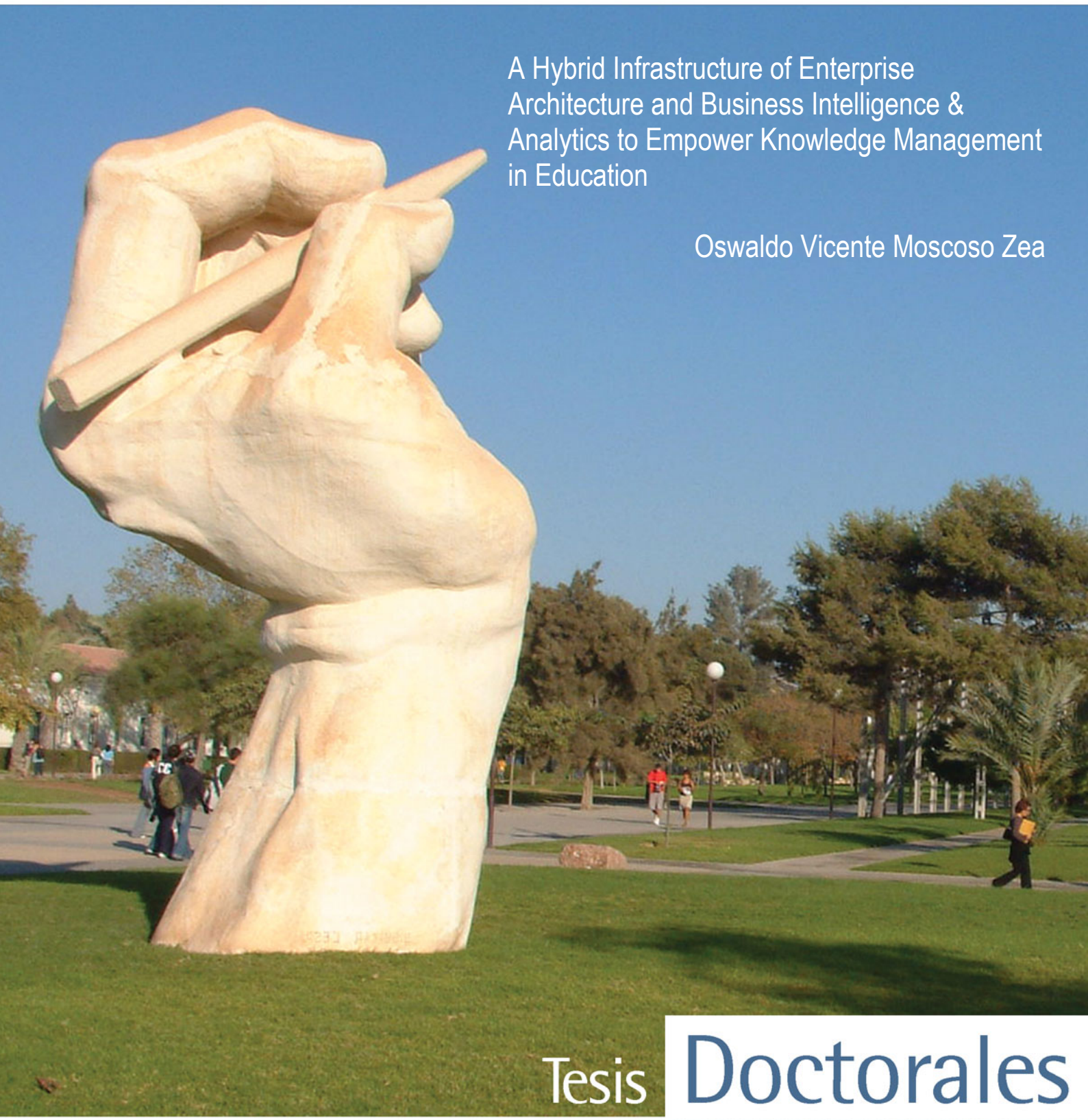
Universitat d'Alacant
Universidad de Alicante

A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics to Empower Knowledge Management in Education

Oswaldo Vicente Moscoso Zea



Tesis Doctorales

www.eltallerdigital.com

UNIVERSIDAD de ALICANTE

Universitat d'Alacant
Universidad de Alicante

Instituto Universitario de Investigación
Informática

Escuela Politécnica Superior

A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics to Empower Knowledge Management in Education

Oswaldo Vicente Moscoso Zea

Tesis presentada para aspirar al grado de

Doctor por la Universidad de Alicante

Doctorado en Informática

Dirigida por:

Dr. Sergio Luján Mora

Mayo 2019

DOCTORAL THESIS IN THE FORM OF COMPENDIUM OF PUBLICATIONS

A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics to Empower Knowledge Management in Education

This document contains a summary of the work done by Oswaldo Moscoso Zea, under the supervision of Dr. Sergio Luján Mora, to opt for the degree of Doctor in Informatics. It is presented at the University of Alicante and is structured according to the regulations established for the presentation of doctoral theses in the form of a compendium of publications. It contains a first part with the background and synthesis, a second part with the scientific publications and a third part with conclusions and future work.

May 2019

TESIS DOCTORAL EN FORMA DE COMPENDIO DE PUBLICACIONES

Una infraestructura híbrida de arquitectura empresarial e inteligencia de negocios y analítica para potenciar la gestión del conocimiento en educación

El presente documento contiene una síntesis del trabajo realizado por Oswaldo Moscoso Zea, bajo la dirección del Dr. Sergio Luján Mora, para optar por el grado de Doctor en Informática. Se presenta en la Universidad de Alicante y se estructura según la normativa establecida para la presentación de tesis doctorales en forma de compendio de publicaciones. Contiene una primera parte con los antecedentes y la síntesis, una segunda parte con las publicaciones científicas realizadas y una tercera parte con las conclusiones y trabajo futuro.

Mayo 2019

Dedication

This thesis is dedicated to God for being my guide at all moment and circumstance; to my kids Sebastian and Emilia because they are the reason and the driving force of my life; to my wife Paula for being the companion that encourages me to be better; to my parents Oswaldo, Cecilia and my siblings Alexandra and Fabian for their unconditional support.

Esta tesis está dedicada a Dios por ser mi guía en todo momento y circunstancia; a mis hijos Sebastián y Emilia porque son la razón y el motor de mi vida; a mi esposa Paula por ser la compañera que me impulsa a ser mejor; a mis padres Oswaldo y Cecilia y a mis hermanos Alexandra y Fabian por su apoyo incondicional.

Acknowledgments

The writing of this thesis and the research papers that are part of it, would not have been possible without all the lessons taught and the guidance of my supervisor Sergio Luján Mora. That is the reason that the main acknowledgments of this work are for Sergio. Thank you for sharing your secrets for an effective research. I would also like to thank the lecturers who have been part of this continuous learning to develop my dissertation and my research. Especially the professors of the University of Münster and Alicante who aroused my interest in scientific research.

A special thanks also to people who seek excellence in their daily work and who rely on an effective knowledge management to innovate and make right and timely decisions. Finally, I would like to thank the people who collaborated in the development of this research and the writing of the papers of the compendium, likewise, to the University UTE for all the support.

La escritura de esta tesis y los artículos que la componen no hubiesen sido posible sin todas las enseñanzas y guía de mi director Sergio Luján Mora, es por ello que el agradecimiento principal de este trabajo es para Sergio. Gracias por compartir los secretos para una investigación efectiva. También quiero agradecer a los docentes que han sido parte del aprendizaje continuo en este trajinar de la investigación, de manera especial a los profesores de la Universidad de Münster y de Alicante que despertaron en mi el interés por la investigación científica.

Un agradecimiento especial también a las personas que buscan la excelencia en su trabajo diario y que se apoyan en una gestión del conocimiento efectiva para innovar y para tomar decisiones acertadas y oportunas. Finalmente, me gustaría agradecer a las personas que colaboraron en el desarrollo de este trabajo de investigación y en la escritura de los artículos del compendio, así como al apoyo brindado por la Universidad UTE.

Quito, May 9, 2019
Oswaldo Moscoso Zea

Abstract

Knowledge has to be improved,
challenged, and increased constantly,
or it vanishes,
PETER DRUCKER

The large volumes of data (Big Data) that are generated on a global scale and within organizations along with the knowledge that resides in people and in business processes makes organizational knowledge management (KM) very complex. A right KM can be a source of opportunities and competitive advantage for organizations that use their data intelligently and subsequently generate knowledge with them.

Two of the fields that support KM and that have had accelerated growth in recent years are business intelligence (BI) and enterprise architecture (EA). On the one hand, BI allows taking advantage of the information stored in data warehouses using different operations such as slice, dice, roll-up, and drill-down. This information is obtained from the operational databases through an extraction, transformation, and loading (ETL) process. On the other hand, EA allows institutions to establish methods that support the creation, sharing and transfer of knowledge that resides in people and processes through the use of blueprints and models. One of the objectives of KM is to create a culture where tacit knowledge (knowledge that resides in a person) stays in an organization when qualified and expert personnel leave the institution or when changes are required in the organizational structure, in computer applications or in the technological infrastructure.

In higher education institutions (HEIs) not having an adequate KM approach to handle data is even a greater problem due to the nature of this industry. Generally, HEIs have very little interdependence between departments and faculties. In other words, there is low standardization, redundancy of information, and constant duplicity of applications and functionalities in the different departments which causes inefficient organizations. That is why the research performed within this dissertation has focused on finding an adequate KM method and researching on the right technological infrastructure that supports the management of information of all the knowledge dimensions such as people, processes and technology. All of this with the objective to discover innovative mechanisms to improve education and the service that HEIs offer to their students and teachers by improving their processes.

Despite the existence of some initiatives, and papers on KM frameworks, we were not able to find a standard framework that supports or guides KM initiatives. In addition, KM frameworks found in the literature do not present practical mechanisms to gather and analyze all the knowledge dimensions to facilitate the implementation of KM projects.

The core contribution of this thesis is a hybrid infrastructure of KM based on EA and BI that was developed from research using an empirical approach and taking as reference the framework developed for KM. The proposed infrastructure will help HEIs to improve education in a general way by analyzing reliable and cleaned data and integrating analytics from the perspective of EA. EA analytics takes into account the interdependence between the objects that make up the organization: people, processes, applications, and technology.

Through the presented infrastructure, the doors are opened for the realization of different research projects that increment the type of knowledge that is generated by integrating the information of the applications found in the data warehouses together with the information of the people and the organizational processes that are found in the EA repositories. In order to validate the proposal, a case study was carried out within a university with promising initial results. As future works, it is planned that different HEIs' activities can be automated through a software development methodology based on EA models. In addition, it is desired to develop a KM system that allows the generation of different and new types of analytics, which would be impossible to obtain with only transactional or multidimensional databases.

Spanish Abstract

El conocimiento debe ser mejorado,
desafiado, y aumentado constantemente,
o desaparece,
PETER DRUCKER

Los grandes volúmenes de datos (Big Data) que se generan hoy en día a escala global y dentro de las organizaciones junto con el conocimiento que reside en las personas y aquel que se plasma en los procesos de negocio hace muy compleja la gestión del conocimiento (GC) organizacional. Una correcta GC puede ser fuente de oportunidades y de ventaja competitiva para aquellas organizaciones que usan sus datos de forma inteligente y posteriormente generan conocimiento con ellos.

Dos de los campos que apoyan la GC y que han tenido un acelerado crecimiento en los últimos años son la inteligencia del negocio (IN) y la arquitectura empresarial (AE). Por un lado, la IN permite sacar provecho de la información guardada en los almacenes de datos multidimensionales (data warehouse) con diferentes operaciones como slice, dice, roll-up y drill-down. Esta información se obtiene de las bases de datos operacionales por medio de un proceso de extracción, transformación y carga (conocido en inglés como ETL por "extraction, transformation and loading"). Por otro lado, la AE permite a las instituciones establecer métodos que permitan crear, compartir y transferir el conocimiento que reside en personas y procesos mediante el uso de mapas y modelos. Uno de los objetivos de la GC es crear una cultura donde el conocimiento tácito (aquel que reside en las personas) se quede en la organización cuando el personal calificado y experto abandone la misma o cuando se requiera realizar cambios en la estructura organizacional, en las aplicaciones informáticas o en la infraestructura tecnológica.

En el caso de las instituciones de educación superior(IES) el no tener un adecuado método de GC es un problema incluso mayor debido a la naturaleza de esta industria. Generalmente en las IES, existe muy poca interdependencia entre los departamentos y facultades. Es decir, existe poca estandarización, redundancia de información, duplicidad de aplicaciones y funcionalidades entre departamentos, lo que ocasiona organizaciones poco eficientes. Es por ello, que este trabajo se ha centrado en buscar un método adecuado de GC e investigar sobre la infraestructura tecnológica que apoye la gestión de la información de todas las dimensiones de conocimiento como son: personas, procesos y tecnología. Todo esto con el objetivo de buscar mecanismos innovadores para mejorar la educación y el servicio que brindan las IES a sus estudiantes y docentes mediante la mejora en sus procesos.

A pesar de existir algunas iniciativas y artículos sobre los marcos de GC, no pudimos encontrar un marco de trabajo estándar que apoye o guíe las iniciativas de GC. Además, los marcos de trabajo de GC encontrados en la bibliografía científica no presentan un

enfoque práctico que permita analizar y consolidar las dimensiones de conocimiento y por ende facilite su implementación. El aporte central de esta tesis es una infraestructura híbrida de GC basada en AE e IN que fue desarrollada completamente desde un enfoque de investigación empírica y tomando como referencia el marco de trabajo desarrollado para la GC. La infraestructura desarrollada ayudará a las IES a mejorar la educación de manera general mediante el análisis de datos educativos confiables y depurados e integrando analítica desde la perspectiva de AE. Tomando en cuenta la interdependencia existente entre los objetos que conforman la organización: personas, procesos, aplicaciones y tecnología.

A través de la infraestructura presentada, se abren las puertas para la realización de distintos experimentos y proyectos de investigación que permitan incrementar el tipo de conocimiento que se genera integrando la información de las aplicaciones que se encuentra en los almacenes de datos junto con la información de las personas y de los procesos organizacionales que se encuentran en los repositorios de AE.

Para validar la propuesta se realizó un caso de estudio dentro de una universidad con resultados iniciales prometedores. Como trabajos futuros se planea automatizar distintas actividades de la IES mediante una metodología de desarrollo de software basada en modelos de AE. Además, se está desarrollando un sistema de gestión del conocimiento que permitirá generar distintos y nuevos tipos de analítica, que solo con bases de datos transaccionales o multidimensionales sería imposible obtener.

Contents

List of Figures

List of Tables

Part I

BACKGROUND AND SYNTHESIS

1 Introduction

Knowledge is power.
Information is liberating.
Education is the premise of progress,
in every society, in every family.
KOFI ANNAN

1.1 Motivation and work undertaken

Nowadays, knowledge is one of the most important assets in organizations. Data is everywhere and it can become a source of opportunities to secure competitiveness if is managed correctly. The ideal scenario is to convert these data into information and information into useful knowledge that can allow organizations to improve their business models and their organizational processes. Moreover, modern organizations need to acquire capabilities to use their organizational knowledge in innovation and change management. These capabilities allow these organizations to reduce operational costs and to have a competitive edge.

Higher education institutions (HEIs) have traditionally been knowledge producers and have established research practices to accomplish this goal. However, most of these organizations do not have a standard or unique approach for knowledge management (KM) practices for their own benefit. Furthermore, the huge amount of information that their academic and administrative systems produce need to be managed properly and with the right technology.

HEIs within their strategies of continuous improvement and with the objective of reaching international quality and excellence standards require to manage the organizational knowledge in an efficient way. The management of educational institutions focuses on achieving greater efficiency and productivity through the use of administrative and technological tools that support measurement and evaluation (Correa, Angélica, & Correa, 2009). One of the key aspects of governance and decision support is based on KM. KM is a discipline that promotes the creation, use, distribution and transfer of knowledge in organizations (Campbell, 2006). An important strategy for continuous improvement in HEIs is the implementation of KM. KM can become a differentiating factor that allows having educational services of excellence. Additionally, KM can improve the efficiency of administrative and academic processes.

One of the main motivations of the work presented in this thesis, which was realized during these last years, was to investigate how the data obtained in HEIs can improve the educational process. For this purpose, we study different technologies like data warehousing (DW), data analytics, educational data mining (EDM), learning analytics,

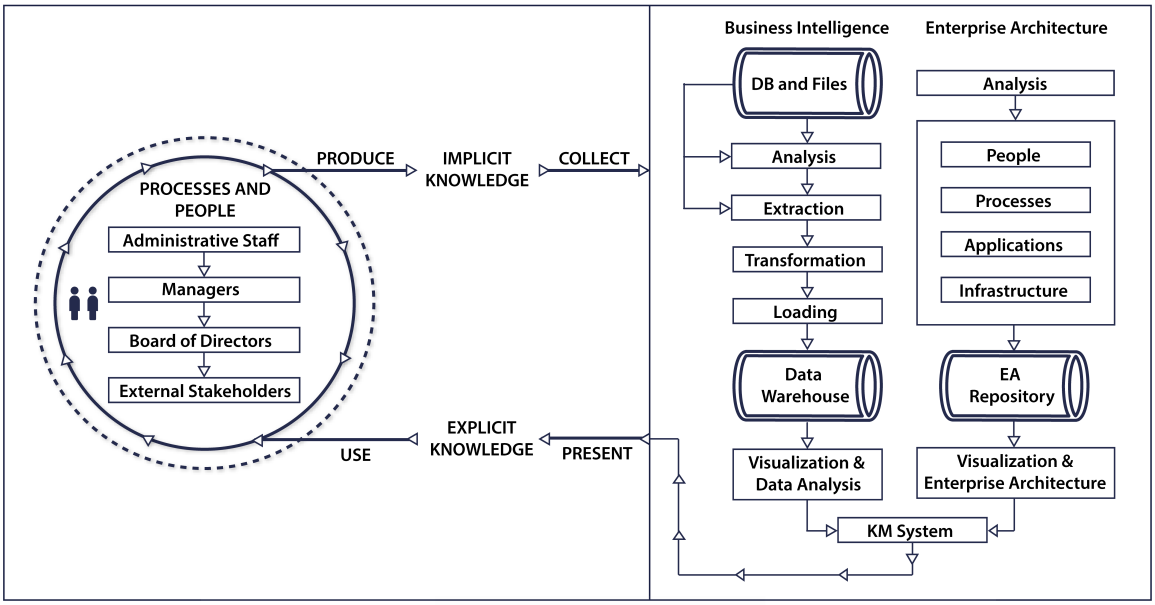


Figure 1.1: Proposed knowledge management framework

business intelligence (BI), enterprise architecture (EA) and KM. From this study, we deduced that KM has become a necessity due to the great amount of information (Big Data) that HEIs generate every day. Data flows from HEIs' application systems and from the mobile devices and gadgets of students, academic and administrative staff. Thus, at first, it was found out that although there were papers describing KM in HEIs, there was not a standard approach to implement KM, nor a practical guide that takes into account all knowledge dimensions. Consequently, we did some research and drew a framework for performing KM (Moscoso-Zea, Luján-Mora, Esquetini Cáceres, & Schweimanns, 2016).

This framework was conceived with the participation of researchers from four universities worldwide: University of Alicante–Spain, Technical University of Berlin–Germany, National Polytechnic School–Ecuador and UTE University–Ecuador. This framework which is shown in Figure 1.1 was designed with a view to becoming a practical guide to implement KM projects in organizations. This development was presented in the International Conference on Enterprise Information Systems (ICEIS) 2016 in Rome, Italy. This paper included a theoretical description of how to implement a KM initiative by taking into account all the knowledge dimensions.

The core idea in the development of the framework was that a successful implementation of KM greatly depends on a well–defined method that supports the creation, capturing, use, distribution, and transfer of knowledge. This paper, which presents the developed framework (Moscoso-Zea, Luján-Mora, et al., 2016), has a guide for performing the main KM tasks. Moreover, the flow of explicit and implicit organizational knowledge is stated. Organizational knowledge is created from different interdependent objects in different enterprise domains. Explicit and implicit knowledge can be derived from these domains. Explicit knowledge is the knowledge that can be formulated, documented and reproduced. Implicit knowledge (also known as tacit knowledge) is the knowledge that is difficult to document or formulate and is normally associated with human knowledge (Nonaka, 1994).

The proposed framework incorporates two tasks which are necessary to implement KM. The first task corresponds to the development of a technological infrastructure to facilitate knowledge capturing and sharing. The second task establishes the mechanisms and procedures for retaining knowledge from people and processes.

In this thesis, the two tasks necessary for KM have been studied and presented in the compendium of research papers under the umbrella of the proposed framework. First, an hybrid infrastructure was developed; second, the mechanisms to capture tacit knowledge from employees and from business processes were established. Although the implementation of the proposed framework seems straightforward, it includes distinct disciplines which do not make the process simple. Therefore, the details of the components of the framework were studied in detail and presented in conference and journal papers around the world.

After studying each of the components and the knowledge flows proposed in the framework, the next step was to look at the mechanisms to hybridize the information and knowledge that is generated in the data warehouse (DWH) and in the EA repository. The hybridization of the components concluded with a proposal of a web knowledge management system that allows the creation of new types of analytics. In addition, the presented articles also propose methods or best practices for the implementation of the main components of the framework.

The box of the right of the proposed KM framework presents a DWH and an EA repository. Within the analysis carried out, the best design methodologies for data storage and the best ways to carry out experiments with EDM were studied. Furthermore, a methodological proposal to evaluate and select the best EA tools was presented.

Since the proposed approach was to improve educational indicators, the investigation started by looking at the possibilities that EDM brought about to accomplish this goal. The fundamental purpose of EDM is to analyze data from educational institutions through the use of different techniques such as prediction, clustering, analysis of time series, classification, among others. This study was presented in the 2016 Iberian Conference on Information Systems and Technologies (CISTI) (Moscoso-Zea & Lujan-Mora, 2016). The paper presented a complete view of EDM that includes a classification of the algorithms, methods, and tools used in the data mining (DM) processes. Furthermore, the paper analyzes the processes and indicators with potential for improvement in HEIs. Within this research, a set of 42 papers which experimented with EDM were studied. From the studied papers, it could be identified that the most used EDM methods were classification, clustering, and association. From the research, it was observed that most of the experiments carried out are focused on the teaching–learning processes, such as prediction of the results of evaluations, prediction of the performance and the profile of the students. The results might also provide feedback to the tutors or teachers to take corrective and preventive actions at an early stage. The specific activities which can be improved by using EDM are the following:

- Characterize the behavior and achievements of the students.
- Improve the teaching–learning process (registration, tutoring, evaluation, graduation).
- Improve management indicators such as decrease drop-out of students, increase the graduation rate, personalized tutoring.

- Improve the infrastructure.
- Optimize the use of classrooms and laboratories.
- Improve organizational efficiency.

The knowledge that is generated by using EDM, supports the authorities of HEIs in making timely decisions and also supports the teachers to analyze the behavior and the way their students learn (Jindal & Borah, 2013). The discipline focuses on the design of models to improve learning experiences and organizational efficiency (Huebner, 2013).

Despite the fact that HEIs are investing in driving EDM initiatives, much of the efforts conducted are not having the expected results due to the lack of a sound process of DW for an educational scenario, therefore, in the next step a study was done to understand the importance of having a DWH prior to experimenting with EDM. The study was presented at the Information Technology Based Higher Education (ITHET) 2016 conference. This paper presented design considerations for the implementation of a DWH in educational institutions. Furthermore, the extraction, transformation, and loading (ETL) of data from operational data sources into a DWH is described. Additionally, as a case study, the paper presents the steps followed for the design of a DWH in a private university (Moscoso-Zea, Sampedro, & Luján- Mora, 2016).

The steps suggested for the implementation of a DWH are:

1. Diagnosis.
2. Information needs analysis.
3. Selection of the methodology.
4. Setting up the technological infrastructure.
5. DWH design.
6. ETL execution.

Regarding the EA component which is drawn on the right of the KM framework, we also performed a study to evaluate the role of EA as a key element for change and KM (Moscoso-Zea, Paredes-Gualtor, & Luján-Mora, 2019). The implementation of EA can be supported by a tool with a modeling language which can be used to describe different businesses processes, organizational structures, and information flow of an organization (Moscoso-Zea & Luján-Mora, 2017). In recent years, EA has become one of the most important fields to consider in information systems studies and has evolved to become an essential task for business management (Bricknall, Darrell, Nilsson, & Pessi, 2006). Consequently, to understand the existing EA tools this paper proposes an approach for the evaluation and selection of EA tools that fit an organizational scenario. After analyzing different software evaluation methodologies, three methodologies are suggested for the evaluation of EA and are shown in Table 1.1.

These methodologies support enterprise architects at the beginning of an EA project. Different activities are suggested in the methodologies to improve the selection process and invest in the most adequate EA tool for the knowledge initiative according to the requirements of the stakeholders in the HEI. Once the tool is installed and configured

Author(s)	Year	Type of software
(Comella-Dorda et al., 2004)	2004	COTS products
(Kontio et al., 1995)	1995	COTS selection
(Morera, 2002)	2002	COTS products

Table 1.1: Suggested methodologies for the evaluation of EA tools

the next step is to define how to manage the EA of the organization. One of the strategies for EA management is to use a tool with a framework. Therefore, we also collaborate in a research project to define and EA framework with an agile approach. The development of an EA in big organizations is complex. Thus, is important to create value at an early stage of the implementation of EA. This framework was developed with a new design based on the architecture development method (ADM) of TOGAF to extract only those deliverables that are important for a specific business or project (Sandoval, Moscoso-Zea, Galvez, & Tutillo, 2016).

During the analysis of EA tools with the research group of the University UTE we study how to migrate a technological enterprise infrastructure to the cloud and what are the data security issues (Saa, Cueva, Moscoso-Zea, & Lujan-Mora, 2017). Recommendations were given for organizations on how to ensure that a comprehensive security strategy is defined before migrating their enterprise applications to the cloud. Some of the given recommendations were that security controls need to be enforced at all levels. This can be done by incorporating a framework that addresses security at the physical, network, data and application level. The conclusion of the paper is that depending on the size of the company more or fewer applications and technology can be migrated to the cloud. The main benefits of cloud-based enterprise infrastructure are its scalability and lower investment costs, creating opportunities for small and medium enterprises. It is interesting to note that our proposal of the hybrid infrastructure can be implemented in a cloud scenario as well. Consequently, this paper suggests the phases and alternatives to implement a cloud project. In Figure 1.2 the framework that was developed for a migration process is presented which is meant to support organizations to implement cloud computing services with best practices at the lowest risk (Moscoso-Zea, Paredes-Gualtor, Saa, & Sandoval, 2018). The main goal of the framework is to improve the migration process by minimizing threats and maximizing opportunities. This framework can support cloud migrations and can be used to migrate the KM infrastructure proposed in this thesis.

The analysis of the components proposed in the framework was encouraging and confirmed their applicability in the process of hybridization and implementation. Therefore, a study on how to use these components in HEIs to produce new organizational knowledge was performed. One of the most important aspects of management and decision support is based on KM. The paper "Knowledge Management in Higher Education Institutions for the Generation of Organizational Knowledge" published in the 12th Iberian Conference on Information Systems and Technologies in 2017, presents a case study with different alternatives for the knowledge creation process (Moscoso-Zea & Lujan-Mora, 2017). Among the important reasons for the realization of this study was the creation of new mechanisms to promote teaching and student development, research, and student welfare. Another reason was to study alternatives on how to improve the institutional processes so they are carried out in an organized and stan-

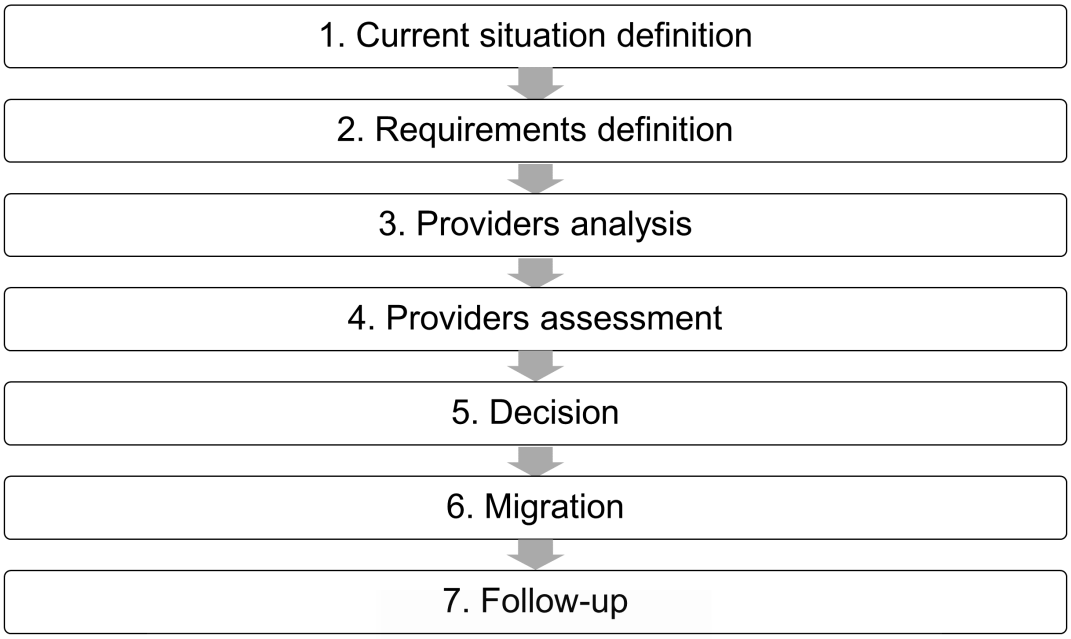


Figure 1.2: Cloud computing adoption framework
(Moscoso-Zea, Paredes-Gualtor, Saa, & Sandoval, 2018)

dardized manner, which will facilitate the achievement of the goals of the institutional vision in the short and medium term. EA and BI are the core technological tools for the creation of repositories and knowledge bases. These knowledge bases can generate organizational intelligence through online analytical processing (OLAP), DM and EA analytics.

In order to establish the possibilities for knowledge creation, we followed the suggestions of Nonaka (1994). Nonaka states that the process to generate knowledge is based on the conversion of tacit knowledge and explicit knowledge (Nonaka, 1994).

Based on this, he describes four ways of knowledge conversion as shown in Figure 1.3 (Nonaka et al., 2000). This figure describes the main activities proposed for knowledge creation and innovation. The model is called SECI due to its components Socialization, Externalization, Combination, and Internalization (SECI). The different knowledge conversion activities from the SECI model are presented in the following list.

- From tacit knowledge to tacit knowledge (Socialization).
- From tacit knowledge to explicit knowledge (Externalization).
- From explicit knowledge to explicit knowledge (Combination).
- From explicit knowledge to tacit knowledge (Internalization).

For the first scenario, knowledge can be created through sharing experiences between new and old employees. Business actors can even create this tacit knowledge just by observing how an expert worker performs tasks. This form of learning and generation of tacit knowledge is called Socialization. Socialization is shown in the upper left quadrant of Figure 1.3. Some tools that can be used for this scenario are brainstorming, observation and imitation. The second scenario corresponds to the upper right

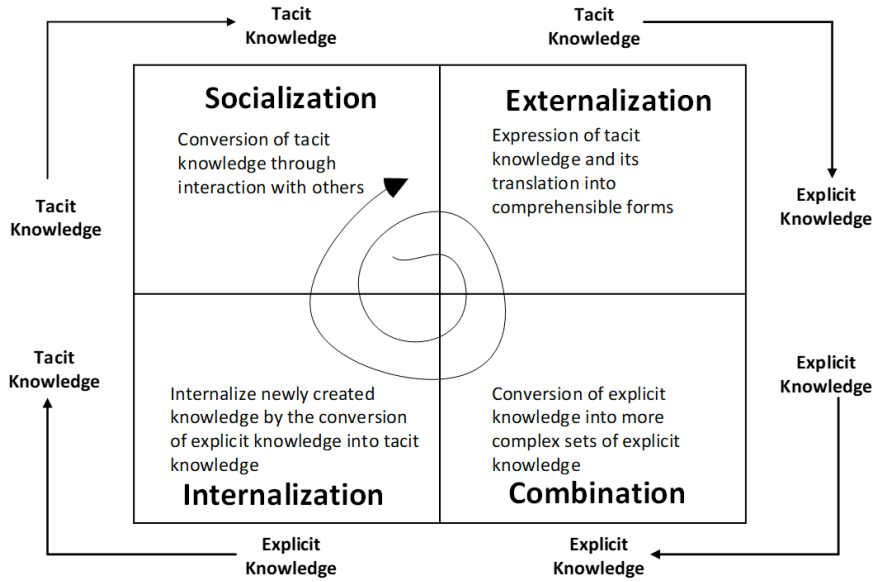


Figure 1.3: SECI model of knowledge creation
(Nonaka, Toyama, & Konno, 2000)

quadrant in the figure and is called Externalization. The conversion of tacit to explicit knowledge occurs through capturing the knowledge of employees in books, databases or information systems. That is to say, tacit knowledge is translated or documented to forms that can be understood and interpreted by other staff.

In the third scenario, the conversion of knowledge occurs through social processes to combine the explicit knowledge of individuals. This knowledge can be shared in meetings or calls. By establishing these interactions, collective and shared knowledge gives rise to new knowledge. This case corresponds to the lower right quadrant that is known as Combination. The combination process is an ideal scenario to create knowledge with technological tools and share the explicit knowledge that resides in the enterprise software applications including databases.

This process of knowledge creation can be based on the technological tools that are proposed in this thesis. Explicit knowledge that resides in the enterprise software applications and in databases can be processed and analyzed using artificial intelligence, BI or EA analytics to create new explicit knowledge.

The fourth scenario is complementary to Externalization and occurs when explicit knowledge is converted to tacit. This is the most similar way to how people learn, that is, to take a manual or a book and internalized this new knowledge to become part of an individual's brain. This form of knowledge generation is called Internalization.

The knowledge management framework that we propose and its components are aligned to the knowledge management cycle. Figure 1.4 shows the suggested KM cycle (McIntyre, Gauvin, & Waruszynski, 2003). This cycle has three main components: management, application, and people. Management is concerned with capturing, organizing and facilitating the organizational knowledge. These activities encompass the externalization and combination quadrants presented in the Nonaka model (Nonaka, 1994). The application component focuses on obtaining relevant content and analysis through the use of DM and advanced analytics. Finally, the people component focuses

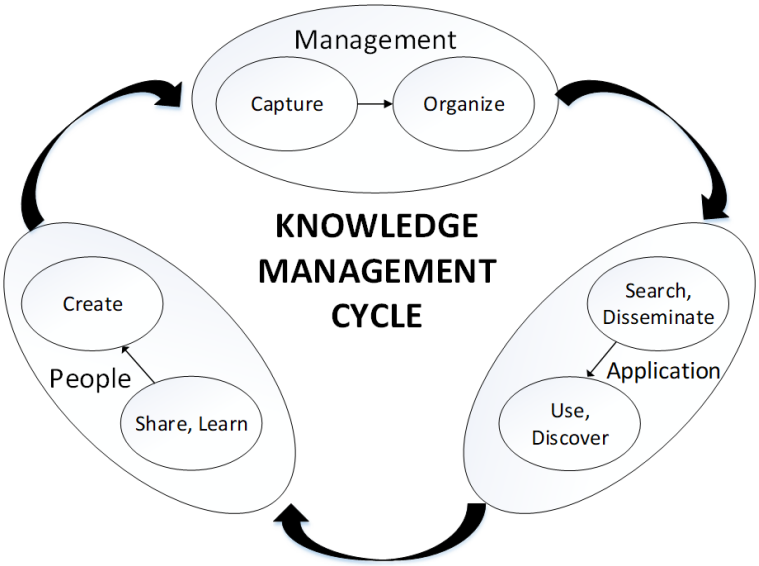


Figure 1.4: Knowledge management cycle
(McIntyre et al., 2003)

on learning, sharing, and collaboration. People component shows the educational process and seeks mechanisms to create tacit knowledge for people who learn and explicit knowledge to be stored in repositories and knowledge bases.

As said previously, one of the main objectives for performing this work was to improve education by using KM. With this in mind, it was also important to investigate the thoughts of students on bringing their laptops or personal devices to classes. The paper of Saa, Moscoso-Zea, and Lujan-Mora (2017) presents the results of a survey which studies on one hand, how universities can be critically affected if they allow students to bring their own devices to classes, on the other hand, it reviews all the concerns that students have from being observed through their devices by the university staff (Saa, Moscoso-Zea, & Lujan-Mora, 2017). The results of the survey show that a total of 57.6 % of students would be willing to bring a laptop or tablet to the lectures and 17.2 % would not. However, there is a 25.2 % that would consider it.

Despite the fact, that with the research done, it was confirmed the importance to include the BI and EA tools in the proposed infrastructure. We found a research gap in how HEIs manage information and knowledge.

The EDM has become a trendy topic on the last years but we have concerns on the validity of the information that is being processed to obtain the results of the analytic process. This is the main reason that we decide to perform a systematic mapping (SM) on the topic DW in education. The results of the SM were presented in the paper "A Holistic View of Data Warehousing in Education" (Moscoso-Zea, Paredes-Gualtor, & Lujan-Mora, 2018). This study shows that DW and BI are widely used in business organizations and have been exhaustively analyzed from the industry standpoint for many years. However, its use is still low in educational institutions. The paper from Shahid et al. (2016) presents different case studies performed to determine the percentage of use of DWHs within different industries.

The results of the case study analysis show that the industry which uses DWH the

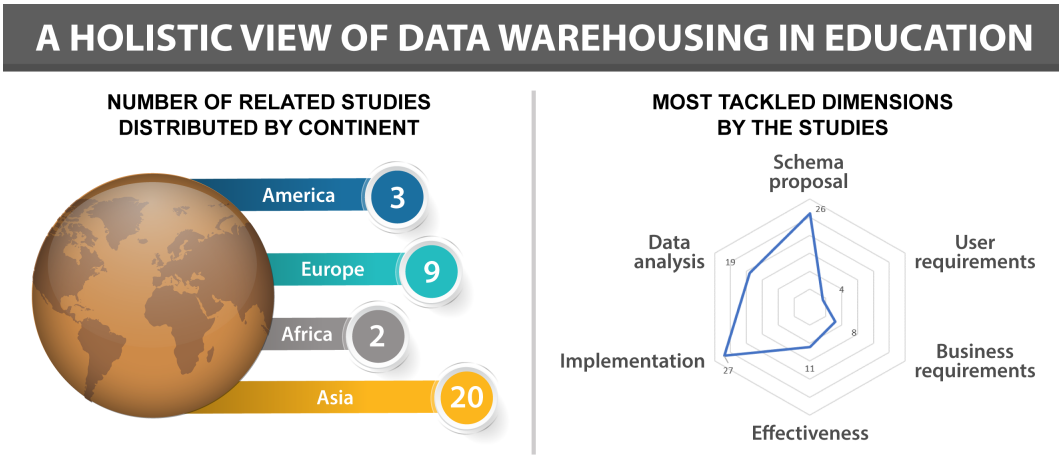


Figure 1.5: Results of the systematic mapping of DW in education (Moscoso-Zea, Paredes-Gualtor, & Lujan-Mora, 2018)

most is the medical industry (hospitals, clinics and physician offices) with 23.3 % of usage. Following that, the finance and banking industry with 6.2 % of DW usage, whereas one of the industries that less use DW is education with only 3.8 % of usage (Shahid et al., 2016). As seen in these facts, only a few educational institutions have adopted a DWH to manage the educational information that is generated from their software applications and learning management systems. Nevertheless, nowadays boards of governance and directors in HEIs are recognizing BI and DW as core technologies to generate organizational knowledge. The implementation of strategies with BI and DW can support HEI to better understand the educational processes. Firstly, it can be a crucial factor to improve education by knowing how students learn and how lecturers teach. Secondly, it can help to identify early enough students with risk to drop out or to fail assignments. Moreover, the modeling or the organizational processes of the institution allowed having a clear vision of the educational processes, which can be of great support to implement innovative change initiatives (Baepler & Murdoch, 2010; Bichsel, 2012; Morris, Wu, & Finnegan, 2015).

The SM carried out on the topic of DW in education was performed between January 2008 and June 2018. In the study, a four-stage method was applied. This study supported the researcher team to answer the proposed research questions. The first research question was targeted to identify details and the number of significant studies with the words “data warehouse” and “education” or “academic”. To accomplish this objective the results of the analysis included the authors, the institutional affiliation, and their country of origin. Furthermore, this analysis listed the journals or conferences in which the resulting papers were published. Figure 1.5 depicts some of the results of this paper. The left of the figure shows the number of related studies distributed by continent. The right of the figure plots the most tackled dimensions by the studies.

The second research question was targeted to six dimensions of analysis that have been tackled in the relevant studies. The results of this analysis shows that the most studied dimensions among researchers are in the following order: Implementation, Schema Proposal, Data Analysis, Effectiveness, Business Requirements, and User Requirements. The results of the research allowed to identify the best practices in DWH design. Moreover, in this paper, a five-step method to implement a DWH is suggested

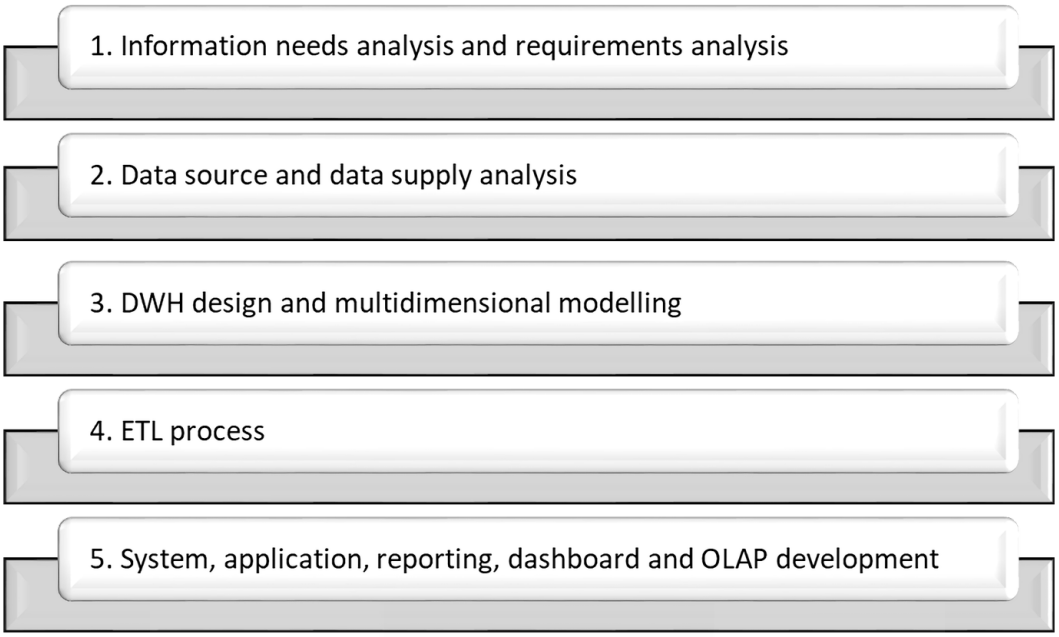


Figure 1.6: Method for DWH project implementation

based on the proposals of the resulting documentary corpus. Figure 1.6 shows the proposed method which can support in future implementation projects. This paper gives also lights that in many cases the implementation of DWHs only takes into account technical requirements and does not pay much attention to the user and business requirements. Failures to obtain the right requirements might be an important cause of the failure of BI and EDM projects. The main problem is that if the data loaded to the database is wrong or inconsistent, the results in the reporting and in the data analytics process will also produce unreliable results.

After studying the components of the framework and while confirming the applicability to support the KM initiative. The next step was to define the technical aspects to implement the KM infrastructure. The paper "A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics for Knowledge Management in Education" (Moscoso-Zea, Castro, Paredes, & Luján-Mora, 2019) presents the information infrastructure formed by BI tools, the educational DWH and the EA repository that allows the digitization of knowledge and empowers the visualization and the analysis of organizational components as people, processes, and technology. Figure 1.7 shows the proposed infrastructure. This infrastructure was designed based on research and will serve to analyze educational data, academic processes and for the creation of explicit knowledge using different algorithms and methods of EDM, learning analytics, online analytical processing (OLAP), and EA analytics.

Finally, the proposed infrastructure was used to discover trends and patterns in educational indicators. The paper "Evaluation of algorithms to predict graduation rate in higher education institutions by applying educational data mining" (Moscoso-Zea, Saa, & Luján-Mora, 2019) compares EDM methods and algorithms and suggests which has the best precision in certain scenarios. The paper concludes that random trees and J48 algorithms are the algorithms with the best performance. The infrastructure has proved to be of great use to capture implicit knowledge and to produce explicit knowl-

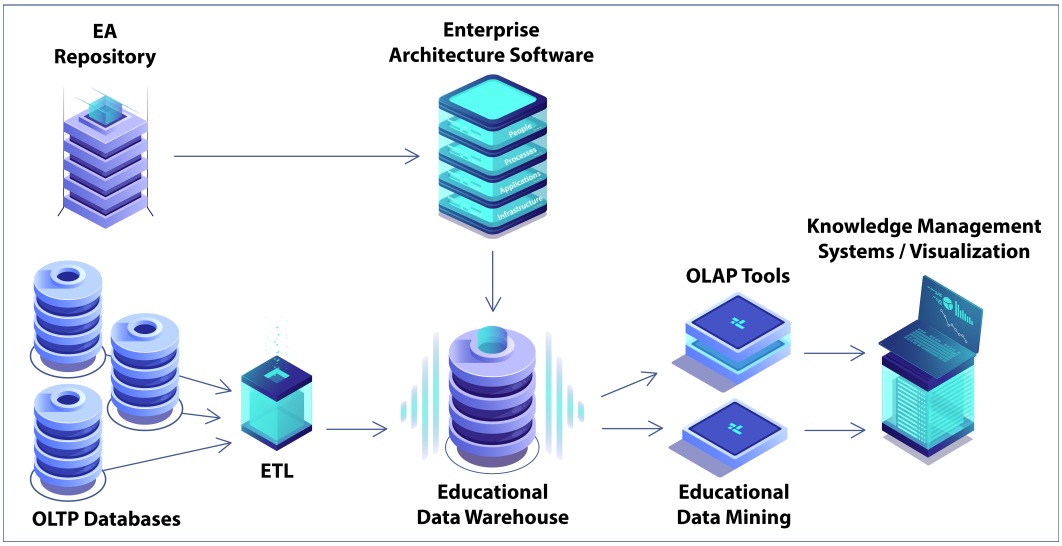


Figure 1.7: Hybrid infrastructure of EA and BI for KM

edge. It is expected that HEIs that implement this proposal will hopefully improve the way KM is handled.

1.2 Objectives

The main objective of the investigation was to design a KM process and an infrastructure that can contribute in improving the educational process. The specific objectives are:

1. Study of the state of the art of KM, EA, DWH, EDM, and cloud computing.
2. Define a method to design a DWH and a method for the selection of EA tools.
3. Study of indicators in HEIs which are susceptible of improvements through a process of EDM.
4. Carry out a study on the importance that HEIs give to the DWH for EDM.
5. Carry out an analysis of algorithms and methods of EDM.
6. Make a methodological proposal for KM in HEIs.
7. Propose the appropriate infrastructure for KM and its implementation.
8. Study mechanisms to transfer the tacit knowledge of employees to the proposed infrastructure.

1.3 Method

While performing the research different methods were used. For example, for the study of the state of the art of this work, we used the methods proposed by Kitchenham, Budgen, and Brereton (2010) and Greenhalgh and Peacock (2005). Moreover, we used the

knowledge discovery in databases (KDD) (Fayyad, Piatetsky-Shapiro, & Smyth, 1996) method for performing experiments with EDM. There were also some methods that we created based on bibliographic research for performing some tasks. For example, a five–step method for DWH design was developed using the results of the systematic mapping performed on DHW on education which is part of this compendium. Additionally, for the design of the infrastructure, we followed the design-science paradigm of Hevner, March, Park, and Ram (2004). This paradigm explains how the creation of innovative artifacts as the one proposed in this thesis seeks to extend the boundaries of organizational capabilities. It is important to observe that the components of the proposed infrastructure have not been used in conjunction for a KM project. Therefore, this works has a strong component of innovation, and the infrastructure represents an innovative artifact.

During the last years, while performing the research work we were able to implement most of the components of the suggested infrastructure. The components of the KM framework were implemented using a case study performed in an HEI. This case study was conducted in conjunction with the planning department of an HEI at the beginning of the project and with the computer science department in the final path of it. There was a previous phase to train the academic and administrative staff previous to perform the interviews and information gathering.

1.4 Results

The main objective of the case study was to define a sound process for knowledge creation and KM. Thus, starting from a scientific point of view and performing different phases of research the hybrid knowledge infrastructure was implemented. The first steps of implementation were the selection and installation of BI tools (DWH, OLAP systems) and EA tools (architecture software). The activities carried out for KM followed the suggestions of the proposed KM framework and are the following:

- Stakeholder analysis and requirement definitions.
- Evaluation and selection of tools.
- Implementation of the hybrid infrastructure.
- Creation of explicit and tacit knowledge.

1.4.1 Stakeholder analysis

The first phase performed within the case study was the stakeholder analysis. The case study was implemented at a private university. The stakeholder's identification was carried out using the power-interest matrix (Ackermann & Eden, 2011). The results of the application of the matrix shown in Figure 1.8 are used for the management and communication of the project. According to this identification, the target group was selected and then interviewed. The interviews were done according to the interest and power they had in the KM project. The main stakeholders identified in the project were departmental managers and directors, students, academic and administrative staff. During this phase, all the requirements of the stakeholders were also elicited and the data sources, where the important information resides, were identified.

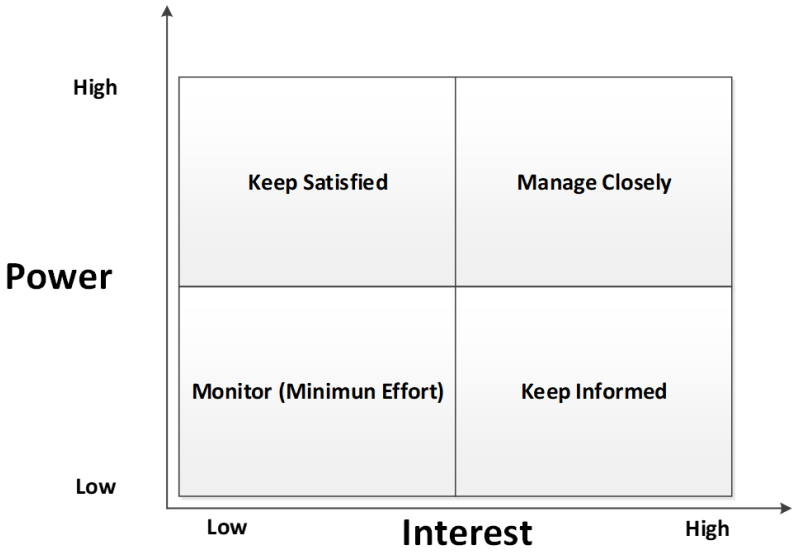


Figure 1.8: Power interest matrix
(Ackermann & Eden, 2011)

1.4.2 Tool evaluation and selection

At this stage, the decisions taken were to select the tools that meet the requirements of the project and the stakeholders. In the present case, the stakeholders carried out an EA software evaluation to choose the appropriate tool (Moscoso-Zea & Luján-Mora, 2017). The methodology used for the selection of tools was the one proposed by Comella-Dorda et al. (2004). After the evaluation process, the Archi tool was selected. Among the most important selection criteria are the following:

- It is based on the archimate standard.
- It allows designing architectures in different domains (Schekkerman, 2011).
- It is a free tool.
- Meets the requirements of the stakeholders.

For the selection of BI tools, a feasibility analysis was carried out. The analysis determined that Microsoft SQL Server and Microsoft Power BI are adequate, mainly because the organization has software licensing and it is not necessary to make an economic investment. Finally, for the DM, Orange Data Miner, and Weka were chosen, for the realization of experiments, because they have the necessary methods for the analysis of educational data of interest.

1.4.3 Implementation of the infrastructure

In order to perform KM, it was necessary to install and implement the tools described in the previous point. The data were extracted from operational databases and transferred to the educational DWH that was created for this purpose. The DWH was designed following Kimball's methodology (Kimball & Ross, 2013) which is the most applied in the

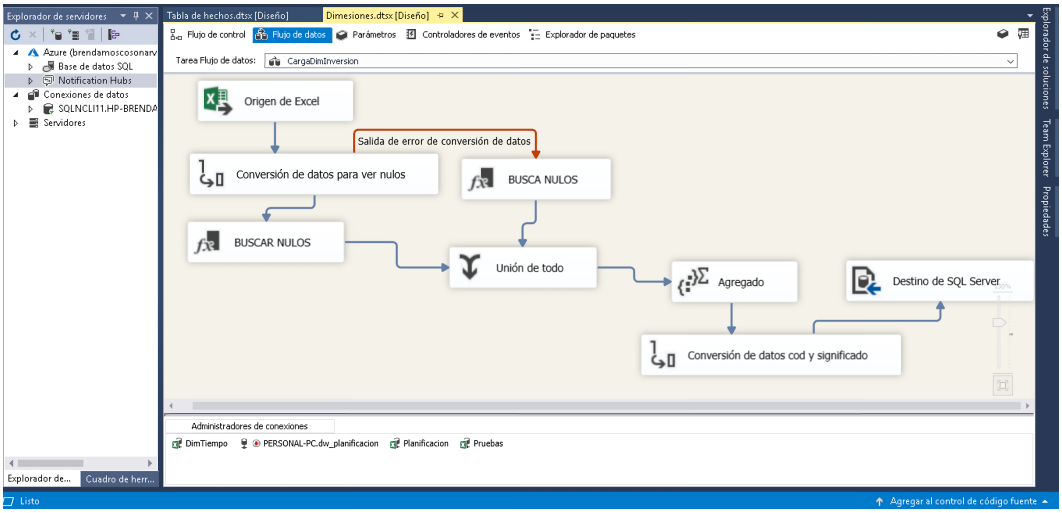


Figure 1.9: Data flow on the dimension tables of the DWH

educational domain (Moscoso-Zea, Paredes-Gualtor, & Lujan-Mora, 2018). Figure 1.9 presents the data flow implementation of the ETL process of a dimension table which exemplifies the different tasks realized to perform this process. Additionally, the EA tool was implemented for process modeling of the different organizational components.

1.4.4 Creation of explicit and tacit knowledge

In order to digitize knowledge through models, interviews were conducted with those responsible for the objective processes of this study (stakeholders defined in the power interest matrix). These interviews served to identify how personnel carries out their assigned activities. In this way, their tacit knowledge (which resides in their minds) is transformed into explicit knowledge and stored in the form of models in the EA tool (externalization process of knowledge creation). To implement the EA process we used the architecture development method (ADM) defined by the Open Group within the TOGAF framework (The Open Group, 2018). The ADM was adapted for the KM project as shown in Figure 1.10. The project consisted in defining the complete EA of a HEI. Figure 1.11 exemplifies the macro process of the HEI. Additionally, Figure 1.12 depicts one of the hundreds of models that we designed from the different departments and campus in the HEI. These models are considered the explicit knowledge of an HEI. Explicit knowledge can then be analyzed for the improvement and the re–engineering of organizational and educational processes. Then, with the data loaded to the DWH, OLAP analysis was performed and data sets were created to experiment with EDM and BI. Moreover, data cubes were defined with the fact and dimension tables to obtain new explicit knowledge.

There are different experiments that can be carried out with the information generated from the proposed infrastructure. In a study, we investigated 42 articles that present experiments with EDM (Moscoso-Zea & Lujan-Mora, 2016). Among the results of this research, it was possible to determine the most applied methods and algorithms that can serve for experimenting and creating knowledge. These methods are classification, clustering, association, and prediction. This served to limit the experimental environment.

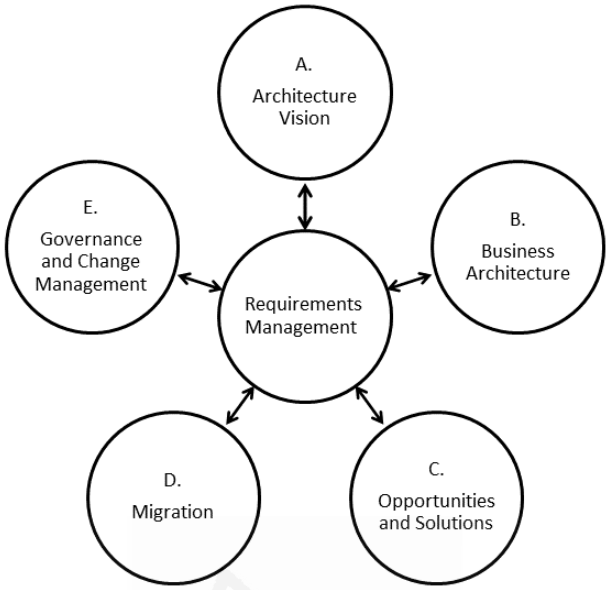


Figure 1.10: ADM adapted for the case study

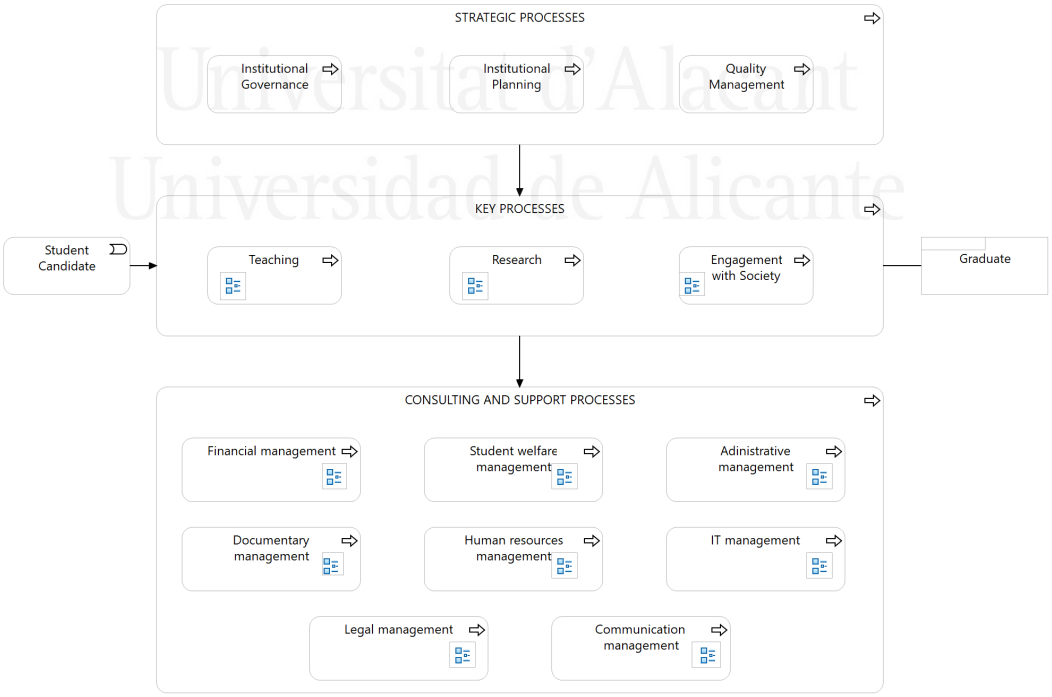


Figure 1.11: Model of a macro–process of an HEI

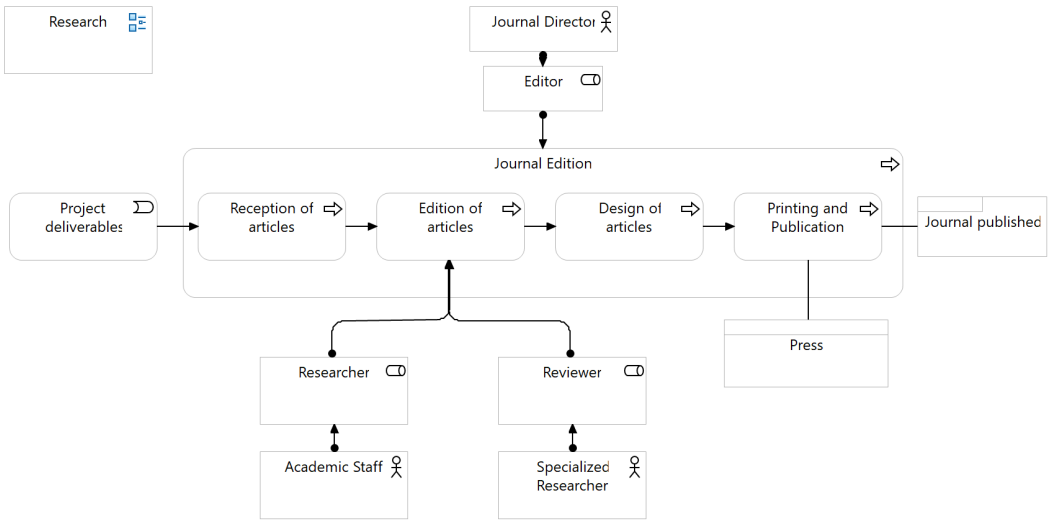


Figure 1.12: Example of a journal edition process in an HEI

Utility Scenarios of KM with EDM	Total
Predict final evaluations or student performance	21
Predict the profile and learning behavior of students	6
Improve teacher support	3
Improve the management of collaboration in educational environments	3
Generation of advice based on historical student data	2
Recommend students navigation links in web systems	2
Discover alternative representations of data in forums	1
Predict student dropout	1
Selection of practices adaptively	1
Predicting how differences between learning environments can increase or decrease the probability of participation	1
Improve decision making processes	1

Table 1.2: Scenarios for improvement with KM and EDM

The priority areas of experimentation that the papers describe to support in the creations of explicit knowledge that is then transformed into tacit knowledge (new knowledge of employees) are shown in Table 1.2. The most repeated scenario where KM is useful is the prediction of evaluations or student performance with 21 experiments and the prediction of the profile and learning behavior of students with 6 experiments.

1.5 Structure of the thesis

After first presenting the background of this dissertation and the synthesis of the research performed, located in Part I, the articles that contributed to the compendium of publications in this thesis are presented in Part II. Following that, in Part III the conclusions of this dissertation are drawn and proposals of potential future research in the KM field are discussed.

The thesis consists of the following chapters:

Chapter 1 Introduction, includes: motivation, objectives, method, results, the structure of the thesis and writing conventions

Chapter 2 Theoretical foundations, includes descriptions on the main topics of research.

Chapter 3 Publications and visibility, includes: publications in journals, conferences and other publications.

Chapter 4 Compendium of publications, includes all the articles published during the doctoral studies.

Chapter 5 Article published in Enfoque Ute Journal, including reference, contribution and full text.

Chapter 6 Article published in the Journal of Information Systems Engineering & Management, including reference, contribution and full text.

Chapter 7 Article published in the IEEE Access Journal, including reference, contribution and full text.

Chapter 8 Article published in the IEEE Access Journal, including reference, contribution and full text.

Chapter 10 Article published in the Australasian Journal of Engineering Education, including reference, contribution and full text.

Chapter 11 Conclusions, includes: conclusions of the work, contributions and future work.

2 Theoretical foundations

This chapter deals with the theoretical foundation of the thesis. It provides the basis for a better understanding of KM, EA, and BI. These fields form the core of the proposed infrastructure and are indispensable for the processes of knowledge generation and transferring in organizations. These topics will be presented in the following sections.

2.1 Knowledge management

KM is as a discipline which enables the efficient creation, use, distribution, and transfer of knowledge in organizations (Campbell, 2006). Innovations in science and technology have led to the emergence of intensively information-based organizations. These organizations need to transform this information into knowledge to secure competitiveness and improve decision making. The core dimensions that need to be examined in a KM project are people, processes, and technology (Edwards, 2011). Knowledge derived from these dimensions should be analyzed and stored using different information repositories. A Knowledge Management Framework (KMF) enables organizations to conduct and implement KM initiatives. KMFs are the foundation for developing information infrastructure and information systems to manage knowledge properly. Karemente, Aduwo, Mugejjera, and Lubega (2009), describe different KMFs; however, none of these integrates and analyzes the three knowledge dimensions as a whole and are difficult to use in practice.

2.2 Enterprise architecture

At present, there is no unanimity in the definition of EA. One of the reasons why there is no agreement in a standard definition could be because the practice of EA is still in a stage of evolution (Ivanov, 2009). An important definition of EA is given by Lankhorst (2013, p. 3): "EA is a coherent set of principles, methods and models that are used in the design, realization and maintenance of an enterprise's business architecture, organizational structure, information architecture and technology architecture with respect to the corporate strategy". This is the definition that is used in this dissertation. As seen in both definitions EA goes beyond Business–IT alignment, EA addresses the need to manage increasing complexity and deal with change management by providing a holistic view of the enterprise. EA is often viewed as a management practice that helps to improve the performance of enterprises (de Vries & van Rensburg, 2008). Moreover, EA addresses the problem of the increasing difficulty in delivering business value with the enterprise's information systems (Roger, 2007). The main methods to

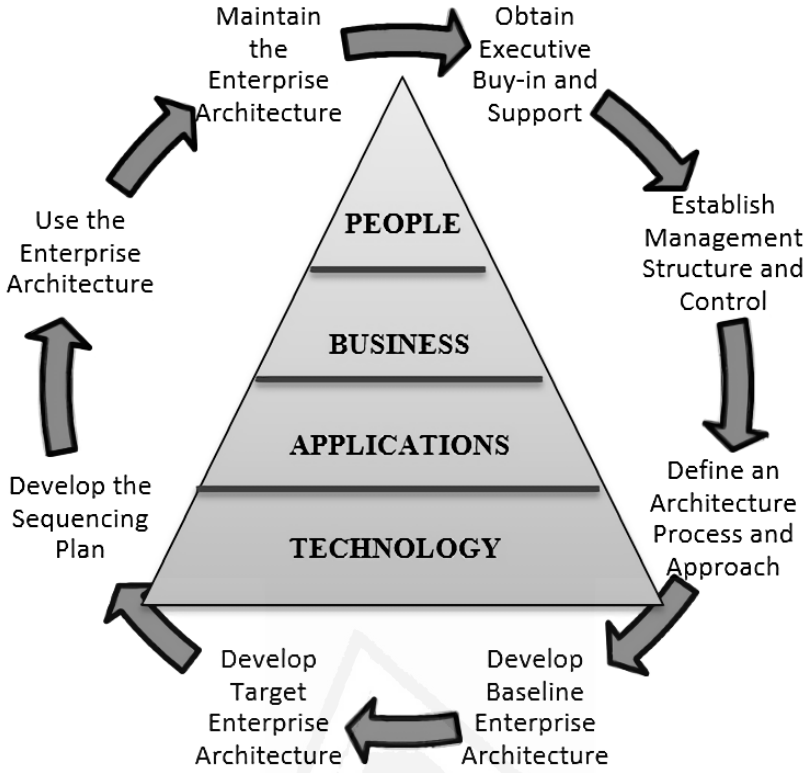


Figure 2.1: Enterprise architecture
(Tucker & Debrosse, 2003)

perform an EA are described in Figure 2.1. These methods are: obtain executive buy–in and support, establish a management structure and control, define an architecture process and approach, develop the EA baseline, develop the EA target, develop the sequencing plan, use and maintain the EA.

2.3 Business intelligence

Business Intelligence (BI) can be defined as a "system comprised of both technical and organizational elements that present historical information to its users for analysis and enables effective decision making and management support, for the overall purpose of increasing organizational performance" (Isik, Jones, & Sidorova, 2011, p. 161). BI builds upon a set of tools and applications that enable the analysis of vast amounts of information (Big Data) to improve decision making and the performance of organizations (Gartner Inc., 2019). To accomplish this objective, decision makers require having access to all organization's data to analyze the business, its requirements and its trends. According to Markets and Markets, the worldwide BI market is forecast to reach $26.88 billion in 2021, an increase of 63.58 % from 2016 (Markets and Markets, 2017). The scope of this market includes all BI and DA platforms, management suites and modern analytics solutions (Gartner Inc., 2019). One of the key applications of BI and DA is the generation of knowledge to enable informed decision-making (Lloyd, 2011).

3 Publications and visibility

3.1 Publications

Throughout the development of this doctoral thesis, and as a result of the research performed, different articles, book chapters, and conference papers have been published. All the publications present the results and the contributions made to the scientific community during this doctoral work. Some of the articles were developed to achieve the objectives of this dissertation, while others have been developed in collaboration with other researchers to fulfill parallel research goals (see Other Publications 3.1.3). The total of articles published during the doctoral period was 24, which included six journal papers (that are presented in the compendium), seven conference papers, and eleven papers considered other publications. The following sections list these work:

Journals, Six journal articles published: (Moscoso-Zea & Luján-Mora, 2017; Saa, Cueva, et al., 2017; Moscoso-Zea, Paredes-Gualtor, & Lujan-Mora, 2018; Moscoso-Zea, Castro, et al., 2019; Moscoso-Zea, Paredes-Gualtor, & Luján-Mora, 2019; Moscoso-Zea, Saa, & Luján-Mora, 2019)

Conferences, Seven conference papers published: (Moscoso-Zea, Luján-Mora, et al., 2016; Moscoso-Zea & Lujan-Mora, 2016; Moscoso-Zea, Sampedro, & Luján-Mora, 2016; Moscoso-Zea & Luján-Mora, 2016; Moscoso-Zea & Lujan-Mora, 2017; Moscoso-Zea, Vizcaino, & Luján-Mora, 2017; Paredes-Gualtor, Moscoso-Zea, & Luján-Mora, 2018)

Other publications, Eleven papers published in this category: (Sandoval et al., 2016; Sandoval, Galvez, & Moscoso-Zea, 2017; Saa, Moscoso-Zea, Costales, & Lujan-Mora, 2017; Saa, Moscoso-Zea, & Lujan-Mora, 2017; Paredes-Gualtor, Moscoso-Zea, Saa, Sandoval, & Rodas, 2017; Saa, Moscoso-Zea, & Lujan-Mora, 2018; Gomez-Torres, Moscoso-Zea, Herrera, & Lujan-Mora, 2018; Moscoso-Zea, Paredes-Gualtor, Saa, & Sandoval, 2018; Inga, Campaña, Hincapié, & Moscoso-Zea, 2018b, 2018a; Velasco et al., 2018)

3.1.1 Journals

This subsection describes the publications on scientific journals. These publications represent the main content of this research work. Therefore, they were included in the compendium of publications of this thesis. The details of the publications are shown in Table 3.1. The first column of the table shows the id of the journal; the second column shows the name of the journal with its ISSN; the third column shows the Journal Citations Report (JCR) impact factor; the fourth column of the table presents the

indexing of the journal, whether the article has been indexed in Scopus (SCO), Web of Science (WOS) or in the Directory of Open Access Journals (DOAJ); and the last column of the table describes the Scimago Journal Ranking (SJR) impact factor.

1. Moscoso-Zea, O., & Luján-Mora, S. (2017). Suggested Methodologies for Evaluation and Selection of Enterprise Architecture Software for Knowledge Digitization. Enfoque UTE, 8 (1), 315-328. (Moscoso-Zea & Luján-Mora, 2017): Chapter 5 explains what is presented in this article.

2. Saa, P., Cueva, A., Moscoso-Zea, O., & Luján-Mora, S. (2017). Moving ERP Systems to the Cloud - Data Security Issues. Journal of Information Systems Engineering & Management, 2 (4), 1-9. (Saa, Cueva, et al., 2017): Chapter 6 explains what is presented in this article.

3. Moscoso-Zea, O., Paredes-Gualtor, J., & Luján-Mora, S. (2018). A Holistic View of Data Warehousing in Education. IEEE Access, 6, 64659-64673.(Moscoso-Zea, Paredes-Gualtor, & Lujan-Mora, 2018): Chapter 7 explains what is presented in this article.

4. Moscoso-Zea, O., Castro, J., Paredes, J., Luján-Mora, S. (2019). A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics for Knowledge Management in Education. IEEE Access, 7, 38778-38788. (Moscoso-Zea, Castro, et al., 2019): Chapter 8 explains what is presented in this article.

5. Moscoso-Zea, O., Paredes-Gualtor, J., & Luján-Mora, S. (2019). Enterprise Architecture, an enabler of change and knowledge management. Enfoque UTE, 10 (1), 247-257. (Moscoso-Zea, Paredes-Gualtor, & Luján-Mora, 2019): Chapter 9 explains what is presented in this article.

6. Moscoso-Zea, O., Saa, P., & Luján-Mora, S. (2019). Evaluation of algorithms to predict graduation rate in higher education institutions by applying educational data mining. Australasian Journal of Engineering Education, 1-10. (Moscoso-Zea, Saa, & Luján-Mora, 2019): Chapter 10 explains what is presented in this article.

3.1.2 Conferences

This subsection describes the publications on conference proceedings. The details of the publications are shown in Table 3.2.

1. Moscoso-Zea, O., Luján-Mora, S., Esquetini Cáceres, C., & Schweimanns, N. (2016). Knowledge Management Framework using Enterprise Architecture and Business Intelligence. In 18th International conference on enterprise information systems (ICEIS) (pp. 244-249). Rome, Italy. (Moscoso-Zea, Luján-Mora, et al., 2016)

2. Moscoso-Zea, O., & Lujan-Mora, S. (2017). Knowledge management in higher education institutions for the generation of organizational knowledge. In 11th Iberian conference on information systems and technologies (CISTI) (pp. 1593-1599). Canary Islands, Spain. (Moscoso-Zea & Lujan-Mora, 2016)

ID	Journal	JCR IF	Indexing	SJR IF
J1	Enfoque UTE Journal EUJ. ISSN: 1390-6542	n/a	DOAJ, WOS	n/a
J2	Journal of Information Systems Engineering & Management JISM. ISSN: 2468-4376	n/a	DOAJ	n/a
J3	IEEE Access Journal IEEEAJ. ISSN: 2169-3536	Q1, 3.557	DOAJ, WOS, SCOPUS	Q1, 0.55
J4	IEEE Access Journal IEEEAJ. ISSN: 2169-3536	Q1, 3.557	DOAJ, WOS, SCOPUS	Q1, 0.55
J5	Enfoque UTE Journal EUJ. ISSN: 1390-6542	n/a	DOAJ, WOS	n/a
J6	Australasian Journal of Engineering Education AJEE. ISSN: 1325-4340	n/a	WOS, SCOPUS	Q3, 0.18

Table 3.1: Journal articles included in the compendium of publications

3. Moscoso-Zea, O., Sampedro, A., & Luján-Mora, S. (2016). Datawarehouse design for Educational Datamining. In 15th Information technology based higher education and training (ITHET) (pp. 1-6). Istanbul - Turkey. (Moscoso-Zea, Sampedro, & Luján- Mora, 2016)

4. Moscoso-Zea, O., & Luján-Mora, S. (2016). Evaluation and Selection of Enterprise Architecture Tools for Knowledge Management. In 1st International conference on information systems and computer science (INCISCOS) (pp. 292-296). Quito, Ecuador. (Moscoso-Zea & Luján-Mora, 2016)

5. Moscoso-Zea, O., & Lujan-Mora, S. (2017). Knowledge management in higher education institutions for the generation of organizational knowledge. In 12th Iberian conference on information systems and technologies (CISTI) (pp. 1593-1599). Lisbon, Portugal. (Moscoso-Zea & Lujan-Mora, 2017)

6. Moscoso-Zea, O., Vizcaino, M., & Luján-Mora, S. (2017). Evaluation of methods and algorithms of educational data mining. In Research in engineering education symposium (REES) (pp. 972-980). Bogota, Colombia. (Moscoso-Zea et al., 2017)

7. Paredes-Gualtor, J., Moscoso-Zea, O., & Luján-Mora, S. (2018). The role of enterprise architecture as a management tool. In 3rd International conference on information systems and computer science (INCISCOS) (p. 306-311). Quito, Ecuador. (Paredes-Gualtor et al., 2018)

3.1.3 Other publications

This subsection describes additional articles published during the doctoral period that target parallel research objectives that are not directly related to the research topic of this thesis. The details of the publications are shown in the following list:

ID	Conference	Location	Dates
C1	18th International conference on enterprise information systems (ICEIS), Indexed in WOS, SCOPUS	Rome, Italy	April 25-28, 2016
C2	11th Iberian conference on information systems and technologies (CISTI), Indexed in WOS, SCOPUS, IEEExplore	Canary Islands, Spain	June 15-18, 2016
C3	15th information technology based higher education and training (ITHET), Indexed in WOS, SCOPUS, IEEExplore	Istanbul, Turkey	September 8-10, 2016
C4	1st International conference on information systems and computer science (INCISCOS), Indexed in WOS, SCOPUS, IEEExplore	Quito, Ecuador	November 24-26, 2016
C5	12th Iberian conference on information systems and technologies (CISTI), Indexed in WOS, SCOPUS, IEEExplore	Lisbon, Portugal	June 21-24, 2017
C6	Research in engineering education symposium (REES), Indexed in SCOPUS	Bogota, Colombia	July 6-8, 2017
C7	3rd International conference on information systems and computer science (INCISCOS) Indexed in WOS, SCOPUS, IEEExplore	Quito, Ecuador	November 14-16, 2018

Table 3.2: Conference papers published during doctoral studies

1. Sandoval, F., Moscoso-Zea, O., Galvez, V., & Tutillo, P. (2016). Enterprise Architecture Framework with Agile Approach based on TOGAF. In 1st International conference on information systems and computer science (INCISCOS) (pp. 77-81). (Sandoval et al., 2016)

2. Sandoval, F., Galvez, V., & Moscoso-Zea, O. (2017). Development of Enterprise Architecture using a Framework with Agile Approach. Enfoque UTE, 8 (1),135-147. (Sandoval et al., 2017)

3. Saa, P., Moscoso-Zea, O., Costales, A., & Lujan-Mora, S. (2017). Data security issues in cloud-based Software-as-a-Service ERP. In Iberian conference on information systems and technologies (CISTI) (pp. 1828-1834). Lisbon - Portugal. (Saa, Moscoso-Zea, Costales, & Lujan-Mora, 2017)

4. Saa, P., Moscoso-Zea, O., & Lujan-Mora, S. (2017). Bring your own device (BYOD): Students perception - Privacy issues: A new trend in education? In 16th International conference on information technology based higher education and training (ITHET) (pp. 1-5). Ohrid, Macedonia. (Saa, Moscoso-Zea, & Lujan-Mora, 2017)

5. Paredes-Gualtor, J., Moscoso-Zea, O., Saa, P., Sandoval, F., & Rodas, P. (2017). Unified cloud computing adoption framework. In 2nd International conference on information systems and computer science (INCISCOS) (pp. 247-252). Quito - Ecuador. (Paredes-Gualtor et al., 2017)

6. Saa, P., Moscoso-Zea, O., & Lujan-Mora, S. (2018). Wearable technology, privacy issues. In Á. Rocha & T. Guarda (Eds.), Advances in intelligent systems and computing (Vol. 721, pp. 518-527). Springer International Publishing. (Saa et al., 2018)

7. Gomez-Torres, E., Moscoso-Zea, O., Herrera, N., & Lujan-Mora, S. (2018). Towards a forensic analysis of mobile devices using android. In Á. Rocha & T. Guarda (Eds.), Advances in intelligent systems and computing (Vol. 721, pp. 30-39). Springer International Publishing. (Gomez-Torres et al., 2018)

8. Moscoso-Zea, O., Paredes-Gualtor, J., Saa, P., & Sandoval, F. (2018). Moving the IT Infrastructure to the Cloud. Enfoque UTE, 9 (1), 79-89 (Moscoso-Zea, Paredes-Gualtor, Saa, & Sandoval, 2018)

9. Inga, E., Campaña, M., Hincapié, R., & Moscoso-Zea, O. (2018). Optimal Dimensioning of Electrical Distribution Networks Considering Stochastic Load Demand and Voltage Levels. In Applications of computational intelligence (Vol. 721, pp. 200-215). (Inga et al., 2018b)

10. Inga, E., Campaña, M., Hincapié, R., Moscoso-Zea, O. (2018). Optimal Deployment of FiWi Networks Using Heuristic Method for Integration Microgrids with Smart Metering. Sensors, 18 (8), 1-22. (Inga et al., 2018a)

11. Velasco, J., Ullauri, R., and Pilicita, L., Jácome, B., Saa, P., & Moscoso-Zea, O. (2018). Benefits of implementing an ISMS according to the ISO 27001 standard in the ecuadorian manufacturing industry. In 3rd International conference on information systems and computer science (INCISCOS) (p. 294-300). Quito - Ecuador. (Velasco et al., 2018)

3.2 Visibility

Scientific visibility is a must for researchers to show their results, increase their citations, receive feedback and criticism and interchange ideas with peers. Therefore, during the period of doctoral studies, some courses were taken in this subject and different academic profiles were created to increase the impact of the research performed. Table 3.3 shows the academic profiles of the author of this thesis. Moreover, one of the decisions prior to sending an article to a journal was to find out the availability of an open access version. That is why five out of the six compendium articles were published as fully open access papers (Moscoso-Zea & Luján-Mora, 2017; Saa, Cueva, et al., 2017; Moscoso-Zea, Paredes-Gualtor, & Lujan-Mora, 2018; Moscoso-Zea, Castro, et al., 2019; Moscoso-Zea, Paredes-Gualtor, & Luján-Mora, 2019).

The visibility strategy followed to disseminate the scientific work has increased the impact of the research performed. Therefore, as of May 2019, there are 95 citations in Google Scholar(GS), 23 citations in SCO and 12 citations in WOS.

ID	Academic Profile	URL
P1	ORCID	https://orcid.org/0000-0003-3233-2133
P2	Google Scholar	https://scholar.google.com/citations?user=c7k9An0AAAAJ&hl=es
P3	Research Gate	https://www.researchgate.net/profile/Oswaldo_Moscoso-Zea
P4	Mendeley	https://www.mendeley.com/profiles/oswaldo-moscoso1/
P5	Scopus	https://www.scopus.com/authid/detail.uri?authorId=57190377269
P6	Publons - Researcher Id	https://publons.com/researcher/2935675/oswaldo-moscoso-zea/
P7	Academia	https://ute.academia.edu/OMoscoso

Table 3.3: Academic profiles

Part II

PUBLISHED WORK

4 Compendium of publications

This chapter presents the main publications of the research performed during the doctoral studies in chronological order. All the publications included in the compendium were journal articles. Three of the publication had an impact factor. Two of them were indexed in a journal classified in the first quartile (Q1) of the Journal Citations Report[1] of WOS (Moscoso-Zea, Paredes-Gualtor, & Lujan-Mora, 2018; Moscoso-Zea, Castro, et al., 2019) and one in the third quartile (Q3) of the Scimago Journal Rank[2] of SCO (Moscoso-Zea, Saa, & Luján-Mora, 2019). The details of the publications are drawn in Figure 4.1.

[1]https://clarivate.com/products/journal-citation-reports/
[2]https://www.scimagojr.com/

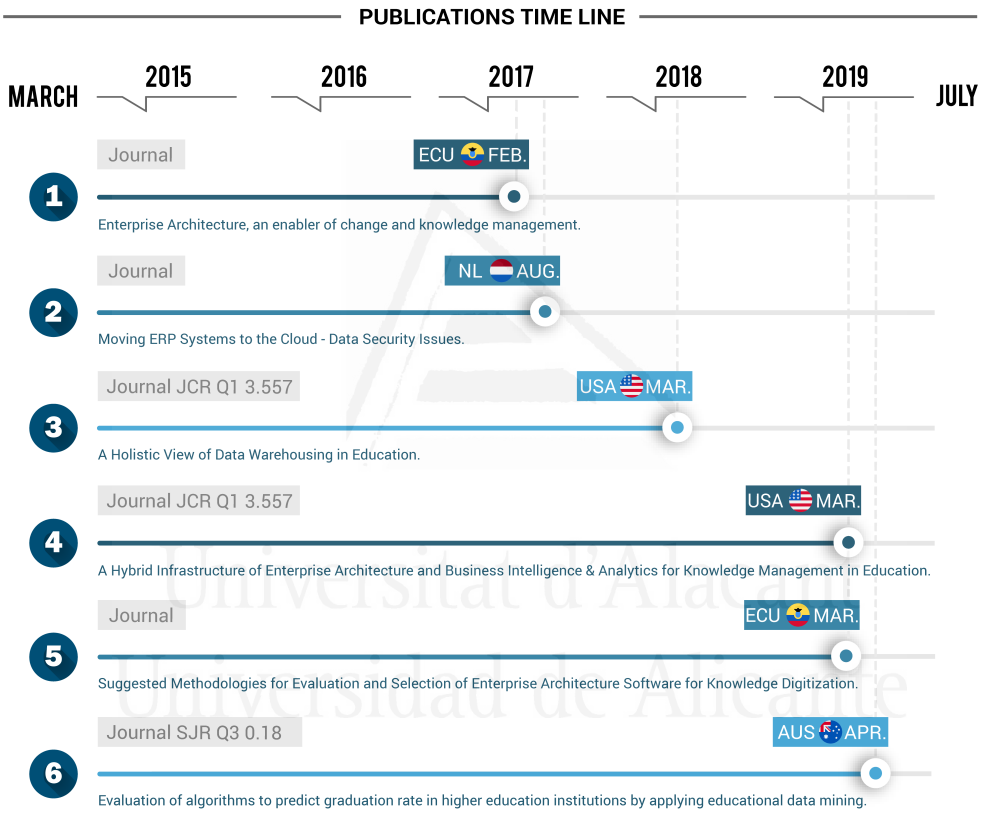


Figure 4.1: Compendium of publications–timeline

5 Suggested Methodologies for Evaluation and Selection of Enterprise Architecture Software for Knowledge Digitization

Reference:

Moscoso-Zea, O., & Luján-Mora, S. (2017). Suggested Methodologies for Evaluation and Selection of Enterprise Architecture Software for Knowledge Digitization. Enfoque UTE, 8 (1), 315-328. (Moscoso-Zea & Luján-Mora, 2017)

Available at:

URL: http://ingenieria.ute.edu.ec/enfoqueute/index.php/revista/article/view/144/151
DOI: https://doi.org/10.29019/enfoqueute.v8n1.144

Topics to which it contributes:

1. State of the art of EA for knowledge digitization.
2. Presentation of a method for knowledge digitization.
3. Suggested methodologies for the evaluation and the selection of EA tools for KM.

Enfoque UTE, V.7-Sup.1, Feb.2017, pp.315 - 328
http://ingenieria.ute.edu.ec/enfoqueute/
e-ISSN: 1390-6542 / p-ISSN: 1390-9363

Recibido (Received): 2017/01/09
Aceptado (Accepted): 2017/02/24

Metodologías Sugeridas de Evaluación y Selección de Software de Arquitectura Empresarial para la Digitalización del Conocimiento

(Suggested Methodologies for Evaluation and Selection of Enterprise Architecture Software for Knowledge Digitization)

Oswaldo Moscoso-Zea[1], Sergio Lujan-Mora[2]

Resumen:
La gestión del conocimiento (GC) es una práctica de administración que permite la creación, uso, distribución y transferencia del conocimiento en organizaciones. Existen marcos de referencia de GC que incluyen componentes de inteligencia de negocios o de arquitectura empresarial (AE) para la implementación de proyectos de GC en organizaciones. Las herramientas de AE son usadas para digitalizar, relacionar y visualizar varias dimensiones de conocimiento como son: estructura organizacional, procesos de negocio, aplicaciones e infraestructura tecnológica. Este artículo tiene como objetivo evaluar el rol de la AE como componente clave de la GC y sugerir metodologías que pueden ser usadas para evaluar herramientas de software de AE. Para esto se realizó una revisión bibliográfica de las metodologías de evaluación de software existente en el mercado y se eligieron aquellas que pueden ser adaptadas para el campo de la AE. Está investigación fue cualitativa y exploratoria a través de un caso de estudio desarrollado en un proveedor de servicios logísticos con presencia internacional. El caso de estudio muestra el proceso realizado para elegir la metodología de evaluación. Además, se describen los pasos para la digitalización del conocimiento organizacional.

Palabras clave: Gestión del Conocimiento; Arquitectura Empresarial, Metodología de Evaluación de Herramientas de Arquitectura Empresarial.

Abstract:
Knowledge Management (KM) is a practice that allows the creation, use, distribution and transfer of knowledge in organizations. Different KM frameworks exist that include business intelligence or enterprise architecture (EA) components for the implementation of KM in organizations. EA tools are used to digitize, relate and visualize the following dimensions of knowledge: organizational structure, business processes, applications and technology. The objective of this paper is to assess the role of EA as a key component in KM and to suggest software evaluation methodologies that can be adapted for the field of EA. For this, an investigation was realized to identify the existing software evaluation methodologies in the market and to filter those that can be adapted for the field of EA. The methodology used for the research was qualitative and exploratory using a case study performed in an international logistic service provider. The case study describes the process done for the selection of the evaluation methodology. Furthermore, it describes the steps for knowledge digitization.

Keywords: Knowledge Management; Enterprise Architecture; Evaluation of Enterprise Architecture Tools.

[1] Universidad Tecnológica Equinoccial, Quito – Ecuador (omoscoso@ute.edu.ec)
[2] Universidad Alicante, Alicante – España (sergio.lujan@ua.es)

6 Moving ERP Systems to the Cloud - Data Security Issues

Reference:

Saa, P., Cueva, A., Moscoso-Zea, O., & Luján-Mora, S. (2017). Moving ERP Systems to the Cloud - Data Security Issues. Journal of Information Systems Engineering & Management, 2 (4). (Saa, Cueva, et al., 2017)

Available at:

URL: https://www.jisem-journal.com/article/8972P1SA
DOI: https://doi.org/10.20897/jisem.201721

Topics to which it contributes:

1. State of the art of cloud computing and data security issues in the cloud.
2. Guidelines on moving the enterprise systems to the cloud.

JOURNAL OF INFORMATION SYSTEMS ENGINEERING & MANAGEMENT

Journal of Information Systems Engineering & Management, 2017, 2(4), 21
ISSN: 2468-4376

lectito

Moving ERP Systems to the Cloud - Data Security Issues

Pablo Saa[1]*, Andrés Cueva Costales[2], Oswaldo Moscoso-Zea[1], Sergio Lujan-Mora[3]

[1] *Faculty of Engineering Sciences, Universidad Tecnológica Equinoccial*, Quito, ECUADOR
[2] *DGPCTI Public Company Yachay*, Quito, ECUADOR
[3] *Dept. of Software and Computing Systems, University of Alicante*, Alicante, SPAIN

*__Corresponding Author:__ psaa@ute.edu.ec

Citation: Saa, P., Costales, A.C., Moscoso-Zea, O. and Lujan-Mora, S. (2017). Moving ERP Systems to the Cloud - Data Security Issues. *Journal of Information Systems Engineering & Management*, 2(4), 21. https://doi.org/10.20897/jisem.201721

Published: August 30, 2017

ABSTRACT

This paper brings to light data security issues and concerns for organizations by moving their Enterprise Resource Planning (ERP) systems to the cloud. Cloud computing has become the new trend of how organizations conduct business and has enabled them to innovate and compete in a dynamic environment through new and innovative business models. The growing popularity and success of the cloud has led to the emergence of cloud-based Software-as-a-Service (SaaS) ERP systems, a new alternative approach to traditional on-premise ERP systems. Cloud-based ERP has a myriad of benefits for organizations. However, infrastructure engineers need to address data security issues before moving their enterprise applications to the cloud. Cloud-based ERP raises specific concerns about the confidentiality and integrity of the data stored in the cloud. Such concerns that affect the adoption of cloud-based ERP are based on the size of the organization. Small to medium enterprises (SMEs) gain the maximum benefits from cloud-based ERP as many of the concerns around data security are not relevant to them. On the contrary, larger organizations are more cautious in moving their mission critical enterprise applications to the cloud. A hybrid solution where organizations can choose to keep their sensitive applications on-premise while leveraging the benefits of the cloud is proposed in this paper as an effective solution that is gaining momentum and popularity for large organizations.

Keywords: ERP, cloud computing, cloud ERP, data security, confidentiality, integrity

INTRODUCTION

Nowadays, "the cloud" has been a buzzword in the last few years and has caused a revolution in the Information and Communication Technologies (ICT) industry. As IBM states "Cloud computing, often referred to as simply 'the cloud,' is the delivery of on-demand computing resources, everything from applications to data centers over the Internet on a pay-for-use basis" (IBM, 2015). This new trend changes the way organizations deploy services, platforms and infrastructure of Information Technologies (IT). The variety of applications and services offered by this new concept affect on one hand organizations and individuals who notice the benefits of cloud services in terms of efficiency, flexibility and reduced investment effort while on the other hand, technology companies and traditional operators see an opportunity to expand their businesses (Lin and Chen, 2012).

According to Gartner, cloud-based services can be defined as "massively scalable system capabilities delivered as a service to external users using Internet technologies" (Gartner, 2015). A study about cloud computing models describes that based on the completeness and abstraction levels of services delivered to the end user, there are

7 A Holistic View of Data Warehousing in Education

Moscoso-Zea, O., Paredes-Gualtor, J., & Luján-Mora, S. (2018). A Holistic View of Data Warehousing in Education. IEEE Access, 6, 64659-64673. (Moscoso-Zea, Paredes-Gualtor, & Lujan-Mora, 2018)

Available at:

URL: https://ieeexplore.ieee.org/abstract/document/8501908
DOI: https://doi.org/10.1109/ACCESS.2018.2876753

Topics to which it contributes:

1. State of the art on DW in education.
2. Proposal of a method for a DWH project implementation.
3. Guidelines for the implementation of a DWH in educational institutions.

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 26, 2018, accepted October 1, 2018, date of publication October 22, 2018, date of current version November 30, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2876753

A Holistic View of Data Warehousing in Education

OSWALDO MOSCOSO-ZEA[1], JOEL PAREDES-GUALTOR[1], AND SERGIO LUJÁN-MORA[2]
[1]Facultad de Ciencias de la Ingeniería e Industrias, Universidad Tecnológica Equinoccial, Quito 170509, Ecuador
[2]Departamento de Lenguajes y Sistemas Informáticos, Universidad de Alicante, 03690 Alicante, Spain

Corresponding author: Oswaldo Moscoso-Zea (omoscoso@ute.edu.ec)

This work was supported by Universidad Tecnológica Equinoccial.

ABSTRACT Data warehousing (DW) is a widespread and essential practice in business organizations that support the data analytic and decision-making process. Despite the importance of DW in complex organizations, the adoption of a data warehouse (DWH) in education is apparently lower compared with other industries. To clarify this situation, this paper presents a systematic mapping that includes the study of empirical research papers from 2008 to 2018 on the topic of DW in education. For this paper, we applied a qualitative and quantitative approach based on a four-stage research method with the objective to have a holistic view of DWHs in education. After filtering and applying the proposed method, 34 relevant papers were identified and studied in detail. The study revealed interesting facts; for example, Kimball's approach is the most applied methodology for DWH design in education. In addition, a mapping between this comprehensive collection of research papers covering educational DW and six dimensions of analysis (schema proposal, analysis of the user requirements, analysis of the business requirements, effectiveness, implementation, and data analysis) was performed. From this analysis, we discovered that the star schema is the most implemented approach. The purpose of the mapping was to explore and identify the priority areas of research and the research gaps within the academic community. These gaps are a source of opportunities to start new lines of research.

INDEX TERMS Business intelligence, data warehouse, educational data warehouse, systematic mapping.

I. INTRODUCTION

Data warehousing (DW) is the process of storing, managing and analyzing large amounts of historical, summarized and non-volatile data. These data are extracted from multiple heterogeneous data sources into a single multi-dimensional repository called data warehouse (DWH). The core objective of DW is to provide greater insights into the performance of an organization and improve decision-making [1]. The complementary fields that study the analysis of the data in this repository are data analytics (DA) and on-line analytical processing (OLAP). On the one hand, DA is the process of analyzing the data in the DWH using technological and statistical tools with the purpose to draw conclusions and generate knowledge from the information it contains [2]. On the other hand, OLAP is the process of exploiting the DWH for multidimensional analysis by applying data cube operations as roll-up, drill-down, slicing and dicing on the dimensions and fact tables [3]. DA and OLAP, plus additional technological tools, are part of what data scientists call business intelligence (BI).

Even though DW and BI are widely used in business organizations and have been exhaustively analyzed from the industry standpoint for many years [4]–[7], its use is still low in educational institutions. The paper from Shahid *et al.* [8] presents a set of case studies performed to determine the percentage of use of DWHs within different industries. The results of the case study analysis show that the industry which uses DW the most is the medical industry (hospitals, clinics and physician offices) with 23.3% of usage. Following that, the finance and banking industry with 6.2% of DW usage, whereas one of the industries that less use DW is education with only 3.8% of usage [8]. Although these facts reveal that not much effort has been conducted to overcome the barriers of adoption of DW in education, nowadays boards of governance and directors in educational institutions are recognizing the potential and the leading role that BI and DW plays in improving educational and organizational processes [9]–[11]. Additionally, some studies suggest to carry out an implementation of a DWH in educational scenarios to improve decision-making and knowledge management [12], [13].

Moreover, different studies are giving major importance to the educational data mining (EDM) topic which bases its analytical process in a well-designed DWH to store and maintain the information of their application systems [14], [15].

8 A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics for Knowledge Management in Education

Moscoso-Zea, O., Castro, J., Paredes, J., Luján-Mora, S. (2019). A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics for Knowledge Management in Education. IEEE Access, 7, 38778 - 38788. (Moscoso-Zea, Castro, et al., 2019)

Available at:

URL: https://ieeexplore.ieee.org/document/8672110
DOI: https://doi.org/10.1109/ACCESS.2019.2906343

Topics to which it contributes:

1. State of the art on KM, BI and EA.
2. Proposal of a hybrid infrastructure for KM.
3. Method for KM and the implementation of the hybrid infrastructure.
4. Provide mechanisms to transfer the tacit knowledge of employees to the knowledge infrastructure.

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 6, 2019, accepted March 12, 2019, date of publication March 20, 2019, date of current version April 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2906343

A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics for Knowledge Management in Education

OSWALDO MOSCOSO-ZEA[1], JORGE CASTRO[1], JOEL PAREDES-GUALTOR[1], AND SERGIO LUJÁN-MORA[2]

[1]Facultad de Ciencias de la Ingeniería e Industrias, Universidad Tecnológica Equinoccial, Quito 170509, Ecuador
[2]Departmento de Lenguas y Sistemas Informáticos, Universidad de Alicante, 03690 Alicante, Spain

Corresponding author: Oswaldo Moscoso-Zea (omoscoso@ute.edu.ec)

This work was supported by the Universidad Tecnológica Equinoccial.

ABSTRACT Advances in science and technology, the Internet of Things, and the proliferation of mobile apps are critical factors to the current increase in the amount, structure, and size of information that organizations have to store, process, and analyze. Traditional data storages present technical deficiencies when handling huge volumes of data and are not adequate for process modeling and business intelligence; to cope with these deficiencies, new methods and technologies have been developed under the umbrella of big data. However, there is still the need in higher education institutions (HEIs) of a technological tool that can be used for big data processing and knowledge management (KM). To overcome this issue, it is essential to develop an information infrastructure that allows the capturing of knowledge and facilitates experimentation by having cleaned and consistent data. Thus, this paper presents a hybrid information infrastructure for business intelligence and analytics (BI&A) and KM based on an educational data warehouse (EDW) and an enterprise architecture (EA) repository that allows the digitization of knowledge and empowers the visualization and the analysis of dissimilar organizational components as people, processes, and technology. The proposed infrastructure was created based on research and will serve to run different experiments to analyze educational data and academic processes and for the creation of explicit knowledge using different algorithms and methods of educational data mining, learning analytics, online analytical processing (OLAP), and EA analytics.

INDEX TERMS Big data, business intelligence, data warehouse, educational data mining, knowledge management.

I. INTRODUCTION

Boards of directors in higher education institutions (HEIs) are recognizing the potential and the leading role that Big Data analysis and knowledge management (KM) plays to improve the decision making processes [1]–[4]. The task of analyzing information in HEIs is currently challenging, mainly due to two factors. Firstly, due to a large amount of information generated every day by different applications and gadgets; and secondly, due to the problem of distributed and heterogeneous information systems. The latter problem is known as islands of information or information silos, which is produced by standalone applications dispersed in different departments and units of most HEIs. This problem has its roots in the disorganized and non-centralized growth of most HEIs due to the autonomy that their units have. These factors increase the complexity to store, integrate, process and analyze data in HEIs.

Moreover, this complexity upsurges due to the emergence and implementation of new technological devices and software as wearables, apps and web applications which increases the volume of data. One of the solutions to tackle this complexity and to perform a sound analysis of this huge amount of information known as Big Data is the development of an information infrastructure, which integrates processed and cleaned data loaded from different internal and external data sources [5]. This infrastructure for HEIs allows carrying

The associate editor coordinating the review of this manuscript and approving it for publication was Saqib Saeed.

9 Enterprise Architecture, an enabler of change and knowledge management

Moscoso-Zea, O., Paredes-Gualtor, J., & Luján-Mora, S. (2019). Enterprise Architecture, an enabler of change and knowledge management. Enfoque UTE, 10 (1), 247-257. (Moscoso-Zea, Paredes-Gualtor, & Luján-Mora, 2019)

Available at:

URL: http://ingenieria.ute.edu.ec/enfoqueute/index.php/revista/article/view/459
DOI: https://doi.org/10.29019/enfoqueute.v10n1.459

Topics to which it contributes:

1. State of the art on EA as a management tool.
2. Proposal for knowledge management using EA with other management tools.

Enfoque UTE, V.10-N.1, Mar.2019, pp. 247-257
http://ingenieria.ute.edu.ec/enfoqueute/
e-ISSN: 1390-6542 / p-ISSN: 1390-9363

Recibido (Received): 2018/11/16
Aceptado (Accepted): 2019/03/06

Enterprise Architecture, an enabler of change and knowledge management

(Arquitectura Empresarial, un habilitador del cambio y la administración del conocimiento)

Oswaldo Moscoso-Zea,[1] *Joel Paredes-Gualtor,*[1] *Sergio Luján-Mora*[2]

Abstract

Organizations around the world require a sound process of change management to innovate and remain competitive over time. Change and knowledge management needs to be supported with the right tools to overcome the challenges of transformations and transitions in the business models and processes of diverse organizations. Steering boards can use enterprise architecture (EA) to implement new knowledge management initiatives in their strategic planning. EA allows companies to model the current situation (as-is models) of the organization and the desired future scenarios (to-be models) and to establish road maps to enable adequate transformations. Different frameworks exist in the market that support the management of organizations, for example: Control Objectives for Information and Related Technologies (COBIT), Information Technology Infrastructure Library (ITIL), quality models such as the one proposed by the European Foundation for Quality Management (EFQM) and systems such as the Balanced Scorecard (BSC) are widely used for the management of business and information technologies (IT). However, EA is not widely used with the other mentioned tools. This paper analyzes EA as a tool for change and knowledge management and compares its functionality with other frameworks in the market. The analysis performed in this paper checks if EA can be used and is compatible with other frameworks. To answer this question, an analysis of the most important processes, good practices, perspectives and tools provided by each framework was performed.

Keywords

Change Management; COBIT; Enterprise Architecture; ITIL, Knowledge Management

Resumen

Las organizaciones de todo el mundo requieren un proceso sólido de gestión del cambio para innovar y seguir siendo competitivas a lo largo del tiempo. La gestión del cambio y del conocimiento debe ser respaldada con las herramientas adecuadas para superar los desafíos de las transformaciones y transiciones en los modelos y procesos de negocios de diversas organizaciones. Las juntas directivas pueden utilizar la arquitectura empresarial (AE) para implementar nuevas iniciativas de gestión del conocimiento en su planificación estratégica. AE permite a las empresas modelar la situación actual de la organización y los escenarios futuros deseados y establecer mapas de ruta para permitir las transformaciones adecuadas. Existen diferentes marcos de trabajo que apoyan la gestión, por ejemplo: Objetivos de control para información y tecnologías relacionadas (COBIT, por su sigla en inglés), Biblioteca de infraestructura de tecnología de la información (ITIL, por su sigla en inglés), modelos de calidad como el de la Fundación Europea para la Gestión de la Calidad (EFQM, por su sigla en inglés) y sistemas como el Cuadro de mando Integral (BSC, por su sigla en inglés). Sin embargo, la AE no es muy utilizada junto con las herramientas mencionadas. Este documento analiza AE como una herramienta para el cambio y la gestión del conocimiento y compara su funcionalidad con otros marcos en el mercado. El análisis realizado en este documento comprueba si se puede utilizar AE y es compatible con otros marcos de trabajo para gestionar el cambio organizativo y la gestión del conocimiento.

Palabras clave

Gestión del Cambio; COBIT; Arquitectura Empresarial; ITIL, Gestión del Conocimiento.

1 Universidad UTE, Quito, Ecuador ({omoscoso, joel.paredesl} @ute.edu.ec).
2 Universidad de Alicante, Alicante, España (sergio.lujan@ua.es)

10 Evaluation of Algorithms to Predict Graduation Rate in Higher Education Institutions by Applying Educational Data Mining

Moscoso-Zea, O., Saa, P., & Luján-Mora, S. (2019). Evaluation of algorithms to predict graduation rate in higher education institutions by applying educational data mining. Australasian Journal of Engineering Education, 1-10. (Moscoso-Zea, Saa, & Luján-Mora, 2019)

Available at:

URL: https://www.tandfonline.com/doi/abs/10.1080/22054952.2019.1601063
DOI: https://doi.org/10.1080/22054952.2019.1601063

Topics to which it contributes:

1. State of the art on EDM.
2. Evaluation and recommendation of algorithms for EDM.
3. Case study with an analysis of educational indicators that can be improved with EDM or BI.

AUSTRALASIAN JOURNAL OF ENGINEERING EDUCATION
https://doi.org/10.1080/22054952.2019.1601063

ARTICLE

Evaluation of algorithms to predict graduation rate in higher education institutions by applying educational data mining

Oswaldo Moscoso-Zea[a], Pablo Saa[a] and Sergio Luján-Mora[b]

[a]Faculty of Engineering Sciences, Universidad Tecnológica Equinoccial, Quito, Ecuador; [b]Department of Software and Computing Systems, University of Alicante, Alicante, Spain

ABSTRACT

Nowadays, researchers analyse student data to predict the graduation rate by looking at the characteristics of students enrolled and to take corrective actions at an early stage or improve the admission process. Educational data mining (EDM) is an emerging field that can support the implementation of changes in the management of higher education institutions. EDM analyses educational data using the development and the application of data mining (DM) methods and algorithms to information stored in academic data repositories. The purpose of this paper is to review which methods and algorithms of DM can be used in the analysis of educational data to improve decision-making. Furthermore, it evaluates these algorithms using a dataset composed of student data in the computer science school of a private university. The core of the analysis is to discover trends and patterns of study in the graduation rate indicator. Finally, it compares these methods and algorithms and suggests which has the best precision in certain scenarios. Our analyses suggest that random trees had better precision but had limitations due to the difficulty of interpretation while the J48 algorithm had better possibilities of interpretation of results in the visualisation of the classification of data and only had slightly inferior performance.

ARTICLE HISTORY
Received 31 January 2018
Accepted 25 March 2019

KEYWORDS
Data mining; data warehouse; educational data mining; academic development

1. Introduction

In today's information era, data are collected and stored in large repositories. The huge amounts of information that educational institutions generate every day call for improved ways of storing and analysing data. The process of converting data into information, and information into knowledge, has to be done by following a comprehensive method to produce the expected outputs.

Higher education institutions are generating large amounts of data from their organisational systems and applications which could be more effectively used to discover trends and predict events in education. In the same manner, as in other industries, the right data management and data visualisation can grant stakeholders with insights to improve organisational processes. Knowledge obtained from data analytics and data mining (DM) are enablers to ensure quality in the educational process, and therefore, it offers directors different viewpoints to improve the education generally. However, DM is not a solution itself at this point, instead, it is a tool which supports the decision-making process through the acquisition of knowledge in order to solve different problems (Buldu, 2010). The production and dissemination of organisational knowledge is a strategic objective that supports higher education institutions in the roadmap for planning, modernisation, and improvement of academic and research indicators (Baepler and Murdoch 2010) (Moscoso-Zea and Lujan-Mora 2017). Thus, it is extremely necessary to establish mechanisms to store data of the highest possible quality and apply EDM methods.

Student success is an essential objective of higher education institutions, so the technological infrastructure of these institutions must include data warehouses to support sound data storage and analysis, as one of the core technologies in this field. A data warehouse is a data repository modelled with a multidimensional design and used specifically for analysis (Moscoso-Zea, Sampedro, and Luján- Mora 2016). The information dispersed in different operational databases is migrated to the data warehouse using extraction, transformation, and loading (ETL) processes. An approach that guides this knowledge creation process is called knowledge discovery in databases (KDD). KDD uses DM as the core element for knowledge creation. Because of this, DM has been applied to different industries and fields of study in the last years with promising results. Some of the fields of analysis of DM are marketing, health, finances, and insurance, among others. The application of DM in educational contexts is known as educational DM (EDM). EDM is a discipline in evolution that focuses on the design of models to improve learning experiences and organisational efficiency (Huebner 2013).

CONTACT Pablo Saa pablo.saa@ute.edu.ec Faculty of Engineering Sciences, Universidad Tecnológica Equinoccial, Quito, Ecuador

Part III

CONCLUSIONS AND OUTLOOK

11 Conclusions

11.1 Discussion

This thesis presents the research performed to define a KM method in HEIs. To accomplish this goal a knowledge infrastructure was proposed. The proposed hybrid information and knowledge infrastructure present a novel approach that includes a DWH and an EA tool. This infrastructure was designed based on empirical research intended to improve HEIs' management. After analyzing the different possibilities we have proposed a method for the implementation of the components of the infrastructure which were evaluated and tested in an HEI. With the infrastructure implemented and by using BI and EDM, different experiments can be carried out to improve academic indicators. The results of these experiments give visibility to the board of directors for timely decision making and are a source of knowledge for improving education in general.

The SM that we performed on the topic of DWH in education shows that although there is an increase in experimentation with EDM and more studies on the topic, researchers and managers in HEIs do not give adequate importance to the implementation of an educational DWH which can enhance the results of the experiments. For that reason, the implementation of the proposed infrastructure can support in the creation and management of the knowledge dimensions generated in the organization. Thus, it can be a source of opportunities to improve organizational processes and a valid source for applying different type of analytics with OLAP tools, dashboards, reporting tools, and EA tools to assure quality outcomes in the BI process.

The implementation of this infrastructure in HEIs is challenging; mainly due to the business model in HEIs, in which the strategic direction can change abruptly. Which is very usual when there is a change of authorities. The proposed infrastructure is being implemented in an HEI with promising results. Nevertheless, it is important to establish mechanisms for the evaluation of the improvements made in knowledge processing.

The hybrid infrastructure design considers the fact of unexpected changes in HEIs and is very flexible to ease disruptive changes in the strategic direction and management. The proposed knowledge infrastructure facilitates the traceability of the students, administrative and academic staff to improve decision making and to improve organizational processes. Besides, the infrastructure is the knowledge base of the organization and a tool to produce explicit knowledge. Moreover, the use of EA, as a part of the solution, can support the externalization of knowledge proposed by Nonaka (1994) in the SECI model for the transformation of tacit knowledge from employees and academic staff into explicit knowledge drawn in the architectural models. These EA models can ease the problems derived from staff turnover which is the main cause

of the loss of implicit knowledge and the increase in operating costs.

The knowledge framework and the knowledge infrastructure proposed within this work provides the basis for the creation, transferring and management of knowledge which can be a cornerstone to become an HEI of excellence. The use of BI and EA tools bridges the gap of capturing all the knowledge dimensions. On the one hand, BI allows the transformation of simple information in valuable knowledge by applying OLAP and DM methods and techniques. On the other hand, EA supports the digitization of implicit knowledge from people and processes by creating architectures in different domains. These architectures facilitate the transfer and distribution of knowledge to different levels of people in the organization. Some benefits of using this framework are reduced training costs of staff turnover, improved decision making processes and the creation of a knowledge repository. All in all, it has been proved with the research done during the doctoral studies that this proposal can be used to improve the knowledge flows of the organization which at the same time are used to improve different areas of HEIs and the learning–teaching process.

11.2 Contributions

This work has contributed in different areas. The main contribution is to give the scientific community a mechanism to improve KM in HEIs. The proposed infrastructure attempts to store the organizational knowledge from all the knowledge dimensions in an hybrid infrastructure. These knowledge can be reused in new projects and can be the basis for knowledge creation by using EA and BI analytics. The contributions of the publications are detailed in the following list:

- Presentation of a method for knowledge digitization.
- Suggested methodologies for the evaluation and the selection of EA tools for KM.
- State of the art of KM, EA, BI, DWH, EDM and cloud computing.
- Guidelines on moving the enterprise systems to the cloud.
- Guidelines for the implementation of a DWH in educational institutions.
- Proposal of a hybrid infrastructure for KM.
- Method for KM and the implementation of the hybrid infrastructure.
- Mechanisms for the externalization process of knowledge creation (convert tacit knowledge to explicit knowledge). This can be done by capturing the tacit knowledge from employees into models in the EA tool.
- Proposal for KM using EA with other management tools.
- Evaluation and recommendation of algorithms to be used in experiments with EDM to improve educational indicators.
- Case study for the application of the infrastructure.

11.3 Future work

As future work, we propose to research more mechanisms to capture the implicit knowledge from the staff of the organization and to automate the ETL process in order to have the repository updated all the time. Moreover, the hybrid infrastructure presented in this work can help improve the efficiency of an HEI. The models captured in the EA tools can be a source for new software development when a labor-intensive process is identified and can be automated.

It is also proposed to generate new modeling objects within the EA tools palette that can support the modeling of educational processes. Moreover, it is important to analyze the information from social networks and from mobile apps that can contribute to enhance the educational process and to use that information to increase student's admissions.

This work also opens the doors for other areas of research, for example, it can be studied the use that can be given to the infrastructure to become the organizational knowledge base that can improve project management. Furthermore, it will be important to test if an organizational process is efficient or not by analyzing the data of the EA. Another example could test if the existing EA models could improve the marketing process by better understanding consumer behavior and their shopping habits.

The proposed infrastructure can also be tested in business organizations for innovation management. In this type of organizations, the proposal of this thesis can serve to improve the design of products and services as well as creating new business models.

Moreover, the proposed infrastructure has clean and quality data that can be used in future experiments with the different algorithms and methods of EDM. The future seems promising, however, it all depends on the support of the authorities of HEIs to maintain the infrastructure working, and the access to the educational data at all levels of the organization.

APPENDIX

A Resumen

En cumplimiento de la normativa de la Universidad de Alicante sobre "TESIS EN LENGUAS DISTINTAS A LAS OFICIALES EN LA COMUNIDAD AUTÓNOMA VALENCIANA"[1] que indica:

En cualquier caso, la tesis deberá contener un resumen en una de las dos lenguas oficiales de esta Comunidad Autónoma. Este resumen deberá contener una introducción general, un resumen global de los resultados obtenidos, de la discusión de estos resultados y de las conclusiones finales. Este resumen deberá dar una idea bastante precisa del contenido de la Tesis. La extensión de la parte escrita en uno de los idiomas oficiales de esta Comunidad Autónoma no será inferior a 5.000 palabras.

Se incluye a continuación este apéndice que es un resumen del trabajo de investigación realizado.

A.1 Introducción

Los grandes volúmenes de datos (Big Data) que se generan hoy en día a escala global y dentro de las organizaciones, junto con el conocimiento que reside en las personas y aquel que se plasma en los procesos de negocio hace muy compleja la gestión del conocimiento (GC) organizacional. Una correcta GC puede ser fuente de oportunidades y de ventaja competitiva para aquellas organizaciones que usan sus datos de forma inteligente y posteriormente generan conocimiento con ellos.

Dos de los campos que apoyan la GC y que han tenido un acelerado crecimiento en los últimos años son la inteligencia del negocio (IN) y la arquitectura empresarial (AE). Por un lado, la IN permite sacar provecho de la información guardada en los almacenes de datos multidimensionales (conocido en inglés como DWH por *data warehouse*) con diferentes operaciones como *slice, dice, roll-up y drill-down*. Esta información se obtiene de las bases de datos operacionales por medio de un proceso de extracción, transformación y carga (conocido en inglés como ETL por *extraction, transformation and loading*). Por otro lado, la AE permite a las instituciones establecer métodos que permitan crear, compartir y transferir el conocimiento que reside en personas y procesos mediante el uso de mapas y modelos.

Uno de los objetivos de la GC es crear una cultura donde el conocimiento tácito (aquel que reside en las personas) se quede en la organización cuando el personal calificado y experto abandone la misma o cuando se requiera realizar cambios en la estructura organizacional, en las aplicaciones informáticas o en la infraestructura tecnológica.

[1] https://edua.ua.es/es/normativas/tesis-en-lenguas-distintas-a-las-oficiales-en-la-comunidad-autonoma-valenciana.html

En el caso de las instituciones de educación superior (IES), el no tener un adecuado método de GC, es un problema incluso mayor ya que debido a la naturaleza de esta industria existe muy poca interdependencia entre los departamentos y facultades, es decir, existe poca estandarización, redundancia de información, duplicidad de aplicaciones y funcionalidades entre departamentos, lo que ocasiona organizaciones poco eficientes.

Es por ello que este trabajo se ha centrado en buscar un método adecuado de GC e investigar sobre la infraestructura tecnológica que apoye la gestión de la información de todas las dimensiones de conocimiento como son: personas, procesos y tecnología. Todo esto con el objetivo de buscar mecanismos innovadores para mejorar la educación y el servicio que brindan las IES a sus estudiantes y docentes mediante la mejora en sus de procesos.

A pesar de existir algunas iniciativas y artículos sobre los marcos de trabajo de GC, no pudimos encontrar un marco estándar que apoye o guíe las iniciativas de GC. Además, los marcos de trabajo de GC encontrados al investigar en la bibliografía científica no presentan un enfoque práctico que permita analizar y consolidar las dimensiones de conocimiento y por ende facilite su implementación.

El aporte central de esta tesis es una infraestructura híbrida de GC basada en AE e IN que fue desarrollada completamente desde un enfoque de investigación empírica y tomando como referencia un marco de trabajo desarrollado previamente para la GC.

La infraestructura desarrollada ayudará a las IES a mejorar la educación de manera general mediante el análisis de datos educativos confiables y depurados e integrando analítica desde la perspectiva de AE, tomando en cuenta la interdependencia existente entre los objetos que conforman la organización: personas, procesos, aplicaciones y tecnología.

A través de la infraestructura presentada se abren las puertas para la realización de distintos experimentos y proyectos de investigación que permitan incrementar el tipo de conocimiento que se genera en las IES, integrando la información de las aplicaciones que se encuentra en los almacenes de datos junto con la información de las personas y de los procesos organizacionales que se encuentran en los repositorios de AE.

Para validar la propuesta se ha realizado un caso de estudio dentro de una universidad con resultados iniciales prometedores. Como trabajos futuros se plantea automatizar distintas actividades de la IES mediante una metodología de desarrollo de software basada en modelos de AE. Además, se está desarrollando un sistema de GC que permita generar distintos y nuevos tipos de analítica, que solo con bases de datos transaccionales o multidimensionales sería imposible obtener.

A.2 Motivación y trabajos realizados

Hoy en día, la información es uno de los activos más importantes en las organizaciones. La información está en todas partes y puede convertirse en una fuente de oportunidades para asegurar la competitividad si se administra correctamente. El escenario ideal es convertir los datos en información y la información en conocimiento útil que permita a las organizaciones mejorar sus modelos de negocios y sus procesos organizacionales. Además, las organizaciones modernas necesitan adquirir capacidades para utilizar su conocimiento organizacional en ser innovadores y reducir los costos de operación de la organización.

Las IES han sido tradicionalmente productoras de conocimiento y han establecido prácticas de investigación para lograr este objetivo. Sin embargo, la mayoría de estas organizaciones no tienen un enfoque estándar o único para las prácticas de GC en su propio beneficio. Además, la gran cantidad de información que producen sus sistemas académicos y administrativos debe gestionarse eficientemente y con la tecnología adecuada.

Las IES dentro de sus estrategias y con el objetivo de alcanzar estándares de calidad, requieren gestionar el conocimiento de la organización de manera eficiente. Uno de los aspectos clave de la gobernabilidad y de apoyo para la toma de decisiones se basa en la GC. La GC es una disciplina que promueve la creación, uso, distribución y transferencia de conocimiento en organizaciones (Campbell, 2006). Una estrategia importante para la mejora continua en las IES es la implementación de la GC como un factor diferenciador que permite brindar servicios educativos de excelencia y mejorar la eficiencia en sus procesos académicos y administrativos.

Una de las principales motivaciones del trabajo realizado durante estos últimos años y que se presenta en esta tesis, fue investigar cómo los datos que se procesan y se obtienen en las IES pueden mejorar el proceso educativo.

Para este propósito, estudiamos diferentes tecnologías como almacenamiento de datos (conocido en inglés como DW por *data warehousing*), analítica de datos, minería de datos educativos (MDE), analítica de aprendizaje, IN, AE y GC. La GC se ha convertido en una necesidad debido a la gran cantidad de información (Big Data) que las IES generan cada día con sus sistemas y que los estudiantes, el personal académico y administrativo generan con diversos dispositivos y aplicaciones móviles. Así, al principio, se entendió que aunque había artículos que describían la GC en las IES, no existía un enfoque estándar para implementar la GC, ni una guía práctica que tome en cuenta todas las dimensiones del conocimiento. En consecuencia, hicimos algunas investigaciones y propusimos un marco para realizar la GC (Moscoso-Zea, Luján-Mora, et al., 2016).

Este marco se concibió con la participación de investigadores de cuatro universidades de todo el mundo: Universidad de Alicante–España, Universidad Técnica de Berlín–Alemania, Escuela Politécnica Nacional–Ecuador y Universidad UTE–Ecuador. Este marco que se muestra en la Figura A.1 se diseñó con el fin de convertirse en una guía práctica para implementar proyectos de GC en las organizaciones. Este desarrollo se presentó en la Conferencia Internacional sobre Sistemas de Información Empresarial (ICEIS) 2016 en Roma, Italia. El documento incluía una descripción teórica de cómo implementar una iniciativa de GC teniendo en cuenta todas las dimensiones del conocimiento.

Este marco se desarrolló con la idea de que una implementación exitosa de GC depende en gran medida de un método bien definido que admita la creación, captura, uso, distribución y transferencia de conocimiento. En el marco desarrollado se presenta una guía para realizar las tareas principales de GC. Además, se muestra el flujo de conocimiento explícito e implícito. El conocimiento organizacional se crea a partir de diferentes objetos interdependientes en diferentes dominios empresariales. El conocimiento explícito e implícito se puede derivar de estos dominios.

El conocimiento explícito es el conocimiento que se puede formular, documentar y reproducir. El conocimiento implícito, también conocido como conocimiento tácito, es el conocimiento que es difícil de documentar o formular, y normalmente se asocia con

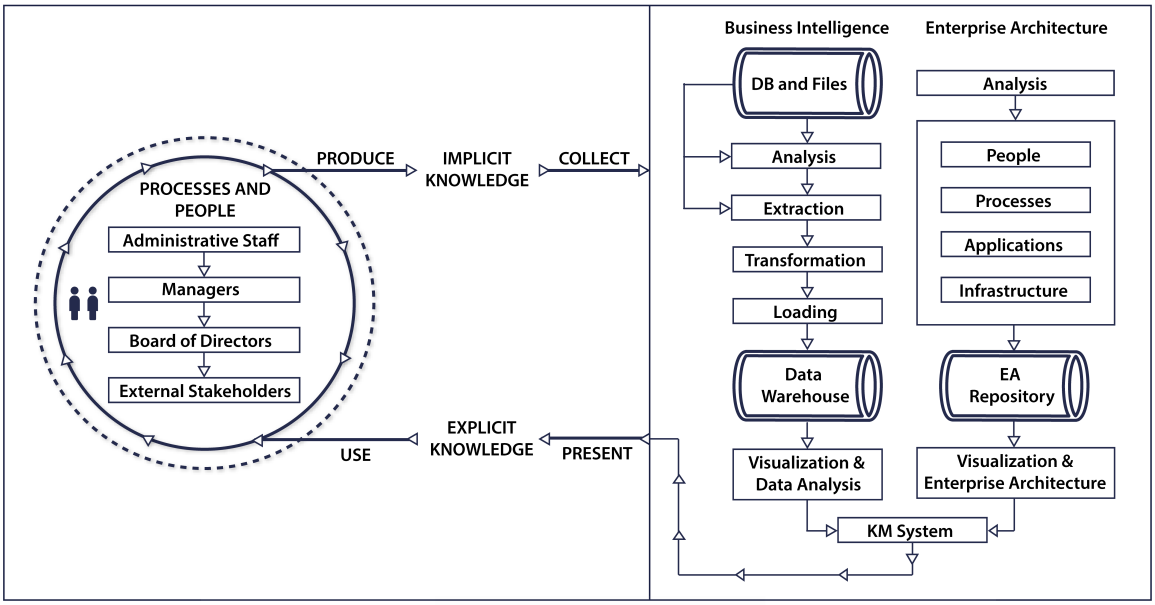


Figura A.1: Marco de trabajo de GC propuesto
(Moscoso-Zea, Luján-Mora, et al., 2016)

el conocimiento humano (Nonaka, 1994).

El marco propuesto incorpora dos tareas que son necesarias para implementar GC. En primer lugar, el desarrollo de una infraestructura tecnológica para facilitar la captura y el intercambio de conocimientos; y, en segundo lugar, el establecimiento de mecanismos y procedimientos para retener el conocimiento de personas y procesos.

En esta tesis, las dos tareas necesarias para el GC se han estudiado y presentado en el compendio de trabajos de investigación bajo el marco de GC propuesto. En primer lugar, se desarrolló una infraestructura híbrida; En segundo lugar, se establecieron los mecanismos para captar el conocimiento tácito de los empleados y de los procesos de negocios. Aunque la implementación del marco propuesto parece sencilla, incluye una variedad de disciplinas que no simplifican el proceso. Por lo tanto, los detalles de los componentes del marco se estudiaron en detalle y se presentaron como artículos de conferencias y revistas alrededor de todo el mundo.

Después de estudiar cada uno de los componentes y los flujos de conocimiento propuestos en el marco, el siguiente paso fue analizar los mecanismos para hibridar la información y el conocimiento que se generan en el DWH y en el repositorio de AE. La hibridación de los componentes concluyó con una propuesta de un sistema de web de GC que permitió la creación de nuevos tipos de analítica. Además, los artículos presentados también proponen métodos o mejores prácticas para la implementación de los componentes del marco.

El recuadro de la derecha del marco de trabajo de GC propuesto presenta un DWH y un repositorio de AE. Dentro de los estudios realizados se estudiaron las mejores metodologías de diseño para el almacenamiento de datos y para realizar una MDE, además se ha presentado una propuesta metodológica para evaluar y seleccionar las mejores herramientas para AE.

Dado que el enfoque propuesto era mejorar los indicadores educativos, la investigación comenzó analizando las posibilidades que la MDE brindó para lograr este objetivo.

El propósito fundamental de MDE es analizar datos de instituciones educativas a través del uso de diferentes técnicas como la predicción, agrupación, análisis de series de tiempo, clasificación, entre otras. Este estudio fue presentado en la Conferencia Ibérica sobre Sistemas y Tecnologías de la Información (CISTI) (Moscoso-Zea & Lujan-Mora, 2016).

El documento presentó una vista completa de MDE que incluye una clasificación de los algoritmos, métodos y herramientas utilizadas en los procesos de minería de datos (MD). Además, el documento analiza los procesos e indicadores con potencial de mejora en las instituciones educativas. Dentro de esta investigación, se estudiaron un conjunto de 42 artículos que presentaron experimentos de MDE. De los artículos estudiados, se pudo identificar que los métodos de MDE más utilizados fueron la clasificación, el agrupamiento y la asociación. A partir de la investigación, se observa que la mayoría de los experimentos realizados se centran en los procesos de enseñanza–aprendizaje, como la predicción de los resultados de las evaluaciones, la predicción del rendimiento y el perfil de los estudiantes. Los resultados también pueden proporcionar retroalimentación a los tutores o maestros para tomar acciones correctivas y preventivas en una etapa temprana. Las actividades específicas que se pueden mejorar utilizando MDE son las siguientes:

- Caracterizar el comportamiento y logros de los alumnos.
- Mejorar el proceso de enseñanza–aprendizaje (registro, tutoría, evaluación, graduación).
- Mejorar los indicadores de gestión, como disminuir la deserción escolar, aumentar la tasa de graduación y mejorar la tutoría personalizada.
- Mejorar la infraestructura.
- Optimizar el uso de aulas y laboratorios.
- Mejorar la eficiencia organizacional.

El conocimiento generado mediante el uso de la MDE apoya a las autoridades de las IES en la toma de decisiones oportunas y también a los maestros para analizar el comportamiento y la forma en que sus estudiantes aprenden (Jindal & Borah, 2013). La disciplina se enfoca en el diseño de modelos para mejorar las experiencias de aprendizaje y la eficiencia organizativa (Huebner, 2013).

A pesar del hecho de que las IES están invirtiendo en impulsar las iniciativas de MDE, muchos de los esfuerzos realizados no tienen los resultados esperados debido a la falta de un proceso sólido de almacenamiento de datos para un escenario educativo, por lo tanto, en el siguiente paso se realizó un estudio para comprender la importancia de tener un DWH antes de experimentar con MDE. El estudio fue presentado en la conferencia de Educación Superior Basada en Tecnología de la Información (ITHET). Este documento presenta consideraciones de diseño para la implementación de un DWH en IES. Además, se describe el proceso ETL desde las fuentes de datos operacionales hacia el DWH. Luego, como estudio de caso, el documento presenta los pasos seguidos para el diseño de un DWH en una universidad privada (Moscoso-Zea, Sampedro, & Luján- Mora, 2016).

Los pasos sugeridos para la implementación de un DWH son:

1. Diagnóstico.
2. Análisis de necesidades de información.
3. Selección de la metodología.
4. Configuración de la infraestructura tecnológica.
5. Diseño del DWH.
6. Ejecución ETL.

Con respecto al componente de AE que se dibuja a la derecha del marco de GC, también realizamos un estudio para evaluar el rol de AE como elemento clave para el cambio y la GC (Moscoso-Zea, Paredes-Gualtor, & Luján-Mora, 2019). La implementación de AE se puede respaldar con una herramienta con un lenguaje de modelado que se puede usar para describir diferentes procesos de negocios, estructuras organizacionales y flujos de información (Moscoso-Zea & Luján-Mora, 2017).

En los últimos años, la AE se ha convertido en uno de los campos de mayor importancia a considerar en los estudios de sistemas de información y ha evolucionado ha considerarse hoy en día en una tarea esencial para la gestión empresarial (Bricknall et al., 2006). En consecuencia, para comprender las herramientas existentes de AE, este documento propone un enfoque para la evaluación y selección de las herramientas de AE que se ajustan a un escenario organizacional. Después de analizar diferentes metodologías de evaluación de software, se sugieren tres metodologías para la evaluación de AE y se muestran en el Cuadro A.1.

Autores	Año	Tipo de Producto
(Comella-Dorda et al., 2004)	2004	Software
(Kontio et al., 1995)	1995	Software comercial
(Morera, 2002)	2002	Productos de software

Cuadro A.1: Metodologías sugeridas para la evaluación de software de AE

Estas metodologías apoyan a los arquitectos empresariales al inicio de un proyecto de AE. En las metodologías se sugieren diferentes actividades para mejorar el proceso de selección e invertir en la herramienta de AE más adecuada para la iniciativa de conocimiento de acuerdo con los requisitos de las partes interesadas en la IES.

Una vez que la herramienta está instalada y configurada, el siguiente paso es definir cómo administrar la AE dentro de la organización. Una de las estrategias para la gestión de AE es utilizar una herramienta de software junto con un marco de trabajo. Por lo tanto, también colaboramos en un proyecto de investigación para definir un marco de trabajo mejorado de AE con un enfoque ágil. El desarrollo de una AE en grandes organizaciones es complejo. Por lo tanto, es importante crear valor en una etapa temprana de la implementación de la AE. Este marco se desarrolló con un nuevo diseño tomando como base el método de desarrollo de arquitectura (ADM) del marco de trabajo de arquitectura de la organización conocida como *The Open Group* (TOGAF por sus siglas en inglés *The Open Group Architecture Framework*). El marco propuesto sugiere extraer solo aquellos entregables que son importantes para un negocio o proyecto específico (Sandoval et al., 2016).

Durante el análisis de las herramientas de AE con el grupo de investigación de la Universidad UTE, estudiamos cómo migrar una infraestructura tecnológica empresarial a la nube y cuáles son los problemas de seguridad de los datos (Saa, Cueva, et al., 2017). Se dieron recomendaciones a las organizaciones sobre cómo garantizar que se defina una estrategia de seguridad integral antes de migrar sus aplicaciones empresariales a la nube.

Algunas de las recomendaciones fueron que los controles de seguridad deben aplicarse en todos los niveles. Esto se puede hacer incorporando un marco que aborde la seguridad a nivel físico, de red, de datos y de aplicaciones. La conclusión del documento es que, dependiendo del tamaño de la empresa, se pueden migrar a la nube más o menos aplicaciones y tecnología. Los principales beneficios de la infraestructura empresarial basada en la nube son su escalabilidad y menores costos de inversión, lo que crea oportunidades para las pequeñas y medianas empresas.

Es interesante observar que nuestra propuesta también se puede implementar en un escenario de nube. En consecuencia, este documento sugiere las fases y alternativas para migrar un proyecto a la nube. En la Figura A.2 se presenta el marco que se desarrolló para un proceso de migración que tiene como objetivo ayudar a las organizaciones a implementar servicios de computación en la nube con las mejores prácticas y el menor riesgo (Moscoso-Zea, Paredes-Gualtor, Saa, & Sandoval, 2018). El objetivo principal del marco es mejorar el proceso de migración minimizando las amenazas y maximizando las oportunidades. Este marco puede guiar migraciones tecnológicas hacia la nube. Por ende, puede servir si se desea implementar la infraestructura de GC propuesta en esta tesis en la nube.

El análisis de los componentes propuestos en el marco de GC fue alentador y confirmó su aplicabilidad en el proceso de hibridación e implementación de los componentes. Por lo tanto, se realizó un estudio sobre cómo usar estos componentes en las IES para producir nuevos conocimientos organizacionales. Uno de los aspectos más importantes de la gestión y el soporte de decisiones se basa en la GC. El documento "Gestión del conocimiento en instituciones de educación superior para la generación de conocimiento organizacional", publicado en la 12va Conferencia Ibérica sobre Sistemas y Tecnologías de la Información, presenta un estudio de caso con diferentes alternativas para la creación de conocimiento (Moscoso-Zea & Lujan-Mora, 2017).

Entre las razones importantes para la realización de este estudio se encuentran los siguientes: la creación de nuevos mecanismos para mejorar la enseñanza y el desarrollo de los estudiantes, la investigación y el bienestar de los estudiantes. Además, se analizó las maneras de mejorar los procesos institucionales para que se lleven a cabo de manera organizada y estandarizada, lo que facilitará el logro de los objetivos de la visión institucional a corto y mediano plazo. La AE y la IN son las herramientas tecnológicas básicas para la creación de repositorios y bases de conocimiento. Estas bases de conocimiento pueden generar inteligencia organizacional a través del procesamiento analítico en línea (conocido en inglés como OLAP por *Online Analytical Processing*), MD y analítica de AE.

Para establecer las posibilidades de creación de conocimiento, seguimos las sugerencias de Nonaka (1994). Nonaka afirma que el proceso para generar conocimiento se basa en la conversión de conocimiento tácito y conocimiento explícito (Nonaka, 1994).

Sobre esta base, describe cuatro formas de conversión de conocimiento como se muestra en la Figura A.3 (Nonaka et al., 2000). Esta figura describe las principales acti-

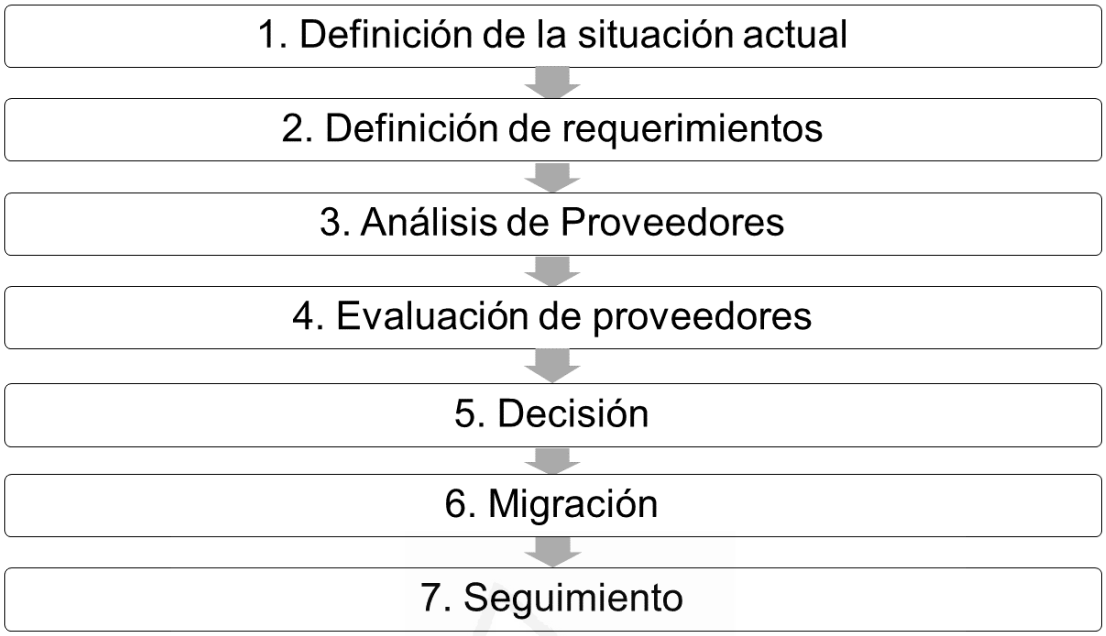


Figura A.2: Marco de adopción de computación en la nube
(Moscoso-Zea, Paredes-Gualtor, Saa, & Sandoval, 2018)

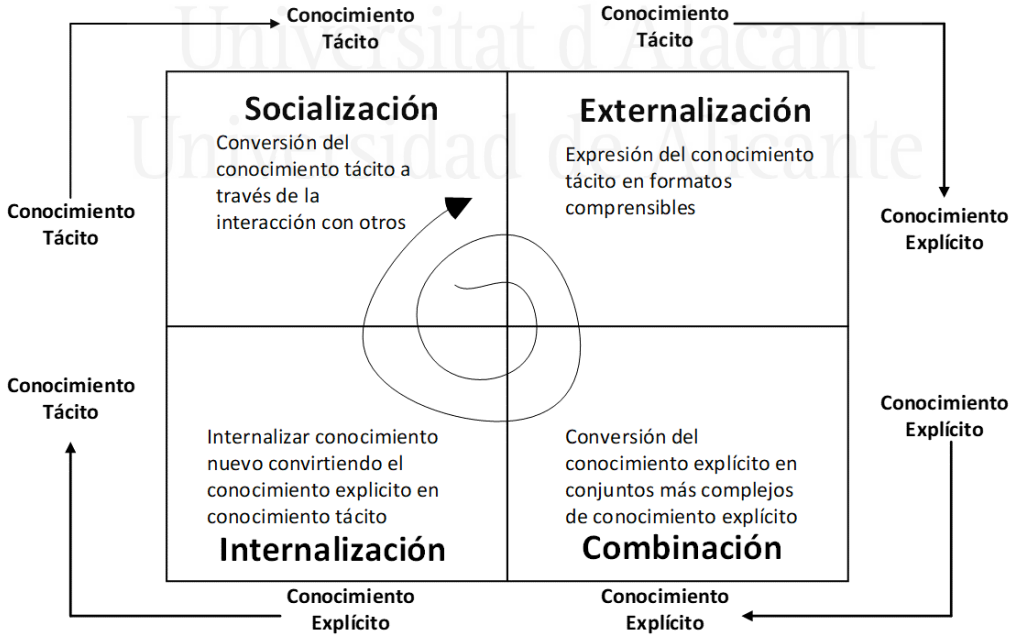


Figura A.3: Modelo SECI de creación del conocimiento
(Nonaka et al., 2000)

vidades propuestas para la creación de conocimiento y la innovación. El modelo se llama SECI debido a sus componentes Socialización, Externalización, Combinación e Internalización (SECI). Las diferentes actividades de conversión de conocimiento del modelo SECI se presentan en la siguiente lista.

- Del conocimiento tácito al conocimiento tácito (Socialización).
- Del conocimiento tácito al conocimiento explícito (Externalización).
- Del conocimiento explícito al conocimiento explícito (Combinación).
- Del conocimiento explícito al conocimiento tácito (Internalización).

Para el primer escenario, el conocimiento se puede crear compartiendo experiencias entre empleados nuevos y antiguos. Los empleados pueden incluso crear este conocimiento tácito con solo observar cómo un trabajador experto realiza las tareas. Esta forma de aprendizaje y generación de conocimiento tácito se denomina socialización, como se muestra en el cuadrante superior izquierdo de la figura. Algunas herramientas que se pueden utilizar para este escenario son la lluvia de ideas, la observación y la imitación.

El segundo escenario corresponde al cuadrante superior derecho de la figura y se denomina Externalización. La conversión de conocimiento tácito a explícito se produce al capturar el conocimiento de los empleados en libros, bases de datos o sistemas de información. Es decir, el conocimiento tácito está traducido o documentado en formularios que puedan ser comprensibles e interpretables por cualquier persona.

En el tercer escenario, la conversión de conocimiento ocurre a través de procesos sociales para combinar el conocimiento explícito de los individuos, este conocimiento puede ser compartido en reuniones o talleres. Al establecer estas interacciones, el conocimiento colectivo y compartido da lugar a un nuevo conocimiento. Este caso corresponde al cuadrante inferior derecho que se conoce como combinación. Este proceso de creación de conocimiento puede basarse en las herramientas tecnológicas que se proponen en esta tesis. El conocimiento explícito que reside en las aplicaciones de software de la empresa y en las bases de datos se puede procesar y analizar utilizando inteligencia artificial, IN o analítica de AE para crear un nuevo conocimiento explícito.

El cuarto escenario es complementario a la externalización y ocurre cuando el conocimiento explícito se convierte en tácito. Esta es la forma más similar a cómo aprenden las personas, es decir, tomar un manual o un libro e internalizar este nuevo conocimiento para que se convierta en parte de la información que reside en el cerebro de un individuo. Esta forma de generación de conocimiento se llama internalización.

El marco de gestión del conocimiento que proponemos y sus componentes están alineados con el ciclo de gestión del conocimiento. La Figura A.4 muestra el ciclo de GC sugerido (McIntyre et al., 2003). Este ciclo tiene tres componentes principales: gestión, aplicación y personas. La gestión se ocupa de capturar, organizar y facilitar el conocimiento organizacional. Estas actividades abarcan los cuadrantes de combinación y externalización presentados en el modelo de Nonaka (1994). El componente de la aplicación se enfoca en obtener contenido relevante y análisis mediante el uso de MD y analítica avanzada.

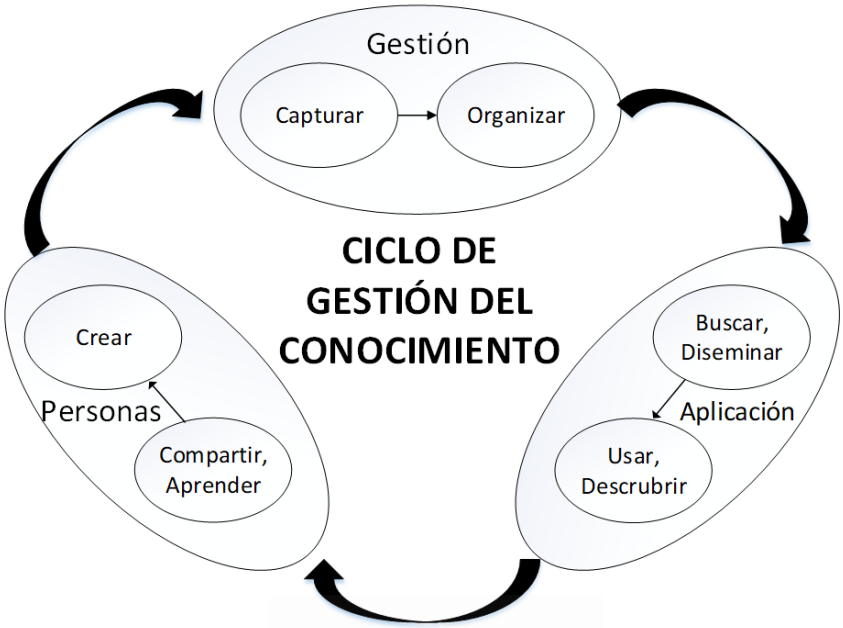


Figura A.4: Ciclo de gestión del conocimiento
(McIntyre et al., 2003)

Finalmente, el componente de personas se centra en el aprendizaje, el intercambio y la colaboración. El componente de personas describe el proceso educativo y busca mecanismos para crear conocimiento tácito para las personas que aprenden y conocimiento explícito que se pueda almacenar en repositorios y bases de conocimiento.

Como se dijo anteriormente, uno de los principales objetivos de realizar este trabajo era mejorar la educación usando GC. Teniendo esto en cuenta, también era importante investigar los pensamientos de los estudiantes sobre cómo llevar sus computadoras portátiles o dispositivos personales a las clases. El documento de Saa, Moscoso-Zea, and Lujan-Mora (2017) presenta los resultados de una encuesta que estudia, por un lado, cómo las universidades pueden verse afectadas de manera crítica si permiten que los estudiantes traigan sus propios dispositivos a las clases, por otro lado, todas las preocupaciones que los estudiantes tienen de ser observados a través de sus dispositivos por parte del personal de la universidad (Saa, Moscoso-Zea, & Lujan-Mora, 2017).

A pesar del hecho de que con la investigación realizada, se confirmó la importancia de incluir las herramientas de IN y de AE en la infraestructura propuesta. Encontramos una brecha de investigación sobre cómo las IES administran la información y el conocimiento.

La MDE se ha convertido en un tema de moda en los últimos años, pero nos preocupa la validez de la información que se procesa para obtener los resultados del proceso analítico. Esta es la razón principal por la que decidimos realizar un mapeo sistemático (MS) sobre el tema DW en educación. Los resultados del MS se presentaron en el documento "Una visión holística del almacenamiento de datos en la educación" (Moscoso-Zea, Paredes-Gualtor, & Lujan-Mora, 2018). Este estudio muestra que los DWH y la IN son ampliamente utilizados en organizaciones empresariales y se han analizado exhaustivamente desde el punto de vista industrial durante muchos años. Sin embargo, su uso es aún bajo en las instituciones educativas. El documento de Shahid et

al. (2016) presenta diferentes estudios de caso realizados para determinar el porcentaje de uso de DWH en diferentes industrias.

Los resultados del análisis del estudio de caso muestran que la industria que más utiliza DWH es la industria médica (hospitales, clínicas y consultorios médicos) con un 23.3 % de uso. Después de eso, la industria financiera y bancaria con 6.2 % de uso de DW, mientras que una de las industrias que menos usan DW es la educación con solo 3.8 % de uso (Shahid et al., 2016). Como se ve en estos datos, solo unas pocas instituciones educativas han adoptado un DWH para administrar la información educativa que se genera a partir de sus aplicaciones de software y sistemas de gestión de aprendizaje.

Sin embargo, en la actualidad, las juntas de gobierno y directores de las IES están reconociendo a la IN y al DW como tecnologías centrales para generar conocimiento organizacional. La implementación de estrategias con BI y DW puede ayudar a HEI a comprender mejor los procesos educativos. En primer lugar, puede ser un factor crucial para mejorar la educación al saber cómo aprenden los estudiantes y cómo enseñan los profesores. En segundo lugar, puede ayudar a identificar con suficiente anticipación a los estudiantes con riesgo de abandonar o fallar en las tareas. Además, el modelado o los procesos organizacionales de la institución permiten tener una visión clara de los procesos educativos, lo que puede ser de gran apoyo para implementar iniciativas de cambio innovadoras (Baepler & Murdoch, 2010; Bichsel, 2012; Morris et al., 2015).

El MS realizado sobre el tema de almacenamiento de datos en educación se realizó entre enero de 2008 y junio de 2018. En el estudio, se aplicó un método de cuatro etapas. Este estudio apoyó al equipo de investigadores para responder las preguntas de investigación propuestas. La primera pregunta de investigación fue dirigida a identificar detalles y el número de estudios significativos con las palabras *"data warehouse" y "education" o "academic"*. Para lograr este objetivo, los resultados del análisis incluyeron a los autores, la afiliación institucional y su país de origen. Además, este análisis enumeró las revistas o conferencias en las que se publicaron los artículos resultantes. Después de estudiar los componentes del marco y al mismo tiempo confirmar la aplicabilidad para apoyar la iniciativa de GC. El siguiente paso fue definir los aspectos técnicos para implementar la infraestructura de GC. El documento "Una infraestructura híbrida de arquitectura empresarial e inteligencia empresarial & Analytics para la gestión del conocimiento en educación" (Moscoso-Zea, Castro, et al., 2019) presenta la infraestructura de información que se compone por herramientas de IN, un DWH educativo y un repositorio de AE que permite la digitalización del conocimiento y faculta la visualización y el análisis de los componentes de la organización como personas, procesos y tecnología.

Esta infraestructura fue diseñada siguiento un proceso de invesrigación y servirá para analizar datos educativos, procesos académicos y para la creación de conocimiento explícito utilizando diferentes algoritmos y métodos de MDE, análisis de aprendizaje, OLAP y análisis de AE.

A.3 Contribuciones

Este trabajo ha contribuido en diferentes áreas. La principal contribución es dar a la comunidad científica un mecanismo para mejorar la GC en las IES. La infraestructura propuesta intenta almacenar todo el conocimiento organizacional de todas las dimen-

siones del conocimiento en una infraestructura híbrida. Estos conocimientos pueden ser reutilizados en nuevos proyectos y puede ser la base para la creación de conocimiento explícito mediante el uso de analítica de AE que apoyará a los directivos de las IES en la mejora de los procesos educativos. Las contribuciones de las publicaciones se detallan en la siguiente lista:

- Presentación de un método para la digitalización del conocimiento.
- Metodologías sugeridas para la evaluación y la selección de herramientas de AE para GC.
- Estado de la cuestión de GC, AE, IN, DWH, MDE y *cloud computing*.
- Directrices para migrar los sistemas empresariales a la nube.
- Pautas para la implementación de un DWH en instituciones educativas.
- Propuesta de una infraestructura híbrida para GC.
- Método para la GC y la implementación de la infraestructura híbrida.
- Mecanismos para el proceso de externalización de la creación de conocimiento (convertir conocimiento tácito en conocimiento explícito). Esto se puede hacer capturando el conocimiento tácito de los empleados en modelos en la herramienta AE.
- Propuesta de GC con AE y con otras herramientas de gestión.
- Evaluación y recomendación de algoritmos para ser utilizados en experimentos con MDE para mejorar los indicadores educativos.
- Estudio de caso para la aplicación de la infraestructura.

A.4 Conclusiones y trabajo futuro

Esta tesis presenta la investigación realizada para proponer un método de GC en las IES que mejore el proceso de creación del conocimiento. Para lograr este objetivo se propuso una infraestructura de GC. La infraestructura de conocimiento propuesta presenta un enfoque novedoso que incluye un DWH y una herramienta de AE. Esta infraestructura fue diseñada en base a una investigación empírica destinada a mejorar la gestión de las IES. Después de analizar las diferentes posibilidades, hemos propuesto un método para la implementación de los componentes de la infraestructura. Estos componentes fueron implementados y se han realizado pruebas iniciales en una IES. Con la infraestructura implementada y utilizando IN y MDE, se pueden llevar a cabo diferentes experimentos para mejorar los indicadores académicos. Los resultados de estos experimentos dan visibilidad a la junta directiva para la toma de decisiones a tiempo y son una fuente de conocimiento para mejorar la educación en general.

El mapeo sistemático que realizamos sobre el tema de DWH en educación muestra que a pesar de que hay un aumento en la experimentación con MDE y más estudios sobre el tema, los investigadores y directores de las IES no otorgan la importancia

adecuada a la implementación de un DWH educativo que puede ayudar a mejorar los resultados de los experimentos. Por esa razón, la implementación de la infraestructura propuesta puede ayudar en la creación y administración de las dimensiones de conocimiento generadas en la organización y puede ser una fuente de oportunidades para mejorar los procesos organizacionales. Además, la infraestructura puede convertirse en una fuente válida para aplicar diferentes tipos de análisis con herramientas OLAP, tableros de mandos, herramientas de IN y herramientas de AE para asegurar resultados de calidad en el proceso de DW.

El diseño de infraestructura híbrida considera el hecho de que pueden existir cambios inesperados de autoridades en las IES y es muy flexible para adaptarse a los cambios en la dirección y gestión estratégica.

La infraestructura de conocimiento propuesta facilita la trazabilidad de los estudiantes, el personal administrativo y académico para mejorar la toma de decisiones y los procesos organizacionales. Además, la infraestructura es la base de conocimiento de la organización y una herramienta para producir conocimiento explícito. Por otro lado, el uso de AE, como parte de la solución, puede apoyar la externalización del conocimiento.

El marco de conocimiento y la infraestructura de conocimiento propuesta en este trabajo proporcionan la base para la creación, transferencia y gestión del conocimiento, que puede ser una piedra angular para convertir a las IES en organizaciones de excelencia. Algunos de los beneficios de usar este marco de trabajo y de la infraestructura propuesta son: reducción de los costos de capacitación, reducción de la rotación de personal, mejora de los procesos de toma de decisiones y la creación de un repositorio de conocimientos.

Bibliography

Ackermann, F., & Eden, C. (2011). Strategic Management of Stakeholders: Theory and Practice. *Long Range Planning*, *44*(3), 179–196. (cited on pages 14, 15)

Baepler, P., & Murdoch, C. J. (2010). Academic Analytics and Data Mining in Higher Education. *International Journal for the Scholarship of Teaching and Learning*, *4*(2), 1–9. (cited on pages 11, 63)

Bichsel, J. (2012). *Analytics in Higher Education Benefits, Barriers, Progress and Recommendations*. Louisville, CO. Retrieved 2019-01-05, from https://library.educause.edu/~/media/files/library/2012/6/ers1207.pdf?la=en (cited on pages 11, 63)

Bricknall, R., Darrell, G., Nilsson, H., & Pessi, K. (2006, 01). Enterprise architecture: Critical factors affecting modelling and management. In (p. 2349-/2361). (cited on pages 6, 58)

Campbell, H. M. (2006). The role of organizational knowledge management strategies in the quest for business intelligence. In *Ieee international engineering management conference (iemc)* (pp. 231–236). Bahia, Brazil. (cited on pages 3, 21, 55)

Comella-Dorda, S., John, D., Lewis, G., Edwin, M., Patricia, O., & Erin, H. (2004). *A Process for COTS Software Evaluation* (Tech. Rep. No. July). Pittsburg, PA: Carnegie Mellon Sofware Engineering Institute. Retrieved from https://resources.sei.cmu.edu/asset_files/TechnicalReport/2004_005_001_14252.pdf (cited on pages 7, 15, 58)

Correa, A., Angélica, A., & Correa, S. (2009). La Gestión Educativa un nuevo paradigma. *Fundación Universitaria Luis Amigó*, *57*(4). Retrieved from http://virtual.funlam.edu.co/repositorio/sites/default/files/6lagestioneducativaunnuevoparadigma.pdf (cited on pages 3)

de Vries, M., & van Rensburg, A. (2008). Enterprise Architecture - New business value perspectives. *Southafrican Journal of Industrial Engeneering*, *19*, 1–16. (cited on pages 21)

Edwards, J. (2011). A Process View of Knowledge Management : It Ain't What you do, it's the way that you do it. *Journal of Knowledge Management*, *9*(4), 297–306. (cited on pages 21)

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). Knowledge Discovery and Data Mining: Towards a Unifying Framework. In *2nd international conference on knowledge discovery and data mining* (pp. 82–88). Menlo Park, California. (cited on pages 14)

Gartner Inc. (2019). *Gartner Business intelligence*. Retrieved 2019-01-20, from http://goo.gl/LmJRG3 (cited on pages 22)

Gomez-Torres, E., Moscoso-Zea, O., Herrera, N., & Lujan-Mora, S. (2018). Towards

a forensic analysis of mobile devices using android. In Á. Rocha & T. Guarda (Eds.), *Advances in intelligent systems and computing* (Vol. 721, pp. 30–39). Springer International Publishing. (cited on pages 23, 27)

Greenhalgh, T., & Peacock, R. (2005). Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *Bmj*, *331*(7524), 1064–1065. (cited on pages 13)

Hevner, A., March, S., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *Management Information Systems Quarterly*, *28*(1), 75–105. (cited on pages 14)

Huebner, R. A. (2013). A Survey of Educational Data-Mining Research. *Research in Higher Education Journal*, *19*, 1–13. (cited on pages 6, 57)

Inga, E., Campaña, M., Hincapié, R., & Moscoso-Zea, O. (2018a). Optimal Deployment of FiWi Networks Using Heuristic Method for Integration Microgrids with Smart Metering. *Sensors*, *18*(8), 1–22. (cited on pages 23, 27)

Inga, E., Campaña, M., Hincapié, R., & Moscoso-Zea, O. (2018b). Optimal Dimensioning of Electrical Distribution Networks Considering Stochastic Load Demand and Voltage Levels. In *Applications of computational intelligence* (Vol. 721, pp. 200–215). (cited on pages 23, 27)

Isik, O., Jones, M., & Sidorova, A. (2011). Business intelligence (BI) success and the role of BI capabilities. *Intelligent systems in accounting, finance and management*, *176*(1), 161–176. (cited on pages 22)

Ivanov, K. (2009). *10 Definitions of Enterprise Architecture.* Retrieved 2019-04-25, from http://www.ariscommunity.com/users/koiv/2009-08-20-10-definitions-enterprise-architecture-which-corresponds-yours (cited on pages 21)

Jindal, R., & Borah, M. D. (2013). A Survey on Educational Data Mining and Research Trends. *International Journal of Database Management Systems*, *5*(3), 53–73. (cited on pages 6, 57)

Karemente, K., Aduwo, J. R., Mugejjera, E., & Lubega, J. (2009). Knowledge Management Frameworks A Review of Conceptual Foundations and a KMF for IT-based Organizations. In *Strengthening the role of ict in development* (Vol. 5, pp. 35–57). (cited on pages 21)

Kimball, R., & Ross, M. (2013). *The Data Warehouse toolkit* (3rd ed.). Indianapolis US: John Wiley & Sons, Inc. (cited on pages 15)

Kitchenham, B., Budgen, D., & Brereton, O. P. (2010). The value of mapping studies – A participant-observer case study. In *14th international conference on evaluation and assessment in software engineering (ease)* (pp. 25–33). Staffordshire, uk. (cited on pages 13)

Kontio, J., Chen, S.-F., Limperos, K., Tesoriero, R., Caldiera, G., & Deutsch, M. (1995). A COTS Selection Method and Experiences of Its Use. In *20th annual software engineering workshop (sew)* (pp. 1–16). Greenbelt, MD. (cited on pages 7, 58)

Lankhorst, M. (2013). *Enterprise Architecture at Work Modelling Communication and Analysis* (3rd ed.). Berlin Heidelberg: Springer-Verlag. (cited on pages 21)

Lloyd, J. (2011). Identifying Key Components of Business Intelligence Systems and Their Role in Managerial Decision making. *Applied Information Management Program*, *1277*(2), 1–76. (cited on pages 22)

Markets and Markets. (2017). *Business Intelligence Market worth 26.88 Billion USD by 2021.* Retrieved 2019-01-15, from http://www.marketsandmarkets.com/PressReleases/social-business-intelligence-bi.asp (cited on pages 22)

McIntyre, S., Gauvin, M., & Waruszynski, B. (2003). Knowledge Management in the Military Context. *Canadian Military Journal*, *4*(1), 35–40. (cited on pages 9, 10, 61, 62)

Morera, D. (2002). COTS Evaluation Using Desmet Methodology & Analytic Hierarchy Process (AHP). In *4th international conference product focused software process improvement (profes)* (pp. 485–493). Rovaniemi, Finland. (cited on pages 7, 58)

Morris, L., Wu, S.-S., & Finnegan, C. (2015). Predicting retention in online general education courses. *American journal of Distance Education*, *19*(1), 23–36. (cited on pages 11, 63)

Moscoso-Zea, O., Castro, J., Paredes, J., & Luján-Mora, S. (2019). A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics for Knowledge Management in Education. *IEEE Access*, *7*, 38778 – 38788. (cited on pages 12, 23, 24, 27, 31, 39, 63)

Moscoso-Zea, O., & Lujan-Mora, S. (2016). Educational data mining: An holistic view. In *11th iberian conference on information systems and technologies, cisti* (pp. 600–605). Canary Islands, Spain. (cited on pages 5, 16, 23, 24, 57)

Moscoso-Zea, O., & Luján-Mora, S. (2016). Evaluation and Selection of Enterprise Architecture Tools for Knowledge Management. In *1st international conference on information systems and computer science (inciscos)* (pp. 292–296). Quito, Ecuador. (cited on pages 23, 25)

Moscoso-Zea, O., & Lujan-Mora, S. (2017). Knowledge management in higher education institutions for the generation of organizational knowledge. In *12th iberian conference on information systems and technologies (cisti)* (pp. 1593–1599). Lisbon, Portugal. (cited on pages 7, 23, 25, 59)

Moscoso-Zea, O., & Luján-Mora, S. (2017). Suggested Methodologies for Evaluation and Selection of Enterprise Architecture Software for Knowledge Digitization. *Enfoque UTE*, *8*(1), 315–328. (cited on pages 6, 15, 23, 24, 27, 33, 58)

Moscoso-Zea, O., Luján-Mora, S., Esquetini Cáceres, C., & Schweimanns, N. (2016). Knowledge Management Framework using Enterprise Architecture and Business Intelligence. In *18th international conference on enterprise information systems (iceis)* (pp. 244–249). Rome - Italy. (cited on pages 4, 23, 24, 55, 56)

Moscoso-Zea, O., Paredes-Gualtor, J., & Lujan-Mora, S. (2018). A Holistic View of Data Warehousing in Education. *IEEE Access*, *6*, 64659–64673. (cited on pages 10, 11, 16, 23, 24, 27, 31, 37, 62)

Moscoso-Zea, O., Paredes-Gualtor, J., & Luján-Mora, S. (2019). Enterprise Architecture, an enabler of change and knowledge management. *Enfoque UTE*, *10*(1), 247–257. (cited on pages 6, 23, 24, 27, 41, 58)

Moscoso-Zea, O., Paredes-Gualtor, J., Saa, P., & Sandoval, F. (2018). Moving the IT Infrastructure to the Cloud. *Enfoque UTE*, *9*(1), 79–89. (cited on pages 7, 8, 23, 27, 59, 60)

Moscoso-Zea, O., Saa, P., & Luján-Mora, S. (2019). Evaluation of algorithms to predict graduation rate in higher education institutions by applying educational data mining. *Australasian Journal of Engineering Education*, 1–10. (cited on pages 12, 23, 24, 31, 43)

Moscoso-Zea, O., Sampedro, A., & Luján- Mora, S. (2016). Datawarehouse design for Educational Datamining. In *15th information technology based higher education and training (ithet)* (pp. 1–6). Istanbul - Turkey. (cited on pages 6, 23, 25, 57)

Moscoso-Zea, O., Vizcaino, M., & Luján-Mora, S. (2017). Evaluation of methods and algorithms of educational data mining. In *Research in engineering education symposium (rees)* (pp. 972–980). Bogota - Colombia. (cited on pages 23, 25)

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. *Organization Science*, *5*(1), 14–37. (cited on pages 4, 8, 9, 47, 56, 59, 61)

Nonaka, I., Toyama, R., & Konno, N. (2000). SECI , Ba and Leadership: a Unified Model of Dynamic Knowledge Creation. *Long Range Planning*, *33*, 5–34. (cited on pages 8, 9, 59, 60)

Paredes-Gualtor, J., Moscoso-Zea, O., & Luján-Mora, S. (2018). The role of enterprise architecture as a management tool. In *3rd international conference on information systems and computer science (inciscos)* (p. 306-311). Quito - Ecuador. (cited on pages 23, 25)

Paredes-Gualtor, J., Moscoso-Zea, O., Saa, P., Sandoval, F., & Rodas, P. (2017). Unified cloud computing adoption framework. In *2nd international conference on information systems and computer science (inciscos)* (pp. 247–252). Quito - Ecuador. (cited on pages 23, 26)

Roger, S. (2007). *A Comparison of the Top Four Enterprise Architecture Methodologies.* Retrieved 2019-04-28, from `http://tiny.cc/zqly6y` (cited on pages 21)

Saa, P., Cueva, A., Moscoso-Zea, O., & Lujan-Mora, S. (2017). Moving ERP Systems to the Cloud - Data Security Issues. *Journal of Information Systems Engineering & Management*, *2*(4), 1–9. (cited on pages 7, 23, 24, 27, 35, 59)

Saa, P., Moscoso-Zea, O., Costales, A., & Lujan-Mora, S. (2017). Data security issues in cloud-based Software-as-a-Service ERP. In *12th iberian conference on information systems and technologies (cisti)* (pp. 1828–1834). Lisbon - Portugal. (cited on pages 23, 26)

Saa, P., Moscoso-Zea, O., & Lujan-Mora, S. (2017). Bring your own device (BYOD): Students perception — Privacy issues: A new trend in education? In *16th international conference on information technology based higher education and training (ithet)* (pp. 1–5). Ohrid, Macedonia. (cited on pages 10, 23, 26, 62)

Saa, P., Moscoso-Zea, O., & Lujan-Mora, S. (2018). Wearable technology, privacy issues. In Á. Rocha & T. Guarda (Eds.), *Advances in intelligent systems and computing* (Vol. 721, pp. 518–527). Springer International Publishing. (cited on pages 23, 27)

Sandoval, F., Galvez, V., & Moscoso-Zea, O. (2017). Development of Enterprise Architecture using a Framework with Agile Approach. *Enfoque UTE*, *8*(1), 135–147. (cited on pages 23, 26)

Sandoval, F., Moscoso-Zea, O., Galvez, V., & Tutillo, P. (2016). Enterprise Architecture Framework with Agile Approach based on TOGAF. In *1st international conference on information systems and computer science (inciscos)* (pp. 77–81). Quito, Ecuador. (cited on pages 7, 23, 26, 58)

Schekkerman, J. (2011). *Enterprise Architecture Tool Selection Guide* (Tech. Rep.). Netherlands: Institute for Enterprise Architecture Developments. (cited on pages 15)

Shahid, M. B., Sheikh, U., Raza, B., Shah, M. A., Kamran, A., & Anjum, A. (2016).

Application of Data Warehouse in Real Life : State-of- the-art Survey from User Preferences' Perspective. *International Journal of Advanced Computer Science and Applications*, 7(4), 415–426. (cited on pages 10, 11, 62, 63)

The Open Group. (2018). *The TOGAF® Standard, Version 9.2.* Retrieved 03/05/2019, from http://pubs.opengroup.org/architecture/togaf9-doc/arch/index.html (cited on pages 16)

Tucker, R., & Debrosse, D. (2003). Enterprise Architecture Roadmap for Modernization. *Enterprise Modernization Issue*, 7(2). (cited on pages 22)

Velasco, J., Ullauri, R., Pilicita, L., Jácome, B., Saa, P., & Moscoso-Zea, O. (2018). Benefits of implementing an isms according to the iso 27001 standard in the ecuadorian manufacturing industry. In *3rd international conference on information systems and computer science (inciscos)* (p. 294-300). Quito - Ecuador. (cited on pages 23, 27)