Theoretical Aspects of The Information and Knowledge Engineering

Maria do Carmo D. Freitas
Department of Science and Information Management
Faculty of Information Management
UFPR - Federal University of Paraná
Curitiba, Brazil
mcf@ufpr.br

Ricardo Mendes Jr
Department of Industrial Engineering.
UFPR - Federal University of Paraná
Curitiba, Brazil
mendesjr@ufpr.br

Guilherme Frederico
Department of Management
Federal University of Parana
Curitiba, Brazil
guilherme.frederico@ufpr.br

Ricardo S. Odorczyk
Department of Science and Information Management
Faculty of Information Management
UFPR - Federal University of Parana
Curitiba, Brazil
odorczyk89@gmail.com

Felisa M. Córdova
School of Industrial Engineering,
Faculty of Engineering and Business.
University Finis Terrae.
Santiago, Chile.
felisa.cordova@gmail.com

Claudia A. Durán
Department of Industrial Engineering.
Faculty of Engineering.
University of Santiago de Chile.
Santiago, Chile.
claudia.duransm@usach.cl

***Abstract*— Deepens the theoretical issues related to the Information and Knowledge Engineering with emphasis on reflection about its importance in the field of scientific research and social contribution to the development of the country. The paper presents these theoretical concepts: Information Engineering as the creation process and systematization of a volume of value-added data to a particular business. The Knowledge Engineering as the organized information that will facilitate recovery, capturing and processing in the form of a specialized knowledge; and the Strategic Competitive Intelligence as the understanding of a lack in the knowledge of the phenomena, identification of trends, risk mapping and discovery of opportunities related to the decision-making process in organizations. A systematic literature review is the method used in this study. Concludes that the information engineering (IE) and the knowledge engineering (KE) are revolutionaries in terms of information management in modern organizations. Active in the process of analysing large volumes of information they will have the power to reduce the cost by understanding the strategic business planning, data and modelling of processes, regardless of technology which will guide decision-making in any type of business.**

Index Terms-information engineering, knowledge engineering, decision-making

I. INTRODUCTION

Information Engineering is an integrated, full lifecycle systems development approach with automated tool support, which can be useful in assisting information systems managers in imposing a rigorous discipline on the systems development process [11]. Furthermore, Information Engineering is based on solid conceptual foundation and has been refined in many ways since its initial proposal [15]. In the past decades, data modelling has become the mainstream information systems approach and practical applications have been developed around the world.

In recent years, information systems become more complex, especially by the increasing use of the Internet and the greater amount of information that should be modelled. The information and its effects started to play an important role as strategies in the management and competitive advantages [26]. Global competition in many industrial and service sectors put high pressures in the need to increase the financial returns on the investments made in engineering expertise. In this context, it was seen as a solution to this information complexity capture people's knowledge, then emerging the Knowledge Engineering as an evolution of Information Engineering. Knowledge Engineering is defined by Pinfold and Chapman [16] as "a key for the organisations attempting to capitalise their expertise and knowhow''. Other authors point as the systematic process in which knowledge is organised in an efficient manner facilitating its exploitation and reuse [17].

Although the organizations and software companies continue to use the Knowledge Engineering currently, there is little updated academic literature dealing with the subject, both theoretical and practical. In this paper, we conducted a literature review that was undertaken to gain insight into the current concepts applied in this subject and mapping the recent literature.

II. SYSTEMATIC LITERATURE REVIEW METHOD

This paper followed a systematic approach to the literature review, and adapted the method proposed by Tranfield [6], divided into three stages: 1) Planning the review, 2) Conducting the review, and 3) Reporting and dissemination.

The first stage occurred to define and limit the scope of the research and define the databases for searching. As the output of the planning stage, we defined that search terms would be "information engineering" and "knowledge engineering" and selected the following databases: ISI Web of Science, Scopus and Science Direct.

For conducting the review, it was performed an initial search with the terms individually to verify what types of results would be returned, and analyse how to refine the search. After this exploratory analysis, it was verified what filters returned the better results to our needs, and defined the next steps, as shared on figure 1.

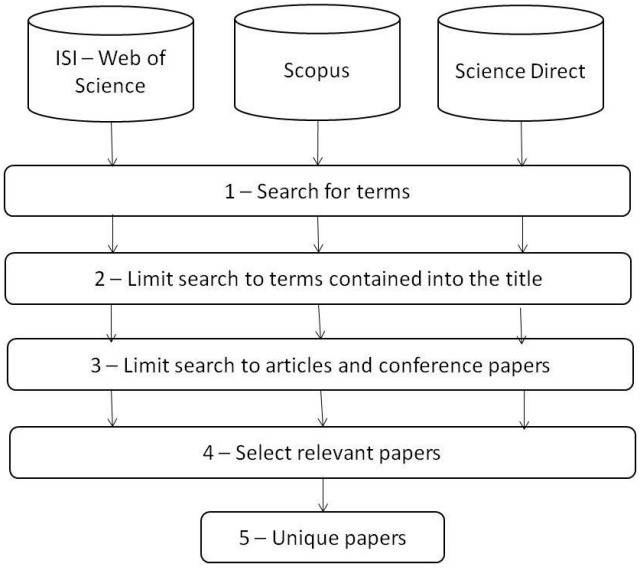

Fig. 1. Research methodology steps – the authors

The review was then conducted on September 15th 2015, using the following search string: ["information engineering"] OR ["knowledge engineering"]. The table 1 shows the results for each step of the search, at the different databases.

TABLE I. SEARCHES RESULTS

Search Step	ISI - Web of Science	Scopus	Science Direct
1	2932	18967	34337
2	729	1054	221
3	291	747	177
4	46	154	102

At the fourth step, it was performed a manual review on the 1215 papers filtered from the previous steps, to verify their relevance and fit to the theme, resulting on 302 papers. Out of those papers, the last stage delivered as output 256 unique papers.

Based on those selected articles, we extracted the key topics, arranged them logically, synthesized them, and compiled the literature review.

After completion of stage 2 of this method, the final stage focused on the report and dissemination, focusing on compiling the key ideas presented on the literature within constructs, and reflecting about those topics.

III. INFORMATION & KNOWLEDGE ENGINEERING

Since the Information Technology was introduced to society, the terms Data, Information and Knowledge are highly mentioned and investigated in researches from several areas of expertise. Souza [4] summarized the concepts from different authors about those terms, at the table 2.

TABLE II. DIFFERENCES BETWEEN DATA, INFORMATION AND KNOWLEDGE [4]

Author	Data	Information	Knowledge
Davenport e Prusak (1998)	Set of different data and objectives relating to events.	They are data that make the difference. It has meaning, relevance and purpose.	Fluid mixture of condensed experience, values, contextual information, which structure the evaluation and incorporation of new information.
Setzer (1999)	A sequence of quantified or quantifiable symbols	An informal abstraction (ie, can not be formalized through a logical theory or mathematics), that represents something meaningful to someone through text, images, sounds and animation.	An interior abstraction, personal, something that has been experienced by someone
Sianes (2006)	Primary form info: unprocessed signal, related, or interpreted without any sense inherent in themselves.	Organized data series in a meaningful way, analyzed and processed, generating hypotheses, suggest solutions, suggestions justifications, critical arguments used to support the decision-making process.	Group evaluated information as to their reliability and relevance and assimilated by the individual or organization, integrating its know earlier and building a picture of the situation.
Sordi (2008)	Collection of relevant evidence on an observed fact.	Interpretation of a set of data according to a relevant purpose and consensus for the target audience (reader).	It is the new knowledge resulting analysis information and reflections according to values and mindset that it develops.
Beal (2008)	Records and facts in its primary form, not necessarily physical.	Records and facts arranged or combined significantly.	Originated from information and added to other elements of the mind as experience, values and context.

Poujuan-Dante [8] explains on the figure 2 the informational pyramid, which proposes the concept of information as matter, based on data and on a hierarchy of quantity versus quality, representing the four concepts illustrated.

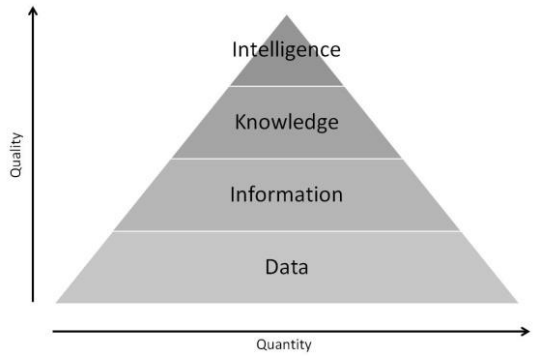

Fig. 2. Information Pyramid - [8]

Barreto [1] refers to the pyramid as a representation of elements that sustain the information, and its elements can be found as stocks or as flows. Those flows allow transitioning among those different dimensions of information. Hence, when talking about flows, we highlight the concept of Information Engineering (IE) to engineer efficient flows, as proposed by Hicks [10].

For Longo [23], "to engineer" means to conceive a product or service, evaluate its feasibility, and design, implement and operate its production. Engineering is a set of activities performed to achieve concrete results throughout a systematic framework. When those frameworks are applied within the domains of information and knowledge management to enhance their performance and results to a business, it can characterized as Information Engineering or Knowledge Engineering (KE).

A. Information Engineering

Information Engineering (IE) was firstly introduced as a methodology by Martin [15], and described by Richmond [12] as a tool for delivering competitive advantage on systems development, focused on: Planning, Data-centred development, Rigorous techniques, User participation and Automation of the methodology. The IE is the usage of several existing techniques in a disciplined way so that improvements and usefulness are leveraged by this methodology [12]. [11] emphasize the IE as an applications development methodology composed by a set of tangible and intangible elements (software, processes, procedures, strategy, tools, etc.) to enable a cross-functional system development. Although it depends on some key factors to be implemented successfully – as shown in figure 3.

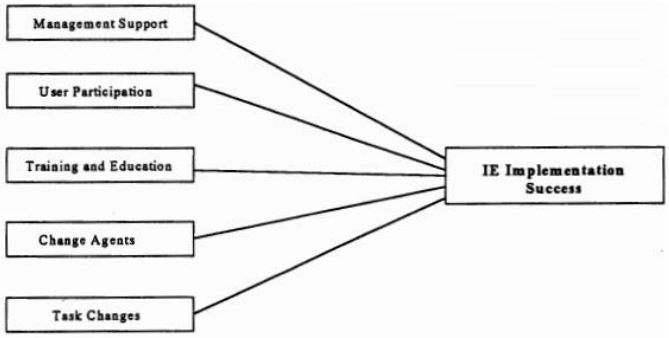

Fig. 3. Factors for Successful IE Implementation – [11]

Galup and Datero [20] describe IE as "a set of tools, techniques, and methodologies used to transform an organization's strategic plan into an information systems architecture. Once new information system architecture is in place, the flow of information is altered leading to organizational change."

Fonseca and Martin [7] present that the construction of ontologies for information systems is one of the key aspects from the Information Engineering, and [10] highlights the importance of having a data-centred system, which requires an strategic analysis of the system components and interactions. The IE will then enable an organized information flow within the system, which might result on benefits and competitive advantages [10]. The figure 4 clearly represents those flows, and how the IE can be used as an instrument with support from the technology, to organize and recover information from different processes within an organization.

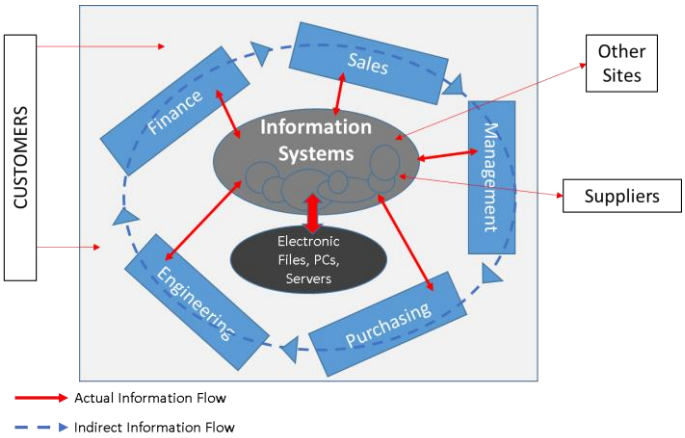

Fig. 4. Organized information flow within an information system – adapted from [10]

An integrated and scalable set that begins with the generation, analysis and use of information processes, tasks and techniques that enhance business communication across the enterprise to enable it to develop people, processes, procedures and systems to achieve their vision [19]

Hovarth and Rudas [13] recall the value added to the information flow of the system by the IE [13], and Teixeira, Freitas and Laurindo [3] state this as an opportunity to increase the competitive advantage at the strategic level of a company [3].

Teixeira defines IE as a tool to provide the means to rapidly react to the changes on the information needs of an organization. It works as an agent to connect the aspects from Information Science, Systems and Management, in a way to leverage efficiency on systems and processes, as shown at figure 5 [2].

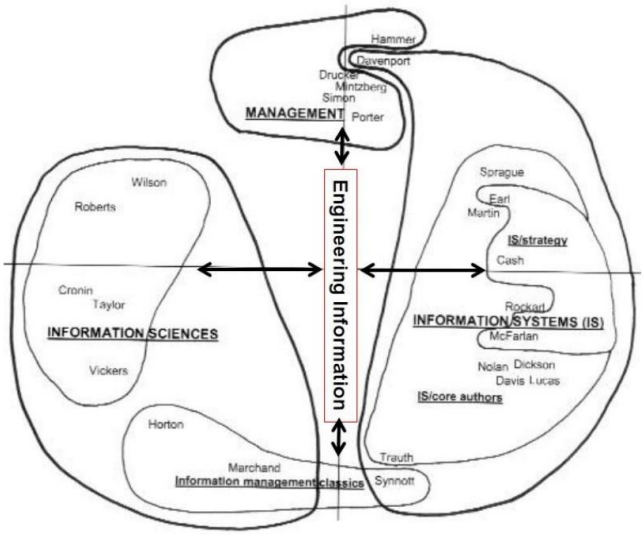

Fig. 5. Relationship between Information Systems, Science, Management and Engineering – [2]

B. *Knowledge Engineering*

Similar to the Information Engineering, the Knowledge Engineering (KE) is described by Wielinga, Sandberg and Schreiber [9] as a set of methods and techniques to develop systems, however with focus on the construction of Knowledge Based Systems (KBS) [24]. Singh, Jagirdar and Basil [21] state the KE as a framework to be used in knowledge based environments [21].

Studer, Benjamins and Fensael [22] share that the KE can be used as a knowledge transferring and knowledge modelling processes, with the goal of implementing the discipline of Software Engineering in the construction of KBS, by "turning the process of constructing KBSs from an art into an engineering discipline."

From another perspective, Lai [14] brings the KE concept as a systematic approach for the Knowledge Management (KM), consisting on a process modelling, verifying, storing, querying and updating knowledge, as shown in figure 6 [14].

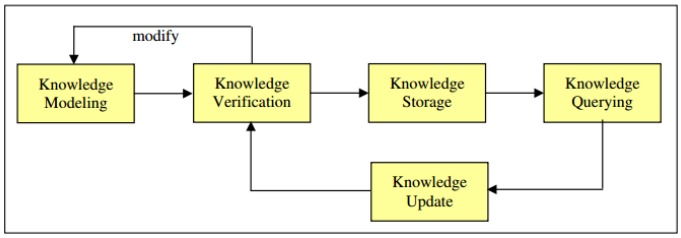

Fig. 6. KMKE aproach workflow – [14]

Wielinga, Sandberg and Schreiber [9] defines a methodological pyramid for the KE, composed by blocks, represented on figure E. The top means the experience when using the methodology, which can provide a continuous flow on the feedback through the pyramid. The tools are the instruments avilable to apply the methods defined. The methods are the procedures to be taken, so that the methodology reaches its objectives. The theories block contains all the scientific concepts that will guide the research methodology. And the base of the pyramid, the principles that will be used as the base for the conduction of the research.

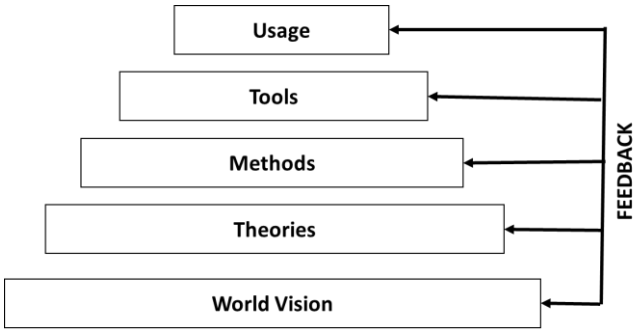

Fig. 7. KE Blocks Pyramid – [5]

Seshasai et al. [18] afirm that the KE is composed by four facets:

- Knowledge Acquisition – all actions related on capturing the knowledge contained in various means.
- Knowledge Management – actions responsible for standardizing and organizing the information captured.
- Knowledge Discovery – actions related to analysing the information and findings the knowledge underlying within them.
- Knowledge Dissemination – action related to the extraction of the relevant knowledge found.

Nazario, Dantas and Todesco [5] share two definitions for KE, as the process to build an specialist system, and the academic area to research models, methods and technologies to represent and process knowledge for the construction of KBS. Several different methods and techniques compose the KE (Table 3).

TABLE III. KE METHODS – ADAPTED FROM [5]

METHODS	
Name	***Description***
MIKE (Model-based and Knowledge Engineering) (1993)	This methodology has to integrate the principle characteristics of the model prototyping and software development life cycle, and in a formal specification language and techniques semiformal
MOKA (Methodology and tools Oriented to Knowledge-Based Engineering Applications) (1998)	It is a framework to structure and provide engineering knowledge, focused for products complex mechanics for aeronautics and automotive industry, mainly routine projects
CommonKADS (2000)	It is a management methodology Knowledge proposed by Wielinga, Sandberg and Schreiber [9] based on models, with support from technical and engineering tools
XP.K (Extreme Programming. Knowledge) (2002)	Methodology principles for using meet the knowledge process modeling, valuing the integration of the project team and specialists knowledge
RapidOWL (2006)	It is an agile methodology that enables the collaborative development of knowledge bases on semantic Web
TECHNIQUES	
Name	***Description***
Specialist System	is a tool that has the ability to understand the knowledge about a problem specific and use this knowledge to intelligently suggest alternative actions. SE is an AI technique developed to resolve problems in a given domain whose knowledge used is obtained from people who They are experts in that domain.
Inteligent Agent	is a computer system located at an environment capable of independently acting on this environment, according to their perception, communication, representation, motivation, deliberation and reasoning learning.
Neural Network	It is a computational model Abstract of the human brain, which mimics the behavior of biological neurons. Similar to the brain when fired at relation to an event, an Artificial Neural Network receives stimuli (signal inputs), processes signals and produces a output.

Genetic Algorythims	are computational models based on the theory of evolution . These implement the selection of solutions based on fitness solution as the answer to a problem, playback solutions and the occasional occurrence of mutation on solutions. With these metaphors, a Genetic Algorithm optimizes the search for an optimal solution among several solutions possible. They are usually employed in problems resource allocation.
Knowledge Discovery in Databases (KDD) and Kowledge Discovery in Texts (KDT)	Discovery knowledge databases (or data mining), if concerned with the development of algorithms and techniques to extract knowledge from large databases and complex. Both the Machine Learning as KDD They share the same goal of finding the data new and useful knowledge, and thus they are most techniques and processes in common. The fundamental difference between Machine Learning and KDD is the volume data to be processed [21]. In [22] the knowledge discovery problem is approached from a non-structured text collection, and describe the Knowledge Discovery System from Text (KDT) which provides for text kinds of KDD operations provided above for structured databases.

Also with the evolution on the computing and information technologies, new opportunities for the application o KE come a long, as an example of the Big Data era, with an enourmous ammount of information and knowledge being created and stored every second, but limited to challenges on its use. [25] states that "This fragmented knowledge is part of the migration puzzle—each piece provides some limited information, but not the whole picture. Traditional knowledge engineering can't obtain and process such fragmented knowledge because it's usually acquired from different sources.", proposing the concept of a knowledge engineering framework to address this issue, called BigKE.

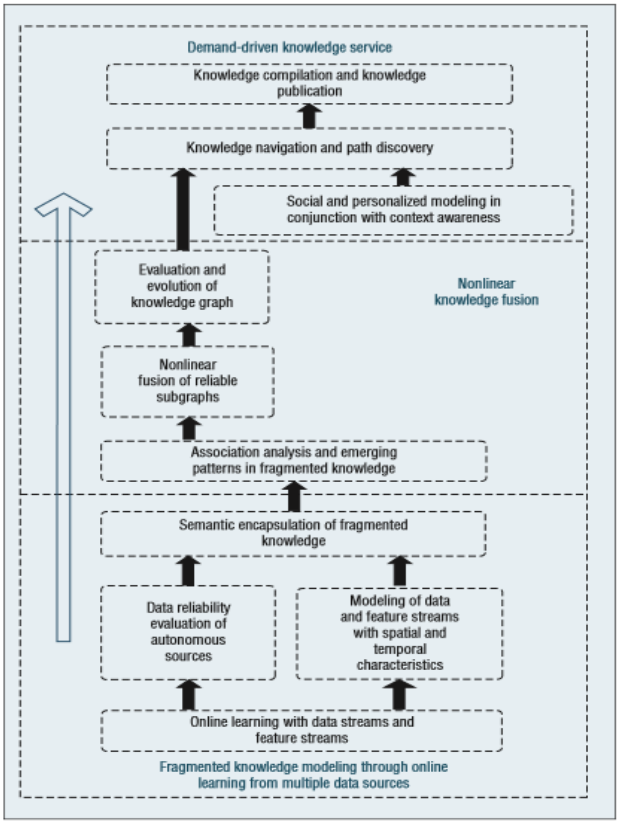

Fig. 8. Big KE – [25]

The figure 8 [25] exemplifies the evolution of the KE as a demand from society needs of information and knowledge usage that is evolving with the Technologies and the digital era.

IV. REVIEW REFLECTIONS

However Data, Information and Knowledge are different elements, there is a connection between them. A single analysed object can be sometimes Data, sometimes Information or even Knowledge, transitioning among these three dimensions depending on the background and perspective of the analysis. Regarding these transitions, it´s then introduced the concepts of Information Engineering and Knowledge Engineering, acting as the agents responsible for converting the elements from one dimension to the other (figure 9).

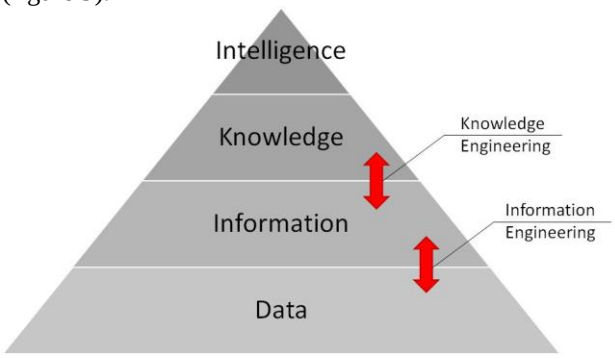

Fig. 9. IE and KE at the Information Pyramid – the authors

The concept of Engineering when applied within the information and knowledge domains acts as a methodology that provides processes and tools to deliver strategy, competitive intelligence and plans, that can support the decision making process of an organization, as describes the figure 10.

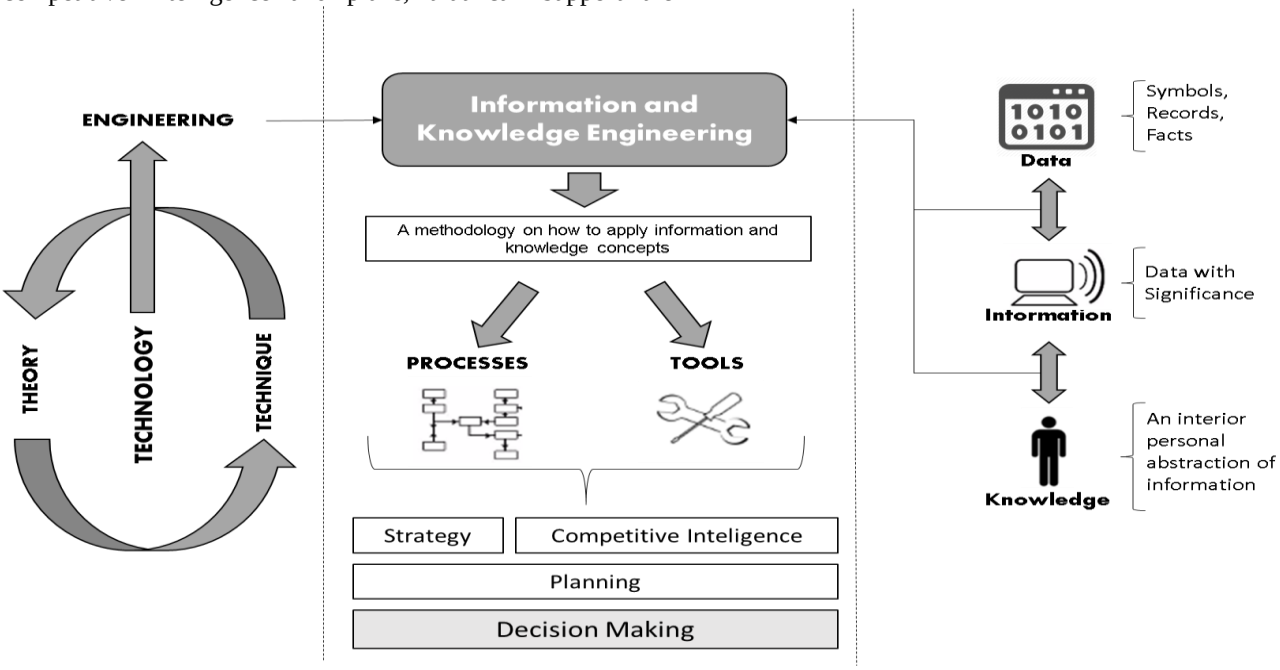

Fig. 10. Information and Knowledge Engineering – the authors

The Information Engineering focuses on the implementation of systems in a way to provide an efficient flow of information between the process, in order to facilitate its organization and allow an easy recovery. With this facilitation on the process of managing the information within the organization systems, the Knowledge Engineering recovers that information in order to process them, and uncover the underlying knowledge.

The result of those combinations will be the efficiency on the process of handling the information to extract strategic knowledge as a base for making decisions.

V. CONSIDERATIONS

Understood as an architectural approach to creation of information from planning, analysis, design, implementation and applications within a company to improve the management of its capital resources and people with no support of information systems focusing on the business vision.

Performance has many purposes, including planning organization, business reengineering, application development, planning and information systems re-engineering systems.

Note that there is still much room for advancement of research in these areas, the use of these concepts in environments beyond the virtual and supported by software and information technologies.

We foresee an opportunity to expand this research focusing on the Big Data era, and the upcoming technologies, in order to combine and structure a framework for IE and KE that can evolve together with the emerging technologies.

References

[1] A. A. Barreto (2002) A condição da informação. São Paulo em Perspectiva, Vol 16, Issue 3. pp 67-74

[2] A.V. Teixeira (2015). Elementos Componentes do Fluxo Informacional em Plataforma de Gerenciamento no Ensino Superior . Dissertation Programa de Pós-graducaçao em Ciëncia e Gestao Informacao, Universidade Federal do Paraná, 134p., 2015.

[3] A.V. Teixeira, M.C. Freitas Duarte, A.M. Laurindo (2014) Engineering Information: Conceptual Elements Related Information Management and Information Systems, EDULEARN14 Proceedings, pp. 6909-6915.

[4] C. G. B. Souza (2015) A Política e o Gerenciamento de Informação na Oficina de Música de Curitiba: Uma Análise de Dados. Universidade Federal do Paraná

[5] D. C. Nazário, M. A. R. Dantas and J. L. Todesco (2014) Knowledge Engineering: Survey of Methodologies, Techniques and Tools. IEEE LATIN AMERICA TRANSACTIONS, VOL. 12, NO. 8, DECEMBER 2014

[6] D. Tranfield; D. Denyer; P. Smart (2003) Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. British Journal of Management, Volume 14, Issue 3. pp 207 - 222

[7] F. T. Fonseca; J. E. Martin (2005) Toward an alternative notion of information systems ontologies: Information engineering as a hermeneutic enterprise. JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, Vol 56, Issue 1. pp 46 - 57

[8] G. Poujuan-Dante (1998) Gestion de Informacion em las organizaciones: principios, conceptos y aplicaciones. Universidad de Chile

[9] G. T. Schreiber, H. Akkermann, A. Anjewierden, R. de Hoog, N. Shadbolt, W. van de Velde, and B. Wielinga, Knowledge Engineering and Management: The CommonKADS Methodology. 2002.

[10] Hicks, B.J.; Culley, S.J.; Mcmahon, C.A. (2006) A study of issues relating to information management across engineering. Innovative Manufacturing Research, University of Bath, UK. pp 269-288

[11] HOGAN, P. T.; RAJA, M. K. (1997) Information engineering implementation in organizations: a study of factores affecting sucess. Journal of Information Technology Management, Volume 8, Numers 3& 4. pp 33 - 44

[12] K. Richmond (1991) INFORMATION ENGINEERING METHODOLOGY: A TOOL FOR COMPETITIVE ADVANTAGE. Telematics and Informatics, Vol. 8, Nos. 1/2, pp 41-57

[13] L. Hovarth; I. J. Rudas (2014) Systems engineering methods for multidisciplinary product definition. SISY 2014 - IEEE 12th International Symposium on Intelligent Systems and Informatics, Proceedings 6923604, pp. 293-298"

[14] L.F. LAI (2007) A knowledge engineering approach to knowledge management. Information Sciences Volume 177, Issue 19, 1 October 2007, Pages 4072–4094"

[15] Martin, James (1990) Information Engineering: Introduction. Prentice Hall: Engelwood Cliffs, New Jersey

[16] C.B Chapman, M. Pinfold (1999) Design engineering—a need to rethink the solution using knowledge based engineering. Knowledge-Based Systems 12 (1999) 257–267.

[17] Quintana-Amate, Bermell-Garcia, & Tiwari, (2015). Transforming expertise into Knowledge-Based Engineering tools: A survey of knowledge sourcing in the context of engineering design. Knowl.-Based Syst.

[18] S Seshasai, A Gupta, A Kumar (2005) An integrated and collaborative framework for business design: A knowledge engineering approach. Data & Knowledge Engineering Volume 52, Issue 1, January 2005, Pages 157–179

[19] S. A. Demurjian; Heidi J. C. Ellis (2012) Information Engineering: Object-Oriented Design and Analyses. Computer Science and Engineering Department The University of Connecticut Storrs, Connecticut 06269-3155

[20] S. D. Galup; R. Dattero (2000) Information engineering methodologies and organizational change: An exploratory study. JOURNAL OF COMPUTER INFORMATION SYSTEMS, Vol 41, Issue 2. pp 48 - 51

[21] Singh, N., Ding, S., Jagirdar, R., Basil, E.A. (1997) A knowledge engineering framework for rapid design. Source of the DocumentComputers and Industrial Engineering 33 (1-2), pp. 345-348"

[22] Studer, R., Benjamins, V.R., Fensel, D. (1998) Knowledge Engineering: Principles and methods. Data and Knowledge Engineering 25 (1-2), pp. 161-197

[23] W. P. LONGO (1987) Conceitos básicos sobre ciência, tecnologia e inovação. Política e Gestão em Ciência e Tecnologia. Escola Superior de Guerra

[24] Wielinga, B., Sandberg, J., Schreiber, G. (1997) Methods and techniques for knowledge management: What has knowledge engineering to offer?.

[25] Xindong Wu, Huanhuan Chen, Gong-Qing Wu, Jun Liu and Qinghua Zheng, Xiaofeng He and Aoying Zhou, Zhong-Qiu Zhao, Bifan Wei, Ming Gao, Yang Li, Qiping Zhang, Shichao Zhang, Ruqian Lu, Nanning Zheng,

[26] Apte and Nath (2004). The Size, Structure and Growth of the U.S. Information Economy in U.M.. Kluwer Academic Publishing.