Model of Knowledge Base for Supporting the Classification of a Psychological Profile in the Context of Social Engineering

Mauricio Amariles, Juan Giraldo, Claudia Durango, and Carla Zapata

Universidad de San Buenaventura, Medellín,
Colombia

{mauricio.amariles, juan.giraldo, claudia.durango, carla.zapata}@usbmed.edu.co

Abstract. Companies are looking for protecting their information implementing security policies and security tool kits in order to save the information, regardless that most thefts or unauthorized access are mainly committed by Human. The social engineer (perpetrator) finds vulnerabilities in employees (victims) who manage the information within Companies. To prevent this issue, is important to identify the profile of the social engineer using a psychometric tests such as 16PF. This psychometric test can identify characteristic on the personality of Social Engineer profile. The profile identified from this psychometric test allows detailing personality variables. Design a knowledge base KB will be a support to prevent and controlling the access to information from social engineering. The purpose of this Knowledge Base is characterize the Social Engineer profile to be applied for evaluate profiles in employees and candidates to job post in various Companies.

Keywords: Model knowledge base, social engineering, information security, psychological test 16PF, human hacking

1 Introduction

Most Companies invest on physical infrastructure, firewalls and software to ensure the transmission and confidentiality of information in order to protect their information. Several security policies, tool kits and hardware are implemented to save the information. However, inside the architecture of information systems, there are security holes where its functionality depends on the surveillance and control from human. The Social Engineering combines psychological techniques and pentesting tool kits to exploit these vulnerabilities on the system.

Social Engineer (perpetrator) uses psychological techniques to manipulate people (victims) in order to perform actions that may be convenient or inconvenient, achieving his goal. Goal can be getting access to confidential information in Companies [1]. Personality test 16PF (Personality Factor Questionnaire) is useful for factor analysis to

identify the basic dimensions of personality in humans. The first publication was in 1949 by Cattell, since 1949 others version are created such as 16PF: A, B, C, D, E and 16PF V -5, all versions measure the same factors, and factors are evaluated using the same methodology to obtain the same information . To characterize the profile of Social Engineer the 16PF Fifth Edition was used.

Social Engineer has patterns behavior characterized by the theft of confidential information using persuasion, deception and shrewdness to manipulate the actions on the targets. Interpret these kind of behavioral factors several physiological tools are proposed from theories of personality factors.

The knowledge base KB is a data repository, where knowledge of a specific domain is stored. Knowledge base is characterized by the ability to upgrade on the time, new knowledge providing from an expert. This knowledge is named as tacit knowledge. When tacit knowledge is transferred it becomes to explicit knowledge. Knowledge is carried on the discovery of valid data indicated objectivity. This discovery is generated by events or actions generating new experiences forward a specific domain [2].

2 State of the Art

Analyzing types of behavioral patterns and predict when a person has characteristics of a Social Engineer, different tools and hypothesis are proposed from the psychological theories to find personality factors. In [3] it is proposed a model to categorize the information in social networks, and then protect it from attacks using social engineering.

In [4] it is proposed a penetration testing methodology within a Company using social engineering. During the test a group of employees were evaluated. This test determines the vulnerabilities found on the targets and the methods to obtain the confidential information. This information is collected in a final report, describing the failed and successful attempts.

In [5] it is designed and implemented specialized expert systems to support decision-making for particular areas. One such system is called Copernicus focused on classification of perceptual skills. This is a tool for computer-aided diagnosis in mental disorders based on Minnesota Multiphasic Personality Inventory test (MMPI).

In [6] it is exposed an analytical study on the predictive power, the locus control and the risk propensity over five personality factors in predicting dishonest decisions in organizational contexts. These factors were measured as: neuroticism, extraversion, openness to experience, agreeableness and responsibility. Neuroticism refers to the tendency to feel nervous, anxious, self-conscious, hostile, vulnerable and impulsive. Extraversion is the tendency to be talkative, gregarious, outgoing, assertive and positive emotions. Openness to experience refers to the receptivity of new ideas, perspectives and experiences, involving tendency to fantasize and being sensitive to the arts. Affability is the propensity to be nice with others, empathy and humility. These features reflect the individual's personality. The analysis defines this profile as the basic component of personality, for measuring it, the 16PF was used to identify, analyze and explain this kind of personality [7].

In [8] authors propose a knowledge-based system (KBS) applied to the analysis of application security system. The knowledge base contains information on regulations, standards and current best practices, and related reports vulnerabilities to take public knowledge in the computing community.

In [9] it is proposed a knowledge-based system (KBS) applied to analysis the security on a system management application. The model is based on a knowledge-based system (KBS) which has a cognitive component that allows incorporate knowledge to the system. The constant threats and cyber-attacks this KBS through dynamic learning will keep updated it-self and is helpful for Information Security Officers in order to preparing requirements specification.

2.1 Concepts

Psychological Test 16PF.

The 16PF test consists of a questionnaire to measure normal personality dimensions [10] and is the only one that is built from a point of view not pathologic [11] (See Table 1).

Table 1. Global Personality Dimensions 16 PF

Dimensions		High and Low Meaning
Extraversion	EXT -	Introverted, Inhibited
	EXT +	Extraverted, Participating
Anxiety	ANS -	Hardy, Stress-resilient.
	ANS +	Stress Prone, Emotionally Unstable
Tough-Mind-edness	DUR -	Receptive, Open-minded
	DUR +	Tough- Minded, Resolute
Independence	IND -	Accommodating, Agreeable
	IND +	Independent, Persuasive, Willful.
Self-Control	AUC -	Unrestrained, Follows Urges
	AUC +	Self- Controlled, Reliable

In Colombia the 16PF V-5 is a valid and reliable test for recruitment, accompanying clinical psychologic processes, vocationally orient and perform legal evidences [12]. Additionally, propose 16PF as a valid instrument for psychological profiles on cyber-criminals [13]. This psychological test supports a set of variables, called primary 16PF scales (see Table 2).

To measure of personality has been proposed multiple instruments. However, the 16PF is the most used nowadays. Due to the acceptability of this instrument (16PF V - 5) in the Colombian context and because studies argue that other psychological test

such as BFQ not measure the total variance of personality [14], for purpose of this case, We opted to use the 16PF.

Table 2. 16PF primary factors

Factors	High and Low Meaning	
Warmth	A -	Reserved, Impersonal, Distant
	A +	Warm, Outgoing, attentive to others
Reasoning	B -	Concrete
	B +	Abstract
Emotional Stability	C -	Reactive, Emotionally Changeable
	C +	Emotionally stable, Adaptive, Mature
Dominance	E -	Differential, Cooperative, Avoids conflicts
	E +	Dominate, Forceful, Assertive
Liveliness	F -	Serious, Restrained, Careful
	F +	Lively, Animated, Spontaneous
Rule- Consciousness	G -	Expedient, Nonconforming
	G +	Rule-Conscious, Dutiful
Social Boldness	H -	Shy, Threat-Sensitive, Timid
	H +	Socially Bold, Venturesome, Tick-Skinned
Sensitivity	I -	Utilitarian, Objective, unsentimental
	I +	Sensitive, Aesthetic, Sentimental
Vigilance	L -	Trusting, Unsuspecting, Accepting
	L +	Vigilant, Suspicious, Skeptical, Wary
Abstractedness	M -	Grounded, Practical, Solution-Oriented
	M +	Abstracted, Imaginative, Idea-Oriented
Privateness	N -	Forthright, Genuine
	N +	Private, Discreet, Non-Disclosing
Apprehension	O -	Self-Assured, Unworried, Complacent
	O +	Apprehensive, Self-Doubting, Worried
Openness to Change	Q1 -	Traditional, Attached to familiar
	Q1 +	Open to change, Experimenting
Self-Reliance	Q2 -	Group-Oriented, Affiliative
	Q2 +	Self-Reliant, Solitary, Individualistic
Perfectionism	Q3 -	Tolerates Disorder, Unexacting, Flexible
	Q3 +	Perfectionistic, Organized, Self-Disciplined
Tension	Q4 -	Relaxed, Placid, Patient
	Q4 +	Tense, High Energy, Driven

Knowledge Management.

The headers in Companies are conscious of the strategic role of knowledge is recognized like generator of added value.

The success background in Companies is recognized by the ability to change according to knowledge evolution. The importance of the study of knowledge is caused by changes in the economy which are represented by the development of strategies to improve competitiveness and achieve sustainable development of actors in the markets [CEDV-1] [15]. Because this discussion, defining that knowledge management like a

dynamic process of pro- creation, storage, transfer, application and use of knowledge , for purposes of improving the performance in Companies [CEDV -2] [16]. Knowledge is a significant resource attached to importance of generate, dissemination and use of information. There are two types of knowledge: explicit and tacit. Explicit knowledge is formal and systematic, easy to communicate and share through products, formulas or software. The knowledge tacit is personal, hard to formalize and communicate it. This is related to the action and function [CEDV -3] [17].

Knowledge management is an organizational systematic and specific process, its purpose is to acquire, organize and communicate tacit and explicit knowledge of employees to be transferred to other employees of the organization in order to improve productivity and efficiency in their work [CEDV -4] [18]. Because universality of knowledge is important to generate knowledge bases. A knowledge base integrates data mining and new knowledge. All of this is based on known knowledge (tacit or explicit). A knowledge base seeks to establish the characteristics and quality of information [CEDV -5] [19].

An expert according to the Royal Spanish Academy (RAE) is "A person who is very skilled and he has great experience in a job or activity" other meaning is "A person with many knowledge of a subject".

3 16 PF Test Results

The psychological test 16PF was applied to 10 people. The profile found is according to people with harder knowledge in computing and exploit security vulnerabilities on the systems, they were referred as social engineer profile. The result was analyzed by an expert (psychologist) and result is shown in table 3.

Table 3. 16PF Test Result Profile

PRIMARY SCALE	FACTOR VALUE	PRIMARY SCALE	FACTOR VALUE
RESERVED	5	OPEN MINDED	3
CONCRET	6	ABSTRACT	9
EMOTIONALLY CHANGEABLE	8	EMOTIONALLY STABLE	2
DIFFERENTIAL	9	DOMINATE	3
SERIOUS	4	LIVELY	7
NONCONFORMING	2	DUTIFUL	7
SHY	2	SOCIALLY BOLD	7
UNSENTIMENTAL	4	SENTIMENTAL	8
TRUSTING	7	SUSPICIOUS	7
SOLUTION-ORIENTED	4	IDEA-ORIENTED	8
GENUINE	8	DISCREET	4
SELF-ASSURED	4	SELF-DOUBTING	7
TRADITIONAL	7	EXPERIMENTING	3
GROUP-ORIENTED	3	SELF-RELIANT	8
FLEXIBLE	8	PERFECTIONISTIC	3
PATIENT	8	TENSE	3
GLOBAL DIMENSION			
STRESS-RESILIENT	8	STRESS PRONE	3,2

PRIMARY SCALE	FACTOR VALUE	PRIMARY SCALE	FACTOR VALUE
INTROVERTED	5	EXTRAVERTED	7
UNRESTRAINED	7	SELF- CONTROLLED	4
ACCOMMODATING	7,8	INDEPENDENT	3
RECEPTIVE	7	TOUGH- MINDED	4

4 System Architecture

Seven components integrated the system functionality. Each employee answers the 16PF test applied by a psychologist. The results are obtained in terms of numerical values are stored in an Excel database, the format allows to specify the primary scales, and global dimensions. These data are taken by the psychologist and is transferred to the system through a graphical interface. This data is compared with values stored previously in the knowledge base, which contains the basic parameters obtained in the laboratory from the 16PF test. It will determine the primary scales and global dimensions for establishing whether an individual evaluated has the profile of social engineer.

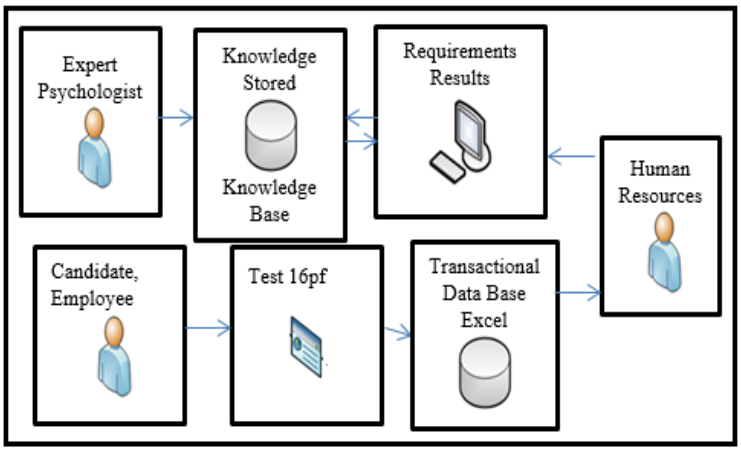

Fig. 1. *Test 16pf* result values are inserted manually into a *transactional data base* built in Excel by the *Human Resource*. System compares the values from *transaction data base* Excel with the data stored in the *Knowledge Base* and the feedback is shown by a graphical user interface.

4.1 Components

Candidate or Employee.
The test 16PF is applied by psychologists to employees or candidates to get a job in Companies. The results are analyzed and stored in a transactional database (Excel).

Psychologist or Expert.
He analyzes the results obtained from the test 16PF. He is in charged to feeds the knowledge base with specific information and general information.

Human Resources.
Use the system to check if the results obtained by applying the test 16PF on employees or candidates, are according to the parameters of primary scales and global dimensions found in the knowledge data base.

Knowledge Database.
In order to obtain the psychometric characteristics that determine whether a candidate or employee has the profile of social engineer. In this case a sample of 20 people have been evaluated by the 16PF test. The results were analyzed in detail by an expert psychologists. With the analysis of results it was possible to obtain quantitative data for each set of variables (primary scales and global dimensions), which are the general items to identify the profile evaluated will meet the characteristics to be a social engineer. Interaction with Knowledge Base is performed by an expert psychologist, who is the actor responsible for maintaining the current information stored in the system. The update consist to feed the knowledge base with the data for the primary scales and global dimensions.

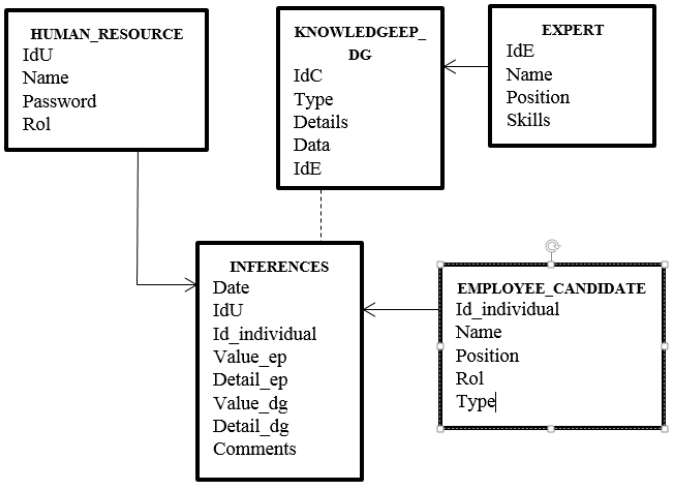

Fig. 2. Knowledge Model for the system

Data Model.
PS: Primary Scales
GD: Global Dimensions

- Human resource area has stored the knowledge with the actors who are responsible for interacting with the system, searching for determine whether an evaluated has a profile of Social Engineer.
- Expert stores data related experts in charged to update the knowledge data base.
- Employee is the data structure that stores information related to 16PF result applied to candidates or employees in the Company. This results determined from the inferences in the system and will determine the profile.

- Inferences saves the results obtained from the application of16PF test. There the results are saved by date, relating the values of primary scales and global dimensions.
- The stored knowledge serves as support for the inferences in the system, added to the results for each candidate or employee evaluated by the 16PF test. This is located in the ep_dg knowledge structure.

Inference Engine.
Facilitate processing of the data, specific knowledge memory ep_dg climbs (see Table 4). In this way the data are loaded into the matrix by reducing the cost of processing between the application and the knowledge base.

Structure Data matrix.
Once the user click on button "infer" ep_dg data are uploaded to memory, (see Table 4). This reduce to call n times to the knowledge base during data processing.

The results of test 16PF applied to each candidate or employee in Company are compared with the data found in ep_dg memory. This comparison is performed through a set of production rules. The data resulting from the application of 16PF test for each candidate or employee in Company are also uploaded to memory, another matrix structure. Almost two symmetric matrices in magnitude. This further processing and inference in the results.

Table 4. 16PF primary factors and global dimension located in memory

MATRIX MEMORY KNOWLEDGE PROCESSING (EPDG)				
TYPE (0,0)	DETAIL (0,1)	DATA1 + - (0,2)	DETAIL (0,3)	DATA2 + - (0,4)
PS	RESERVED	5	OPEN MINDED	3
PS	CONCRET	6	ABSTRACT	9
PS	EMOTIONALLY CHANGEABLE	8	EMOTIONALLY STABLE	2
PS	DIFFERENTIAL	9	DOMINATE	3
PS	SERIOUS	4	LIVELY	7
PS	NONCONFORMING	2	DUTIFUL	7
PS	SHY	2	SOCIALLY BOLD	7
PS	UNSENTIMENTAL	4	SENTIMENTAL	8
PS	TRUSTING	7	SUSPICIOUS	7
PS	SOLUTION-ORIENTED	4	IDEA-ORIENTED	8
PS	GENUINE	8	DISCREET	4
PS	SELF-ASSURED	4	SELF-DOUBTING	7
PS	TRADITIONAL	7	EXPERIMENTING	3
PS	GROUP-ORIENTED	3	SELF-RELIANT	8
PS	FLEXIBLE	8	PERFECTIONISTIC	3
PS	PATIENT	8	TENSE	3
GD	STRESS-RESILIENT	8	STRESS PRONE	3,2
GD	INTROVERTED	5	EXTRAVERTED	7
GD	UNRESTRAINED	7	SELF- CONTROLLED	4
GD	ACCOMMODATING	7,8	INDEPENDENT	3
GD	RECEPTIVE	7	TOUGH- MINDED	4

Graphical Interface User

This interface allows users to register the result of 16 pf test. System infer and feedback with result of profile. Depends of values, the result could matched with a social engineer profile.

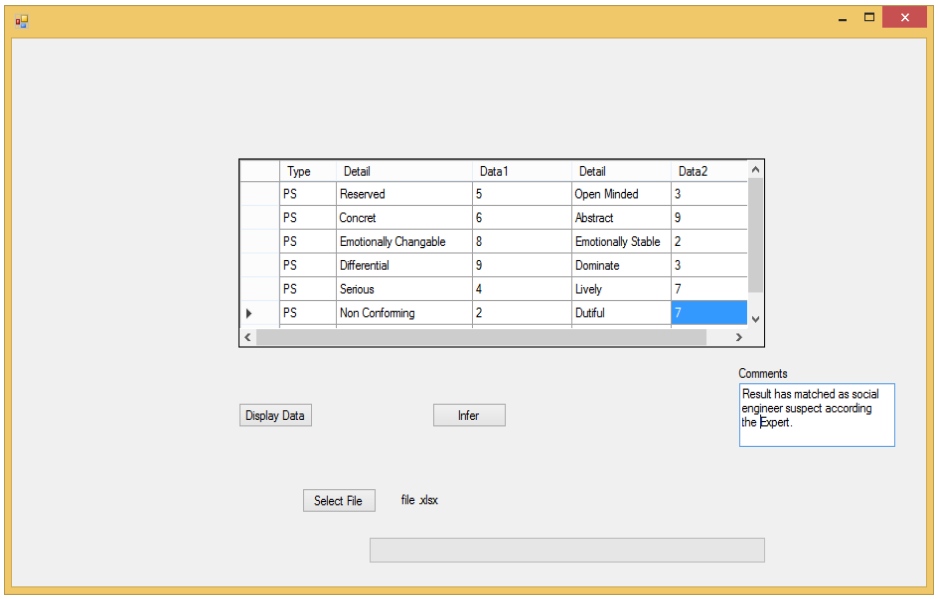

Fig. 3. Knowledge Base Graphical Interface. The Personal Human Resource, use this interface to upload the result of 16PF test. System infer with the information stored in data base knowledge. Feedback depends of information registered by Expert previously.

4.2 Base Rule

Matrix_epdg: contains the general knowledge base to characterize an evaluated as social engineer profile.

Matrix_r16pf contains the results of each candidate of employee evaluated using the 16PF test.

```
Profile Employee Test Program (Output)
  Memory_payload  file_results_16pf (employee)
  Memory_payload file_results_16pf (knowledge base ep_dg)
  For  i=1;i<=21;i++
    Get matrix_epdg(i,2), matriz_r16pf(i,2)
    If  matrix_r16pf(i,2) > = matrix_epdg(i,2) then
      Print matrix_epdg(i,3)
    Else
      Print matrix_epdg(i,1)
    End If
    If  matrix_r16pf(i,4) >= matrix_epdg(i,4) then
```

```
      Print matrixz_epdg(i,1)
    Else
      Print matrix_epdg(1,3)
    End If
  End For
  Close_  file_results_16pf (employee)
  Close_ file_results_16pf (knowledge base ep_dg)
End.
```

5 Case Study

To test the knowledge base system five employees in a Law Firm (name reserved) were evaluated using the 16PF. In order to keep the confidentiality of information, non all data are disclosed.

Case Data File 1.
User ID: private
Name: Private
Position: Office Manager
Role: Information System Manager
Employment relationship: Employee

Human Resource.
Human Resource ID: Private
Name: Private
Password: Private
Role: Human Office Manager

Inference.
Date: 25/04/2015
Human Resource ID: Private
User ID: private
Value_ep, Detail_ep, Value_dg, Detail_dg: see Table 5.
Comments: The profile result has matched as social engineer suspect according the Expert.

Table 5. 16 PF test result from an evaluated employee (matrix of memory knowledge processing (EPDG))

TYPE (0,0)	DETAIL (0,1)	DATA1 + - (0,2)	DETAIL (0,3)	DATA2 + - (0,4)
PS	RESERVED	5	OPEN MINDED	3
PS	CONCRET	6	ABSTRACT	9
PS	EMOTIONALLY CHANGEABLE	8	EMOTIONALLY STABLE	2
PS	DIFFERENTIAL	9	DOMINATE	3
PS	SERIOUS	4	LIVELY	7
PS	NONCONFORMING	2	DUTIFUL	7

TYPE (0,0)	DETAIL (0,1)	DATA1 + - (0,2)	DETAIL (0,3)	DATA2 + - (0,4)
PS	SHY	2	SOCIALLY BOLD	7
PS	UNSENTIMENTAL	4	SENTIMENTAL	8
PS	TRUSTING	7	SUSPICIOUS	7
PS	SOLUTION-ORIENTED	4	IDEA-ORIENTED	8
PS	GENUINE	8	DISCREET	4
PS	SELF-ASSURED	4	SELF-DOUBTING	7
PS	TRADITIONAL	7	EXPERIMENTING	3
PS	GROUP-ORIENTED	3	SELF-RELIANT	8
PS	FLEXIBLE	8	PERFECTIONISTIC	3
PS	PATIENT	8	TENSE	3
GD	STRESS-RESILIENT	8	STRESS PRONE	3,2
GD	INTROVERTED	5	EXTRAVERTED	7
GD	UNRESTRAINED	7	SELF- CONTROLLED	4
GD	ACCOMMODATING	7,8	INDEPENDENT	3
GD	RECEPTIVE	7	TOUGH- MINDED	4

References

1. C. Hadnagy: Social engineering: The art of human hacking. John Wiley & Sons (2010)
2. R. Pieraccini, L. Rabiner: Artificial Intelligence versus Brute Force The Voice in the Machine: Building Computers That Understand Speech. pp. 83–107 (2012)
3. I. Kotenko, M. Stepashkin, E. Doynikova: Security Analysis of Information Systems Taking into Account Social Engineering Attacks. In: 19th International Euromicro Conference on Parallel, Distributed and Network-Based Processing, pp. 611–618 (2011)
4. T. Dimkov, W. Pieters, P. Hartel: Two methodologies for physical penetration testing using social engineering. In: Proc. 26th Annu. Comput. Secur. Appl. Conf. - ACSAC'10, pp. 399–408 (2010)
5. D. Jachyra, K. Pancerz, J. Gomula: Multiway classification of MMPI profiles. In: International Conference on Digital Technologies, DT 2013, pp. 90–98 (2013)
6. E. E. Kausel, P. I. Leiva, M. Sanfuentes, E. Barros: Más allá de los cinco grandes: Disposiciones y personalidad en la predicción de decisiones deshonestas en el contexto organizacional. Rev. Innovar J., vol. 22, pp. 109–122 (2012)
7. B. Sandin, P. Chorot: Evaluación de rasgos psicológicos. Rev. Psicodiagnóstico, vol. 2, pp. 287–326 (1990)
8. María Victoria Bajarlía, Jorge Eterovic, Jorge Ierache: Modelo de Sistema Basado en Conocimiento para el Análisis de la Seguridad de la Información en el Contexto de los Sistemas de Gestión. Cacic - xvi congreso argentino de ciencias de la computación (2010)
9. M. Bajarlia, J. Eterovic, J. Ierache: Modelo de Sistema Basado en Conocimiento en el Dominio de la Seguridad de Aplicaciones. Revista Latinoamericana de Ingeniería de Software, 1(6): 241–252 (2013)
10. S. Karson, J. W. O'Dell: A guide to the clinical use of the 16PF. 160 p. (1976)
11. M. Torras: Instrumentos usuales en la evaluación clínica de adultos. España (1994)
12. A. M. Álvarez S., N. M. Giraldo H.: Perfil psicológico del personal que desempeña el cargo de cajero de un hipermercado de la ciudad de Medellín. (2009)
13. S. González L., V. Peralta P., M. Reyes B., E. Reza S.: Perfiles psicológicos de delincuentes informáticos. (2009)

14. S. R. Briggs: The optimal level of measurement for personality constructs. Nueva Jersey: Springer, pp. 246–260 (1989)
15. J. C. Acosta, M. Longo, A. L. Fischer: Capacidades dinámicas y gestión del conocimiento en nuevas empresas de base tecnológicas. Cuad. admon. ser. organ, vol. 26, no. 47, pp. 35–62 (2013)
16. J. J. Tarí, M. García: ¿Puede la gestión del conocimiento influir en los resultados empresariales? Cuaderno de Gestión, vol. 13, no. 1, pp. 151–176 (2013)
17. M. C. Muñoz, F. S. Marco: Gestión del conocimiento: representación y métricas. utilizando el método DACUM. Rev. Ing. Ind., vol. 1, no. 1, pp. 5–14 (2002)
18. M. D. Gil-Montelongo, G. López-Orozco, C. Molina-García, C. A. Bolio-Yris: La gestión de la información como base de una iniciativa de gestión del conocimiento. Ing. Ind., vol. XXXII, no. 3, pp. 231–237 (2011)
19. B. Qin: Reductions in a knowledge base. Inf. Sci., vol. 320, pp. 190–205 (2015)