Information and Knowledge Management www.iiste.org
ISSN 2224-5758 (Paper) ISSN 2224-896X (Online)
Vol.3, No.11, 2013

Information and Knowledge Valuation using the Information Theory and Informative Matrices

Juan Carlos Piedra-Calderón[1]* Rubén González Crespo[2] J. Javier Rainer[2]

1. Pontifical University of Salamanca, Av. de la Merced 108, 37005, Salamanca, España.
2. Pontifical University of Salamanca, Paseo Juan XXIII, 3 - 28040 - Madrid. España.

*E-mail of the corresponding: juankpc@gmail.com

Acknowledgement: This research is supported by the Society and Information Technology Research Group (GISTI), Pontifical University of Salamanca.

Abstract

For a more concrete analysis of the quantification of knowledge, one must normalize the quantification of data and information. One must, without entering into excessive detail, resort to the information theory, recognizing three features that data, information and knowledge have in common: weight, content and quantity. The following article is an introduction to and conceptualization, in the light of the Information Theory (IT), of the quantification of information and knowledge by means of informative matrices and a normalization valid for every IT organization.

Key words: Data, information, knowledge, informative quantity, informative weight, informative content, informative matrix, information theory, generate.

Starting With The Information Theory

To quantify abstract concepts such as information and knowledge, one must begin with concrete models, such as the Information Theory (Textos Científicios, 2008), which this is directly related to these concepts. Nevertheless, in order to have a complete monitoring of this development, it is necessary to perform a standardization adapted to this work that will allow us to have a better understanding of the results. This standardization will develop an analogy to find a way to measure the information received from any system, not necessarily from a computer based system (López, 1998).

Since the 40's (note 1) there have been many normalization proposals, some more complex than others, but most of them taking the information as a set of mediated communication transmitted by technologies with several elements such as: receiver, transmitter, channel, interference and others (Hency, 2011). But it is not easy to find a proposal that only take into account the quantity of information and knowledge gathered from a set of specific data. From a more concrete measurement, one could reach a more definitive conclusion towards the measurement of knowledge a person can achieve from a processed and given context. For example: if one studies an orange, all that matters is what the orange provides, without taking into account that previously a tree was planted and watered until it grew and bore fruit. This means that ultimately what matters is the information as the unique actor, and not as part of a set of actors framed in communication. This implies that it doesn't matter if there is a difference between the original and the received message, because only the final information that has been received will be processed.

Data

There are many ways to define data although it is always seen as a synonymous of information. According to the Royal Spanish Academy Dictionary data can described as follows:

- "Data is a necessary premise to reach the exact knowledge of something" (Diccionario, 1986). Note that in this concept there are three important aspects to underline: history, process (implicit) and knowledge.
- "Data is a knowledge that has no meaning by itself when it is out of context. It is something incomplete that needs a complement such us other data or elaboration process that will give it more sense" (De Pablo, 1989).
- "Data are numeric variables, qualitative values, phrases, words, symbols, etc.., that brings a real knowledge of the system or the fact that it is being studied" (Minguet & Read, 2008).

Of these concepts can be underline items such as: it is a premise, has no meaning by itself, it must be accompanied by other information to put it into context, and it is anything from which someone can get knowledge. And so on, one can find plenty of definitions that always mingle the three interlinked concepts: data, information and knowledge, without effectively focusing on what is data.

All data by itself can be considered as a source of information, but in order to have a meaning, it must be necessarily associated to a system that contextualizes it. That means that it is essential to have another data and compare these two, in order to reach a conclusion that will give a meaning to that information and produce knowledge.

Data is a basic entity which has information, but it depends on a feature of contextualization. This means that data as an entity has value, which corresponds to the informative weight (TorredeBabel, 2012). This weight revolves around its entity, which would correspond to the informative content, and at the same time it gives it a context, which would correspond to the informative quantity. From this ontological conception a better definition can be reached that better fit what data really is (TorredeBabel, 2012a):

"Data is an entity whose characteristics are to have informative weight, informative content, and informative quantity, but by itself has no value or meaning. Because it has not been contextualized its contribution does not make sense or is null. The combination of data forms an informative system".

While speaking of data the word information is introduced, not referring to the semantic meaning of the word, but rather the substance that data has by itself, that is to say, its content. One must take into account though, that both information and knowledge are concepts that have an individual entity, and therefore can be considered as data at a given time.

Data Type

There are two types of data: simple and complex:

- ***Simple*** → Their characteristics are those of the element and do not depend on any other complement.
- ***Complex*** → Their characteristics are given by the addition of characteristics of other simple or complex data that complement each other.

Informative Systems

The most basic systems (note 2) always have two data, for example: on and off (appliance), to ring or no to ring (ring bell), dot-dash (Morse), one and zero (binary), to beat or not to beat (heart). If one of these data units happens to be missed, then the system would be meaningless, because it needs a partner to have sense (Singh, 1982).

Another example: given the datum "15", it is easy to think that it is a number that could represent a quantity, but it could be several things because it has not been contextualized. If another data unit is added, “years old” and yet we add a third one, "girl", all these pieces of data together start to form a system that is beginning to make sense. Therefore the operation of the information systems would be as follows: various data are received, which together provide a contextualization, from which one can obtain information and can contribute with knowledge.

Table 1: Informative weight, informative content and informative quantity of d1, d2 y d3

datum$_1$	***datum$_2$***	***datum$_3$***	***Information***	***Knowledge***
“15”	“years old”	“girl”	“15 years old girl”	“age”

Informative Weight, Informative Content And Informative Quantity Of A Datum

Assuming a datum: d_1, which contributes with a weight, a content and a quantity of information. To contextualize d_1, will be assumed to be an orange. This orange has a weight (50g), a content (the pulp, which consist of several liquid drops) and a quantity (the juice which would be left after squeezing it out).

In the following table, it is possible to see the information values that has a datum by itself. It has no units, because they do not belong to any system:

Table 2: Informative weight, informative content and informative quantity, of a datum.

Datum	***Individual weight relative (d_p)***	***Individual content relative (d_c)***	***Individual quantity relative (d_T)***
d_1	1	¿1?	¿1?

Information

Following the same procedure as when defining “data”, our first resource will be the dictionary of the Royal Academy of the Spanish language where it analyzes several concepts of information to reach a better understanding of this entity:

- Information is a: "Communication or knowledge acquisition that allows to expand or precise the knowledge on a particular subject" (Diccionario, 1986).
- "Information is an elaborated set of data placed in a context so that it has a meaning to someone in a particular time and place" (De Pablo, 1989).
- "Information is a concept by which men represent events and facts" (Minguet & Read, 2008).

Of these concepts, one can take: information is exchanged, in other words, it has to be interpreted in order to reach to an understanding. At the same time it means it would be pointless to have unused information, because it would not be able to have a feedback. Additionally, it is a data set, which is under the same context, and from which one can extract some knowledge. Out of all this one can say that:

Information is an entity whose characteristics are to have informative weight, informative content and informative quantity, within a given context. It is the gathering of several data existing into a same context and forming an informative group.

Information is the key element of knowledge, with the characteristic that it can be grouped with more informative systems, within the same context. The orderly combination of informative systems with the same entity's features, allows it to be part of a larger system where it behaves such as complex data. Information itself is a simple system, but when it is combined with other information it becomes a complex system.

Informative Weight, Informative Content And Informative Quantity Calculation

Assuming two data d_1 and d_2, which form an informative system, that for instance on the computer world could represent a "0" or "1" (Minguet & Read, 2008), each will have a weight, content and quantity of information. Contextualized in the following manner: two data (e.g., flour and water mixture), the weight is the addition of individual weights of each element (e.g., wheat or corn flour; fresh water or sea water), which will leave as a result the total system weight. And finally the quantity will be what we make of it (e.g., bread).

3.1.1. Informative Weight

The informative weight is the global maximum value from single or complex data that can by itself an informative system. Then to calculate the weight of data, it is necessary to start from the conception of that each one will donate its total weight, that is:

$$d_1 = 1 \qquad \text{and} \qquad d_2 = 1$$

Keeping in mind that although they have the same weight, these represent two completely different data but belong to the same system, therefore d_1 will always be different from d_2. Making an association with those two data in a system, the informative weight of these association, called "Ip", could be obtained. The result will come out from the addition of the individual weight values of d_1 and d_2, that is:

$$I_p = d_1 + d_2$$

In general, one would obtain the following equation to find the weight of the information of any system:

$$I_p = \sum_{i=0}^{m} d_i \tag{1}$$

Substituting for the values of their weights:

$$I_p = 1 + 1$$

The result of the informative weight would be:

$$I_p = 2 \text{ units of informative weight}$$

3.1.2. Informative Content

It is the maximum quantity of information that can bring each permutation series of an information system. To find the value of d_1 and d_2, the following possibilities should be considered:

- The informative content "I_c", can be obtained through the addition of the weights of data:

$$\boldsymbol{I_c = d_1 + d_2}$$

If the weights of information are the same, it could be simplified to:

$$\boldsymbol{I_c = n \cdot I_p} \qquad (2)$$

Which in this case, after substituting in (2) for the values of their weights would leave as follows:

$$I_c = 1 + 1 = 2 \qquad o \qquad I_c = 2 \cdot (1) = 2$$

Therefore, the result of the informative content would be:

I_c = 2 units of informative content

- The informative content "Ic" can be achieved by multiplying the weights of data, as follows:

$$\boldsymbol{I_c = I_{p1} \cdot I_{p2}}$$

If the weights of information received are the same, it could be simplified to:

$$\boldsymbol{I_c = (I_p)^n} \qquad (3)$$

Substituting in (3) for the values of their weights:

$$I_c = 1 \cdot 1 = 1 \qquad o \qquad I_c = (1)^2 = 1$$

The result of the informative content would be:

I_c = 1 unit of informative content

As it is not possible that an operation with the same number of data with constant weights, gives as a result two different informative contents, therefore it is necessary to find another mathematical form that contains the properties of addition and multiplication.

There is a mathematical property of the logarithms, which says "the logarithm of a multiplication is equal to the addition of the logarithms of the factors" (Profesoenlinea, 2011), that is:

$$\log A \cdot B = \log A + \log B$$

Substituting with the values of the example, the general equation to find the informative content would be the following:

$$\boldsymbol{I_c = \log_a (I_{p1} \cdot I_{p2}) = \log_a I_{p1} + \log_a I_{p2}}$$

In general, when the informative contents "I_p" are the same, the equation will be as follows:

$$\boldsymbol{I_c = \log_a (I_p)^n} \qquad (4)$$

$$\boldsymbol{I_c = n \cdot \log_a I_p} \qquad (4')$$

Before replacing the weight values of d_1 and d_2, it is important to know what will be the base "a" of the logarithm to be used in the operation. Ultimately, and to sum it up, it is necessary to make an analogy:

- Given that a datum in itself has an informative content equal 1, and it has been proven that to obtain the informative content is necessary to use a logarithmic function, then "d_c" would have to comply with the following:

$$d_c = \log_a d_p$$

Resorting to the concept of logarithm, as "logarithm of a number to a given base is the exponent to which the base must be raised to obtain the number" (Profesoenlinea, 2011). Solving the equation, the subscript "a", would be:

$$a^{dc} = d_p$$

Knowing that the minimum unit that brings a data value is $d_c = 1$, then the final answer to the logarithm base, will be:

$$a = d_p$$

This means that the logarithmic base that is used to obtain the informative quantity that is given by the information of the received data, will depend on the system upon which they are contextualized, that is to say, for a system with two data, the base shall be two and so on (Arrendondo, 2011).

Therefore, to find the value of the informative content of the data pair d1 and d_2, the value "I_p" should be replaced, and in this case the base of the logarithm is "2" because there are two data, and would be as follows:

$$\boldsymbol{I_c = \log_2 (I_p)} \qquad (5)$$

Where

$$I_c = log_2 (d_1 + d_2)$$

Substituting in (5) for the values of their weights:

$$I_c = \log_2 (1 + 1)$$

That is to say:

$$I_c = \log_2 2$$

The final result is:

$$I_c = 1 \text{ unit of informative content}$$

3.1.3. Informative Quantity,

Until now all seems to be going fine under the Information Theory, but in the next step is where the first change shows up by introducing the new concept of "informative quantity", which is not directly related to the weight of the system, but with the informative content (Singh, 1982).

In order to obtain the informative quantity from the data pair, it is necessary to consider the probability of occurrence "P", which the data has. Whereby informative quantity "d_t" would be obtained by multiplying the probability of the occurrence "Pp" of each value, with the informative content:

$$d_t = P_p \cdot log_a (I_c) \qquad (6)$$

According to Table 2, the value of "d_t", would have to be "1", which is consistent with the law of probabilities that says: "the probability of a safe process is equal to 1". In this case, the datum is an entity that exists by itself as a whole. Therefore its probability to be one datum is 100% or "1".

$$d_t = 1 (\log_2 1)$$

Therefore the result would be:

$$d_t = 0$$

This indicates that the Table 2 would have to be modified in the part that corresponds to the informative weight, content and quantity, than had been assumed at the beginning for one individual datum value, as follows:

Table 3: Informative weight, informative content and informative quantity, in base of a datum.

Datum	***Relative Individual Weight (d_p)***	***Relative Individual Content (d_c)***	***Relative Individual Quantity (d_t)***
d_1	1	1	0

As an example, if this unit (quantity) of Minimum information "d_t" is concretized and contextualized in the computer world, it would be known as "bits".

Now, returning to a minimum system with a set of data d_1 and d_2, in order to obtain the informative quantity "I_t", it must be related with the probability of occurrence of data, as well as with its informative content. Therefore the information "I_t" will be the addition of the products of probabilities "P_c" multiplied by the weight of the corresponding data, as follows:

$$I_t = P_c \cdot (I_c)$$

Replacing "I_c"

$$I_t = P_c \cdot log_a (I_p)$$

Replacing "I_p"

$$I_t = P_c \cdot [\, log_a \,(\, d_1 + d_2 \,)\,]$$

Making a general rule to get "I_t", keeping in mind the possibility that the probabilities of occurrence of data are different, being the logarithmic base "a", the formula would be:

$$I_t = P_{ci} \cdot log_a \,(\, \sum_{i=1}^{m} d_{p\text{-}i} \,) \qquad (7)$$

As d_1 and d_2, have the same probability to be received, and cannot be received separately, that is to say 100% or 1 for each one given that the contextualization is very wide. Replacing in (7) for their numerical values:

$$I_t = 1 \cdot \log_2 (1 + 1)$$

The result of the informative quantity would be:

$$I_t = 1 \text{ unit of informative quantity}$$

This leads to the measuring unit of the informative quantity of systems in general.

Contextualized in the world of computers, and grouping them into a set of eight bits, the result is the "byte". That is also known as the unit of information, and shows that contextualizing data, can provide a first sense of it and provide some information (Singh, 1982).

Finally an informative system consists in the grouping of permutation of the received data. As a first analysis, it will give us as a result, an informative weight, an informative content and an informative quantity of the information system.

3.2. Informative Weight, Informative Content And Informative Quantity, Of A Simple System

As previously seen, a pair of data gives limited information, therefore it is necessary to think of a more complex system. Taking the same example of the data d_1 and d_2, and doing permutations with them to conform new groups (pairs in this case) of data d_1 and d_2, it will give some additional information, shown in the following table:

Table 4: Permutations of data d1 and d2.

Permutation "n"	***Permutations of data d_1 y d_2***	
I	d_1	d_1
II	d_1	d_2
III	d_2	d_1
IV	d_2	d_2

The above table shows that with two data it is possible to obtain four data combinations, which means four pairs of data. This means having more information from the whole set. The next step is to make operations between the weights of the data, their content, their quantities and their respective probabilities of occurrence.

3.3. Probabilities And Relative Weights Of The Data In Permutations In A System

The values of probability of occurrence, are the result from the calculation of the relative weights of individual data. They are used to display the occurrence that a datum has in respect to others in the same group or permutation where it is located. That is, if it is expected two different data but instead it had received a pair with the same datum, the probability of that datum will be 100%. The following table presents the probability of occurrence of each of the data, regardless of the data that is expected.

Table 5: Relative probability of data d1 and d2

Permutation "n"	***Individual probability to be only d_1***	***Individual probability to be only d_2***	***Individual relative probability permutation***
I	100,00%	0,00%	100,00%
II	50,00%	50,00%	100,00%
III	50,00%	50,00%	100,00%
IV	0,00%	100,00%	100,00%

Considering that the addition of the weight of a pair of data is one unit, and that the individual probability that a pair of data could be received is 100%, that leaves the relative weight of data as follows:

Table 6: Relative weight of a pair of data d1 and d2 in their permutation

Permutation "n"	***Relative weight of a pair of data*** d_1 *y* d_2		***Weight of the permutation***
I	0,5	0,5	1
II	0,5	0,5	1
III	0,5	0,5	1
IV	0,5	0,5	1
			Total weight = 4

From the table above it can be deduced that the value of the relative weights will also be equal to the value of their relative probabilities, because if a datum X is expected, and a datum Y is received, the relative probability that either one or the other is received will be reduced to the half that if it just waits to get the datum X. That is, if it waits d_2 and d_2 is received, the probability is 100%, but if it waits d_2 and d_1 is received, then the probability to receive d_2 has been reduced to 50%. In larger systems, the probability of permutation will be reduced according to the number of data that the group has. This is shown in the following table:

Table 7: Relative probability of a pair of data d1 and d2 in their permutation

Permutation "n"	***Relative probability of a pair of data*** d_1 *y* d_2		***Permutation Probability***
I	50,00%	50,00%	100,00%
II	50,00%	50,00%	100,00%
III	50,00%	50,00%	100,00%
IV	50,00%	50,00%	100,00%

3.4. Probability And Absolute Weight Of Data In Permutations On A System

Given that the analysis of relative weights and relative probabilities does not comply with the total weight of the system that is one unit, and the total probability of the system which is 100%, it is necessary to find the absolute values of probability and weight for the system. The following premises should be taken into account:

- All permutations form a system.
- The addition of probabilities for each permutation, should result in the probability of the whole system that is always 100%.
- Having four permutations, the absolute probability of each permutation will be the result of dividing the total probability of the system between the numbers of permutations that the system has.
- The values for the individual absolute probabilities of each data in the pair, will be a result obtained by dividing the probability of the permutation with the number of data in the series but considering whether the data arrives or not.

Then to fill the following table, first the initial probability of each data is calculated by dividing the total weight of the system (which is one unit of weight with 100 % of probability), by the number of permutations (there are four of them in the example).

Once the value of the probability of each permutation has been obtained, this becomes a new total that must be divided by the value of data the permutation has. Taking into account the location where each data is expected. The probability of obtaining an unexpected data is zero, which will leave the table as follows:

Table 8: Individual probability of a pair of data d1 y d2

Permutation "n"	***Absolute individual probability of data d_1 and d_2 in the pair***		***Absolute Probability of the permutations***
	d_1	***d_2***	
I	25,00%	0,00%	25,00%
II	12,50%	12,50%	25,00%
III	12,50%	12,50%	25,00%
IV	0,00%	25,00%	25,00%
Relative probability of data	**only d_1 50,00%**	**only d_2 50,00%**	**System probability 100,00%**

For absolute weights that a specific data has in regards of another, the same previous analogy of additions and divisions should be followed. , Once the table of probabilities has been obtained, it would only be necessary to extrapolate it directly to the concept of weight, as they have a one to one ratio (Arrendondo, 2011), obtaining the following table: (Quintanilla, 2012)

Table 9: Individual weight of a pair of data d1 y d2

Permutation "n"	***Absolute individual weight of data d_1 and d_2 in the pair***		***Absolute weight of the permutations***
I	0,25	0	0,25
II	0,125	0,125	0,25
III	0,125	0,125	0,25
IV	0	0,25	0,25
Relative weight of data	**Relative weight of d_1 0,5**	**Relative weight of d_2 0,5**	**System weight 1**

Finally, to obtain the real absolute values of the entire system, one should have into consideration that the system is always formed by a certain series of permutation, that is, a group of expected information to be more likely than other permutations. Which means that once the division has been made to get the weight values for the permutations, as well as the division for the weight values of each data, the following must be considered:

- In the box of d_1 datum, if the same datum is obtained for it, the value obtained should be put in that box.
- In the box of d_2 datum, if another datum is obtained, a zero is placed on the box, and its value is added to the data d_1, which is found in the same position and permutation as d_1.
- The value is null when in the d_1 box, is received d_2 and vice versa, (if they are in the same permutation). And the value of this permutation is added to the permutation value that the expected series has. In this permutation the individual calculation is repeated.
- For a system of "n" data, the same analogy will be followed.

Table 10: Absolute individual Weight / Probability of data d1 and d2 in the permutation

Permutation "n"	***Expected series (d_1, d_2)***		***Absolut Weight / Probability of data d_1, d_2***		***Absolut Weight / Probability***
	Permutations		***Expected d_1***	***Expected d_2***	
I	d_1	d_1	0,250	0,00	0,250
II	d_1	d_2	0,250	0,250	0,5
III	d_2	d_1	0,00	0,00	0,00
IV	d_2	d_2	0,00	0,250	0,250
			0,50	**0,50**	**System Weight / Probability 1**

3.5. Informative System Weight, Calculation

As seen previously, within a given system, one can find two kinds of weight to calculate the system, absolute and relative:

- **Absolute** → it is the real weight of a datum within the data series. The absolute weight of the permutation is the addition of the absolute weights of the data series. The addition of the absolute weight, result in the absolute weight as a whole.
- **Relative** → it is the addition of the absolute weights of a same datum in the same data line where it is expected to be received, as it is shown in the next table. The addition of the relative weights results in the absolute weight of the whole.

To differentiate the weight of the system, from the rest of weights that are within the system, it will be represented as "D", and its result will be the addition of all the individual weights of each group of system data.

$$D = \sum_{i=1}^{m} I_p = \sum_{i=1}^{m} d_p \tag{8}$$

Table 11: Absolute y Relative weight of data d1 y d2

Permutation "n"	***Absolute individual weight of data d_1 and d_2 in the pair***		***Absolut weight of information I_p***
I	d_1	d_1	$d_1 + d_1$
II	d_1	d_2	$d_1 + d_2$
III	d_2	d_1	$d_2 + d_1$
IV	d_2	d_2	$d_2 + d_2$
	Relative weight of d_1 in the system d_{p1}	**Relative weight of d_2 in the system d_{p2}**	**Informative System Content** D

3.6. Informative System Content, Calculation

To get the absolute value of the information content of the system, the logarithmic function is used in each permutation value, as indicated above, , and later on adding all these values to obtain the total result of the system. To differentiate the informative system content, from other informative content, it will be represented as "C", and its result will be the multiplication of the logarithm with base "a" (which depends on the system, but for our case, due to our calculations being with decimal ending, in our case will have base 10), with the weight of the entire system information.

$$A = \sum_{i=1}^{m} log_a\ I_c \tag{9}$$

Table 12: Informative System content of data d1 and d2

Permutation "n"	***Absolute individual weight of data d_1 and d_2 in the pair***		***Absolut informative weight "I_p"***	***Absolut informative content "I_c"***
I	d_1	d_1	$d_1 + d_1$	$\log_a I_{p1}$
II	d_1	d_2	$d_1 + d_2$	$\log_a I_{p2}$
III	d_2	d_1	$d_2 + d_1$	$\log_a I_{p3}$
IV	d_2	d_2	$d_2 + d_2$	$\log_a I_{p4}$
	Relative System weight of d_1 **d_{p1}**	**Relative System weight of d_2** **d_{p2}**	**System weight** D	**System content** **A**

3.7. Informative Quantity System, Calculation

To obtain the informative quantity system, there are two ways:

- The informative quantity in partials series of the system is represented as "I"", and it is the result of the addition of all partial informative quantities. This result represents the minimum informative quantity that can be obtained out of the system.

$$I'' = \sum_{i=1}^{m} P \cdot I_t \tag{10}$$

- The informative quantity in partials content of the system is represented as "I'", and it is the result of the multiplication of the probability "P'" by logarithm base "a" of the total informative partial content of the entire system. This result represents the value of informative quantity that can be obtained out of the system.

$$I' = P' \cdot A' \tag{11}$$

- The informative quantity of the system will be represented with the letter "I" and it is the result of the multiplication of the probability "P" of the system, by logarithm base "a" of the total informative content of the entire system. This result represents the maximum informative quantity that can be obtained out of the system.

$$I = P \cdot A \tag{12}$$

Table 13: System informative quantity of data d1 y d2

Permutation "n"	***Absolute individual weight of data d_1 and d_2 in the pair***		***Absolute informative weight "I_p"***	***Absolute informative content "I_c"***	***Absolute informative quantity "I_t"***
I	d_1	d_1	$d_1 + d_1$	$\log_a I_{p1}$	$p \cdot \log_a I_{c1}$
II	d_1	d_2	$d_1 + d_2$	$\log_a I_{p2}$	$p \cdot \log_a I_{c2}$
III	d_2	d_1	$d_2 + d_1$	$\log_a I_{p3}$	$p \cdot \log_a I_{c3}$
IV	d_2	d_2	$d_2 + d_2$	$\log_a I_{p4}$	$p \cdot \log_a I_{c4}$
	Relative System weight of d_1 d_{p1}	**Relative System weight of d_2 d_{p2}**	**System weight** D	**System content by partial series A'**	**Quantity by partial series I''**
				System content A	**Quantity by partial content I'**
					System quantity I

3.8. *Permutation Matrix Types That Can Be Obtained In A System*

One thing to keep in mind is that three types of permutation matrices can be found (two of which belong to permutation series and one to a complex series) for system calculations for "n" data:

- Permutation matrices with absolute data values.
- Permutation matrices with absolute and relative data values.
- Permutation matrices with relative data values.

3.9. *Permutation Matrices To Find D, A, I For A System With Absolute Data*

For a system with "n" data, D, A, and I can be calculated by taking the absolute values of the data, that is, if it is expected a system of "n" data, it always receives the "n" data without considering the permutations where there are any missing data. Then, the number of series of permutations that can be found for this calculation, would be given by "n · (n - 1)!" permutations, up to (n - 1) = 2.

For the following table, it is considered that the system receives three data (d_1, d_2 and d_3), and so, its permutations would be:

$$n \cdot (n - 1)! \qquad (13)$$

$$3 \cdot (2) = 6 \text{ permutations}$$

3.10. *Permutation Matrices To Find D, A, I For A System With Absolute And Relative Data*

For a system with "n" data, D, A, and I can be calculated by taking the relative values of the data. If a system of "n" data is expected, it always receives the "n" data, but considering the possibility of combinations where some of the data does not appear. For this case, the calculation would be given by "$n \cdot (n - 1)! + n^n$" permutations.

On the following table, a three data system is considered (d_1, d_2 and d_3), thus its permutations would be:

$$n \cdot (n-1)! + n^n \tag{14}$$

$$3 \cdot (2) + 3^3$$

$$6 + 21 = 27 \text{ permutations}$$

3.11. Permutation Matrices To Find D, A, I For A System With Relative Data

It is possible to work with data that certainty is unknown. For this case, all the data that will be worked with, will be relative because it does not belong to a formal informative system. It could be said that its starting point are assumptions from which some knowledge wants to be obtained. In this case is given a particularity, as it is not part of a certain and formal system, the system weight is not the unit, because of the uncertainty of the data that will be worked with. This concept is better explained on the following table:

Table 15: Matrix example with the relative values of d1, d2 y d3.

Permutation "n"	***Relative individual weight of data d_n in the system***		
	d_1	***d_2***	***d_3***
I	0,33	0,08	0,1
II	0	0,7	0
III	0	0,21	0,17
IV	0,62	0	0
V	0,02	0	0,5
VI	0	0,04	0

3.12. Informative Weight, Informative Content And Informative Quantity, For A Complex System

Certainly, the work with simple systems is very limited because of the small amount of information received. However, they are the basis for the calculations of complex systems. For a more down to earth job, a more precise solution will be the use of complex systems calculations. An example is shown in the following table:

Table 16: matrix of complex systems

"n" Series	**Complex Matrix System**				
	1	***2***	***3***	***4***	***5***
I	John	is	tall	-	-
II	Beer	is	always	a	liquid
III	Lisa	and	Anna	are	friends
IV	He	is	talking	with	others
V	Our	voices	are	listened	-

It is observed that as in the matrix with relative values, there is no need to make permutations, because these will come given by series of a different value each. These values can be obtained in two ways: using simple systems or adapted systems.

- **Through simple systems** → values that fit any system can be obtained, that is, it can be calculated the value for each word according to its position in the system, and also according to the probability to appear that each letter has, compared with other letters of the alphabet, which would be very complicated, tedious and for little use.

For the following example, it has been taken into account that the information once analyzed, can be transformed again into data, therefore its informative quantity will become the informative weight to be analyzed by the new system. For this example, take the minimum amount “I” will always be taken.

Then every word in the following table would represent a simple system that will have to be analyzed by performing permutations to obtain the absolute values of the system for I_c, I_p, I_t, C, A', A and finally I', I'' and I to obtain new knowledge. As follows:

Table 17. example of a complex system matrix based on simple systems

John = sist.1	is = sist.2	tall = sist.3	-	-
Beer = sist.4	is = sist.2	always = sist.5	a = sist.6	liquid = sist.7
Lisa = sist.8	and = sist.9	Anna = sist.10	are = sist.11	friends = sist.12
He= sist.13	is = siste.2	talking = sist.14	with = sist.15	others = sist.16
Our = sist.17	voices = sist.18	are = sist.11	listened = sist.19	-

- **Through customized systems** → a more practical solution is given, and adjusted to any need that a complex informative system analysis might have. For this case, if the information is known, it could be possible to assign values to data for the analysis.

The values here used would not be random, as they will come from a previous analysis that highlights the needs. In this example it can highlight the grammatical results from the sentences, that is, it can give a value to each grammatical figure, as follows:

Table 18: weights for a complex system

Subject = 5 units	Article = 2 units
Verb = 10 units	Preposition /Adverb = 2 units
Predicate = 7 units	Noun = 3 units

It could drill down to obtain the best fit to the needs of each analysis. In this case, if one word in the sentence does the function of subject, predicate, and at the same time is noun, then the values are divided by the number of words that share the same feature and are added to the values of the proper figure, for example: “John” is a subject but also a noun, then the values will be 5 + 1,5 (because there are two nouns in the sentence), “is” does the function of the verb, and it has a value of 10, and “tall” is the predicate and a noun and thus its values will be 7 + 1,5.

Table 19: example of a complex system matrix based on a particular analysis

Series "n"	Matrix for complex system				
	1	***2***	***3***	***4***	***5***
I	John = 5 + 1.5	is = 10	tall = 7 + 1.5		
II	Beer = 5 + 1.5	is = 10	always = 2	a = 2 + 3.5	liquid = 3.5 + 1.5
III	Lisa = 1.67 + 1	and = 1.67	Anna = 1.67 + 1	are = 10	friends = 7 + 1
IV	He = 5 + 1.5	is = 5	talking = 5	with = 3.5 + 2	others = 3.5 + 1.5
V	Our = 2.5	voices = 2.5 + 3	are =10	listened = 7	

3.13. Complex System Matrix To Obtain D, A, I For A Complex Data System

In general the analysis for a complex system is the same as for simple systems with the following particularities:

- The informative weight is obtained by adding the values of the series to obtain the weight of the series and then adding those values to obtain the minimum system weight.
- The informative content is obtained proceeding with the same operation for simple systems, increasing it to a concept of informative content by partial series that would come out from the addition of partial informative content. The total informative content would be the multiplication of the logarithm by the total value of the informative system weight obtained
- The informative quantity is obtained with the same operation for simple systems, adding the concept of informative quantity of the content by partial series that would come out as a result from the multiplication of the probability of the informative content by the logarithm of the value of the particular informative content. In the case of complex systems, three concepts have to be pointed out: informative quantity by partial series, informative quantity by informative content and the informative quantity of the system.

It is necessary to keep in mind that although these matrices seem similar to those formed with relative data, in this case, they are not formed by permutations, but instead from received data series.

Table 20: Summary of all concepts for a complex system

Series "n"	Absolute weight for each element in the series	Absolute informative weight "I_p"	Absolute informative content "I_c"	Absolute informative quantity "I_t"
I	Serie 1	$I_{p1} = \sum d_I$	$I_{c1} = \log_a I_{p1}$	$I_{t1} = p \cdot \log_a I_{c1}$
II	Serie 2	$I_{p2} = \sum d_{II}$	$I_{c2} = \log_a I_{p2}$	$I_{t2} = p \cdot \log_a I_{c2}$
III	Serie 3	$I_{p3} = \sum d_{III}$	$I_{c3} = \log_a I_{p3}$	$I_{t3} = p \cdot \log_a I_{c3}$
IV	Serie 4	$I_{p4} = \sum d_{IV}$	$I_{c4} = \log_a I_{p4}$	$I_{t4} = p \cdot \log_a I_{c4}$
...	...	...	...	...
V	Serie n	$I_{pn} = \sum d_n$	$I_{cn} = \log_a I_{pn}$	$I_{tn} = p \cdot \log_a I_{cn}$
	Relative system weight of each element	**System weight** $\mathbf{D = \sum_{i=1}^{m} I_{pi}}$	**Content by partial series** $\mathbf{A' = \sum_{i=1}^{m} I_{ci}}$	**Quantity by partial series** $\mathbf{I'' = \sum_{i=1}^{m} I_{ti}}$
			System content $\mathbf{A = \log_a D}$	**Quantity by partial content** $\mathbf{I' = P' \cdot \log_a A'}$
				System quantity $\mathbf{I = P \cdot \log_a A}$

4. The Knowledge

To speak about knowledge is to make reference to a system much more complex than that of information, through which a person acquires an added value for his knowledge, it allows a person to acquire two elements of the human reasoning which are: experience, as a set of skills that are acquired through the process to gain knowledge, and wisdom as automatic knowledge (Miranda, 2008).

Knowledge is the product of two systems: biological and social. It can be studied from many perspectives based on multiple empirical sciences. One of these perspectives is given through scientific knowledge. However, in spite of many attempts there is not a specific theory of knowledge, only philosophical preconceptions, drawn from cognitive processes that explain what knowledge could be, but that do not quantify it (Quintanilla, 2012).

Two dictionary definitions, are:

- Knowledge comprises of: "Understanding, intelligence, natural reason and each of the sensory faculties of man in so far as they are active" (Diccionario, 1986).
- "It is the conscious perception of the meaning and significance of the information" (Miranda, 2008).

From these concepts one can take out very few things that indicate a path for the quantification of knowledge, just ideas that put together various subjects, such as science, sociology, and psychology.

An important factor to consider in the formation of knowledge is experience, which is directly responsible for the growth of it in any subject. The great philosophers affirm that the addition of knowledge on the same topic, generates the experience. That is, tacit knowledge, explicit and implicit give true value to global knowledge. And experience is the factor through which one can pass from one knowledge to another, acting directly on the empowerment of information, thereby generating an exponential growth of knowledge.

For good transmission of knowledge, in a technological way, ICT are the most useful tools, especially to transmit explicit knowledge, but also there are some others that help in the transmission of tacit knowledge and implicit, such as video conferencing, webinars, and others (Contreras, 2010).

Finally, based on these previous explanations, and putting into relation all the concepts already seen about data and information, taking a more technological approach, one can determine an approximation to a concept that describes in a better way what knowledge is:

Knowledge is an entity that results from the interaction of cognitive processes, skills and self-experiences, acquired by transmission and shared to give a bigger value, that are able to transform any information or knowledge shared through a particular and global vision of it, and at the same time enhancing its value.

4.1. Knowledge Quantification

It is difficult to theoretically quantify the amount of knowledge of a person, because this would only serve exclusively for that person. But generalizations and deductions from common factors can give a better idea of quantification. Actually the idea of achieving the quantification of knowledge comes from the economy studies (Miranda, 2008). This is because the current growth of a country is measured by the technological resources, as well as information, knowledge, apprenticeship, and more recently the use of networks and collaboration between organizations as elements pointing directly to an improvement in the production of a company (Argüelles & Benavides, 2010). This capitalist vision of the knowledge society, puts the knowledge as a determinant factor in economic growth, and thus it is often called "knowledge economy".

The quantifying of knowledge is an attainable concept if studied in general terms because in particular form, it presents itself as a non-measurable item. This is because of the various structured cognitive processes through which knowledge is developed in each brain. Evidence shows a need to qualify and quantify the knowledge of a person. As a matter of fact, an approximation is made indirectly in job interviews, where the results give off a value of knowledge and skills that a person has (from a more psychological rather than technical point of view). For example, psychometric tests used to interview staff, using a range of situations to which it has been assigned a score that helps to quantify some skill or experience (Contreras, 2010). These ranges are made largely to follow a market criteria and the needs of the company.

4.2. Calculating Knowledge Quantification

It would be possible to obtain a value for knowledge from the evaluation of data and information or from retrofitted knowledge (that works as new information) and that generates knowledge value. Therefore, within the concept of resolution of complex matrices of information, can be found a development of a concept for the quantification of knowledge.

In order to obtain a quantification of knowledge, an important key stands is the concept of "information processing", which, associated with different concepts of knowledge, provides learning, and through it, experience.

Going back to the data system example, there was the data "girl", and "15", and it was complemented with the datum "years old". If on top of it is added another data system that consists of "girl", " the party" and "enjoy", these two systems would provide us with the following information: on one hand, "girl is 15 years old" and in the other hand,: "girl enjoy the parties." From the process and analysis of these two systems, a knowledge can be taken out of "girls who are 15 years old enjoy the parties".

Of course, knowledge is not about taking parts of sentences and summing it up in another sentence, but this simple example shows briefly how does the extraction of knowledge take place through an easy process of information. All this goes to show that knowledge can represent concepts generated by the processing of an informative system that are received, understood and become facts. Furthermore, the more particular that knowledge becomes, the more accurate, and more complex it will be, due to the processing of concrete data as well as specific information that generate specific knowledge.

The following considerations should be taken into account:

- The first information to be drawn from the data, gives a first knowledge, also known as unitary knowledge and represented by C' (note 3). This is derived from the contextualization and assessment of the information that is contained within the first primary data, that is:

$$C' = I \tag{15}$$

- The knowledge obtained from this information, does not remain static, but serves to generate a global knowledge or "C′′", which is generated through a cognitive (information) and experimental (experience) process that adds a new value of knowledge. C'' is the addition of tacit knowledge and explicit knowledge, that is:

$$C'' = I^e \tag{16}$$

- The value of tacit knowledge "C_t", will be the result of the division of the years of specific experience "e" in a subject, divided by the specific formative years "f" in a subject, that is:

$$C_t = \frac{e}{f} \tag{17}$$

- The value to explicit knowledge "C_e", will result from raising the specific formative years "f", to the specific years of experience "e" in a subject, that is:

$$C_e = f^e \tag{18}$$

Then following what the concept of knowledge says and its subsequent analysis, the following equation can be deduced:

C = information + tacit knowledge + explicit knowledge

The result will be:

$$C = I + \frac{e}{f} + f^e$$

Substituting for the values of information "I", would be:

$$C = P \cdot \log_a A + \frac{e}{f} + f^e \tag{19}$$

Conclusions

The usefulness of information generated from patterns that add bigger value to organizations, can be seen as a competitive factor. This is because the results obtained from all calculations done through the series of information received, generate knowledge useful for any organization.

For example, it can be of a great help in the election of the best person to lead and manage a project. Taking the information from the candidates, and making an individual analysis of each one using informative matrices, it can give a comparison of values of skills, abilities, knowledge and experience, that will give a clue as to whom will perform best at the job.

An example could be to search the best result of knowledge, abilities and experience of professional inside the enterprises, in order to select the better director or manager for the next project development. As seen in the next example:

Table 21: Knowledge – Abilities – Experience Valuation (values min 0 to max 10)

	Partial Experience by field				**(max 40)**		
	Programmer	***Analist***	***Project Architect***	***Project Manager***	***D***	***A/A'***	***I/I'/I''***
Java	7	3	0	1	*11*	*1,041*	*0,018*
Html	9	1	1	2	*13*	*1,114*	*0,052*
Uml	5	2	0	0	*7*	*0,845*	*-0,062**
Xml	2	2	0	0	*4*	*0,602*	*-0,133**
BD	4	1	0	0	*5*	*0,699*	*-0,109**
CSS	6	2	0	1	*9*	*0,954*	*-0,020**
Content management	4	3	0	3	*10*	*1*	*0*
Partial Experience (max. 70)	**37**	**14**	**1**	**7**	***D=59***	***A'=6,255***	***I''=0,394***
						A=1,771	***I'=4,980***
							I=0,440

*Every negative value must be change to positive, because any information give an amount of knowledge

Table 22: Percentage of answer obtained

	Max	**Min**	**Answer %**
D	280,00	59	21,071
A	2,447	1,771	72,374
A'	11,214	6,255	55,778
I	0,964	0,440	45,643
I'	11,772	4,980	42,304
I''	2,295	0,394	17,168

When these results are compared with other evaluation of other consultants, it is possible to obtain a comparative answer for any search for that have used this method.

From this first analysis, one can further investigate to find a more useful and better approach to quantification of information and knowledge, in any system.

Notes

Note 1. The year when Shannon y Weaver published the "Information Theory".
Nota 2. In this case a system is conceived as an association of data, information or knowledge, always associated with those of the same context.
Note 3. C' → first knowledge of the first information obtained from the primary data.

References

Argüelles V, M. & Benavides G, C. (2008). Conocimiento y Crecimiento Económico: Una Estrategia Para Los Países En Vías De Desarrollo. *Revista de Economía Mundial*. Nº 18. ISSN: 1576-0162. pp. 65 – 77. Retrieved

February 09, 2012, from http://rabida.uhu.es/dspace/bitstream/handle/10272/553/b1513600.pdf?sequence=1

Arredondo, T. (2011). *I*ntroducción A La Teoría De La Información. *Universidad Técnica Federico Santa María*. Chile. pp. 5 – 29. Retrieved October 10, 2012, from http://profesores.elo.utfsm.cl/~tarredondo/info/soft-comp/Introduccion%20a%20la%20Teoria%20de%20la%20Informacion.pdf

Contreras, E. (2010). Gestión del Conocimiento: Del tácito al explícito 20 años después. *Trend Management.* Special Edition. Mayo 2010. pp. 2 – 3. Retrieved February 09, 2012, from http://www.dii.uchile.cl/wp-content/uploads/2011/06/UCH-Contreras.pdf

De Pablo L, I. (1989), *El Reto Informático: La gestión de la Información en la Empresa*. Editorial Pirámide. España. ISBN: 84-368-0464-3.

Diccionario de la Lengua Española, (1986), *Diccionario de la Lengua Española*, Distribuciones Mateos. Editorial Antalbe, S.A. Madrid. pp. 144, 181, 182, 248, 396

Hercy, (2011), *Teorías De La Información Y Comunicación*. Taking form the Book Manual Técnico de Alfabetización Digital 2011. Editorial del Libro LANIA. México. pp. 1 Retrieved October 31, 2012, from http://www.slideshare.net/hercyb/teorias-de-la-informacin-y-comunicacin

López R, (1998). Crítica A La Teoría De La Información. *Chile University.* Retrieved October 31, 2012, from http://www.periodismo.uchile.cl/cursos/psicologia/criticainformacion.pdf.

Minguet, J M., & Read, T., (2008). *Informática Fundamental*. Editorial Universitaria Ramón Areces S.A. Madrid. Primera reimpresión 2008. pp. 52 - 61, 66 – 69.

Miranda, E. (2008). Epistemología Y Gnoseología. *Teoría del Conocimiento*. Retrieved Decembre 24, 2012, from http://www.slideshare.net/guest975e56/epistemologa-y-gnoseologia-ok

Profesorenlinea, (2011). Logaritmos. *Web of Profesor en Línea.* Retrieved October 12, 2012, from http://www.10.cl/matematica/logaritmo.html

Quintanilla, M A. (2012). Teoría Del Conocimiento. *Fundación Gustavo Bueno. Proyecto de Filosofía en Español.* Retrieved Decembre 24, 2012, from http://www.filosofia.org/enc/dfc/conocimi.htm

Singh, J., (1982). *Teoría De La Información, Del Lenguaje Y De La Cibernética*. Editorial Alianza Universidad. Madrid. pp. 24 – 33.

Textos Científicos, (2008), Teoría de la Información. *Web Textos Científicos.* Retrieved October 25, 2012, from http://www.textoscientificos.com/informacion/teoria

Torredebabel. (2012). Definición De Ente. *Web Definición.de*. Retrieved November 02, 2012, from http://www.e-torredebabel.com/Psicologia/Vocabulario/Ontologia.htm

Torredebabel. (2012a). Ontología. *Web Definición.de*. Retrieved November 02, 2012, from http://www.e-torredebabel.com/Psicologia/Vocabulario/Ontologia.htm

Juan Carlos Piedra Calderón received his Advance Study Diploma (DEA) part of doctorate program and received his Master degree in Business Process Management and Master in Information and Knowledge Management from Universidad Pontificia de Salamanca (UPSAM – Spain). Actually is full time working on his doctoral thesis for PhD degree in Information and Knowledge Society (UPSAM). He has been working as a Knowledge Management Expert in several projects.

Rubén González Crespo received a PhD in Software Engineering from Universidad Pontificia de Salamanca (UPSAM – Spain). He is working in the department of Languages and Computer Systems. Director of the Research Group of Society and Information Technologies. Director adjunct of the engineer area in Universidad de La Rioja. Technical Adviser of the Panama Government.

J. Javier Rainer received a PhD in Robotics and Automation from Universidad Politécnica de Madrid, UPM (Spain). He is Director of Research and Director of Engineering Area at Bureau Veritas Business School, and researcher at the Intelligent Control Group of the UPM. He received the best paper Award in IARIA Cognitive 2010.