



HAL
open science

Integrating Document Usage with Document Index in Competitive Intelligence Process

Lukman Akanbi, Emmanuel Rotimi Adagunodo, Amos David

► **To cite this version:**

Lukman Akanbi, Emmanuel Rotimi Adagunodo, Amos David. Integrating Document Usage with Document Index in Competitive Intelligence Process. *International Journal of Computer Applications*, 2015, 132, pp.975 - 8887. hal-02306070

HAL Id: hal-02306070

<https://hal.science/hal-02306070v1>

Submitted on 4 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Integrating Document Usage with Document Index in Competitive Intelligence Process

Lukman A. Akanbi
Obafemi Awolowo University
Ile-Ife, Nigeria

Emmanuel R. Adagunodo
Obafemi Awolowo University
Ile-Ife, Nigeria

Amos David
Université de Lorraine
Nancy, France

ABSTRACT

The amount of information in term of documents, available to users as a result of information retrieval process for the purpose of resolution of decision problems is a major factor that determines whether economically viable decisions would be made or not. Various works in the literature had addressed the challenges of representing the documents with key terms (generated from the document) as well as the variations in the meaning of each key terms. In this work, a document representation scheme that is based on the key terms generated from the documents and their usage was developed. To realize this document representation scheme, a computational model for capturing document usage was designed with the use of attribute value pair technique of document annotation. The document usage model designed was applied in the development of a Competitive Intelligence based Document Usage Creation and Exploration system that is currently under development. A preliminary evaluation of the document usage model based on cosine similarity function between user query and documents set was carried out. The result obtained shows that representing documents in terms of their usage can enhance the quality of information search results as documents that would hitherto be considered not relevant to user query are found to be ranked very relevant based on previous usages.

General Terms

Content Analysis and Indexing Methods.

Keywords

Document usage, document representation, document index, usage modelling, decision problem, attribute value pair.

1. INTRODUCTION

In information retrieval for the purpose of solving a decision problem, the amount of information in terms of documents as well as detailed information about their previous usage that are available to users is a major factor that influences result of the problem resolution. Okunoye et al. in [16] identified that discovery of relevant information for solving decision problems is pivotal to making right decision and that the information has to be sought, collected, processed with a view to elicit knowledge relevant to the problem as well as representing both the collected information and elicited knowledge in a form that will facilitate information reuse.

The modern information retrieval (IR) models (i.e. Boolean, Vector and Probabilistic) employ terms in the documents to produce the index terms to represent the documents in the document space [1], [13]. This documents representation technique based on the terms available in the document is employed by most systems that provide search facility.

Therefore such systems always return relevant documents based on the terms found in the query. However, experience has shown that there are relevant documents whose terms are orthogonal to the terms available in the query submitted by the user. The reason for this situation is express in the hypothesis that every search activity for information is associated with a decision problem (DP). This problem may be expressed explicitly or is implicit to the information seeker. It is however believed that, keeping track of the kind of DP that documents have been used to resolve would ease the process of resolving similar problems in the future.

There are various works in the literatures that had addressed the challenge of representing the documents with key terms as well as the variations in the meaning of each key term (for example, [7] [9]). However, the concept of document usage has not been given attention. A document usage model based on attribute-value-pair (AVP) technique of document annotation proposed in [18] is therefore presented in this work.

The rest of the paper is organized as follows: Competitive intelligence is described in section two; document annotation and its relevant to competitive intelligence is described in section three. In section four, the document usage model and the architecture of the CI-DUCE system are presented. In section five the evaluation of the model and the preliminary results are discussed and section six concludes the work.

2. COMPETITIVE INTELLIGENCE

As a special case of application of IR process, CI is a systematic process of information gathering, processing, analysis and decomposition, which is conducted within the context of the external environment of an organization's activities. This process is with the major goal of supplying the right information, at the right moment, and in the correct structure, to the right person, in order to support the best decision possible [22]. The Strategic and Competitive Intelligence Professionals (SCIP) in [21] defines CI as the legal and ethical collection and analysis of information regarding the capabilities, vulnerabilities, and intentions of business competitors. The information legally and ethically collected is to be used by the decision makers in an organization to support them in arriving at the best possible decision. Dishman and Calof in [8] described CI as a process involving the gathering, analyzing and communicating of environmental information to assist in strategic decision-making, and as such it is the fundamental basis of the strategic decision-making process. Camelo et al. [4] pointed out that the CI process should aim to provide information on technologies and general commercial tendencies in order to facilitate the decision making process, helping to achieve the strategic goal of the company. According to Odumuyiwa [14],

CI is linked to various similar concepts such as economic intelligence, business intelligence and knowledge management and it is a process that embodies decision making. CI as an information process is made up of the following phases [6]:

- i) identification of decision problem
- ii) transformation of decision problem to information search problem
- iii) identification of relevant sources of information
- iv) collection of relevant information
- v) analysis of collected information to extract indicators for decision making
- vi) interpretation of indicators
- vii) decision making

According to [4], in CI process, the time needed for execution is of crucial importance, as in case of short or medium term work, decision making depends on vast quantity of information and this precise lack of time available to process an unpredictable volume of information, with selection of only that which is useful for decision making can reasonably be seen as one of the main challenges encountered in CI.

The focus of the present work is on how to improve items (iii) and (iv) of the CI process listed above in order to enhance the process of decision making in the shortest time possible thereby addressing the concern raised in [4]. These two items is basically an IR process.

3. DOCUMENT ANNOTATION

According to [15], Annotation can be seen as simply information about the document, assigned by a process or human, after the original creation of the document, such as information provided by a named-entity tagger or an automatic layout analysis tool. This information may be less certain than structural markup, as the process or person creating the annotations may not know the author's original intent. Annotation is also described as a tool or technique used to describe a set of information. These set of information could range from title of a document to a section of a document and/or the whole chapter in a book. As identified in [20] annotation tools are becoming very important for document interpretation. Its importance as a tool in information interpretation can be seen with its popularity in text processors like Microsoft word and Adobe Acrobat that integrate features that enable users to annotate document. More than the normal use of annotation for document interpretation, annotation tool can be designed to assist in information research [20].

The traditional domain of document annotation covers the annotation of arbitrary textual documents, or parts of them, which can be manual (i.e. performed by one or more people), semi-automatic (based on automatic suggestions), or fully automatic. Manual annotation tools allow users to add annotations to web pages or other resources, and share these with others. Automatic tools can perform similar annotations (such as named-entity recognition) without manual intervention [19].

The meaning of annotation can vary significantly when considered in a specialized context. For example, in historical and religious scholarship, annotations can provide contextual detail about primary sources, or describe interpretations or differing perspectives. Whereas in the legal and governmental

domains, annotations often provide references to relevant instances of an abstract concept in practice, such as court decisions associated with particular statutes, but the term can also refer to quite lengthy documents about specific cases [11].

From the fore going, one thing stand out about the concept of annotation, i.e. extracting or adding information to an information object of interest. This work is therefore a specific application of the annotation concept to add usage to an information object and made such usage part of document representation. This is believed to enhance the process of DP resolution in CI environment. Some of the examples of uses of annotation in CI process are discussed in the next subsection.

3.1 Some Applications of Document Annotation

In [20], Robert and David presented annotation model that will assist in information research for decision making. The Annotation Model for Information Exchange (AMIE) was reported to have been conceived with the objective of information sharing and reuse. The use of annotation process as an indexing approach that allows the users to identify and regroup documents or their sub-elements under a particular use context was employed in [12]. [17] is a work on knowledge capitalization through annotation among economic intelligence actors in a collaborative environment that uses annotation concept to develop a model for eliciting knowledge among CI actors.

Another interesting work that made use of annotation concept is the knowledge and information management (KIM) system presented in [10]. The work introduces the authors' vision for a holistic architecture for semantic annotation, indexing, and retrieval of documents with regard to extensive semantic repositories. Another similar work is [5], which proposed a model for the exploitation of ontology-based knowledge bases to improve search over large document repositories. Their approach includes an ontology-based scheme for the semiautomatic annotation of documents and a retrieval system. The retrieval model is based on an adaptation of the classic vector-space model, including an annotation weighting algorithm, and a ranking algorithm.

4. THE DOCUMENT USAGE MODEL

The document usage is modelled as a quadruple that comprises of the following components. Users represented as U, document object represented as D, decision problem represented as P and the environment represented as E.

$$DU = \{U, P, D, E\} \quad 1$$

Where

DU = Document usage

U = Users attributes.

P = Decision problem.

D = Document

E = Environment.

R = User's rate of document relevance to DP.

The U component of the equation 1 has the attributes that includes id, status, exp and estab. The id is a unique identity allocated to the user during the course of registering to use the system. The status is the current position of the user in the establishment or organisation. The exp is the attribute of the

user that captures the experience of the user. This experience is considered in respect of the number of years the user had spent, participating in the core business of the establishment. Note that this number of year is not necessarily the same as the number of year spent in the establishment. For example, somebody moving from one company to another company in the same line of business cannot be said to be novice in the business activity of the new company. Such a user already had cognate experience on the core business of the company and as such the year of experience in the previous company is also considered. The users' establishment information such as type of establishment (private or public) and name of the establishment are captured with the estab attribute of the user component of the document usage model. In other word, for clarity purpose the user attribute could be represented as a tuple as shown in equation 2.

$$U = \{id, status, exp, estab\} \quad 2$$

The decision problem represented as P in the DU equation is an attribute used to capture the description of the decision problem by the user. It comprises the stake's object, signal and hypothesis. This representation of the decision problem with the stake attribute is inspired by the work on explanation and modelling of decision problem presented in [3]. The stake object is used to capture the user's description of the decision problem at hand, while the stake signal and stake hypothesis allows the user to supply the level of information he has about the problem and what the establishment stands to gain or loss respectively. The P component of the document usage model can therefore be describe as a tuple as shown in equation 3.

$$P = \{Obj, Signal, Hyp\} \quad 3$$

where

Obj = stake object

Signal = stake signal

Hyp = stake hypothesis

The D attribute of the document usage model contains information about the document to which usage is to be created. It comprise of document identity (id), and three document information elements namely document title (doc_t), document abstract or summary (doc_abs) and the rest part of the document (d_body). Worthy of mention at this point is the fact that only text documents are considered in this work. The findings in this work will be extended to other form of documents. Equation 4 presents the document attribute of the usage model as a tuple.

$$D = \{id, doc_t, doc_abs, d_body\} \quad 4$$

The environment (E) component of DU equation describes the environment in respect of the source of the document. It takes either of the two values internal or external. If the establishment that is creating the usage for a document is the producer of that document, then the value of environment attribute of the document usage model is internal. For example, in the case of a head of a unit in a hospital that uses

medical records of patience in the hospital to examine the performance of a group of drug expended in the hospital over a period of time. The E attribute's value will be external if the source of the document is outside of the establishment.

The flow of activity in the process of usage creation is shown in figure 1. The document usage model presented in this paper is being implemented as part of a Competitive Intelligence base Document Usage Creation and Evaluation (CI-DUCE) system that is currently being developed. Therefore the flowchart in figure 1 is a part in the whole system design. From the figure, when the user log in to create usage for a particular document, the system determine whether the logged user is an information watcher or not, if the user is an information watcher (i.e. person charged with the responsibility of information search and system update in a CI system), then the user is prompted to select the user that has authorised the creation of usage for the selected document. The user then supplies appropriate values for the decision problem attribute and the environment attribute. Finally the document usage index is updated accordingly. If the user however is not Information watcher, then the system captures the logged user information and allows the user to go straight to supply the values for the decision problem and environment attributes.

4.1 Document Usage Model Evaluation

As stated in [13], the representation of a set of documents as vector in a common vector space is known as the vector space model (VSM) and is fundamental to a host of information retrieval operations ranging from scoring documents on a query, document classification and document clustering. To evaluate the document usage model presented in this work, the evaluation was based on the VSM of the classic IR model which has been adjudged to outperform the Boolean and Probabilistic models of IR [1].

Two document representations were considered for the evaluation. The first representation is based on the key terms in the document (i.e. document index generated based on the terms in the document). This representation is the basis of document index for the existing search engines. The second representation is based on the terms in the document and the usage index (i.e. term index + usage index). Each decision problem (i.e. user query terms) is then compare with the document collection to calculate the similarity between it and the documents in the collection. Figure 2 shows the formal description of two document representation schemes.

As demonstrated in [2], a document collection comprised of n documents which are indexed by m terms can be represented as an m x n term-by-document matrix A. The n (column) vectors representing the n documents forms the columns of the matrix. Thus, the matrix element a_{ij} is the weighted frequency at which term i occurs in document j. Using the VSM, the columns of A are interpreted as the document vectors, and the rows of A are considered the term vectors.

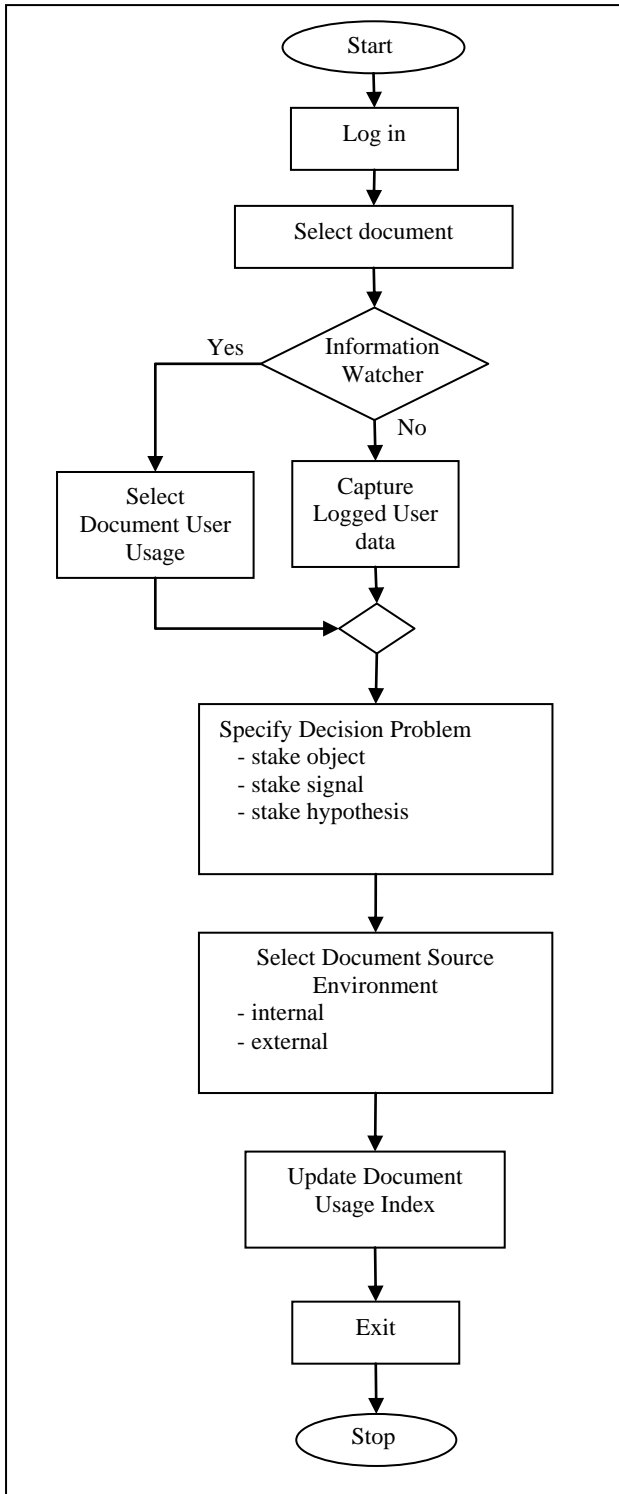


Figure 1: Flowchart of Document Usage Creation Module

Term based document index

A document is represented as a vector $\vec{d} \in \bar{D}$ where \bar{D} is the vector of all the documents in the document collection space. \vec{d} is a subset of \bar{D} .

$$\bar{D} = (d_1, d_2, d_3, d_4, \dots, d_n)$$

$d_i = \text{document } i \text{ in document collection } D.$

$$\vec{d} = (t_1, t_2, t_3, t_4, \dots, t_n)$$

$t_i \in \vec{d} = \text{terms in the document index}$

Term + usage based document index

Two components to represent document D.

a) $d = (t_1, t_2, t_3, t_4, \dots, t_n), t_i \in D$

b) decision problem stake object represented as vector \vec{P} containing terms k_i used to describe the problem.
 $\vec{P} = (k_1, k_2, k_3, k_4, \dots, k_n)$

New document representation dr is given by

$$dr = (d . P)$$

$$= (t_1, t_2, t_3, t_4, \dots, t_n . k_1, k_2, k_3, k_4, \dots, k_n).$$

Figure 2: Formal description of the Key term based and Usage based document Index

4.2 Data for Usage Model Evaluation

Data for evaluation of the document usage model was obtained from 20 postgraduate students of the Obafemi Awolowo University, Nigeria. The data was collected with the use of questionnaire and interview. The respondents were asked to describe their research work in less than 80 words. The descriptions were used to form the stake object Obj of the decision problem attribute P. Title of six documents that the respondents found very useful to their work were also collected. The documents titles and the DP description are shown in Figure 3 for one of the respondents. From the figure, d1 through d6 represent the documents and dp represents the DP. After the pre-processing of the documents to extract the key terms, the document collection represented with only the key terms in document is shown in figure 4 and the document collection represented based on document usage model (i.e. document representation based on terms in the document and the document usage) is shown in figure 5. These two document collection are transform to document vectors as shown in equations 7 and 8.

To calculate the similarity between the documents collection represented as term-by-document matrix in equations 7 and 8, the DP stake object represented as dp in figure 3 is also represented in form of a vector in the document collection space after pre-processing task to extract the key terms. The resulting vector is shown in equation 9. The similarity between the decision problem and the documents in the document collection is determined by calculating the cosine of angle between decision problem vector P (equation 9) and the document vectors (equations 7 and 8). d_j is defined as the jth document vector (i.e. the jth column of the term-by-document matrix A), then the cosines between the decision problem vector P and the $n = 6$ document vectors are defined by equation 10.

d1 = AMIE: An annotation model for information research

d2 = AMTEA: Tool for Creating and Exploiting Annotations in the Context of Economic Intelligence (Competitive Intelligence)

d3 = What Is a "Document"?

d4 = CI Spider: a tool for competitive intelligence on the Web

d5 = Dynamic Knowledge Capitalization through Annotation among Economic Intelligence Actors in a Collaborative Environment

d6 = Design and Development of a Model for Generating and Exploiting Annotation in the Context of Economic Intelligence

dp = Development of model and system to create usage descriptors for documents. Then writing of Doctorate thesis on the model and the system.

Figure 3: Sample Documents and DP for Analysis

The result of the cosine similarity analysis between the DP and each of the documents in the document vector space in both cases is shown in Table 1. The usual practice in the literature is to set a threshold at which document will be considered as relevance to the query and return ranked base on the similarity value. Note that the higher the similarity value, the closer the document and the query (DP in this case). From the result in Table 1, for the term-based index, documents d4 and d5 would not be part of the returned document for a search based on the terms in the DP because their cosine similarity is 0.00000 each.

However in the term+usage based index, documents d4 and d5 not only have non-zero value for the cosine similarity, their values are even greater than that of some documents that were considered similar (i.e. relevant) to the DP based on term-based index. The is due to the level of involvement of the document in the problem resolution process even though the terms in the document may not correlate to any term in the DP description. The result in Tables 1 is depicted graphically in Figure 6. It is clear from the figure that the similarity between documents and DP increased drastically from an average value of 0.10000 for term-based to 0.80000 for usage based index. Although, for this analysis, only the titles of the documents were considered, the result will also follow the same pattern when the full texts in the documents are considered.

d1 = AMIE: annotation model information research

d2 = AMTEA: Tool Creating Exploiting Annotations Context Economic Intelligence Competitive Intelligence

d3 = Document

d4 = CI Spider tool competitive intelligence Web

d5 = Dynamic Knowledge Capitalization Annotation Economic Intelligence Actors Collaborative Environment

d6 = Design Development Model Generating Exploiting Annotation Context Economic Intelligence

Figure 4: Documents Representation based on only

d1 = AMIE: annotation model information research
Development model system create usage descriptors documents. Doctorate thesis model system.

d2 = AMTEA: Tool Creating Exploiting Annotations Context Economic Intelligence Competitive Intelligence Development model system create usage descriptors documents. Doctorate thesis model system.

d3 = Document Development model system create usage descriptors documents. Doctorate thesis model system.

d4 = CI Spider tool competitive intelligence Web Development model system create usage descriptors documents. Doctorate thesis model system.

d5 = Dynamic Knowledge Capitalization Annotation Economic Intelligence Actors Collaborative Environment Development model system create usage descriptors documents. Doctorate thesis model system.

d6 = Design Development Model Generating Exploiting Annotation Context Economic Intelligence Development model system create usage descriptors documents. Doctorate thesis model system

Figure 5: Documents Representation based on terms in the document and usage

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

7

$$A_n = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad 8$$

$$P = (00000000011011100000000020021010)^T \quad 9$$

$$\begin{aligned} sim(d_j, q) &= \frac{\bar{d}_j \cdot \bar{q}}{|\bar{d}_j| \times |\bar{q}|} \quad 10 \\ &= \frac{\sum_{i=1}^t a_{ij} \times P_i}{\sqrt{\sum_{i=1}^t a_{ij}^2} \times \sqrt{\sum_{i=1}^t P_i^2}} \end{aligned}$$

Where j = 1 n (n= number of document in the document collection space).
t = the number of terms in the vector space.

Table 1: Result of Cosine Similarity Analysis between DP and Documents

Documents	Term-based Index	Term + Usage Based Index
d ₁	0.23094	0.86141
d ₂	0.07454	0.76714
d ₃	0.25820	0.97373
d ₄	0.00000	0.84515
d ₅	0.00000	0.79057
d ₆	0.25820	0.84853

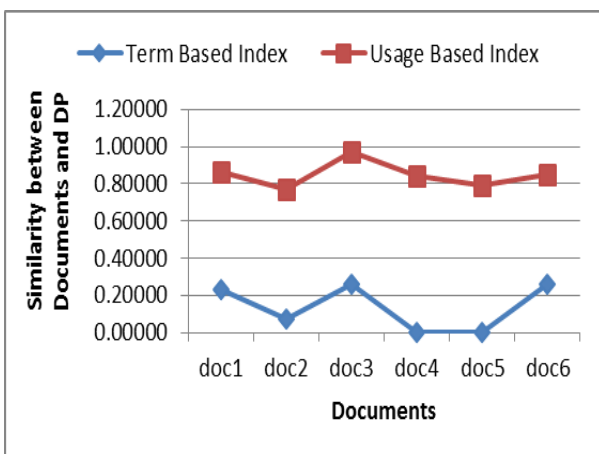


Figure 6: Similarity between DP and Documents

5. CONCLUSION

The IR process component of the CI process is very crucial to speedy and easy resolution of DPs facing an organisation. This work has sought to keep track of what documents have been used for and incorporate it into the document

representation scheme to enhance the quality of IR process at a later time. The document usage model presented is used to preserve the effort of the decision makers (users) in discovering relevant documents for the resolution of DPs. The result obtained shows that representing documents in terms of their usage can enhance the quality of information search results as documents that would hitherto be considered not relevant to user query are found to be ranked very relevant based on previous usages.

The idea presented in this work can be used to develop a Decision Support System (DSS) for various industries that need to resolve one problem or the other. For example, the concept can be used to develop DSS for various units in the hospitals. For example, medical doctors can use the system to keep track of the process of treating patients (from diagnosis to drug prescription). The system can later be used by another doctor that have similar case at hand, it could also be used by the administrators to examine how doctors are faring in the hospital. A pharmacy unit of the hospital could use the system to keep track of the kind of medical conditions each drug is used to address. This will enhance the quality of decisions to be made on each drug when there is need for such.

6. ACKNOWLEDGEMNET

This work was supported by the French Government Scholarship through the French Embassy in Nigeria. The Obafemi Awolowo University is hereby acknowledge for fully supporting this research effort.

7. REFERENCES

- [1] Baeza-Yates, R. and Ribeiro-Neto, B. (1999). Modern Information Retrieval ACM Press New York.
- [2] Berry, M. W. and Browne, M. (2005) Understanding Search Engines: Mathematical Modelling and Text Retrieval. Second Edition. *Society for Industrial and Applied Mathematics (SIAM)*. Philadelphia
- [3] Bouaka, N. (2004) Développement d'un modèle pour l'explicitation d'un problème décisionnel: un outil d'aide à la decisiodans un contexte d'intelligence économique. Doctorant Thesis, Université Nancy 2, Nancy. France.
- [4] Camelo, C., Trigo, M. R., Quoniam, L. and Cardoso, J. C. (2011) Competitive Intelligence and the Development of Corporate Universities. In *Competitive Intelligence and Decision Problems* (ed A. David), John Wiley & Sons, Inc., Hoboken, NJ, USA. Chapter 14. Pp 281-298.
- [5] Castells, P., Fernandez, M. and Vallet, D. (2007). An Adaptation of the Vector-Space Model for Ontology-Based Information Retrieval. *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, no. 2, Pp 261-272.
- [6] David A. (2009) Relevance Information in Economic Intelligence. In *Information Science* (ed F. Papy). ISTE Ltd London and John Willey & Son New York
- [7] Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K. and Harshman R. (1990): Indexing by Latent Semantic Analysis. *Journal of American Society for Information Science*. Vol 41. No 6. Pp 391-407.
- [8] Dishman, P. L. and Calof, J. L. (2008) Competitive intelligence: a multiphasic precedent to marketing strategy, *European Journal of Marketing*, Vol. 42, No 7/8. Pp 766 – 785.

- [9] Gupta, P. and Sharma, A. K. (2010): Context Based Indexing in Search Engines using Ontology. *International Journal of Computer Applications (0975-8887)*. Vol. 1 No. 14. Pp 49-52.
- [10] Kiryakov, A., Popov, B., Terziev, I., Manov, D. and Ognyanoff, D. (2004). Semantic annotation, indexing, and retrieval. Elsevier Journal of Web Semantics: Science, Services and Agents on the World Wide Web 2, Vol. 2, Iss. 1, Pp49 - 79.
- [11] MacMullen, W. J. (2005). Annotation as Process, Thing, and Knowledge: Multi-domain studies of structured dataannotation. *SILS Technical Report TR-2005-02*. University of North Carolina, School of Information and Library Science. Technical Report series, Chapel Hill.
- [12] Maghreb, H. and David A. (2008) Open System for Indexing and Retrieving Multimedia Information, *36th Canadian Association for Information Science Conference - CAIS 2008*, University of British Columbia, Vancouver.
- [13] Manning, C. D., Raghavan, P. and Schütze, H. (2009) An Introduction to information Retrieval, *Cambridge University Press, Cambridge, England*. Online Edition.
- [14] Odumuyiwa, V. (2011) Collaborative Information Seeking in the Competitive Intelligence Process. In *Competitive Intelligence and Decision Problems* (ed A. David), John Wiley & Sons, Inc., Hoboken, NJ, USA. Chapter 4. Pp 69-91.
- [15] Ogilvie, P. (2010) Retrieval using Document Structure and Annotations. Ph.D. Thesis Carnegie Mellon University, Pittsburgh. U.S.A.
- [16] Okunoye, O. B., David, A., and Uwadia, C. (2010a). Amtea: Tool for creating and exploring annotations in the context of economic intelligence (Competitive Intelligence). In *11th IEEE International Conference on Information Reuse and Integration (IRI 2010)*, Pp 249- 252, Las Vegas, United States.
- [17] Okunoye, O., Oladejo, F. and Odumuyiwa, V. (2010b) Dynamic Knowledge Capitalization through Annotation among Economic Intelligence Actors in a Collaborative Environment. *Colloque International Veille Stratégique Scientifique et Technologique - VSST 2010*, October 2010, Toulouse, France. Pp. 1-17.
- [18] Okunoye, O. B. and Uwadia, C. O. (2011) Design and Development of a Model for Generating and Exploiting Annotation in the Context of Economic Intelligence. In *Competitive Intelligence and Decision Problems* (ed A. David), John Wiley & Sons, Inc., Hoboken, NJ, USA. Chapter 9. Pp 183-197.
- [19] [19] Oren, E. Moller, K., Scerri, S., Handschuh, S. and Sintek M. (2006). What are Semantic Annotations? *Technical Report DERI Galwat*. Available at: <http://www.siegfriedhandschuh.net/pub/2006/whatissemanticannot2006.pdf> (downloaded 06102012)
- [20] Robert, C. A. and David, A. (2006) AMIE: An Annotation Model for Information Research. In L. Barolli, B. A. Abderazek, T. Grill, T. M. Nguyen, and D. Tjondronegore (eds.): *Frontiers in Mobile and Web Computing*, Austrian Computer Society. Vol. 216, ISBN 3-85403-216-1, pp 129-137.
- [21] SCIP, (2012) About SCIP, *Strategic and Competitive Intelligence Professional*. Available at <http://www.scip.org/content.cm?itemnumber=2214&navItemNumber=492>, Retrieved 15/10/2012
- [22] Trigo, M. R., Gouveia, L. B., Quoniam, L., and Riccio, E. L. (2007) Using competitive intelligence as a strategic tool in a Higher Education context, *8th European Conference on Knowledge Management (ECKM)*. *Consorci Escola Industrial de Barcelona (CEIB)*, Barcelona, Spain. 6-70 September 2007.