

A QOS-Aware Semantic Web Services Selection Model for Tourism

Ekkachai Naenudorn¹, Suphakit Niwattanakul², Nisachol Chamnongsri³

¹School of Information Technology, Suranaree University of Technology,
NakhonRatchasima, 30000, Thailand

²School of Information Technology, Suranaree University of Technology,
NakhonRatchasima, 30000, Thailand

³School of Information Technology, Suranaree University of Technology,
NakhonRatchasima, 30000, Thailand

Abstract

Current searches web services can be found at Public UDDI but the results of the search may not be complete. Because used keyword matching for searching and some web services may relativity with keywords it was not discovered this relationship is called semantically of relation. On the Web service was discovery. How can we know that the quality of web services. This research presents the model semantic web services selection. By thinking into the quality of service and a web service-related travel. Take a concept ontology to development semantic searching. And calculate the quality service of website that is make web search data more efficiently.

Keywords: Web services selection; Semantic web service; Quality of Service (QoS).

1. INTRODUCTION

Currently we have developed the technology quickly. Causing social change to the information society. As that required is accurate and fast. A lot of information is stored on the distribution internet network and may be to increase. Access to the information can be done by the search engines. In general, doing by search engines that be using the words or keywords. The processor will sent the keywords to compare with the document data stored in the database. Consider from Keyword that be appears to search or not. If have not the system that is mean the document is not relevant. In fact, some documents may be relevant. Although the document does not have the keywords in the document that is call semantically related. The semantic search has become a new way to help find information more efficiently because you can understand the true meaning of the document.

Web services [1] is a technology that makes any application software that' can developed on the basis of language and different architecture. It can communicate and changing information by XML language because that is be central of language to communication through sharing information Protocol HTTP. That can use other

standard Web services with web WSDL languages to describe a form of data service. Currently may be we are using website in the form of a SOAP. By the traditional standards or REST and Simple but uncomplicated. OASIS and W3C are a main agency to control standards and the architecture of web services.

Currently, we can search a lot of web services from Public UDDI. Mostly the system of search it is the keywords to compare the data with Web services, documents stored in the database if the results can be found only a list of Web services that match the keyword directly. Therefore, the results may not be accurate and complete because some web services that may be relevant semantic but can not found it. How can know using to use the web service or how much quality when we found it. This research of semantic web services was designed and developed models selection. The web service using only the information relevant for tourism if regard to the quality of service. Bringing ontologies technology to assist in the interpretation of data. Search a means of stalling the weight of words. Web services are grouped before measuring the similarity of Web services in each group. And calculating the quality of service of Web services. All this makes web search more effective. And the results to meet the needs of more users.

2. THEORY AND RELATED LITERATURE

2.1 Ontology

Ontology [2] that is a concept that appears in many disciplines. Each major was given a deferent definition about Ontologies. The computer also has a lot of meaning but can concluded that is a specification about the concept. This concept is trying to describe the idea of domain or interest in one side of things within that can the classification of information and links associated with them. That can present by Class, Instance, Relationship, Property and Rule. Further, that was built with the ability to share information data can be reuse and has the ability

to convey the property. (Inheritance) applying to use it as an alternative way of sharing information. And extracting knowledge from databases. Using language to know that there is a clear and more accurate to describe using natural language vocabulary link together a sentence to describe things in a different aspect, necessitating the use of symbols to convey (Semantics) to software or computer can understand the meaning and apply it to practical use. Only by finding out what they want from the Internet, a virtual data warehouse large. If we can not find the meaning of a word from the search. It will help us get the results that meet the needs and faster. This corresponds to the concept of the semantic web.

Currently, there are applications or tools for supporting the development of Ontologies (Ontology Editor) to popular programs Protégé developed by Stanford University, Oxford. And Hozo developed by Osaka University, and supported by these tools is a process engineering knowledge gives users the knowledge engineers or specialists. Storage and transfer of knowledge in the form of that convenient and easier.

2.2 SKOS

Simple Knowledge Organization System (SKOS) [3] as a model for creating a Store of words or standard data sharing based SKOS consists relations following keywords (skos: preLabel) Synonym (skos: altLabel) words. wider (skos: broader) than the narrow (skos: narrower) and related words. (skos: related) An example of the definition of the term structure of SKOS as shown in Figure 1.

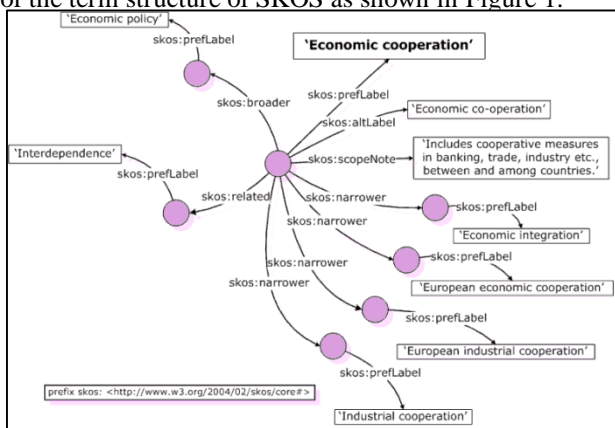


Figure 1SKOS Core Vocabulary Structure [3]

2.3 WordNet

WordNet [4] it is a database of Large English vocabury consist by nouns, verbs, adjectives and adverbs. Are grouped of words on the meaning. It will be linked by the relationships of concepts, and meaning of words to Relate in WordNet definition that can see at below.

- 1.Synonymy is the set of words with similar meaning (Synsets, Synonym sets) to replace the words with meaning.

- 2.Antonymy is a word that means the opposite of symmetry, which focuses on the words. Of adjectives and verbs magic.
- 3.Hyponymy and Hypernymy is the relationship of the Concept the appearance of the concept and super concept by Hyponym the Concept above Concept other parts Hypernym the Concept under Concepts such as hotel there. Hypernymy is building and building, the hotel is a Hyponymy.
- 4.Meronymy Holonymy and the relationship some parts like arms, legs, hands and every part of the body and the steering wheel is part of the braking and so on.
- 5.Troponymy verb or action to show symptoms. Which is used to display the names of the hierarchy concept.
- 6.Entailment is a relationship between a verb, defined by WordNet.

2.4 K-means

K-means [5] that is a technique used to group data. The technique is to learn without a teacher. (Unsupervised Learning) by grouping divides data into cluster will provide information that is similar in the same cluster. And address information of each cluster will have a similar minimum. Grouping information with the use of the information that was provided in the same cluster as the representative of all the data in that cluster. The group started out as a parameter, where k is the number of clusters to search. The algorithm will be randomly selected starting number k series, each series has come to be the starting point of each cluster (centroid), then grouped with the remaining data. Information will be provided in the same cluster are similar to a representative of that cluster most. And then calculated the average of the new cluster. And carry out the same process with the next. Until all the data Information is grouped perfectly and no longer change the function of the k-means data can be highly effective when densely packed. Each group separately and clearly. And the density of each group of similar strength of k-means is simple and can be used with multiple data types. And is also effective in terms of speed. But the weakness of the k-means that it is not appropriate to all types of information. And can not handle the group has a spherical shape or size or density difference is also k-means there is limited information available for agent vague or unclear.

2.5 RAP-API

RAP-API [6] is a program of PHP used to interconnect structure of Ontologies with a search system that RAP will find information on OWL or RDF through language, SPARQL the results of the search will be. find a relationship that matches the keyword directly. Also find about keywords that are related in structure Ontologies. RAP is responsible to Parsing, Searching, Manipulating, Serializing and Serving RDF models.

Table 1: To shows the details of web services ontology.

ID	Name	Type	Description
1	Element	Class	Class for the information of elements
2	Interface	Class	Class for the information of interfaces
3	Operation	Class	Class for the information of operations
4	Parameter	Class	Class for the information of parameters
5	Service	Class	Class for the information of web services
6	Classification	Class	Class for the information of classification of web service
7	Category	Class	Class for the information of category of web service
8	Version	Class	Class for the information of versions
9	Description	Class	Class for the information of descriptions of web service
10	Endpoint	Class	Class for the information of endpoints
11	Address	Class	Class for the information of addresses
12	skos:Concept	Class	Class for the information of terms
13	Input	Class	Class for the information of inputs
14	Output	Class	Class for the information of outputs
15	Owner	Data type	The constant indicated value showing owner
16	Cost	Data type	The constant indicated value showing cost
17	hasInterface	Object	Showing the relationship between class:Service and class:Interface on the interfaces in each service
18	hasOperation	Object	Showing the relationship between class:Interface and class:Operation on the operations in each interface

19	hasParameter	Object	Showing the relationship between class:Operation and class:Parameter on the parameters in each operation
20	hasCategory	Object	Showing the relationship between class:Service and class:Category on the services in each category
21	hasDescription	Object	Showing the relationship between class:Service and class:Description on the description in each web service
22	hasEndpoint	Object	Showing the relationship between class:Service and class:Endpoint on the endpoint in each web service
23	hasAddress	Object	Showing the relationship between class:Endpoint and class:Address on the address in each endpoint
24	hasTerm	Object	Showing the relationship between class:Service and Class:skos:Concept on the terms in each web service
25	ClassHasTerm	Object	Showing the relationship between class:Category and Class:skos:Concept on the terms in each category

3. Collect the vocabulary related to travel services from glossary' WordNet containing the keywords of vocabulary as Synonyms, wider, narrower and related words.

4. Save the vocabulary words from WordNet Fill in a store words in the ontology using SKOS in Protégé programs.

3.3 Web Services Clustering

1. Data Preparation

- Word Segmentation is the process of cutting the separate words in sentences, such as "GetWeatherByZip" also can be four words including Get | Weather | By | Zip or sentence "Expedia is a world leader in travel and technology" and then 9 words such as Expedia | is | a | world | leader | in | travel | and | technology.

- The word cleaning wrap stop words or stop words are steps cut out unnecessary words stop words are words that extend to the other, but there have not meaning

itself or linking words as an example that a, an, and, at, do, for, in, is, of, than, the, to, what, while etc.

2. Data Transformation This step is the process of converting data from Nominal type to a Numerical type. (Convert Nominal to Numerical).
3. It is web services group using an algorithm K-means. After converting information in the form Numerical. The group of web service using an algorithm K-means the k=5 identifying the type of measurement. Numerical measure and use method measuring Cosine Similariry.

3.4 Calculating the similarity of web services (Web Services Similarity)

After complete the grouping Web services. This procedure is the calculation to determine the similarity of Web services in each group. By using Vector Space Model theory for the similar value that is Cosine Similarity [9] [10] By comparison, consider the name of the web service is Operation Name, Description, Input and Output.

$$\text{similarity} = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (1)$$

3.5 Determination the weight of words

Determination the weight of words divided into 5 parts for example.

1. Determination the weight of the structure SKOS. The determination the weight of the structure SKOS can show at Table 2.

Table 2: Example determination the weight of the structure SKOS.

Type of Term	SKOS Label	Weight
Keyword	prefLabel	1.00
Synonym term	altLabel	1.00
Broader term	broader	0.75
Narrower term	narrower	0.75
Related term	related	0.50

For example, the definition of the relationship with the value weight that have main words is “hostel” similar word is “inn” broader “hotel” narrower is “posthouse” and “roadhouse” Finally, related a word on the "lodge" is shown in Figure4.

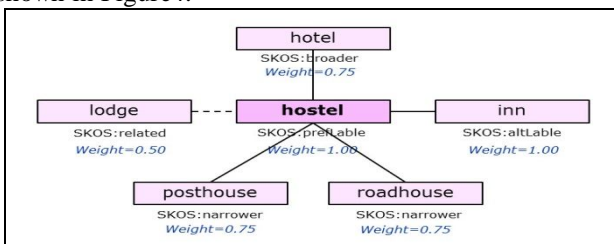


Figure 4 SKOS Vocabulary Relationship Example

2. Determination of weight according to position structure of web services.

In the determination of the weight of the position structure of web services. Which of these words is stored in the form of a class in ontology, including Web service name, Description, Operation, Input and Output shown in Table 3.

Table 3: shows an example of the weight according to position structure of web services

Position Name	Weight
Web service name	0.50
Description	1.00
Operation	0.50
Input	0.50
Output	0.75

3. Determination the weight of the structure for search

To determine the weight of the structure for search Is determined by the order of the keywords the user entered from left to right as the user enters the word “hostel booking” The hostel will be important and worth its weight over the booking and so the test is given. the first order is worth its weight is 1.00, and the second word is worth its weight is 0.75, the weight can be adjusted as appropriate, in the event that there are two more keywords.

For example, to calculate the similarities between keywords and Ontologies Web services technology. Follow these steps:

- a) The key word is “hostel booking” So the weight of the words is important. Keywords in the first word "hostel" is worth its weight is 1.00, and the second word in the term “booking” the weight is 0.75.
- b) Leads a group of keywords that are then compared to the weighted index that is stored in the warehouse, according to the SKOS structure of words with the weight.
 - i) hostel = {hostel (1.00), inn (1.00), hotel (0.75), posthouse (0.75), roadhouse (0.75), lodge (0.50)}.
 - ii) booking = {booking (1.00), reservation (1.00), employment (0.75), reserve (0.50), book (0.50)}.
- c) To calculate the TF-IDF [11] to every word in Ontologies the structure position of a web service on the equation.

$$tf(t, d) = 0.5 + \frac{0.5 * f(t, d)}{\max \{f(w, d) : w \in d\}} \quad (2)$$

where tf (t, d) is the weight of word' t on web service' d
f (t, d) is the frequency of the word' t on web service' d

max {f(w, d): w ∈ D} is the maximum frequency of the word 'w' on web service 'd'

$$idf(t, D) = \log \frac{|D|}{|\{d \in D: t \in d\}|} \quad (3)$$

where idf(t, d) is the back to weight of word 't' all on web service

|D| is the number of all web services

|\{d ∈ D: t ∈ d\}| is the number of web services with word 't'.

$$WSP = \sum_{k=0}^n (tf_{idf_k} * WePoWS_k) \quad (4)$$

where WSP is the weight of the words in each location, based on the structure of the web service.

tf_idf is the value of weight in important of words.

WePoWS (Weight of Position Web Service) is the weight of the structure of the web service. shown in Table 3.

According calculated to the above equation, the results reflect the IF-IDF weight as shown in Table 8. in the Appendix A.

4. Calculate the value of weight, every word in group from 2) when compared the value of weight on structure SKOS that shown in Table 2 can be represented by the following equation.

3.2 Web Services Clustering

1. Data Preparation

- Word Segmentation is the process of cutting the separate words in sentences, such as "GetWeatherByZip" also can be four words including Get | Weather | By | Zip or sentence "Expedia is a world leader in travel and technology" and then 9 words such as Expedia | is | a | world | leader | in | travel | and | technology.
- The word cleaning wrap stop words or stop words are steps cut out unnecessary words stop words are words that extend to the other, but there have not meaning itself or linking words as an example that a, an, and, at, do, for, in, is, of, than, the, to, what, while etc.

2. Data Transformation This step is the process of converting data from Nominal type to a Numerical type. (Convert Nominal to Numerical).

3. It is web services group using an algorithm K-means. After converting information in the form Numerical. The group of web service using an algorithm K-means the k=5 identifying the type of measurement. Numerical measure and use method measuring Cosine Similariry.

3.6 Calculating the similarity of web services (Web Services Similarity)

After complete the grouping Web services. This procedure is the calculation to determine the similarity of Web services in each group. By using Vector Space Model theory for the similar value that is Cosine Similarity [9]

[10] By comparison, consider the name of the web service is Operation Name, Description, Input and Output.

$$\text{similarity} = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (1)$$

3.3 Determination the weight of words

Determination the weight of words divided into 5 parts for example.

4. Determination the weight of the structure SKOS.

The determination the weight of the structure SKOS can show at Table 2.

Table 2: Example determination the weight of the structure SKOS.

Type of Term	SKOS Label	Weight
Keyword	prefLabel	1.00
Synonym term	altLabel	1.00
Broader term	broader	0.75
Narrower term	narrower	0.75
Related term	related	0.50

For example, the definition of the relationship with the value weight that have main words is "hostel" similar word is "inn" broader "hotel" narrower is "posthouse" and "roadhouse" Finally, related a word on the "lodge" is shown in Figure4.

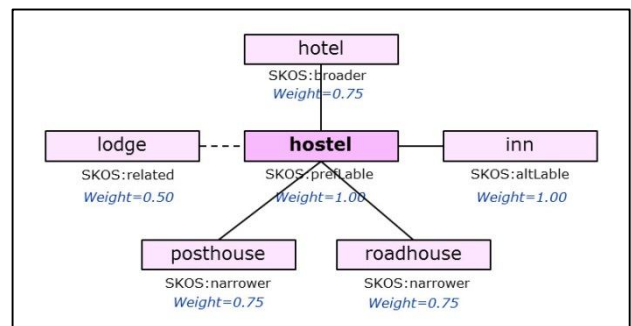


Figure 4 SKOS Vocabulary Relationship Example

5. Determination of weight according to position structure of web services.

In the determination of the weight of the position structure of web services. Which of these words is stored in the form of a class in ontology, including Web service name, Description, Operation, Input and Output shown in Table 3.

Table 3: shows an example of the weight according to position structure of web services

Position Name	Weight
Web service name	0.50
Description	1.00
Operation	0.50
Input	0.50
Output	0.75

6. Determination the weight of the structure for search

To determine the weight of the structure for search is determined by the order of the keywords the user entered from left to right as the user enters the word "hostel booking" The hostel will be important and worth its weight over the booking and so the test is given. the first order is worth its weight is 1.00, and the second word is worth its weight is 0.75, the weight can be adjusted as appropriate, in the event that there are two more keywords.

For example, to calculate the similarities between keywords and Ontologies Web services technology. Follow these steps:

- a) The key word is "hostel booking" So the weight of the words is important. Keywords in the first word "hostel" is worth its weight is 1.00, and the second word in the term "booking" the weight is 0.75.
- b) Leads a group of keywords that are then compared to the weighted index that is stored in the warehouse, according to the SKOS structure of words with the weight.
 - i) hostel = {hostel (1.00), inn (1.00), hotel (0.75), posthouse (0.75), roadhouse (0.75), lodge (0.50)}.
 - ii) booking = {booking (1.00), reservation (1.00), employment (0.75), reserve (0.50), book (0.50)}.
- c) To calculate the TF-IDF [11] to every word in Ontologies the structure position of a web service on the equation.

$$tf(t, d) = 0.5 + \frac{0.5 * f(t, d)}{\max\{f(w, d) : w \in d\}} \quad (2)$$

where $tf(t, d)$ is the weight of word' t on web service' d
 $f(t, d)$ is the frequency of the word' t on web service' d
 $\max\{f(w, d) : w \in d\}$ is the maximum frequency of the word' w. on web service' d

$$idf(t, D) = \log \frac{|D|}{|\{d \in D : t \in d\}|} \quad (3)$$

where $idf(t, d)$ is the back to weight of word' t all on web service

$|D|$ is the number of all web services
 $|\{d \in D : t \in d\}|$ is the number of web services with word' t.

$$WSP = \sum_{k=0}^n (tf_idf_k * WePoWS_k) \quad (4)$$

where WSP is the weight of the words in each location, based on the structure of the web service.

tf_idf is the value of weight in important of words.
 WePoWS (Weight of Position Web Service) is the weight of the structure of the web service. shown in Table 3.

According calculated to the above equation, the results reflect the IF-IDF weight as shown in Table 8. in the Appendix A.

7. Calculate the value of weight, every word in group from 2) when compared the value of weight on structure SKOS that shown in Table 2 can be represented by the following equation.

$$SKOSPOS = \sum_{k=0}^n (WSP_k * WePoSK_k) \quad (5)$$

where SKOSPOS is the value of weight in structure SKOS.

WSP is the value of weight, words in each position on the structure of the web service.

WePoSK (Weight of Position SKOS) is the value of weight in structure of the web service.

8. The calculation of the similarity semantic between important words ontologies on web services. Can be represented the equation.

$$SIM = \sum_{k=0}^n (SKOSPOS_k * WPQ_k) \quad (6)$$

where SIM is the value of score the semantic each web services.

SKOSPOS is the value of weight in structure' SKOS.

WPQ (Weight of Position Query) is the value of weight of the position of key words obtained from the user.

$$SEMSIM = \frac{SIM_k}{\max(SIM)} \quad (7)$$

where SEMSIM is the value of score the semantic similarity after passing Normalization, The score ranges from 0 to 1.

SIM is the value of score the semantic on the web services.

$\max(SIM)$ is the maximum score of semantic all the web services.

3.7 Determination of the quality of web services (QoS)

How to choose an appropriate service. Consider on 4 factors [12]. (1) The response time of the server provider (Response time) is the period from the service sends a request to a server that provides the data received in return. (2) The availability of services. (Availability) is a Web service that is displayed and users can access the service immediately. The calculation of the response time for each time a user can access the given period. (3) The stability or reliability of the system. (Reliability) is the ability to serve up Web service. Represented by the ratio of the number of calls serviced successful. And (4) The cost of a service call (Cost) is the cost of running the service of the service. In some of the web service may be charged.

How to evaluate your web service to web service options that are appropriate for each service. Scores are based on a maximum. For details on the determination of the four factors and calculate the quality of service of web services are as follows.

1. The response time of the server service. Measured from the time the applicant submitted a request to get the data back from the server of the service provider. Which is calculated by the following equation.

$$T = \frac{\sum_{i=1}^n (T_r - T_s)}{n} \quad (8)$$

$$RT = 1 - \frac{T}{\sum T_i}$$

where T is the average of response time, Milliseconds
 T_r is the time to get the data back, Milliseconds
 T_s is the time to start sending a request. Milliseconds
 RT is the probability of the average response time.

2. Ready to used which Measured by the result on all time, users can access the service has been achieved by the total time all users to run web services in a given period. Which is calculated by the equation.

$$A = \frac{N_{suc}}{N_{all}} \quad (9)$$

where A is ready to use.

N_{suc} is the amount result of all time, users can access the web service successfully. In the given period that' called milliseconds

N_{all} is the amount result of all time, users can access the web service. In the given period that' called milliseconds.

3. Reliability that can Measured by the amount of time and users run successfully serviced with the total number of

times it is users run a service. In the given period Which is calculated by the equation.

$$R = \frac{T_{suc}}{T_{all}} \quad (10)$$

where R is reliable.

T_{suc} is the amount of times that is users run web services successfully. In the given period

T_{all} is the amount of times that is users run web services. In the given period

For example, Who is user requests a service, a amount of four times in a given period. And has requested to use of the succeeding for three times but failed the first time, so the reliability is equal to 3/4 or equal to 0.75.

4. The expend of running the service. Measured by the cost of charge with group together to provide the same service provider. Which is calculated by the equation.

$$C = 1 - \frac{P}{\sum P_i} \quad (11)$$

where C is the probability of the expend for service provider.

P is the rate of charged for the service provider.

$\sum P_i$ is the amount of the rates to charged for service all providers.

5. The calculation of the service quality of web services. The calculation of the service quality of web services using equation weighted. Calculated as the quality of service of the four factors, including the response time of the server for service. Ready to used, Credibility And the expend of running the service. The equation is:

$$QoS_{Si} = \sum_{x \in V} W_x QoS_{Si,x} \quad (12)$$

where QoS_{Si} is the quality of service of Web Si.

$QoS_{Si,x}$ is the value of quality provider for web services, x factor of Si.

W_x is the value of weight given determine to the quality of the service by the x factor by $\sum_{x \in V} W_x = 1$

V is a set of factors that determine the quality of service provided by the $V = \{\text{Response time, Availability, Reliability, Cost}\}$.

3.8 Design and development the user interface for semantic searching

To search for web services data users must fill the keyword to query the system. Then the system will compare between the keywords with standard knowledge about Ontologies of web services. Therefore, it must be developed with the user interface. As intermediaries in

contacts between users and standard knowledge that can using by the PHP as the primary language for developing user interfaces using CSS to format the output. The query from yeast or OWL Ontologies using SPARQL [7] for more information (Query) and extract data from OWL through RAP API for PHP [6].

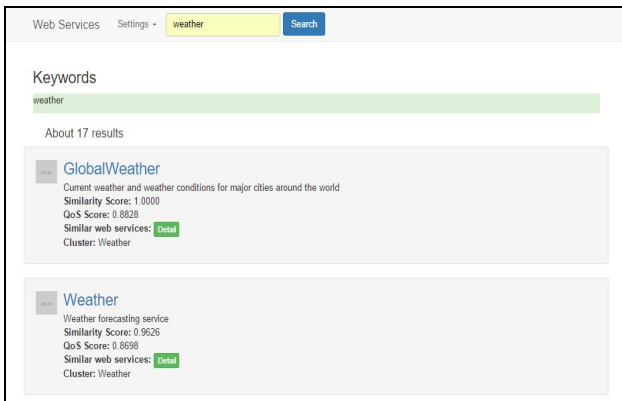


Figure 5 Example of User Interface for Index Page

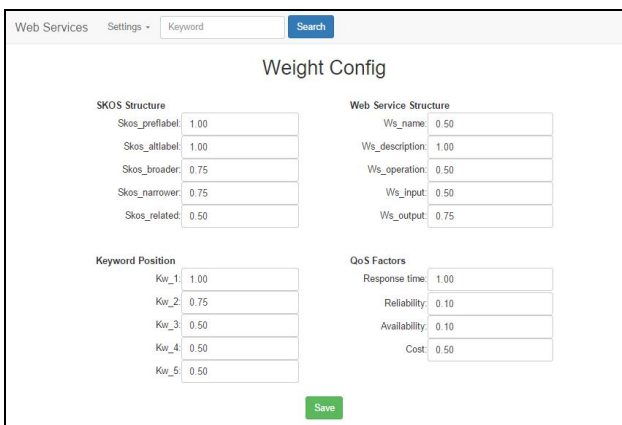


Figure 6 Example of User Interface for Settings Page

6. EXPERIMENTAL RESULT

The results can found the Examples of key word that is the word "hostel booking" when searched into Ontologies with SPARQL language technology and RAP-API. That result can discover the list name of web services as a keyword with a similar meaning. Most of the top three web services include the name EasyTobook, Expedia and HotelClub. The semantic similarity scores are 1.00, 0.89 and 0.83.

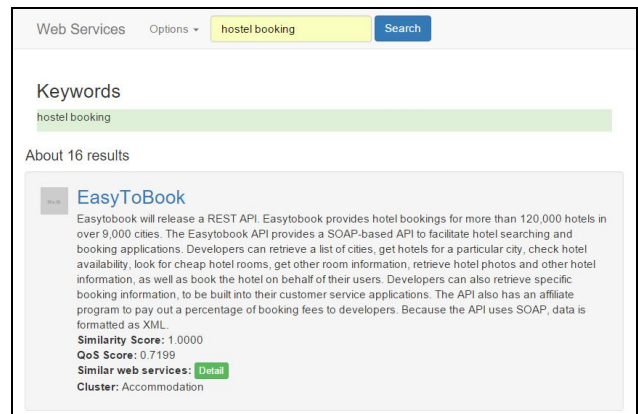


Figure 7 Results for Web Services Semantic Searching

Web Service Name
 Description
 Operation
 Input
 Output

Submit

Web Service Similarity Table (Accommodation)

	Jasons	Zumata Hotels	Expedia	Cleartrip Hotel	EasyToBook	Global Matrix	Resort Data Processing	Majestic International Group	Alpharooms Affiliate	Adriagate
Jasons	1.0000	0.9583	0.9729	0.9798	0.9601	0.9806	0.9553	0.9277	0.9730	0.9770
Zumata Hotels	0.9583	1.0000	0.9439	0.9600	0.9329	0.9554	0.9336	0.9026	0.9457	0.9507
Expedia	0.9729	0.9439	1.0000	0.9704	0.9498	0.9715	0.9425	0.9094	0.9605	0.9653
Cleartrip Hotel	0.9798	0.9600	0.9704	1.0000	0.9563	0.9767	0.9490	0.9212	0.9692	0.9730
EasyToBook	0.9601	0.9329	0.9498	0.9563	1.0000	0.9573	0.9296	0.8957	0.9501	0.9517
Global Matrix	0.9806	0.9554	0.9715	0.9767	0.9573	1.0000	0.9476	0.9193	0.9700	0.9738
Resort Data Processing	0.9553	0.9336	0.9425	0.9490	0.9296	0.9476	1.0000	0.9486	0.9451	0.9489
Majestic International Group	0.9277	0.9026	0.9094	0.9212	0.8957	0.9193	0.9486	1.0000	0.9156	0.9209
Alpharooms Affiliate	0.9730	0.9457	0.9605	0.9692	0.9501	0.9700	0.9451	0.9156	1.0000	0.9658
Adriagate	0.9770	0.9507	0.9653	0.9730	0.9517	0.9738	0.9489	0.9209	0.9658	1.0000

Figure 8 Example of web Services Similarity Scores

Calculated equation in Section 3.5 results show the values in Table 4. and Table 5.

Table 4: shows the value of words from the keyword by "hostel"

Term (Weight)	Hostel (1.00)	Hotel (0.75)	Inn (1.00)	lodge (0.50)	posthouse (0.75)	roadhouse (0.75)
Adriagate	0.90	0.04	0.53	0.38	0.34	0.00
Alpharooms	0.45	0.04	0.53	0.38	0.34	0.00
Cleartrip	0.45	0.03	0.98	0.23	0.59	0.25
EasyTobook	0.45	0.04	1.39	0.30	0.85	0.25
Expedia	0.45	0.04	1.02	0.41	0.59	0.25
HotelClub	0.45	0.04	0.87	0.30	0.59	0.25
Jasons	0.45	0.04	0.64	0.23	0.59	0.51

Table 5: shows the value of words from the keyword by “booking”

Term (Weight)	book (0.50)	employment (0.75)	booking (1.00)	reserve (0.50)	reservation (1.00)
Adriagate	0.26	0.34	0.15	0.34	0.43
Alpharooms	0.26	0.34	0.15	0.56	0.43
Cleartrip	0.30	0.34	0.23	0.23	0.21
EasyTobook	0.26	0.34	0.15	0.34	0.43
Expedia	0.34	0.34	0.23	0.45	0.21
HotelClub	0.38	0.34	0.30	0.34	0.21
Jasons	0.15	0.68	0.30	0.23	0.21
Majestic	0.15	0.34	0.15	0.23	0.21

By calculating the similarity of semantic web services which is the equation in section 3.5 and the calculation of the quality of service. The equation in section 3.6 of the results shown in Table 6.

Table 6: shows the semantic similarity

Web Service Name	Semantic Similarity Score	QoS Score
Adriagate	0.75	0.86
Alpharooms	0.69	0.84
Cleartrip	0.79	0.94
EasyTobook	1.00	0.72
Expedia	0.89	0.72
HotelClub	0.83	0.24
Jasons	0.82	0.77
Majestic	0.56	0.85

The result of evaluation to search web service precision value is 0.8274, Recall value is 0.8186, F-measure value is 0.8127 and Accuracy value is 0.9488. The details show in table 7.

Table 7: To shows the results valuation of search for web services

Method	Value
Precision	0.8274
Recall	0.8186
F-measure	0.8127
Accuracy	0.9488

7. CONCLUSION

This research shows the model selected semantic web services. By regard to quality of web services for tourism. This model consists of 1) The web service ontology that can describe the relationship and structure of data using by OWL language which can be linked together semantically. 2) How to find the similarity of web services. 3) How to select quality Web services. And 4) How to calculate the weight of words. This model can be made or selected web services have effectively increased and the results of the search to meet the needs of more users.

References

- [1] Christopher Ferris and Joel Farrell, What are Web services?, Communications of the ACM, v.46 n.6, 2003.
- [2] OWL 2 Web Ontology Language Document Overview - W3C Recommendation, available at <http://www.w3.org/TR/owl-overview/>, 2012.
- [3] SKOS Simple Knowledge Organization System Reference - W3C Recommendation, available at <http://www.w3.org/TR/skos-reference/>, 2009.
- [4] George A. Miller, WordNet: a lexical database for English, Communications of the ACM, v.38 n.11, p.39-41, 1995.
- [5] Tapas Kanungo , David M. Mount , Nathan S. Netanyahu , et al., An Efficient k-Means Clustering Algorithm: Analysis and Implementation, IEEE Transactions on Pattern Analysis and Machine Intelligence, v.24 n.7, p.881-892, 2002.
- [6] Radoslaw Oldakowski , Christian Bizer, Daniel Westphal. RAP: RDF API for PHP, in proc. International workshop on interpreted languages, 2004.
- [7] SPARQL 1.1 Overview - W3C Recommendation, available at <http://www.w3.org/TR/sparql11-overview/> , 2013.
- [8] Jos de Bruijn, Christoph Bussler, John Domingue, et al., Web service modeling ontology (WSMO) - W3C Member Submission, 2005.
- [9] Singhal, Amit. “Modern Information Retrieval: A Brief Overview”. Bulletin of the IEEE Computer Society Technical Committee on Data Engineering 24 (4): pp. 35–43, 2001.
- [10] Sidorov Grigori, Gelbukh Alexander, Gómez-Adorno Helena, et al., “Soft Similarity and Soft Cosine Measure: Similarity of Features in Vector Space Model”. Computación y Sistemas 18 (3): 491–504, 2014.
- [11] Stephen Robertson. “Understanding inverse document frequency: On theoretical arguments for IDF”. Journal of Documentation 60 (5): 503–520, 2004.
- [12] Ekkachai Naenudorn and Suphakit Niwattanakul, “Web Services Selection Algorithm Based on the Quality of Service,” Lecture Notes on Software Engineering vol. 2, no. 1, pp. 26-30, 2014.

AUTHOR



Ekkachai Naenudorn is a PhD student at the School of Information Technology, Suranaree University of Technology, Thailand. Currently he is lecturer at the Mahasarakham Business School, Mahasarakham University, Thailand. He completed his MSc in Information Technology from the Khon Kaen University, Thailand in 2006. His research areas include web service, semantic searching and data mining.



Suphakit Niwattanakul is assistant professor at the school of Information Technology, Suranaree University of Technology, Thailand. He received his PhD in computer science from the University of La Rochelle, France in 2008. His current research is about

Semantic Web technologies applied to information extraction and retrieval system.



Nisachol Chamnongsriis lecturer at the school of Information Technology, Suranaree University of Technology, Thailand. He received his PhD in Information Studies from the Khon Kaen University, Thailand in 2009. His current research is about Digital

Libraries, Information Organization, Metadata Development and Digital Preservation.

Appendix A.

Table 8: Show some of the results TF-IDF weight of every word on each web services obtained from calculations according to equation 4

Web Service Name	book	car	employment	flight	hostel	hotel	inn	booking	lodge	reserve	posthouse	reservation	roadhouse	search
Adriagate	0.53	0.30	0.45	0.30	0.90	0.06	0.53	0.15	0.75	0.68	0.45	0.43	0.00	0.20
Alpharooms	0.53	0.30	0.45	0.60	0.45	0.06	0.53	0.15	0.75	1.13	0.45	0.43	0.00	0.20
Cleartrip	0.60	0.64	0.45	0.64	0.45	0.04	0.98	0.23	0.45	0.45	0.79	0.21	0.34	0.49
EasyTobook	0.53	0.64	0.45	0.64	0.45	0.06	1.39	0.15	0.60	0.68	1.13	0.43	0.34	0.44
Expedia	0.68	0.98	0.45	0.79	0.45	0.06	1.02	0.23	0.83	0.90	0.79	0.21	0.34	0.83
HotelClub	0.75	0.94	0.45	0.98	0.45	0.06	0.87	0.30	0.60	0.68	0.79	0.21	0.34	0.44
Jasons	0.30	0.64	0.90	0.64	0.45	0.06	0.64	0.30	0.45	0.45	0.79	0.21	0.68	0.54
Majestic	0.30	0.60	0.45	0.30	0.45	0.03	0.30	0.15	0.45	0.45	0.90	0.21	0.00	0.10