Application of Ontology in Semantic Web Searching of Flight Ticket as a Study Case

Togu Novriansyah Turnip, Ephraim Kesaba Silalahi, Yuli Arantxa Vicario Sinulingga and Valentina Siregar
Faculty of Informatics and Electrical Engineering, Institut Teknologi Del Sitoluama, Laguboti, Tobasa 22381 Indonesia

Abstract. The development of technology allows people to obtain information easily. Information about goods or services that society needs can be obtained easily as many websites are built to provide that information. But this is also a challenge for the community. People should search and compare price of the cheapest goods or services according to their needs as well as in purchasing flight tickets. People also should look for flight ticket prices on many websites and compare the tickets themselves for each website. Thus, we need a website that can integrate information from heterogeneous website as a source of data and can help people comparing the price of flight tickets easily. Semantic web is considered as one solution to the problem. Semantic Web requires appropriate scrapping of data to retrieve the online data from specified websites. By using ontology, the semantic web can collect the data and enhance the search process of flight ticket.

1. Introduction
Nowadays, internet becomes a basic human need. The trend of internet usage is people tend to do buying and selling products or services in online shopping. This phenomenon is gradually replacing the traditional way of transaction between seller and buyer to be more modern. World Wide Web (WWW) has quickly become a huge storage containing a certain amount of information about all the interesting things. One of favorite products highly consumed and known for every people is flight tickets.

There are the growing number of websites providing online flight tickets as information sources such as Pigi, Pegi, Utiket, nusatrip, etc. Due to many website of flight tickets make consumer confused to compare which one is the cheaper in term of price. Most of the web content is primarily designed to be read by humans, machines can only parse web pages for layout. Machine is not able to automatically process data from any web site without understanding its semantics. This is become a problem because consumer can spend much times to compare one by one from many websites.

Semantic Web presents to answer the problem. Semantic web is defined as a set of technologies which allows computer understand the meaning of information based on metadata, i.e., information about the contents of the information [1]. It can be said by implementing web semantic, it can integrate information from heterogeneous sources and organize those information for enhancing the search processes.
Semantic technologies require a metadata refer to the ontology. Ontology is a representation of a certain domain knowledge that contains terms in the domain with the relationship between existing terms [1]. The term of Ontology actually comes from philosophy of “Ontology” which mean something that really exists and how to describe it. In computer term, Ontology is used to specify a conceptualization. Zebua & Mustikasari [1] said “Nowadays, Ontology is widely used to support web semantic”. Ontology will serve as metadata; it becomes the fundamental to develop integration function semantically such as conceptual searching and navigation of semantics [1]. The language structures that can construct Ontologies include: XML, XML Schema, RDF, RDF Schema and OWL [1]. The development ontology in study case flight tickets used OWL as structure language because OWL has advantages compared to other structure languages that have more facilities than XML and RDF/s in expressing meaning and semantic.

Thus, web semantic can understand the meaning of search by the costumer and showing the best results. It can be said, used ontology to group data while web semantic to speed up the process search that make customer know which websites that provide the cheapest flight tickets and reduced the problem. It is expected that the combination of ontology and semantic web can make search process of the cheapest flight tickets easily.

2. Study Literature
The following are the related study to Semantic Web Development as a service to provide information about flight tickets.

A. Ontology
According to article from Natalya and Deborah, the Artificial-Intelligence literature contains many definitions of an ontology; many of these contradict one another. For the purposes of this guide an ontology is a formal explicit description of concepts in a domain of discourse (classes (sometimes called concepts)), properties of each concept describing various features and attributes of the concept (slots (sometimes called roles or properties)), and restrictions on slots (facets(sometimes called role restrictions)). Ontology connects the symbol that humans understand with the following forms that can be processed by machines (Antoniou & Harmelen, 2004). Ontology is built from several interconnected components to produce domain knowledge. The main components of the ontology compilers are classes, properties, slots and instances. Ontology has a formal and undefined language structures. The languages structures that can construct ontology are XML, XML Schema, RDF, RDF Schema and OWL (Zebua & Mustikasari, 2012).

B. OWL
OWL (Ontology Web Language) is ontology language recommended by W3C to represent the meaning of terms of a domain knowledge explicitly and the relation between the terms that cannot separated with document and web application. OWL is an extension of the RDF vocabulary. Each RDF graph can form OWL ontology. The meaning of an RDF graph equals to the meaning given by the OWL graph. Information in OWL form is incorporated in ontology, which can be stored as a document on the web (Nurkamid, 2009).

C. Methodology Methontology
Methontology is an ontology development proposing the expression of ideas as a set of Intermediate Representations (IR) and generating ontologies using translators (Saf, 2015). There are 7 stages that are done in methontology method that is:
1. Build Glossary of Terms
2. Building Concept Taxonomies
3. Build Ad Hoc Binary Relation Diagram
4. Building a Concept Dictionary
5. Describe Component Domain
6. Describe Formal Axioms
7. Describe Instance

D. Semantic Web

Semantic Web is the development of the generation or evolution of the World Wide Web (WWW). Semantic Web is a very large amount of information that is connected globally in a certain way and understood by the machine, so that it can be processed directly by the machine into knowledge to be displayed to the user (Ferdila & Mustikasari, 2011). Using semantic web methods, HTML-based data can be transformed into formats that can be understood by machines so that machines can gather information and understand the relationship between information (Ferdila & Mustikasari, 2011). Semantic web can make this change with the help of XML and data language standards like RDF and OWL. Semantic web principles are implemented in web and standard technology layers can be seen in Figure 1.

![Web Semantic Architecture](image)

**Figure 1.** Web Semantic Architecture

E. SPARQL

RDF data model is a statement in triple form consisting subject, predicate and object (Zebua & Mustikasari, 2012). In getting information from an RDF graph it takes a query. SPARQL is a query language that can be used to access data on semantic web (Zebua & Mustikasari, 2012). The use of SPARQL allows to:
1. Taking values from structured data and semi-structured data
2. Develop data by querying an unknown relationship
3. Can performs query operations of complex join operations on different databases more simply
4. Convert an RDF data into another vocabulary Results from SPARQL queries can return values in several data formats including: XML, RDF and HTML

3. Analysis and Design

System needs several websites supporting as a source of information which will displayed on a system. The data have html tags that can contain one or more data. This html tag is utilized by the system to retrieve data required by the system. The data retrieval process uses scrapping technique which allows a system to retrieve data from a website by referring to the html tag of the data. By knowing which html tags contain the required data the scrapping technique can be easier to run.

The authors need a survey which aims to know and explore html tags what representing data which require by the system. This survey can use inspect element technique on websites that provide flight tickets booking services. Based on technasia.com, the authors got website providers of flight booking services such as traveloka.com, nusatrip.com, Tiket.com, pegipegi.com, utiket.com, Tiket2.com/Indonesia, catallya.com, airpaz.com/en and Ezytravel.co.id (Setyanti, 2014). Based on the survey conducted by doing investigation of an element from those websites, there is a website that can’t generate the tag data which is catallya.com. Catallya.com can’t generate a tag data because the
domain has been expired and replaced with reservasi.com. Then the other eight websites allow for the process of analysis with the inspect element. However, traveloka.com is not included because the tags data have been encrypted to prevent the data scrapping. It is similar to utiket.com, which is a comparison website that doesn’t provide bookings through itself. Utiket.com provides ticket through website previously existing, so not included to reduce redundancy data. So, the data of flight tickets will be taken as an information source from five websites such as Tiket.com, pegipegi.com, reservasi.com, airpaz.com and ezytravel.co.id.

The design of ontology used Methontology method because this method is the best method to build ontology. The design of ontology applied seven stages in Methontology method. This ontology design produces 8 class object properties and 6 data properties. For classes that are formed can be seen on Figure 2.

![Figure 2. Class in Domain Concept](image)

Design of flight ticket search website was done after design of ontology has been finished. Design of flight tickets search process is applying OWL for data storage that has been generated into semantic web. The data stored in the OWL can be accessed by using SPARQL as query language for RDF/OWL. This semantic web is built using Jena Java RDF API and toolkit which is a Java-based framework for constructing semantic web. The use case diagram of flight tickets search can be seen in Figure 3.

![Figure 3. Use Case Diagram](image)
Use case diagram as shown in the picture allows the customer to do *Pencarian Dasar* for searching flight tickets. After customer performed *Pencarian Dasar*, user got information about the price of flight tickets from the cheapest to the highest based on the various websites. When searching the flight tickets, system loaded Ontology to retrieve the information.

### 4. Implementation

Based on the results of ontology analysis and design conducted, the ontology implementation consists of three stages:

- **a. Class Implementation**
  
  Domain of Flight Tickets consists of eight classes: Baggage class, Airport class, Price class, Classroom Class, city class, flight class, Travel class and Website class. Flight tickets domain class hierarchy can be seen in Figure 4.

- **b. Properties Implementation**
  
  Flight Ticket Ontology uses two types of properties: object properties and data properties. Implementation of object properties generates 16 object properties. Implementation of data properties generates 6 data properties.

- **c. Individuals Implementation**
  
  Every class have individual. Implementation of individual in Bagasi class can be seen in Figure 5.

Based on the results of analysis, the development team implement semantic web. Semantic Web Ticket Search System is a web-based application built using Java programming language with Spring framework and Jena plugin. This semantic web provides the function of searching flight tickets information from various websites by entering keywords in the field provided on the web. Flight tickets data were retrieved online from the specified websites which using data scrapping technique.
The process of scrapping data from the internet took 5-15 minutes. The initial view of the Ticket Search System is shown in Figure 6.

![Figure 6. Home Page](image)

The flight search parameter consists of:
- a. Departure Airport
- b. Destination Airport
- c. Date of departure

When the user has entered the keyword in based on the above parameters, then the system displays the results of the search. The view of the flight search results can be seen in Figure 7.

![Figure 7. Result Page](image)

5. Conclusion
The conclusions based on the assessment and applications of Ontology on Semantic Web in the case study of flight tickets are:
1. The application of ontology in the process of searching flights ticket information in five websites has provide facilities for users in search of the ticket where users no longer need to compare the price of flight ticket from the five websites. This is to say, customers do not spend much time compare the price of the flight tickets with open several website flight ticket sales.
2. Ontology designed and implemented in building semantic web includes quality ontology because semantic web meets user needs and solves the existing problems.
3. Internet connection is an essensial factor on the process of data retrieval. If the internet connection is fast, then the process of taking data took 5-15 minutes. While, if the speed is slow, it can took more or less 2 times than the data retrieval when the internet connection fast and also the amount of data retrieved is less.
6. Suggestions

There are some further development:
1. The implementation can focus on the development of the Semantic Web and Ontology development for case international flights.
2. Required API features of the website to make the data collection process can be done more quickly.
3. The web semantic should host in the Internet in order to make the data collection process more quickly.

References

[1] T. Berners-Lee, "W3C Architecture," 2001. [Online]. Available: https://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html. [Accessed 04 January 2017].
[2] Ferdila and M. Mustikasari, "APLIKASI WEB SEMANTIK UNTUK PENCARIAN MATERI PERKULIAHAN," pp. 1-16, 2011.
[3] M. R. A. Saf, "Pemodelan Ontologi untuk Sistem Informasi Proyek (SIMPOA) di Politeknik Caltex RIau," Jurnal Sains, Teknologi dan Industri, vol. XIII, no. 1, pp. 46-52, 2015.
[4] E. P. Setyanti, "Techinasia," 2014. [Online]. Available: https://id.techinasia.com/website-booking-ticket-pesawat-travel-liburan-indonesia. [Accessed 3 January 2017].
[5] J. Zebua and M. Mustikasari, "APLIKASI PENCARIAN BUKU PERPUSTAKAAN DENGAN MENGGUNAKAN RDF DAN SPARQL," UG Jurnal, vol. VI, p. 04, 2012.
[6] M. Nurkamid, "APLIKASI BIBLIOGRAFI PERPUSTAKAAN BERBASIS TEKNOLOGI WEB SEMANTIK," pp. 1-22, 2009.
[7] G. Antoniou and F. v. Harmelen, A Semantic Web Primer, 2nd ed., London: The MIT Press, 2004.
[8] K. D. P. Novianti, N. A. Setiawan and S. S. Kusumawardani, "Perancangan Ontologi Sebagai Teknologi Penyimpanan Informasi untuk Penulisan Pustaka pada SIRREF JTETI UGM," Jurnal Nasional Pendidikan Teknik Informatika, vol. 3, no. 3, p. 6, Desember 2014.
[9] Natalya F. Noy and Deborah L. McGuinness, "What is in an ontology?". Available: https://protege.stanford.edu/publications/ontology_de velopment/ontology101-noy-mcginness.html [Accessed 04 January 2017].