إقــــــرار

أنا الموقع أدناه مقدم الرسالة التي تحمل العنوان:

تصنيف المستندات العربية استناداً إلى الأنطولوجيا

Ontology-Based Arabic Documents Classification

أقر بأن ما اشتملت عليه هذه الرسالة إنما هو نتاج جهدي الخاص، باستثناء ما تمت الإشارة إليه حيثما ورد، وإن هذه الرسالة ككل أو أي جزء منها لم يقدم من قبل لنيل درجة أو لقب علمي أو بحثي لدى أي مؤسسة تعليمية أو بحثية أخرى.

DECLARATION

The work provided in this thesis, unless otherwise referenced, is the researcher's own work, and has not been submitted elsewhere for any other degree or qualification

Student's name: **Mohammed M. N. Abu Jasser** – اسم الطالب: **محمد مروان ناجي أبو جاسر**

Signature: التوقيع:

Date: التاريخ:

بسم الله الرحمن الرحيم

The Islamic University – Gaza
Deanery of Graduate Studies
Faculty of Information Technology

الجامعـــــة الإســـــلامية – غـــــزة
عمـــــادة الدراســـــات العليـــــا
كليـــــة تكنولوجيـــــا المعلومـــــات

Ontology-Based Arabic Documents Classification

By:

Mohammed M. Abu Jasser

Supervised by:

Dr. Rebhi S. Baraka

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master in Information Technology

Jumada I 1436 H / March 2015 AD

بِسۡمِ ٱللَّهِ ٱلرَّحۡمَٰنِ ٱلرَّحِيمِ

الجامعة الإسلامية – غزة
The Islamic University - Gaza

مكتب نائب الرئيس للبحث العلمي والدراسات العليا هاتف داخلي: 1150

الرقم: ج س غ/35/ .... Ref

التاريخ: 2015/04/07م .... Date

نتيجة الحكم على أطروحة ماجستير

بناءً على موافقة شئون البحث العلمي والدراسات العليا بالجامعة الإسلامية بغزة على تشكيل لجنة الحكم على أطروحة الباحث/ **محمد مروان ناجي أبوجاسر** لنيل درجة الماجستير في **كلية *تكنولوجيا المعلومات*** **برنامج تكنولوجيا المعلومات** وموضوعها:

تصنيف المستندات العربية استنادًا إلى الأنطولوجيا
Ontology-Based Arabic Documents Classification

وبعد المناقشة التي تمت اليوم الثلاثاء 18 جمادى الآخر 1436هـ، الموافق 2015/04/07م الساعة الحادية عشرة صباحاً، اجتمعت لجنة الحكم على الأطروحة والمكونة من:

د. ربحي سليمان بركة	**مشرفاً ورئيساً**	....
أ.د. علاء مصطفى الهليس	**مناقشاً داخلياً**	....
د. يوسف نبيل أبو شعبان	**مناقشاً خارجياً**	....

وبعد المداولة أوصت اللجنة بمنح الباحث درجة الماجستير في **كلية *تكنولوجيا المعلومات*/ برنامج تكنولوجيا المعلومات.**

واللجنة إذ تمنحه هذه الدرجة فإنها توصيه بتقوى الله ولزوم طاعته وأن يسخر علمه في خدمة دينه ووطنه.

والله ولي التوفيق ،،،

مساعد نائب الرئيس للبحث العلمي والدراسات العليا

أ.د. فؤاد علي العاجز

بسم الله الرحمن الرحيم

﴿ سُبْحَانَكَ لَا عِلْمَ لَنَا إِلَّا مَا عَلَّمْتَنَا
إِنَّكَ أَنتَ الْعَلِيمُ الْحَكِيمُ ﴾ (سورة البقرة، الآية: 32)

صدق الله العظيم

To my beloved parents

To my brothers and sisters

To my dear wife and children

To those who gave me support

To all of them I dedicate this work

Acknowledgment

First of all, I thank Allah for giving me the strength and courage to persevere throughout the duration of this thesis and made all of this and everything else possible.

I am greatly indebted to my family: my dear father, my dear mother, my wife, my children (Marwan, Yamen, Rakan and Kenan) and my brothers and sisters for their encouragement and support during my course studies and during my thesis work .

I am deeply grateful to my supervisor Dr. Rebhi S. Baraka for his continued encouragement, unlimited efforts, persistent motivation, support, and great knowledge throughout my thesis, without his help, guidance, and follow-up, this thesis would never have been.

Last but not least, I extend my thanks to all my friends.

Abstract

Automatic documents classification is an important task due to the rapid growth of the number of electronic documents. Classification aims to assign the document to a predefined category automatically based on its contents. In general, text classification plays an important role in information extraction and summarization, text retrieval, question answering, e-mail spam detection, web page content filtering, and automatic message routing.

Most existing methods and techniques in the field of documents classification are keyword based without many features. Even methods that ontology-based classification is limited to English language support.

In this research, we propose an approach to investigate the role of ontology (an Arabic news domain ontology) in Arabic documents classification. The results show that the proposed ontology-based approach achieves improvement in the process of documents classification based on the different evaluation criteria. The experimental results show that the accuracy of the approach is 92%. These results prove that the ontology contribute effectively in the process of Arabic documents classification.

Keywords: Arabic Language, Text Mining, Documents Classification, Documents Annotation, Ontology, News Ontology.

الملخص العربي Arabic Abstract

تصنيف المستندات العربية استناداً إلى الأنطولوجيا

التصنيف التلقائي للمستندات ذو أهمية عالية نظراً للنمو السريع في عدد الوثائق الإلكترونية، ويهدف التصنيف إلى تحديد فئة المستند طبقاً لإحدى الفئات المحددة مسبقاً بشكل تلقائي بناءً على محتوياته. بشكل عام، تصنيف النصوص يلعب دورا هاماً في استخراج المعلومات وتلخيص النصوص واسترجاعها، وكذلك الإجابة على الأسئلة، والكشف عن البريد المزعج في خدمات البريد الإلكتروني، وبرامج تحليل وفلترة المحتويات للصفحات الإلكترونية، والتوجيه التلقائي للرسائل الإلكترونية.

وتعتمد معظم الطرق والأساليب القائمة في مجال تصنيف المستندات على التصنيف المعتمد على الكلمات الموجودة في محتوى تلك المستندات دون النظر إلى المعنى الدلالي لتلك الكلمات، وحتى الأساليب التي تستخدم التصنيف على أساس الأنطولوجيا تقتصر على دعم اللغة الإنجليزية.

هذا البحث يقدم طريقة للتحقق من دور الأنطولوجيا في تصنيف الوثائق العربية، والتي تم بناؤها في مجال الأخبار ، حيث أظهرت النتائج أن الطريقة المستخدمة والمستندة إلى الأنطولوجيا تحقق تحسناً في عملية تصنيف الوثائق حسب معايير التقييم المختلفة. أظهرت النتائج دقة في التصنيف بنسبة 92%، وبذلك يتبين أن استخدام الأنطولوجيا يساهم بشكل فعال في عملية تصنيف الوثائق العربية.

الكلمات الدلالية: اللغة العربية، تنقيب البيانات، تصنيف المستندات، الأنطولوجيا، أنطولوجيا الأخبار.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CPSL	Common Pattern Specification Language
DAML	DARPA Agent Markup Language
DC	Document Classification
GATE	General Architecture For Text Engineering
JAPE	Java Annotation Pattern Engine
KDD	Knowledge Discovery from Data
KNN	K-Nearest Neighbors
LR	Language Resource
OIL	Ontology Inference Layer
OWL	Web Ontology Language
PR	Processing Recourse
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
SVM	Support Vector Machines

Chapter 1: Introduction

This chapter introduces the thesis by describing the document classification, the thesis problem, the research objectives, importance of the research, the scope and limitations of the thesis work, the methodology, and finally the thesis structure.

The rapid growth of Internet has increased the number of online documents. Manual classification of documents is time consuming and expensive, which makes it infeasible for handling the huge number of Internet documents [60]. This has led to the development of automated text and document classification systems that are capable of automatically classifying documents [34].

Document classification (categorization) is the process of classifying documents into a predefined set of categories based on their content [34]. A document must pass through preprocessing steps including conversion of document into plain text, stop word removing, and words must be stemmed after stop words removal [32] [36]. Then the document is passed to the classification system.

Arabic language is a Semitic language that has a complex and much morphology than English; it is a highly inflected language. This complex morphology needs a set of preprocessing routines to be suitable for classification [32] [66].

In traditional machine-learning algorithms a classifier is automatically built by learning the characteristics of the categories from a set of reclassified documents, then the classifier is used to classify documents into predefined categories. However, in order to train classifier, human must collect large number of training text terms, which considered a hard process [10].

In addition, most of these traditional methods are keyword based without sufficient features. They do not considered the semantic relations between words, therefore, providing inaccurate results and limited accuracy. In various applications, intelligent text classification systems apply the computational knowledge model, such as ontology, to enhance the classification algorithms.

Ontology is a conceptual description, including concept, attribute, entity and association description with the main purpose of knowledge sharing and reusing knowledge [22]. Ontology-based text classification methods consider the semantic relations between the terminology information extracted from texts and ontology categories, thus improve the performance, in terms of its accuracy, over traditional approaches to gain effectiveness amid current information environment [45] [12].

In this research, we develop an ontology-based document classification approach to support the process of Arabic document classification. The current classification systems, approaches and traditional methods is studied. A news domain ontology is constructed. A prototype of the proposed classification approach which depends on the constructed Arabic news ontology is implemented. Using the proposed approach the documents is annotated with the constructed ontology components, then a classifier is built to classify the documents based on these annotations. A series of experiments on a set of documents is conducted and results are collected. Evaluation is performed based on the results and shows that the proposed ontology-based approach achieves a significant improvement in the process of documents classification based on the basic evaluation measures: accuracy, precision, recall and F-measure.

1.1. Statement of the problem

In traditional classification algorithms, human needs to train the classifier. Most of these algorithms have not considered the semantic relations between words. In order to solve these problems, some ontology-based classification methods are researched. These methods use ontologies, which consider the semantic relation to represent characteristics of the categories, so documents are classified with more understanding leading to improve accuracy of the classification. However, work in ontology-based classification is limited and does not support the Arabic language.

Therefore, the problem of this research is to investigate the role of ontology in Arabic documents classification.

1.2. Objectives

1.2.1 Main Objective

To investigate the role of ontology in Arabic documents classification.

1.2.2 Specific Objectives

The specific objectives of this research are:

- Gain insight and vision of current document classification systems and approaches to determine its advantages, shortcomings and requirements, through studying and analysis.
- Construct a valid domain specific ontology which includes domain terms, relations, properties and rules.
- Develop an approach that uses the constructed ontology to perform the document classification process.
- Implement a prototype of the ontology-based classification approach.
- Evaluate the approach for accuracy based on a chosen evaluation strategy.
- Explore the role of ontology on documents classification.

1.3. Importance of the research

- The importance of this research stems from the rapid growth of web documents and the need of classifying them to improve knowledge acquisition.
- Little work carried out on researches in the field of semantic web especially on Arabic language; therefore, by this research we can contribute in Arabic ontology building.
- Document classification supported by ontology produce more understandable results through reasoning.

- Classification also can be used in document indexing, e-mail filtering, web browsing, and personal information agents.
- The developed ontology can also be used as a basis for other applications.

1.4. *Scope and Limitation of the Research*

- The classification process depends on documents that exist in a corpus but collected from the web.
- The classification process focuses only on documents in Arabic language.
- A domain specific ontology is constructed. Namely, from Arabic news domain.
- A prototype rather than a system is implemented to provide a proof of concept for the proposed approach.

1.5. *Methodology*

To accomplish the objectives of the research, the following methodology is followed:

phase 1. Studying and analyzing current systems and applications that are related to document classification in both Arabic and English language.

phase 2. Constructing the domain ontology, which involves the following steps:

- Decide whether construction of ontology has to be done from scratch, reuse an existing ontology, use English to Arabic mapping, and or automatic ontology building.
- Specify the scope and purpose of the ontology. Also reveals the relationship among classes and subclasses.
- Validate the ontology through reasoning.

phase 3. Developing an approach that uses the constructed ontology in the document classification process including:

- Specify the requirements of the approach.

- ▪ Specify the interaction between its components.

phase 4. Implementing a prototype of the proposed approach using GATE software and JAPE rules for documents annotation, Eclipse for code editing, and JENA API for programming.

phase 5. Evaluate the approach and verify that the goal is achieved. There are different measures that we can use to evaluate the classification approach. The basic measures that we can use are: accuracy, precision, recall and F-measure

1.6. Thesis Structure

This thesis consists of six chapters organized as follows:

- Chapter 1 (Introduction): introduces the thesis problem, objectives, importance, limitations, methodology, and finally the thesis structure.
- Chapter 2 (Background and State of the Art): focuses on the background and theoretical concepts related to the document classification, Arabic DC, semantic web, and semantic annotation.
- Chapter 3 (Related Work): reviews the related work in the domain of Arabic document classification, Arabic ontology building, and the used of ontology in DC.
- Chapter 4 (The Proposed Classification Approach): present the steps of implementing the methodology. It describes the proposed approach, the construction of domain ontology, the process of documents annotation and category assignments.
- Chapter 5 (Experimental Results and Evaluation): presents an evaluation of the proposed approach and discusses the results.
- Chapter 6 (Conclusion and Future Work): concludes the thesis and presents future directions.

Chapter 2: Background and State of the Art

This chapter provide a brief introduction about *data mining*, *documents classification*, and its methods, then introduces the classification in Arabic documents. It also introduces the semantic web and its technologies, ontology, ontology building, semantic annotation, and their roles in documents classification.

2.1. Data Mining

Data mining refers to extracting or "mining" knowledge from large amounts of data [33]. It's a process of nontrivial extraction of implicit, previously unknown and potentially useful information [48]. Many other terms carry a similar or slightly different meaning to data mining, such as *knowledge mining from data*, *knowledge extraction*, *data/pattern analysis*, *data archaeology*, and *data dredging*.

The process of discovering patterns in data must be automatic or (more usually) semiautomatic. The patterns discovered must be meaningful in that they lead to some advantage, usually an economic advantage. The data is invariably present in substantial quantities [64].

Many people treat data mining as a synonym for another popularly used term, *Knowledge Discovery from Data*, or KDD. Alternatively, others view data mining as simply an essential step in the process of knowledge discovery.

Knowledge discovery as a process is depicted in Figure 2.1 and consists of an iterative sequence of the following steps [33]:

1. Data cleaning (to remove noise and inconsistent data)
2. Data integration (where multiple data sources may be combined)
3. Data selection (where data relevant to the analysis task are retrieved from the database)

4. Data transformation (where data are transformed or consolidated into forms appropriate for mining by performing summary or aggregation operations, for instance)
5. Data mining (an essential process where intelligent methods are applied in order to extract data patterns)
6. Pattern evaluation (to identify the truly interesting patterns representing knowledge based on some interestingness measures)
7. Knowledge presentation (where visualization and knowledge representation techniques are used to present the mined knowledge to the user)

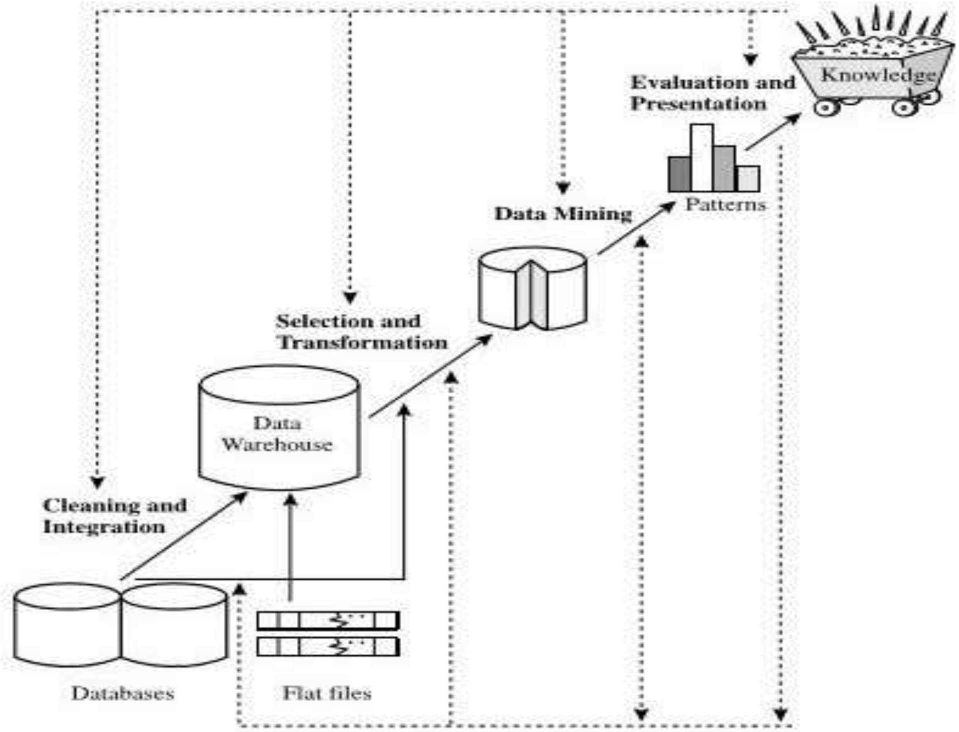


Figure 2.1 Data mining as a step in the process of knowledge discovery [33]

Data mining functionalities are used to specify the kind of patterns to be found in data mining tasks. In general, data mining tasks can be classified into two categories: descriptive and predictive. Descriptive mining tasks characterize the general properties of the data in the database. Predictive mining tasks perform inference on the current data in order to make predictions [33].

The most famous data mining tasks is *description, estimation*, *prediction, classification, clustering*, and *association* [33]. Perhaps the most common one is that of *classification*. Which is the focus of our research.

2.2. Document Classification

Document classification is the task of classifying natural language documents into a predefined set of categories based on their content [31]. This assignment can be used for classification, filtering, and retrieval purposes. As the first and a vital step, text representation converts the content of a textual document into a compact format so that the document can be recognized and classified by a computer or a classifier [11].

Documents Classification Methods

A number of classifiers have been proposed in the last few years, and a wide variety of techniques have been designed for text classification. The documents can be classified by three ways unsupervised, semi supervised and supervised classification. Some key methods, which are commonly used for text classification are as follows:

- **Probabilistic Methods**: Probabilistic methods are the most fundamental among all data classification methods. Probabilistic classification algorithms use statistical inference to find the best class for a given example. In addition to simply assigning the best class like other classification algorithms, probabilistic classification algorithms will output a corresponding posterior probability of the test instance being a member of each of the possible classes. The posterior probability is defined as the probability after observing the specific characteristics of the test instance. On the other hand, the prior probability is simply the fraction of training records belonging to each particular class, with no knowledge of the test instance. After obtaining the posterior probabilities, we use decision theory to determine class membership for each new instance [1].

- **Decision Trees**: Decision trees are designed with the use of a hierarchical division of the underlying data space with the use of different text features. The hierarchical division of the data space is designed in order to create class partitions which are more skewed in terms of their class distribution. For a given text instance, we determine the partition that it is most likely to belong to, and use it for the purposes of classification [42].
- **Pattern (Rule)-based Methods**: Rule-based methods are closely related to decision trees, except that they do not create a strict hierarchical partitioning of the training data. Rather, overlaps are allowed in order to create greater robustness for the training model. Any path in a decision tree may be interpreted as a rule, which assigns a test instance to a particular label [1].
- **Instance-Based Learning**: In instance-based learning, the first phase of constructing the training model is often dispensed with. The test instance is directly related to the training instances in order to create a classification model. Such methods are referred to as lazy learning methods, because they wait for knowledge of the test instance in order to create a locally optimized model, which is specific to the test instance. The advantage of such methods is that they can be directly tailored to the particular test instance, and can avoid the information loss associated with the incompleteness of any training model. An example of a very simple instance-based method is the nearest neighbor classifier [1].
- **SVM Classifiers**: SVM (Support Vector Machines) Classifiers attempt to partition the data space with the use of linear or non-linear delineations between the different classes. The key in such classifiers is to determine the optimal boundaries between the different classes and use them for the purposes of classification. The major downside of SVM methods is that they

are slow [42]. However, they are very popular and tend to have high accuracy in many practical domains [1].

- **Neural Network Classifiers**: Neural networks are used in a wide variety of domains for the purposes of classification. In the context of text data, the main difference for neural network classifiers is to adapt these classifiers with the use of word features. Neural network classifiers are related to SVM classifiers; indeed, they both are in the category of discriminative classifiers, which are in contrast with the generative classifiers [1].
- **Bayesian Classifiers**: Bayesian classifiers (also called generative classifiers), attempt to build a probabilistic classifier based on modeling the underlying word features in different classes. The idea is then to classify text based on the posterior probability of the documents belonging to the different classes on the basis of the word presence in the documents [1].

Of course there are many other classification methods, we cannot enumerate and summarize all of them. Most of these traditional methods are keyword based without many intelligent features, it have not considered the semantic relations between words, providing inaccurate results. Therefore, it is difficult to improve the accuracy. In various applications, intelligent text classification system applies the computational knowledge model, such as ontology, to enhance the classification algorithms [65].

2.3. Arabic and Document Classification

Arabic is one of the most widely used languages in the world. It’s the mother language of more than 300 million people. Unlike Latin-based alphabets, the orientation of writing in Arabic is from right to left; there are 28 characters in Arabic. The characters are connected and do not start with capital letter as in English. Furthermore, most of the characters differ in shape based on their position in the sentence and adjacent letters [6] [51] [58].

Arabic words have two genders, feminine and masculine; three numbers, singular, dual, and plural; and three grammatical cases, nominative, accusative, and genitive. A noun has the nominative case when it is subject; accusative when it is the object of a verb; and the genitive when it is the object of a preposition. Words are classified into three main parts of speech, nouns (including adjectives and adverbs), verbs, and particles [57].

Arabic language has a very complex morphology, words have affluent meanings and contain a great deal of grammatical and lexical information [63]. Majority of words have a tri-letter root. The rest have a quad letter root, penta-letter root or hexa-letter root. In addition, Arabic scripts do not use capitalization for proper nouns, which are necessary in classifying documents [34].

Arabic is a challenging language for a number of reasons [58]:

1. Orthographic with diacritics is less ambiguous and more phonetic in Arabic, certain combinations of characters can be written in different ways.
2. Arabic has a very complex morphology recording as compare to English language.
3. Broken plurals are common. Broken plurals are somewhat like irregular English plurals except that they often do not resemble the singular form as closely as irregular plurals resemble the singular in English. Because broken plurals do not obey normal morphological rules, they are not handled by existing stemmers.
4. In Arabic we have short vowels which give different pronunciation. Grammatically they are required but omitted in written Arabic texts.
5. Arabic synonyms are so many.

A few researchers have applied a number of classification approaches which are solely applicable for the problem of Arabic text classification. Some of these works is listed on the next chapter. Researchers conclude that the Arabic text documents are required extensive pre-processing routines to build accurate classification systems.

2.4. Semantic Web and Documents Classification

The representation of text as a bag-of-words has been disadvantaged by the ignorance of any relationship between terms thus the importance of the integration of a semantic web technologies is to improve the process of document classification by solving the problem of words ambiguity.

The semantic Web would be an extension of the current one, in which information would be given well-defined meaning, better enabling computers and people to work in cooperation. Its aim is to allow much more advanced knowledge management [71].

Figure 2.2 illustrate the layered architecture of the semantic Web (due to Tim Berners Lee).

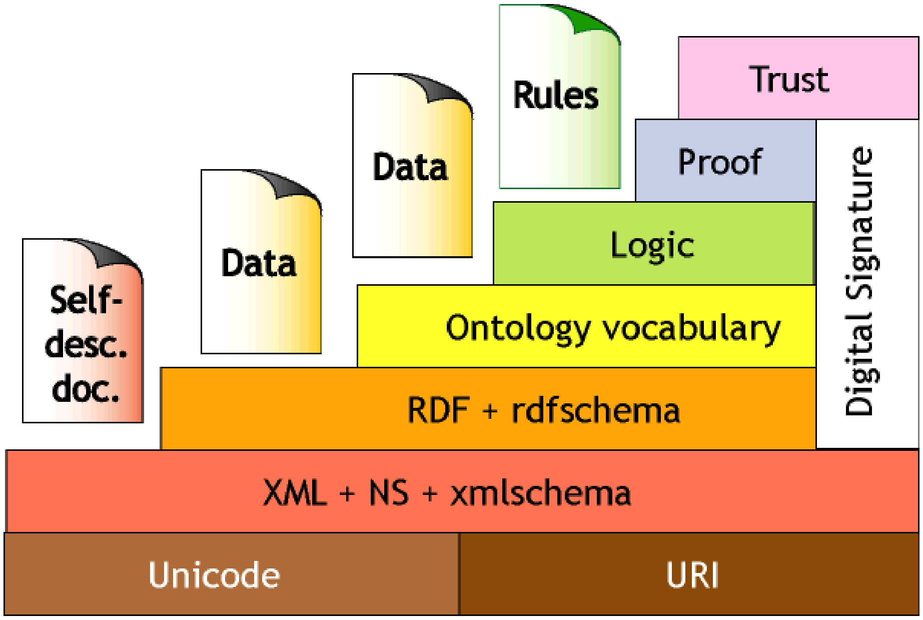


Figure 2.2 Semantic Web Stack [19]

The architecture includes UNICODE and URI (Uniform Resource Identifier), XML (Extensible Markup Language), NS (Name Space), XMLSchema, RDF (Resource Description Framework), RDFS (Resource Description Framework Schema), Ontology, Logic, Proof, and Trust.

At the bottom we find XML, a language that lets one write structured Web documents with a user-defined vocabulary. XML is particularly suitable for sending documents across the Web. RDF is a basic data model, like the entity-relationship model, for writing simple statements about Web objects (resources). RDF Schema provides modeling primitives for organizing Web objects into hierarchies. Key primitives are classes and properties, sub class and sub property relationships, and domain and range restrictions [7].

RDF Schema can be viewed as a primitive language for writing ontologies. But there is a need for more powerful ontology languages that expand RDF Schema and allow the representations of more complex relationships between Web objects. The Logic layer is used to enhance the ontology language further and to allow the writing of application-specific declarative knowledge [19].

The Proof layer involves the actual deductive process as well as the representation of proofs in Web languages (from lower levels) and proof validation. Finally, the Trust layer will emerge through the use of digital signatures and other kinds of knowledge, based on recommendations by trusted agents or on rating and certification agencies and consumer bodies [7].

The semantic web extends the original web with technologies that provide syntactic structure and semantic meaning permitting the development of taxonomies and inference rules. When combined together as description logics languages, they enable the compilation of knowledge representations or information collections known as ontologies [62].

2.3.1 Ontology

An ontology is a basic building block for the Semantic Web, it's a formal, explicit specification of a shared conceptualization [67]. "*Formal*" refers to the fact that the ontology should be machine understandable. "*Explicit*" means that the type of concepts used and the constraints on their use are explicitly defined. "*Shared*" reflects the notion that an ontology captures consensual knowledge, that is, it is not restricted to some individual but accepted by a group. A "*conceptualization*" refers to an abstract model of some phenomenon in the world that identifies the relevant concepts of that phenomenon [28].

An ontology can be viewed as a declarative model of a domain that defines and represents the concepts existing in that domain, their attributes and relationships between them [67]. It plays a vital role in the semantic web and tries to capture the semantics of a domain by deploying knowledge representation primitives, enabling a machine to understand the relationships between concepts in a domain. It is typically represented as a knowledge base which then becomes available to applications that need to use and/or share the knowledge of a domain [68], which the news domain in our thesis.

Ontologies specifies a set of constraints that declare what should necessarily hold in any possible world. It used to identify what "is" or "can be" in the world. It is the intention to build a complete world model for describing the semantics of information exchange. Especially in the area of artificial intelligence, ontologies are being used to facilitate knowledge sharing and reuse [21].

Ontology is comprised of concepts, properties, relationships between concepts and constraints. Figure 2.3 represents a simple ontology also called lightweight ontology containing classes and its taxonomical relations.

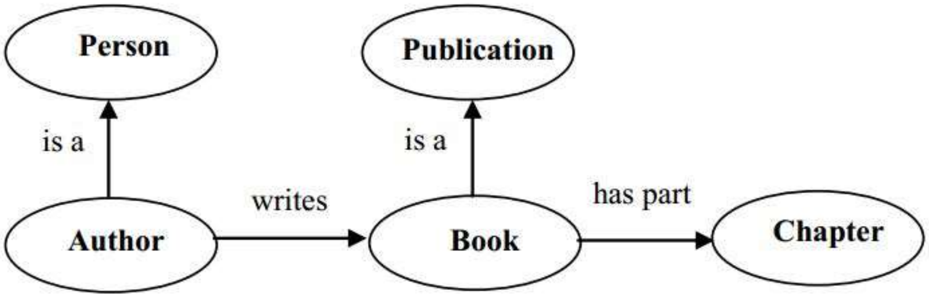


Figure 2.3 Example of a small ontology [2]

Based on Language Expressivity and Formality there are *Informational*, *Linguistic/Terminological*, and *Formal* ontologies [26], linguistic ontologies are large-scale lexical resources that cover most words of a language and have a hierarchical structure based on the relations between concepts. These ontologies can cover specific or general domains. Their top levels are described by the primitives that are the generic terms that include other terms [20]. An example of a primitive is computer science that includes software, hardware, networks, and so forth.

In natural language texts, the meaning of a term is usually not defined explicitly, but strongly depends on the context in which the term occurs. Humans are able to disambiguate them using their knowledge about the context the term is used in. Current automatic disambiguation approaches fail frequently due to missing commonsense knowledge or appropriate ontology models [20].

The advantages of an ontology-based classification approach over the existing ones are that the nature of the relational structure of ontology provides a mechanism to enable machine reasoning; also, the conceptual instances within ontology are not only a bag of keywords but have inherent semantics and a close relationship with the class representatives of the classification schemes [46].

Ontology Building

The creation of ontologies presents a tedious task, because it requires specialized skills and involves various stakeholders with personalized, depends on a variety of factors (such as software building tool, the implementation language, the development methodology, the applications in which the ontology will be used, the type of the ontology under construction, the available informal and formal existing knowledge resources, etc.) [59].

In addition, there is no single correct way to model a domain, there are always viable alternatives. Therefore, there is no one correct way to develop an ontology [13], but the quality of the solution depends on the skills of the people who will participate in the ontology development process. Several research groups have proposed various methodologies for building ontologies. The skilled knowledge engineer can look up the different methodologies before selecting, or adapting one that fits his needs [59].

Several methodologies for ontology building have been reported, includes *Cyc*, *Uschold and King's* method, *KACTUS*, *Methontology*, *SENSUS*, *On-to-Knowledge*, *Grüninger and Fox*, *TOVE*, *CommonKADS*, *DILIGENT* [29] [16] [61]. The most complete ones are Methontology and On-to-Knowledge. All these methodologies are composed of several activities. The development process is not a linear process but a refinement one where each activity can be repeated several times. Among all the activities the most important are: Ontology specification, Knowledge acquisition, Conceptualization, Formalization, Implementation, Evaluation, Maintenance, and Documentation [26].

In our research we used the manual construction method for constructing the News ontology, and we follow mainly *Noy and McGuinness* [52], which consist of 7 steps as illustrated in table 2.1 below.

Table 2.1 Construction processes of domain ontology [37]

Step	Process
1.	Determine the domain and scope of the ontology.
2.	Consider reusing existing ontologies.
3.	Enumerate important terms in the ontology.
4.	Define the classes and the class hierarchy.
5.	Define the properties of classes (slots).
6.	Define the facets of the slots.
7.	Create instances.

Like any development process, ontology construction is iterated, not a linear, and may be necessary to backtrack to earlier steps [7].

All the previous methods and methodologies were proposed for building ontologies. However, many other methods and methodologies have been proposed for other tasks, such as ontology reengineering, ontology learning, ontology evaluation, ontology evolution, ontology merging, etc. [16].

- ***Ontology Building Tools***

There are a lot of software tools related to semantic web. Many ontology editors could be found on internet. Some of them (like: Apollo, OntoStudio, Protégé, Swoop, and TopBraid). All these tools are popular in the ontology design and development sector. They are accepted by relatively large semantic web communities [2].

These tools can be applied to several stages of the ontology life cycle including the creation, implementation, and maintenance of ontologies. It’s used for building a new ontology either from scratch or by reusing existing one, which usually supports editing, browsing, documentation, export and import from different formats, and they may have attached inference engines [13].

Ontology Building Languages

There are several languages used in ontology building like XML, RDF/RDFS DAML (DARPA Agent Markup Language) + OIL (Ontology Inference Layer) and OWL (Web Ontology Language). Many ontology tools have been developed for implementing metadata of ontology using these languages.

XML provides syntax for structured documents, without semantic meaning constraints on the documents. RDF is a data model for representing objects and relations between them. It provides simple semantics for the model and can be represented in XML syntax. RDF-Schema is a language for defining vocabulary for describing properties and classes of RDF resources. RDFS is used to define graphs of trio RDF, with semantics of generalization/prioritization of such properties and classes. OWL adds vocabulary for describing properties and classes, relations between classes (e.g. disjointness), cardinality and characteristics of properties (e.g. symmetry). OWL is developed as an extension of RDF vocabularies, and it is derived from the ontology DAML + OIL [2].

Ontology Evaluation

In order to build high quality ontologies, ontology evaluation technologies are needed. The primary goal of these evaluation methods is to prevent applications from using inconsistent or incorrect ontologies [48].

A variety of researches of ontology evaluation have been established depends on the perspective of what should be evaluated. Most of them focus on the evaluation of the whole ontology; others focus on partial evaluation of the ontology, for reuse it in an ontology engineering task [14].

Ontology evaluation can be divide basis of: *Corpus-based evaluation*: is used to estimate empirically the accuracy and the coverage of the ontology. *Gold-Standard-based evaluation*: that compares candidate ontologies to gold-standard ontology that serves as a reference. *Task-based evaluation*: looks at how the results of the ontology-based application are affected by the use of an ontology. *Expert-based evaluation*: where ontologies are presented to human experts who have to judge in how far the developed ontology is correct. *Criteria-based evaluation*: measures in how far an ontology adheres to desirable criteria [14].

There are various methodologies to evaluate ontologies; most of them based on one of the following categories:

- Fitting or coverage techniques between an ontology and a domain of knowledge that the ontology is created for.
- The effort done by human experts who try to assess how well the ontology meets a set of predefined criteria, standards, and requirements.
- Using the ontology in the context of an application or project to evaluate its effectiveness. The use of the system may reveal weakness or strength points in the ontology.
- Comparing the ontology with other ontologies in the same domain.
- Studying ontology relationships considering some measures.
- Studying and comparing the formal representation of the ontology with other ontologies formal representations, criterions, or measures [41].

2.3.2 Semantic Annotation

The construction of metadata which annotates the documents is one of the major tasks for making data understandable to the machine on the Semantic Web. Semantic annotation is the process of inserting metadata, which are concepts of an ontology (i.e. classes, instances, properties and relations), in order to assign semantics [53] [15], and tells us what exactly the concepts annotated mean in the context. The traditional process of annotation is to associate text with the corresponding element model. The reference process can be implemented by calculating the similarity degree of concepts [67].

Annotating data can help provide better classification facilities, since queries will be based not only on traditional keywords, but also well-defined concepts described by the domain ontology [53].

To achieve that and create this annotations, technologies should be available for a common format, such as RDF, which allows anyone to make statements about any resource using an XML-based syntax [15].

Manual annotation is difficult, slow, time-consuming, tedious and costly [15]. It's neither efficient nor practicable. The only way to bypass all these drawbacks is to automate the process of semantic annotation [39].

Many systems that can lead the process automatically have been proposed to automate the generation of semantic annotations [53]. Examples are GATE, SHOE Knowledge Annotator, AeroDAML, AKTiveMedia, GoNTogle, KIM, Magpie, MnM, Annozilla, SMORE, SemanticWord, Melita, Yawas, OntoAnnotate, OntoMat Annotizer etc. [15][53].

These tools are designed to enable users with limited knowledge of semantic Web technologies such as RDF, OWL to markup documents with semantics. With these tools, authoring linked data is mainly a matter of dragging in data and binding it together through ontology using a graphical interface [15].

2.5. Summary

In this chapter, we have presented a background for this research. We have discussed the document classification process, and reviewed the commonly used traditional classification methods, which are keyword based without many intelligent features. So we had to address the concept of Semantic Web and its technologies, the ontology and its building process, the ontology building tools and languages, and the ontology evaluation. Then we presented the semantic annotation and how we can use ontology in the process of annotation and the importance of it in the field of document classification.

Chapter 3: Related Work

This chapter presents a review of existing related work in the field of documents classification in general, and classifying Arabic documents in particular, also introduces and analyzes the related work in the field of using ontology and semantic annotation in the field of document classification.

3.1. Arabic Documents Classification

Because the subject of documents classification is linked to many fields, and the solution to this problem contributes in development of many fields and applications. Many researchers have applied several researches based on a number of classification approaches, models and techniques. We illustrates some of these works below.

El-Halees [23] uses maximum entropy, which is a probabilistic supervised learning method for text classification on Arabic data sets collected from Aljazeera Arabic news website. His method focuses in natural language processing techniques to preprocess the documents before applying the classification. The author has tested the proposed method with and without pre-processing phase. Using the preprocessing techniques increases the f-measure from 68.13% to 80.41%.

Harrag et al. [35] develop an Arabic text classification model using neural network and singular value decomposition method. Introducing the singular value decomposition method improved the categorization performance, the reduced size of the vectors also decreased the computational time in the back-propagation neural network. The method achieves performance of 88%.

Hadi [31] applies a methodology in a proposed expert system based on learning the rules from the database rather than inputting the rules by the knowledge engineer from the domain expert, so the accuracy and the processing time are improved. The proposed automated expert system contains a learning method based on Association Classification mining. Five different classification approaches: Decision trees (C4.5, KNN, SVM, MCAR and NB) and the proposed automated expert system, which called EMCAR, have been tested on the Islamic data set to determine the suitable method in classifying Arabic texts. The basis of the comparison in the experimentation are different text evaluation metrics, including error-rate, precision, and recall. The results indicated that the least applicable learning algorithm towards the chosen Arabic data set is KNN. Moreover, the most applicable algorithm to the Arabic data set is EMCAR in which it derived higher results in all evaluation.

Elberrichi et al. [25] use Arabic WordNet as a lexical and semantic resource for categorizing Arabic texts. To comprehend its effect, they incorporate it in a comparative study with the other usual modes of representation (bag of words and N-grams), and they use the K-Nearest Neighbors learning scheme with different similarity measure. Their result show the benefits and advantages of this representation compared to the more conventional method, and demonstrate that the addition of the semantic dimension is promising ways for the automatic categorization of Arabic texts.

El-Monsef et al. [24] propose a method that consider a specific word related to the field called Field Association words by considering their ranks or levels. They built a Java software system to make classification on Arabic text using keyword, FA words and compound FA words. Furthermore, a comparative study of keywords, FA words and compound FA words on Arabic text were done using experimental results generated by their software. The methods estimated by simulation results of 1819 files and 16 super fields. And the classification results, F-measure is 72% of classification using FA words, F-measure is 82% of classification using compound FA words.

Al-Harbi et al. [57] investigate the effect of stemming for improving Arabic text categorization. Three stemming techniques (root based stemming, ETS2 stemming, and light stemming) were employed. Their performance was assessed in text classification exercises for an Arabic corpus to compare and contrast the effect of these Arabic stemming algorithms on improving text mining. The results showed significant classification accuracy improvement when using the ETS2 stemmer, accuracy varies among classes between 73.33% and 82.19%.

Also **Saad et al.** [58] study the impact of text pre-processing and different term weighting schemes on Arabic text classification. They develop a new combinations of term weighting schemes to be applied on Arabic text for classification purposes. They use C4.5 decision tree algorithm to classify their Arabic dataset. Empirical results showed Term stemming and pruning, document normalization, and term weighting dramatically reduce dimensionality, enhance text representation and directly impact text mining performance.

3.2. Ontology and Documents Classification

De Luca et al. [20] present a search system that uses ontologies to classify search results online in order to disambiguate result sets with respect to given search terms. Thus, the user can select directly a subset of the search results (“folder of sense”) which reflects his search context without the need to scan the list of all retrieved documents.

Mu-Hee et al. [46] focus on document classification based on the similarities of documents already categorized by ontology using terminology information extracted from the documents. The document classification technique proposed by this paper does not involve any learning processes or experimental data and can be performed in real time. Their classification results, the precision, recall, and F1 measures 89.68%, 95.43%, 92.39%, respectively. And the F1 measurements is compared with TF-IDF and Bayesian method which got 79.87% and 82.45%.

Zakaria et al. [70] propose a medical document classification method based on Mesh (Medical Subject Headings) domain ontology, in their proposed method they uses a mapping terms to concepts strategy, to enrich the representation vector, to reduce its dimensions. The approach was tested with the KNN and the C4.5 and the result have a significant performance upgrading of 30%.

Marina et al. [47] present an ontology-based web content mining methodology that contains such main stages as collecting a training set of labeled documents from a given domain, building a classification model above this domain given the domain ontology, and classification of new documents via the induced model. They tested the proposed methodology in a specific domain, namely web pages containing information about production of certain chemicals. Using their methodology, they are interested to identify all relevant web documents while ignoring the documents that do not contain any relevant information. Their system receives as input an OWL file built in Protégé tool, which contains the domain-specific ontology, and a set of web documents classified by a human expert as ”relevant” or ”non-relevant”. They use a language-independent key-phrase extractor with integrated ontology parser for creating the database from input documents and use it as a training set for the classification algorithm. The system classification accuracy using various levels of ontology is evaluated. The current version of their system supports web content mining in English, Arabic, Russian, and Hebrew languages.

Prabowo et al. [54] propose an automatic classifier, which focuses on the use of ontologies for classifying Web pages. They use shared classes to define a set of common class representatives, and to link the class representatives with their associated domain ontologies. Then to correctly identify the main topic of a Web page a term weighting strategy is applied. The experiment results show a statistical improvement in terms of accuracy.

Jian et al. [46] present an ontology-based text-mining methods for grouping of research proposals. A research ontology is constructed to categorize the concept terms in different discipline areas and to form relationships among them. It facilitates text-mining and optimization techniques to cluster research proposals based on their similarities and then to balance them according to the applicants' characteristics. The experimental results at showed that the proposed method improved the similarity in proposal groups, as well as took into consideration the applicants' characteristics. Also, the proposed method promotes the efficiency in the proposal grouping process.

Fang et al. [27] propose an ontology-based web documents classification and ranking method. Firstly, weighted terms set are extracted from web documents, and ontology is build up by clarifying and augmenting an existing ontology; then similarity score between documents and ontology is computed based on WordNet by using Earth Mover's Distance (EMD) method; finally, web documents are assigned to categories according to the similarity score, and a ranking method is used to sort the documents in the same categories. The experiment result shows that the classification algorithm achieves better precision 82.1% and recall 93.3% compare with adaptive KNN method, and is competitive with SVM method, the ranking method also has good performance. They explained that the main reason for this improvement is that ontology-based method considers semantic relations between words when calculating similarity between documents and ontologies.

From previously related work we notice that there are some studies in the field of document classification use ontology, but they are limited to English language. The results show significant performance in document classification, so that it supports research in the field of document classification using ontology, and encourage to take its advantages to classify Arabic documents.

3.3. Arabic Ontology Building

A few researchers have studied Arabic ontology building, most of these ontologies are domain specific. Researchers in most cases used these ontologies in the field of searching and information retrieval.

Sina institute for knowledge engineering and Arabic technologies [9] work in a project of Arabic ontology, which is a long term project. In [38] they clarify that the project aims to build the upper levels of the Arabic ontology, which forms the basis for the Arabic ontology. They create a database containing approximately thirty thousand Arabic terms and their semantic meaning. Also they create a computer program which works on the link between the concepts of Arab ontology with its corresponding one in English ontology.

Al-Safadi et al. [5] describe the development of an Arabic ontology in a computer technology domain to serve semantic-based search and retrieval of Arabic blogs on the web. They analyze the Arabic language on the web and investigate the existing Arabic support offered by semantic web applications and research. The analysis showed weak support for Arabic language. Thus, the need for developing Arabic domain-based ontologies.

Zaidi et al. [69] describe a web-based multilingual tool for Arabic information retrieval based on ontology in the legal domain. They illustrate the manual construction of the ontology and the way it is edited using Protégé. They identify the legal terms and the semantic relations between them before mapping them onto their position in the ontology.

Aliane et al. [3] present a project of building an ontology centered infrastructure for Arabic resources and applications. They named their project *Al-*

Khalil in the sake of the famous grammarian AL-Khalil Ibn Ahmad Alfarahidi because they consider that he was the first to have built an ontology for the Arabic language trough his book “Kitab al-'Ayn”. The core of the infrastructure is a linguistic ontology that is founded on Arabic Traditional Grammar. The methodology they have chosen consists in reusing an existing ontology, namely the Gold linguistic ontology. They discuss the development of the ontology and present their vision for the whole project. And because of ontology is intended to be a reference for linguists and NLP researchers in different areas of the field, They aim the ontology to contain exhaustive knowledge about standard Arabic, formal and NLP works on Arabic, dialects and linguistic phenomena relating to Arabic, And linking their ontology to projects on Arabic corpus for instance the Algerian Arabic treasury project and building significant applications that use the ontology.

Dalloul [18] builds an Isnad judgment system that automatically generates a suggested judgment of Hadith Isnad based on ontology in the Hadith domain. He illustrate the construction steps of the ontology using Protégé. A prototype of the approach implemented to provide a proof of concept for the requirements and to verify its accuracy. The results prove that the ontology supports the process of Isnad judgment.

Kaloub [40] proposes an approach for enhancing the process of information retrieval for Arabic language that depends on the ontology in prayer jurisprudence domain. He uses GATE software as annotation tool to annotate documents based on the constructed ontology to enhance the process of information retrieval. The results of the approach show significant improvement in the process of documents retrieval depending on the two common evaluation criteria precision and recall.

3.4. Summary

In this chapter we presented a review of some related works in the field of documents classification, in both English and Arabic languages. Then we reviewed related works that use ontology in the classification process. We found that most researches and developments are in English language. Also we reviewed the related works in Arabic ontology building which mostly are domain specific and are used in the field of searching and information retrieval. Finally by reviewing most of works in the field of document annotation we found that most of them is for retrieval purposes, and do not cover the field of document classification.

Chapter 4: The Proposed Classification Approach

This chapter describes and discusses the stages of developing the Arabic documents classification approach which depends on the constructed Arabic news ontology. The ontology building process will be described in details, then the use of the news ontology to annotate documents is illustrated. After that the classifier which will classify the documents based on these annotations will be presented.

4.1. Classification Approach Structure

The ontology-based Arabic document classification approach consists of the following parts as shown in Figure 4.1.

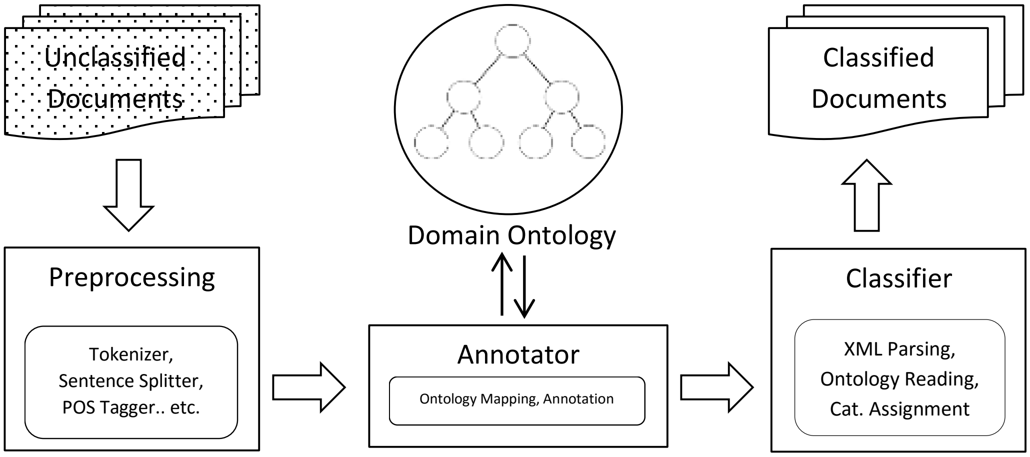


Figure 4.1 Classification Approach Structure

Document classification components:

- **Preprocessing**: The unclassified documents are passed to the system, then a preprocessing stage is performed, which include text tokenizing, sentence splitting, POS tagging, and morphological analyzing.
- **Domain Ontology**: Arabic News ontology which used in the annotation and classification process.

- **Annotator**: which maps terms to the corresponding classes, that is, the sets defining the different meanings of a term and the linguistic relations from the used ontology.
- **Classifier**: which classifies documents depending on the annotation process.

The process of building the classification approach have to be performed based on our methodology. Each stage will be explained in a separate section. The main required stages are shown in Figure 4.2 and include preparing the corpus, constructing the domain ontology, documents annotation, and documents category assignment.

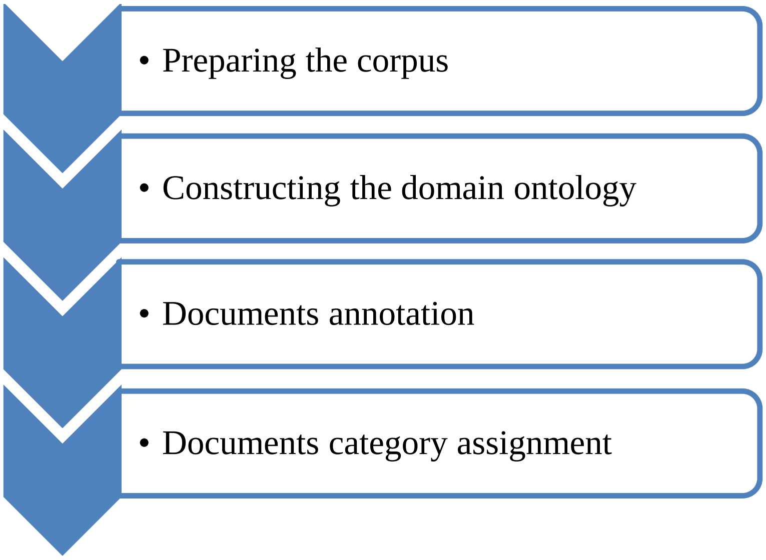

• Preparing the corpus
• Constructing the domain ontology
• Documents annotation
• Documents category assignment

Figure 4.2 Approach Building Stages

4.2. Preparing the Corpus

Preparing the corpus is one of the most important stages in the research project. The corpus is a collection of documents in a selected domain. In our work we collect nearly 100 documents related to News domain. It's collected manually from Aljazeera news web site [4], which is one of the largest Arabic news websites. The collected documents are chosen from the news field of *Politics* “سياسة”, *Economy* “اقتصاد”, *Sport* “رياضة”, *Health* “صحة” and *Science and Technology* “علوم وتكنولوجيا”. Then all documents is converted to xml format to facilitate the processing of documents annotation.

4.3. Constructing the Domain Ontology

Building ontology is the main theme in this study, we used a top-down approach in building the ontology. Most abstract concepts are identified first, then specialized into more specific concepts to build our domain Arabic ontology which represents the basic knowledge in our work. We construct the ontology manually.

We have developed the ontology contents for News domain, collected from a number of relevant research papers. The ontology is implemented with Protégé tool in OWL format.

As we have previously illustrated the ontology building (section 2.3.1), the manual development of ontology consists of the following steps:

- Determine the domain and scope of the ontology.
- Consider reusing existing ontologies.
- Enumerate important terms in the ontology.
- Define the classes and the class hierarchy.
- Define the properties of classes (slots).
- Define the facets of the slots.
- Create instances.

Step 1: Determine the domain and scope of the ontology

The definition of ontology domain and scope is the first step of ontology development. To determine the ontology domain and scope we should know which domain will the ontology cover, the purpose of the ontology, and the type of question should the information in the ontology be able to provide answers, and who will use the ontology.

By answering these questions, we can say that the domain of the ontology will cover news “الأخبار” domain, which is used in our approach to annotate documents for the purpose of classification.

Step 2: Consider reusing existing ontologies

This step is to ascertain if there exists ontology that is developed previously in the same subject area. If such ontology exists, it is easier to modify the existing ontology to suit ones needs than to create a new one. And because we have not found a previously ontology created on the news area neither in Arabic nor in English language, we omitted these step.

Step 3: Enumerate important terms in the ontology

This step can be viewed as a brainstorming activity, in which we list the words that we want to use, to demonstrate the ontology terms, and the properties that may have.

We also benefited from the collected News documents to get the knowledge about News terms.

Step 4: Define the classes and the class hierarchy

This step defines classes (concepts) used in our ontology domain. We define classes and sub-classes related to our domain. Table 4.1 depicts the Ontology classes, News “الأخبار” is the most general concept. Politics “سياسة”, Economy “اقتصاد”, Sport “رياضة”, Health “صحة” and Science and Technology “ علوم وتكنولوجيا” are general top level concepts. The remaining are the most specific classes (or the bottom level classes).

Table 4.1 Ontology Classes and Sub-classes

No.	Classes/Subclasses In Arabic	Classes/Subclasses In English	No.	Classes/Subclasses In Arabic	Classes/Subclasses In English
1.	**سياسة**	**Politics**	21. `	مباراة	Match
2.	انتخابات	Election	22.	دوري	periodic
3.	أمة	Nation	23.	نادي	Club
4.	دبلوماسية	Diplomacy	24.	**صحة**	**Health**
5.	مفاوضات	Negotiation	25.	أمراض	Diseases
6.	نزاعات	Conflicts	26.	علاج	Treatment
7.	قضايا	Issues	27.	مجاعة	Famine
8.	اعلام	Media	28.	موت	Death
9.	**اقتصاد**	**Economy**	29.	**علوم وتكنولوجيا**	**Science and Tech.**
10.	بناء	Building	30.	أبحاث	Researches
11.	تجارة	Trade	31.	أجهزة	Devices
12.	زراعة	Agriculture	32.	اتصال	Communication
13.	سياحة	Tourism	33.	اختراع	Invention
14.	شركات	Company	34.	الطقس	Weather
15.	صناعة	Industry	35.	الفلك	Astronomy
16.	طاقة	Energy	36.	الكترونيات	Electronics
17.	نقل	Transport	37.	باحث	Researcher
18.	**رياضة**	**Sport**	38.	عالم	Scientist
19.	رياضات فردية	Individual sports	39.	علوم	Science
20.	رياضات جماعية	Collective sports	40.	تكنولوجيا	Technology

Choosing these concepts has a direct relationship with the process of documents classification where they represent the general accepted classification categories. In Table 4.1 we mention every ontology concept in Arabic and its synonym in English. Some of these concepts have relations with other concepts which help in the classification process. Also, most of these concepts have synonym words and they contain instances to help in the process of documents classification.

Figure 4.3 depicts the top level ontology classes. The Thing, which represent the class of all things. The News "الأخبار" class which is the root class, and others are the sub-classes.

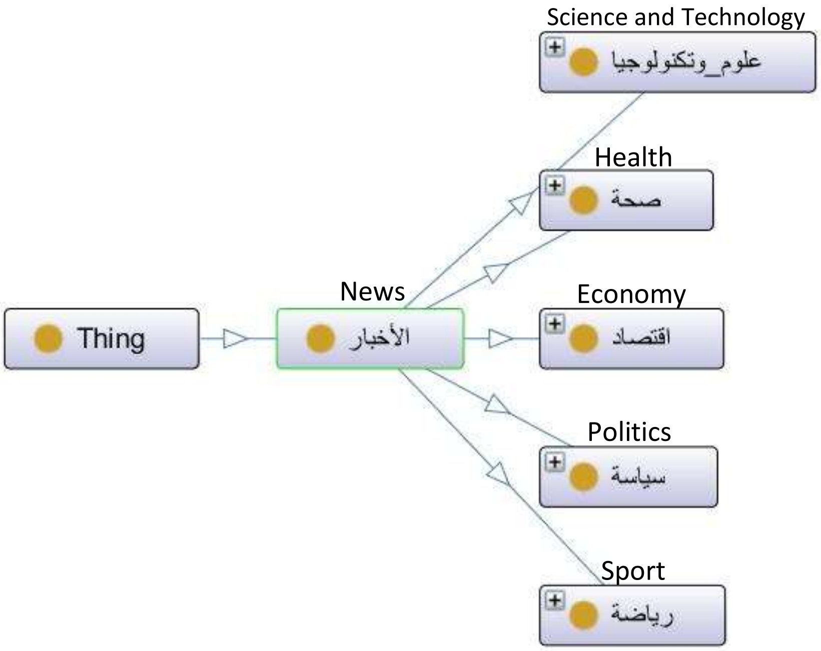


Figure 4.3 Top Level Ontology Classes

Step 5: Define the properties of classes (slots).

Define object properties (relations) among classes, which role is connecting concepts of the ontology.

Because of the nature of the ontology, which used for classification purposes there is rarely relations between ontology concepts.

We used 4 object properties that connect the important concepts which have relations with each other. Table 4.2 depicts these 4 object properties in our

Arabic ontology and its synonym in English. Also depicts properties domain and range classes.

Table 4.2 Ontology Object Properties

No.	Lang.	Object Properties	Domain	Range
1.	AR	تحتاج إلى	أمراض	علاج
	EN	Need a	Diseases	Treatment
2.	AR	تستخدم بواسطة	تقنية	أجهزة
	EN	Used by	Technique	Devices
3.	AR	يلعب في	لاعب	فريق
	EN	Plays in	Player	Team
4.	AR	تحل بواسطة	نزاعات	مفاوضات
	EN	Solved by	Conflicts	Negotiation

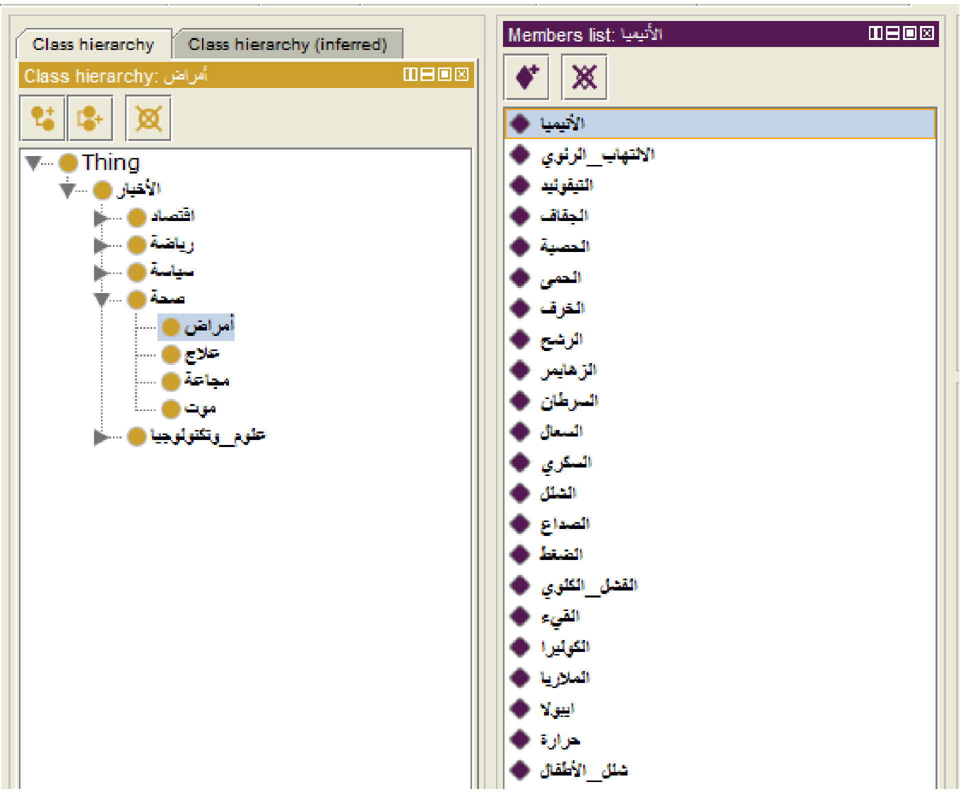
Step 6: Define the facets of the slots.

Slots (sometimes called roles or properties) have different facets (sometimes called role restrictions) that describe the value type, allowed values, the number of the values (cardinality), and other features of the values the slot can take. In our case most of the slot values are string. For example, the value of a *synonyms* slot is one string. That is, *synonyms* is a slot with value type String. A slot *used by* “يستخدم بواسطة” can have multiple values and the values are instances of the class Devices “أجهزة”. That is, *used by* “يستخدم بواسطة” is a slot with value type Instance with Devices “أجهزة” as allowed class.

Step 7: Create instances.

Creating instances (individuals) is a very important step to enrich the ontology with direct relation with classes and sub-classes.

For example, the class *Diseases* “أمراض” have several instances, which include Alzheimer's “الزهايمر”, Diabetes “السكري”, Pressure “الضغط”, Anemia “الأنيميا”, Polio “شلل الأطفال” etc. Figure 4.4 and Figure 4.5 depicts some of these instances.



Class hierarchy: أمراض

- Thing
  - الأخبار
    - اقتصاد
    - رياضة
    - سياسة
    - صحة
      - أمراض
      - علاج
      - مجاعة
      - موت
    - علوم_وتكنولوجيا

Members list: الأنيميا

- الأنيميا
- الالتهاب_الرئوي
- التيفوئيد
- الجفاف
- الحصبة
- الحمى
- الخرف
- الرشح
- الزهايمر
- السرطان
- السعال
- السكري
- الشلل
- الصداع
- الضغط
- الفشل_الكلوي
- القيء
- الكوليرا
- الملاريا
- ايبولا
- حرارة
- شلل_الأطفال

Figure 4.4 List of Some Ontology Instances

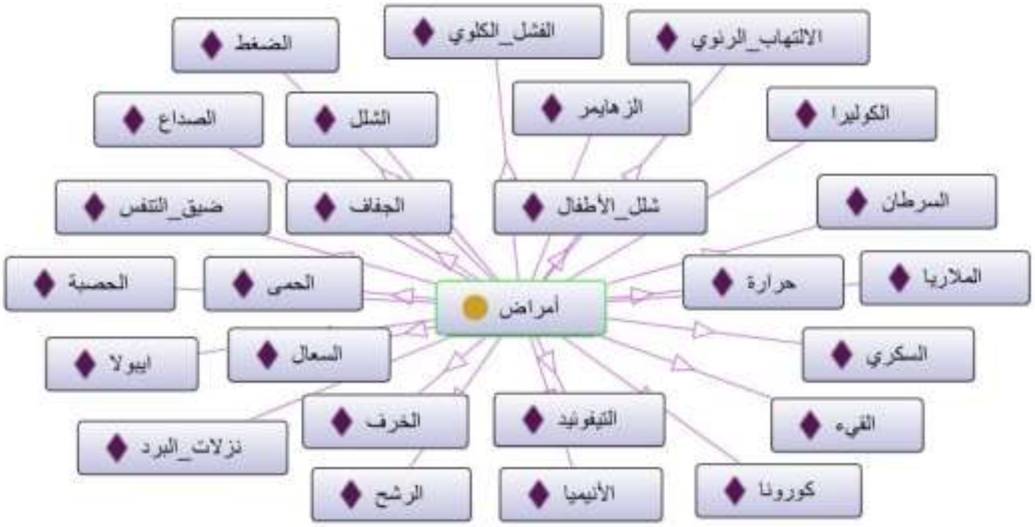


Figure 4.5 Graph of Some Ontology Instances

Ontology Evaluation

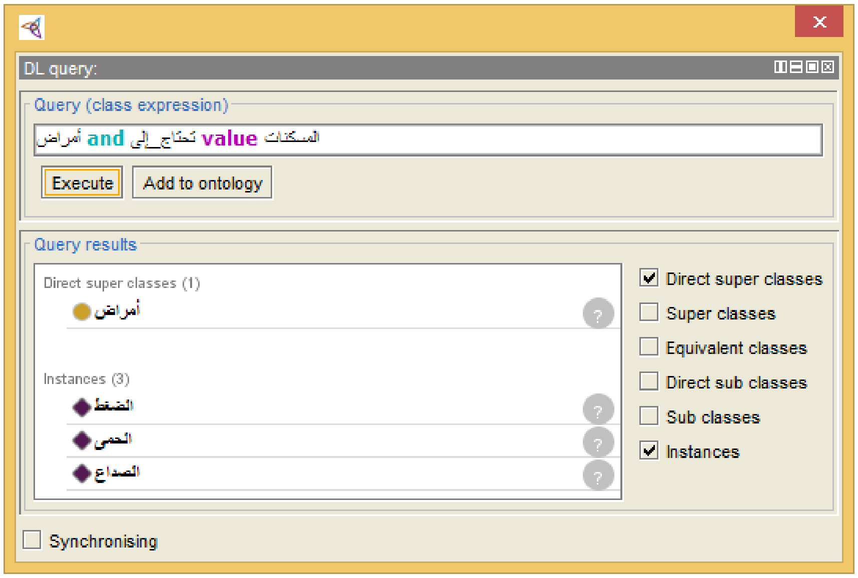
In order to verify and validate the ontology with regards to competency questions, we use the Description Logic Query (DL-Query) that is a standard Protégé plugin and it is based on the Manchester OWL syntax with HermiT OWL reasoner.

An example of the querying function that answers the questions that are asked in the development process of the ontology is: *What are the diseases that need painkillers?*, which is illustrated in Table 4.3 in DL-Query format.

Table 4.3 Question in DL-Query Format.

Lang.	DL-Query			
AR	أمراض	and	تحتاج إلى	المسكنات
EN	*Diseases*		Need a	*painkillers*

Figure 4.6 depicts the result of DL-Query which illustrates the individuals of *Diseases* “أمراض” class that are needs *painkillers* “المسكنات” as *Treatment* “علاج”.



DL query:

Query (class expression)

أمراض and تحتاج_إلى value المسكنات

Query results

Direct super classes (1)

● أمراض

Instances (3)

◆ الضغط
◆ الحمى
◆ الصداع

☑ Direct super classes
☐ Super classes
☐ Equivalent classes
☐ Direct sub classes
☐ Sub classes
☑ Instances

☐ Synchronising

Figure 4.6 Result of DL-Query

The results of DL-Query example show that the ontology are successfully portrays the body of knowledge related to the News ontology.

4.4. Documents Annotation

To perform document annotations, we use GATE software. More information about GATE can be found in Section 5.2.2. The annotation process is performed on the documents stored in the GATE Corpus. A Corpus in GATE is a Java set whose members are documents. Both ontology and documents are types of *LanguageResource* (LR), which is an entity that holds linguistic data such as documents, corpora, ontologies.

GATE supports various document formats such as: Plain text, HTML, SGML, XML, RTF, and PDF. In our work we use Plain text documents, then we saved them in XML formats after making annotation.

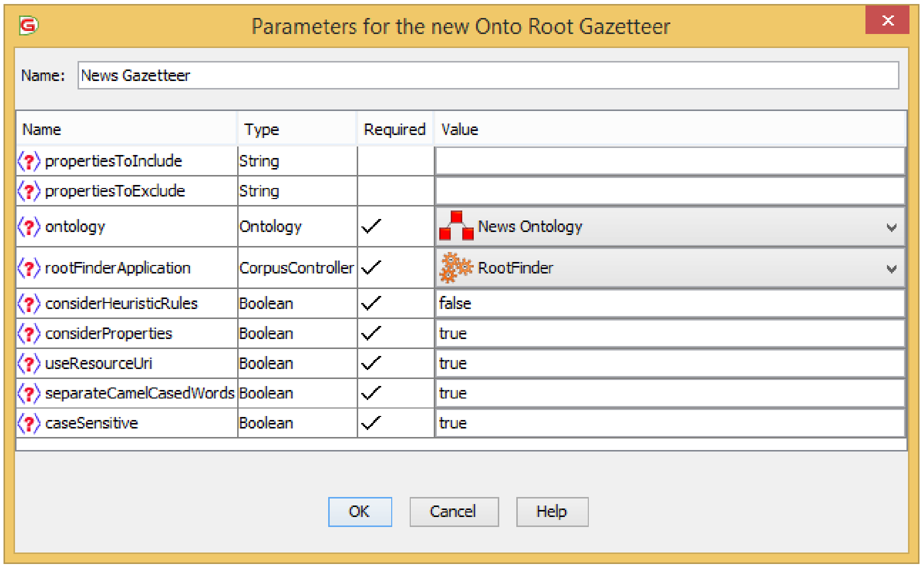
We use the OntoRoot Gazetteer *ProcessingRecourse* (PR) which is a type of dynamically created gazetteer that is, in combination with few other generic GATE resources, capable of producing ontology-based annotations over the given content with regards to the given ontology. PR represents entities that are primarily algorithmic, such as parsers or generators, which will do some sort of processing on text.

A gazetteer consists of a set of lists containing names of entities such as cities, organizations, days of the week, etc. These lists are used to find occurrences of these names in text. The word ‘gazetteer’ is often used interchangeably for both the set of entity lists and for the processing resource that makes use of those lists to find occurrences of the names in text.

OntoRoot Gazetteer links text to an ontology by creating Lookup annotations which come from the ontology rather than a default gazetteer.

The OntoRoot Gazetteer needs a few mandatory parameters to be initialized as shown in Figure 4.7 which are:

- Ontology LR.
- Root Finder Application.



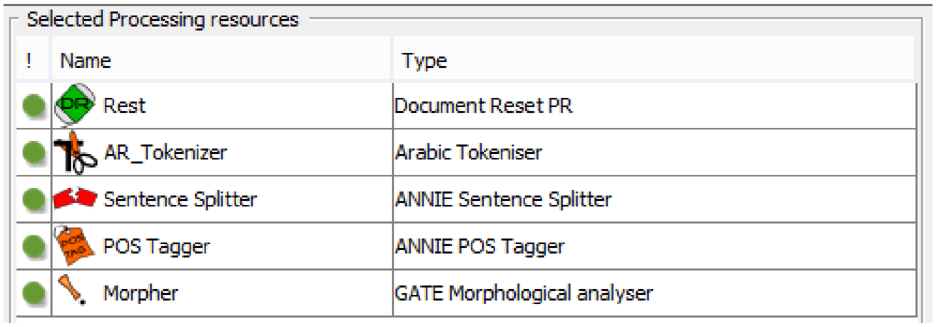
Name: News Gazetteer

Name	Type	Required	Value
propertiesToInclude	String		
propertiesToExclude	String		
ontology	Ontology	✓	News Ontology
rootFinderApplication	CorpusController	✓	RootFinder
considerHeuristicRules	Boolean	✓	false
considerProperties	Boolean	✓	true
useResourceUri	Boolean	✓	true
separateCamelCasedWords	Boolean	✓	true
caseSensitive	Boolean	✓	true

Figure 4.7 OntoRoot Gazetteer Parameters

The root finder application consists of the following Processing Resources (PRs), which is illustrated in Figure 4.8 and are states as follows:

- Document Reset PR: which enables the document to be reset to its original state, by removing all the annotation sets and their contents.
- Arabic Tokenizer: which tokenizes the text with kinds of Tokens as Word, Number, Punctuation and Space Token.
- ANNIE Sentence Splitter: which segments the text into sentences.
- ANNIE POS Tagger: Part of Speech Tagger which produces a part-of-speech tags as an annotation on each word or symbol.
- GATE Morphological Analyzer: which finds the root and affix values of a token and adds them as features to the tokens.

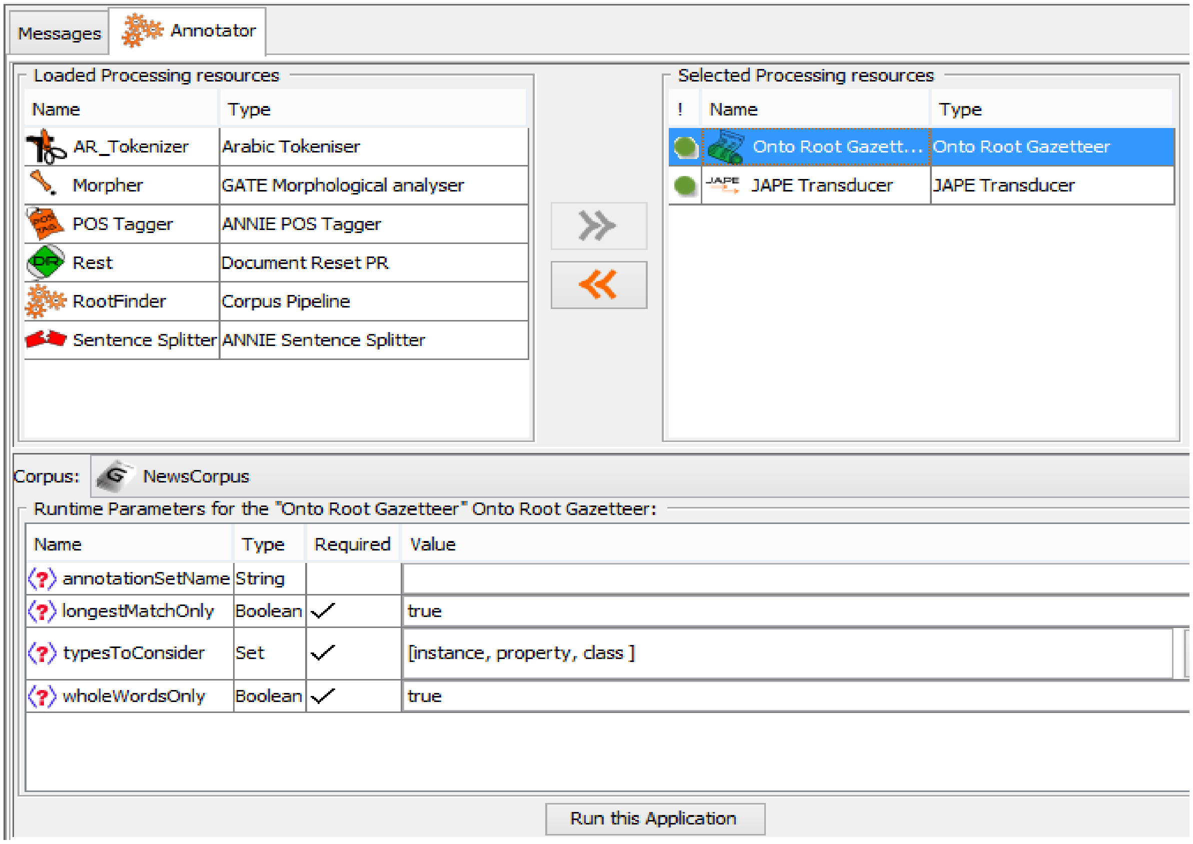


Selected Processing resources

!	Name	Type
●	Rest	Document Reset PR
●	AR_Tokenizer	Arabic Tokeniser
●	Sentence Splitter	ANNIE Sentence Splitter
●	POS Tagger	ANNIE POS Tagger
●	Morpher	GATE Morphological analyser

Figure 4.8 Root Finder Application PRs

By running the OntoRoot Gazetteer PR on the corpus, which contain the Arabic news field documents, and based on the constructed News ontology, the documents will annotate with these ontology components and the result will be annotation with ontology classes, instances, and properties. Figure 4.9 depicts using OntoRoot Gazetteer annotator.



Loaded Processing resources

Name	Type
AR_Tokenizer	Arabic Tokeniser
Morpher	GATE Morphological analyser
POS Tagger	ANNIE POS Tagger
Rest	Document Reset PR
RootFinder	Corpus Pipeline
Sentence Splitter	ANNIE Sentence Splitter

Selected Processing resources

!	Name	Type
●	Onto Root Gazett...	Onto Root Gazetteer
●	JAPE Transducer	JAPE Transducer

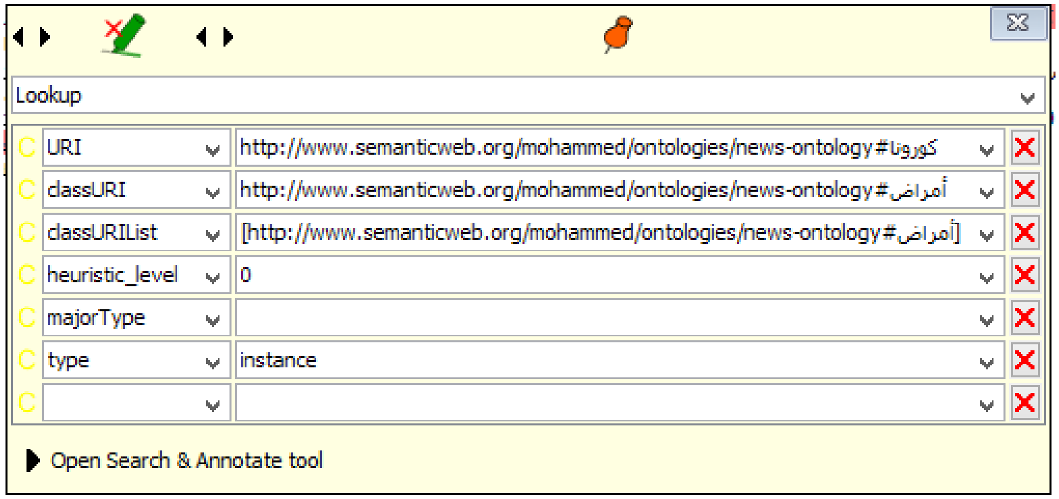
Corpus: NewsCorpus

Runtime Parameters for the "Onto Root Gazetteer" Onto Root Gazetteer:

Name	Type	Required	Value
annotationSetName	String		
longestMatchOnly	Boolean	✓	true
typesToConsider	Set	✓	[instance, property, class]
wholeWordsOnly	Boolean	✓	true

Figure 4.9 OntoRoot Gazetteer PR

As a results of running OntoRoot Gazetteer, a Lookup annotations are generated. Feature URI refers to the URI of the ontology resource, while *type* identifies the type of the resource such as class, instance, or property. Figure 4.10 depicts the Lookup annotation, in which *corona* “كورونا” is annotated with the corresponding ontology instance, which is an instance of the ontology class *Diseases* “أمراض”.

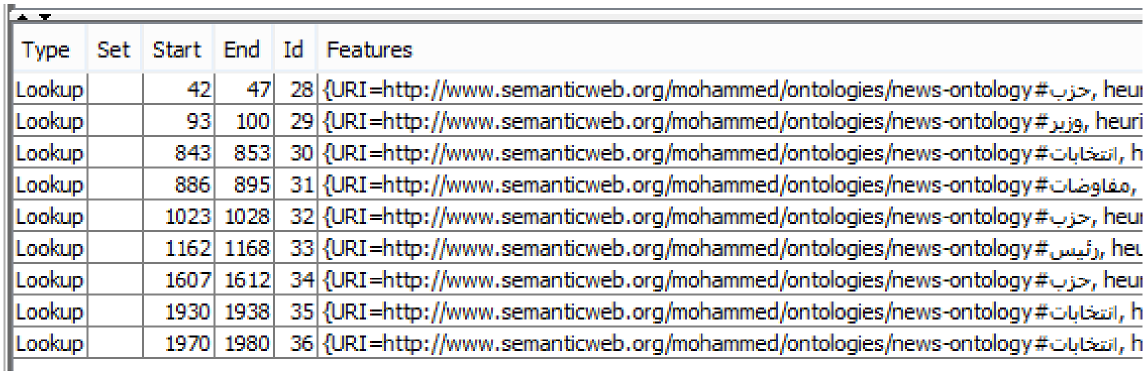


Lookup

URI	http://www.semanticweb.org/mohammed/ontologies/news-ontology#كورونا
classURI	http://www.semanticweb.org/mohammed/ontologies/news-ontology#أمراض
classURIList	[http://www.semanticweb.org/mohammed/ontologies/news-ontology#أمراض]
heuristic_level	0
majorType	
type	instance

Figure 4.10 Sample Result of Lookup Annotation

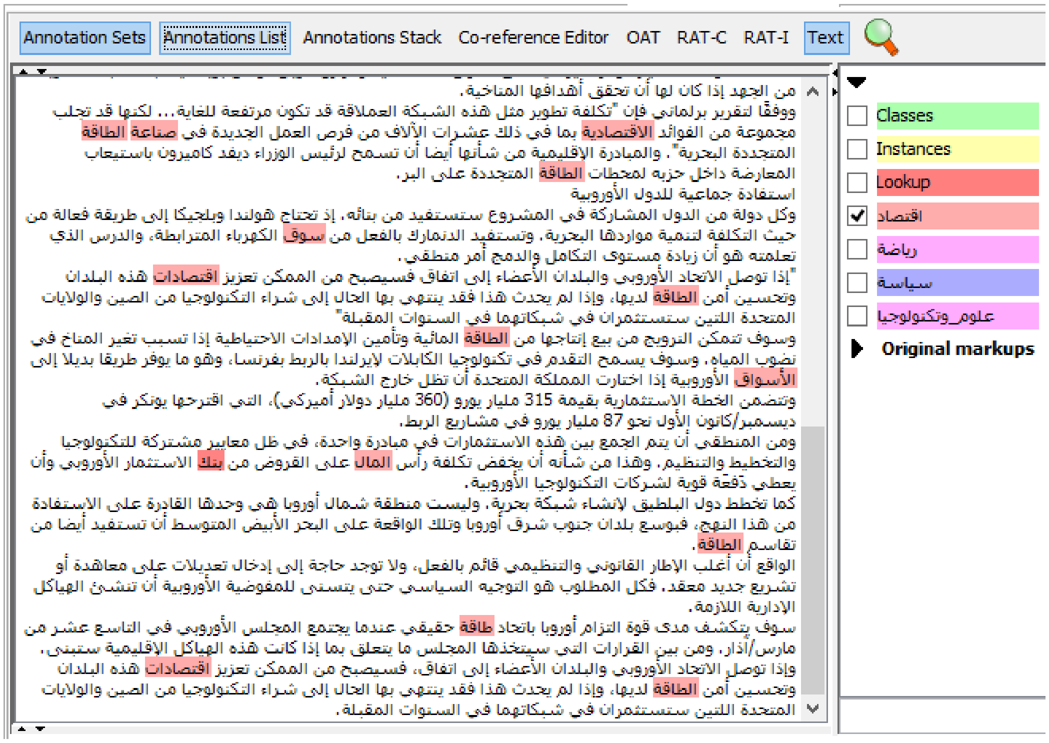
Figure 4.11 illustrates the Lookup annotation list which includes the start and end positions of the annotation in the annotated document, and the features which contain URI, classURI, classURIList, and type, is which related to the corresponding ontology component.



Type	Set	Start	End	Id	Features
Lookup		42	47	28	{URI=http://www.semanticweb.org/mohammed/ontologies/news-ontology#حزب, heu
Lookup		93	100	29	{URI=http://www.semanticweb.org/mohammed/ontologies/news-ontology#وزير, heuri
Lookup		843	853	30	{URI=http://www.semanticweb.org/mohammed/ontologies/news-ontology#انتخابات, h
Lookup		886	895	31	{URI=http://www.semanticweb.org/mohammed/ontologies/news-ontology#مفاوضات,
Lookup		1023	1028	32	{URI=http://www.semanticweb.org/mohammed/ontologies/news-ontology#حزب, heu
Lookup		1162	1168	33	{URI=http://www.semanticweb.org/mohammed/ontologies/news-ontology#رئيس, heu
Lookup		1607	1612	34	{URI=http://www.semanticweb.org/mohammed/ontologies/news-ontology#حزب, heu
Lookup		1930	1938	35	{URI=http://www.semanticweb.org/mohammed/ontologies/news-ontology#انتخابات, h
Lookup		1970	1980	36	{URI=http://www.semanticweb.org/mohammed/ontologies/news-ontology#انتخابات, h

Figure 4.11 Sample Result of Lookup Annotation

A sample of the result is shown in Figure 4.12, after running the OntoRoot Gazetteer processing recourse the document is annotated with the corresponding ontology component, which appears in the annotation sets. When we select the type of the annotation from the set, all annotated words is highlighted.



من الجهد إذا كان لها أن تحقق أهدافها المناخية.
ووفقًا لتقرير برلماني فإن "تكلفة تطوير مثل هذه الشبكة العملاقة قد تكون مرتفعة للغاية... لكنها قد تجلب
مجموعة من الفوائد الاقتصادية بما في ذلك عشرات الآلاف من فرص العمل الجديدة في صناعة الطاقة
المتجددة البحرية". والمبادرة الإقليمية من شأنها أيضا أن تسمح لرئيس الوزراء ديفد كاميرون باستيعاب
المعارضة داخل حزبه لمحطات الطاقة المتجددة على البر.
استفادة جماعية للدول الأوروبية
وكل دولة من الدول المشاركة في المشروع ستستفيد من بنائه، إذ تحتاج هولندا وبلجيكا إلى طريقة فعالة من
حيث التكلفة لتنمية مواردها البحرية، وتستفيد الدنمارك بالفعل من سوق الكهرباء المترابطة، والدرس الذي
تعلمته هو أن زيادة مستوى التكامل والدمج أمر منطقي.
"إذا توصل الاتحاد الأوروبي والبلدان الأعضاء إلى اتفاق فسيصبح من الممكن تعزيز اقتصادات هذه البلدان
وتحسين أمن الطاقة لديها، وإذا لم يحدث هذا فقد ينتهي بها الحال إلى شراء التكنولوجيا من الصين والولايات
المتحدة اللتين ستستثمران في شبكاتهما في السنوات المقبلة"
وسوف تتمكن النرويج من بيع إنتاجها من الطاقة المائية وتأمين الإمدادات الاحتياطية إذا تسبب تغير المناخ في
نضوب المياه، وسوف يسمح التقدم في تكنولوجيا الكابلات لإيرلندا بالربط بفرنسا، وهو ما يوفر طريقا بديلا إلى
الأسواق الأوروبية إذا اختارت المملكة المتحدة أن تظل خارج الشبكة.
وتتضمن الخطة الاستثمارية بقيمة 315 مليار يورو (360 مليار دولار أميركي)، التي اقترحها يونكر في
ديسمبر/كانون الأول نحو 87 مليار يورو في مشاريع الربط.
ومن المنطقي أن يتم الجمع بين هذه الاستثمارات في مبادرة واحدة، في ظل معايير مشتركة للتكنولوجيا
والتخطيط والتنظيم. وهذا من شأنه أن يخفض تكلفة رأس المال على القروض من بنك الاستثمار الأوروبي وأن
يعطي دَفعة قوية لشركات التكنولوجيا الأوروبية.
كما تخطط دول البلطيق لإنشاء شبكة بحرية. وليست منطقة شمال أوروبا هي وحدها القادرة على الاستفادة
من هذا النهج، فبوسع بلدان جنوب شرق أوروبا وتلك الواقعة على البحر الأبيض المتوسط أن تستفيد أيضا من
تقاسم الطاقة.
الواقع أن أغلب الإطار القانوني والتنظيمي قائم بالفعل، ولا توجد حاجة إلى إدخال تعديلات على معاهدة أو
تشريع جديد معقد. فكل المطلوب هو التوجيه السياسي حتى يتسنى للمفوضية الأوروبية أن تنشئ الهياكل
الإدارية اللازمة.
سوف يتكشف مدى قوة التزام أوروبا باتحاد طاقة حقيقي عندما يجتمع المجلس الأوروبي في التاسع عشر من
مارس/آذار. ومن بين القرارات التي سيتخذها المجلس ما يتعلق بما إذا كانت هذه الهياكل الإقليمية ستبنى.
وإذا توصل الاتحاد الأوروبي والبلدان الأعضاء إلى اتفاق، فسيصبح من الممكن تعزيز اقتصادات هذه البلدان
وتحسين أمن الطاقة لديها، وإذا لم يحدث هذا فقد ينتهي بها الحال إلى شراء التكنولوجيا من الصين والولايات
المتحدة اللتين ستستثمران في شبكاتهما في السنوات المقبلة.

Figure 4.12 Sample Result of Document Annotation

Figure 4.13 depicts the xml annotation tags, <*Annotation*> and <*Feature*> tags, which the OntoRoot Gazetteer added to the document after annotation.

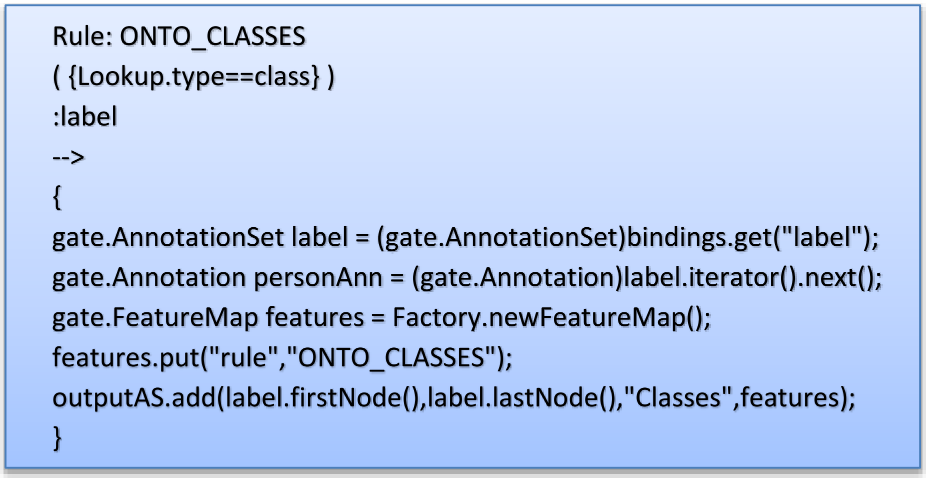
```
<AnnotationSet>
    <Annotation Id="2383" Type="Lookup" StartNode="545" EndNode="550">
       <Feature>
                <Name className="java.lang.String">URI</Name>
                <Value className=
              "java.lang.String">http://www.semanticweb.org/mohammed/
              ontologies/news-ontology#اقتصاد </Value>
       </Feature>
    </Annotation>
</AnnotationSet>
```

Figure 4.13 XML Sample Result of the Annotation

4.5. Processing the Annotated Documents

In this step after annotating documents using OntoRoot Gazetteer annotator, we pass annotated documents to JAPE Transducer plugin, which needs a

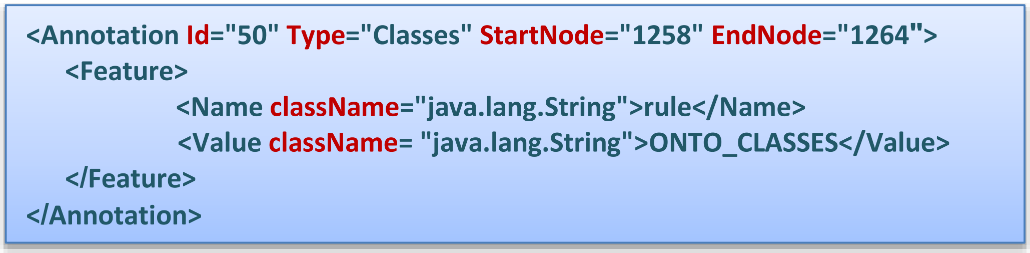
Grammar parameter URL. The User can use this parameter to specify the JAPE rules that consider files written with the extension ".jape". The JAPE Transducer parses and compiles the JAPE rules at run-time to execute them over the GATE document. Figure 4.14 depicts some JAPE rules used in processing the annotated documents.



```
Rule: ONTO_CLASSES
( {Lookup.type==class} )
:label
-->
{
gate.AnnotationSet label = (gate.AnnotationSet)bindings.get("label");
gate.Annotation personAnn = (gate.Annotation)label.iterator().next();
gate.FeatureMap features = Factory.newFeatureMap();
features.put("rule","ONTO_CLASSES");
outputAS.add(label.firstNode(),label.lastNode(),"Classes",features);
}
```

Figure 4.14 JAPE Rule Example 1

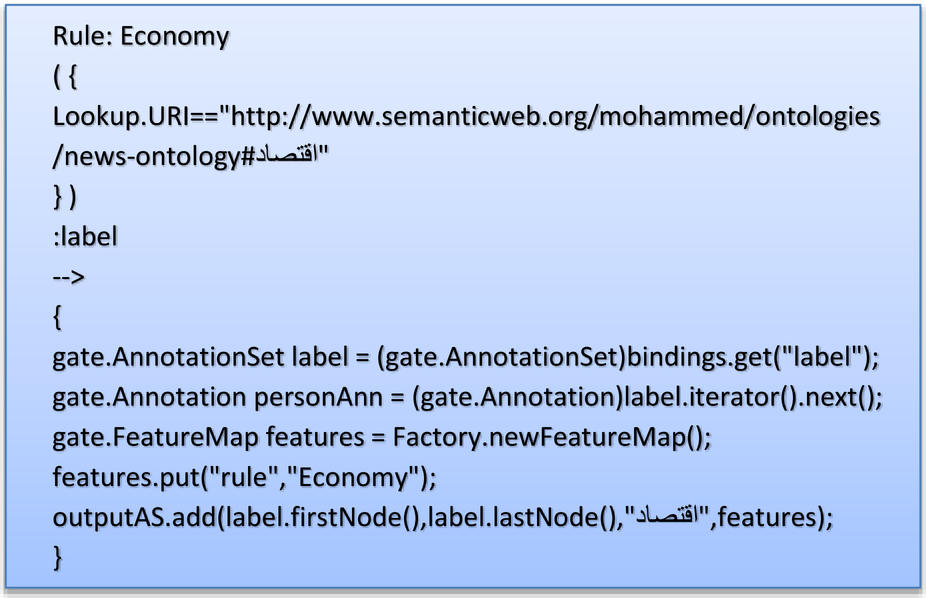
In the previous JAPE rule example, the rule name ONTO_CLASSES, the left hand side of the rule indicates searching about all GATE lookup annotations with type “class” and passes them to the right hand side for processing. The RHS adds annotation with type “Classes” and with feature rule and name as we defined. The XML annotation of JAPE rules is illustrated in Figure 4.15.



```
<Annotation Id="50" Type="Classes" StartNode="1258" EndNode="1264">
    <Feature>
                <Name className="java.lang.String">rule</Name>
                <Value className= "java.lang.String">ONTO_CLASSES</Value>
    </Feature>
</Annotation>
```

Figure 4.15 XML Annotation of JAPE Rule

In Figure 4.16 the JAPE rule adds annotation with type Economy “اقتصاد” and feature name rule with value “Economy”.



```
Rule: Economy
( {
Lookup.URI=="http://www.semanticweb.org/mohammed/ontologies
/news-ontology#اقتصاد"
} )
:label
-->
{
gate.AnnotationSet label = (gate.AnnotationSet)bindings.get("label");
gate.Annotation personAnn = (gate.Annotation)label.iterator().next();
gate.FeatureMap features = Factory.newFeatureMap();
features.put("rule","Economy");
outputAS.add(label.firstNode(),label.lastNode(),"اقتصاد",features);
}
```

Figure 4.16 JAPE Rule Example 2

In the same way we can process all documents with JAPE Transducer. The resulting annotation facilitates the process of document classification based on these rules as we explain next.

4.6. Documents Category Assignment

In the final stage, an algorithm is used to determine the document class based on its annotation set. The document may contain annotations belong to various classes. Therefore, the document which contains the highest matching score with the corresponding general top level ontology class is assigned to this class, which is considered as the most suitable category for the given document.

In this step the received xml document is parsed to reach the annotation sets and determine the annotation category. Then by reading the ontology we can reach to the top class level which we want to map documents to. Finally by computing each annotation category rank we can assign documents to their categories.

4.7. Summary

In this chapter, we have discussed the steps to execute our methodology. In the first section, we talked about the structure of the proposed approach. In the second section, we illustrate the process of preparing the corpus, the number of documents and how we obtain them. In the third section, ontology building stages are explained. In the fourth section documents annotation steps are explained depending on OntoRoot Gazetteer plugin. In the fifth section we explained how we processed annotated documents using JAPE Rules. Finally in the sixth section the documents category assignment is explained.

Chapter 5: Experimental Results and Evaluation

This chapter presents and analyzes the experimental results and the evaluation of our approach. Firstly we explain the work environment, tools and programs used in our research to develop the proposed approach. Then we evaluate the classification approach performance. Finally at the end of the chapter we discuss our results.

5.1. Implementation of the Classification Approach

As we presented the structure of the classification approach in Section 4.1, its main components are: the domain ontology, the annotator and the classifier. We develop a prototype for the ontology-based classification approach that automatically classify documents based on the News domain ontology. To satisfy this goal we divide the prototype into the following two stages:

First Stage: **Documents Annotation**, which is annotating documents using GATE software based on the Arabic News ontology.

Second Stage: **Category Assignment**, which is assign document to category using JENA API based on the annotation phase.

The classification approach is implemented using Protégé, GATE software, JAPE Rules, Eclipse with Java SDK and JENA. We use Protégé tool to build the Arabic News ontology, GATE software and JAPE Rules used for documents annotation. The classifier which assigns documents to categories is built using Eclipse and JENA API.

5.2. Tools, Framework and API

To implement the ontology-based classification approach, we need different components at different stages for classifying documents. So various kinds of open source software tools have been used, which are Protégé for ontology building, GATE and JAPE rules for documents annotation, Eclipse for code editing, and JENA API for programming. These tools are described below:

5.2.1 Protégé

For ontology building, we used Protégé 5 [55] which is a knowledge base ontology editor providing graphical user interface. It is chosen for our ontology building because it provides better flexibility for meta-modeling, enables the construction of domain ontologies, and customizes data entry forms to enter data. It is typically targeted at the knowledge engineering and conceptual modeling without knowing or thinking about syntax of output language [13], also it supports the construction of the ontology in Arabic language. Figure 5.1 shows a Protégé screenshot that include the classes' hierarchy and its corresponding graph.

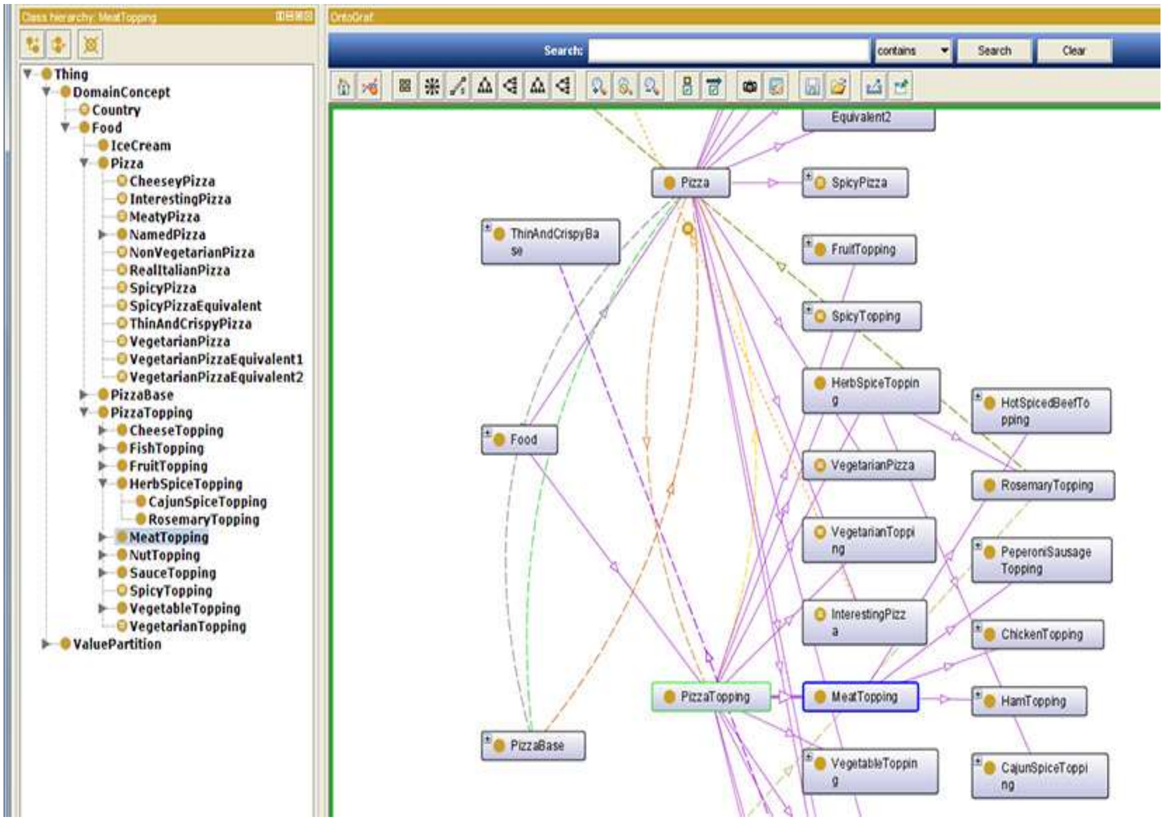
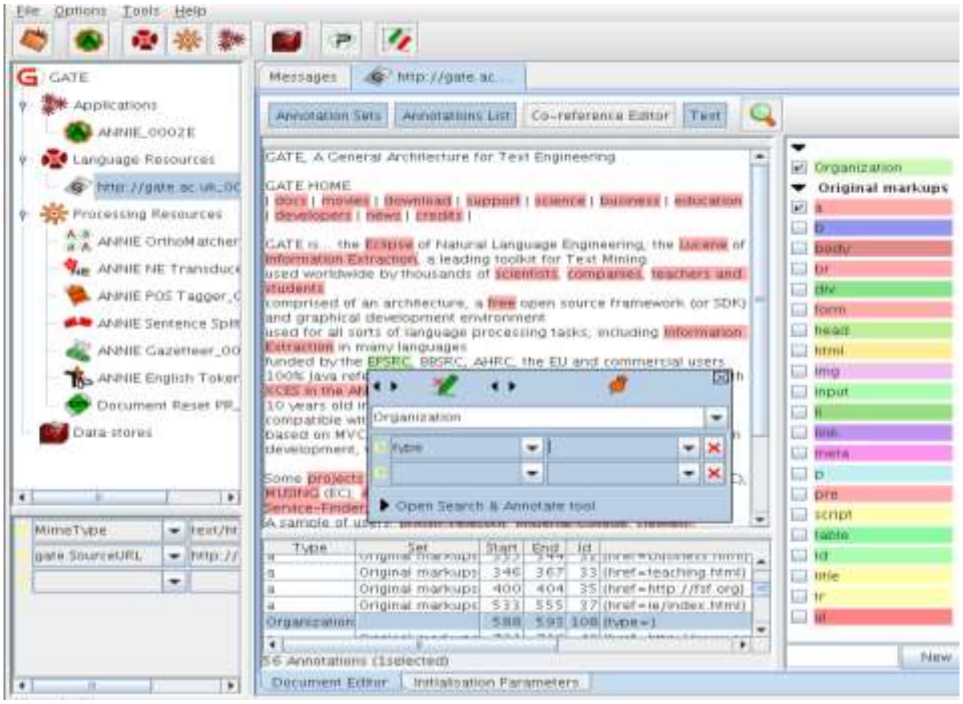


Figure 5.1 Protégé Screenshot [55]

5.2.2 GATE

GATE [30] (General Architecture for Text Engineering) is an open source software developed by the University of Sheffield. It’s an infrastructure for developing and deploying software components that process human language. GATE is an effective tool used for performing some NLP (Natural Language Processing), it has many features, such as manual annotation, automatic annotation, using variety of gazetteer, information retrieval, ontology-based processing [56].

GATE includes an information extraction system called ANNIE (A Nearly-New Information Extraction System) which is a set of modules comprising a tokenizer, a gazetteer, a sentence splitter, a part of speech tagger, a named entities transducer and a coreference tagger. ANNIE can be used as-is to provide basic information extraction functionality, or provide a starting point for more specific tasks.



GATE, A General Architecture for Text Engineering

GATE HOME
| docs | movies | download | support | science | business | education
| developers | news | credits |

GATE is ... the Eclipse of Natural Language Engineering, the Lucene of Information Extraction, a leading toolkit for Text Mining
used worldwide by thousands of scientists, companies, teachers and students
comprised of an architecture, a free open source framework (or SDK) and graphical development environment
used for all sorts of language processing tasks, including Information Extraction in many languages
funded by the EPSRC, BBSRC, AHRC, the EU and commercial users
100% Java refe
XCES in the AN
10 years old in
compatible wi
based on MVC
development,

Some projects
MUSING (EC),
Service-Finder
A sample of users:

Type	Set	Start	End	Id	
a	Original markups	346	367	33	{href=teaching.html}
a	Original markups	400	404	35	{href=http://fsf.org}
a	Original markups	533	555	37	{href=ie/index.html}
Organization		588	593	108	{type=}

56 Annotations (1selected)

Figure 5.2 GATE Screenshot [30]

Figure 5.2 shows GATE screenshot that illustrates its components which come in three types:

- Language resources (LRs): This is a number of linguistic data such as documents, Corpora, Ontologies.
- Processing Resources (PRs): These are programs or algorithms which will do some sort of processing on text i.e. Tokenizing or dictionary lookup, parsing etc.
- Visual Resources (VRs): These are components for graphical user interface and allows viewing and editing of other types of resources.

5.2.3 JAPE Rules

JAPE is a Java Annotation Patterns Engine, a component of the GATE platform. JAPE provides finite state transduction over annotations based on regular expressions. JAPE is a version of Common Pattern Specification Language (CPSL).

A JAPE grammar consists of a set of phases, each of which consists of a set of pattern/action rules. The phases run sequentially and constitute a cascade of finite state transducers over annotations. The left-hand-side (LHS) of the rules consist of an annotation pattern description. The right-hand-side (RHS) consists of annotation manipulation statements. Annotations matched on the LHS of a rule may be referred to on the RHS by means of labels that are attached to pattern elements [17].

5.2.4 JENA API

A free and open source Java framework for building Semantic Web and Linked Data applications [8]. It provides an API for the creation and manipulation of RDF repositories. Also provides classes/interfaces for the management of OWL-based ontologies. JENA includes a rule-based inference engine and a various reasoners can be set up to work with it. The Apache Jena architecture is shown in figure 5.3. We use JENA API to read and manipulate our constructed News ontology.

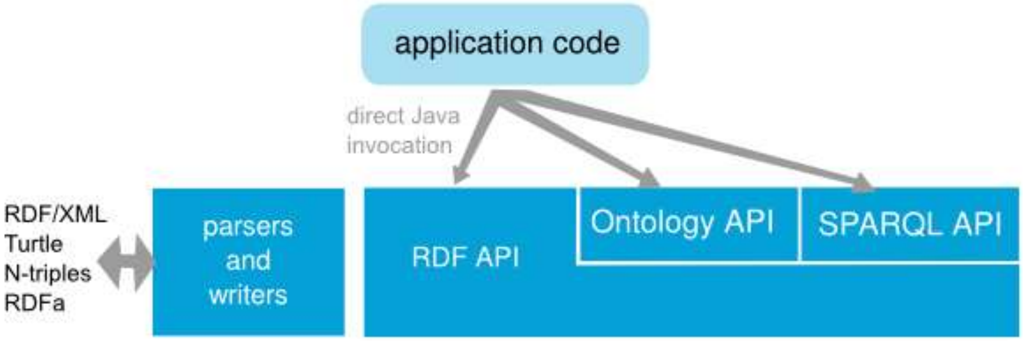


Figure 5.3 The Apache Jena Architecture [8]

5.3. Experiments

We performed a series of experiments to demonstrate the ability of our approach to classify documents based on the constructed Arabic News ontology. In the first stage documents are gathered in a GATE corpus with UTF-8 encoding type as a language resource. Also the Arabic News ontology which built in protégé is added to GATE as another language resource. Using OntoRoot Gazetteer the documents are annotated depend on the annotation type's ontology classes, instances, and properties. JAPE Transducer plugin then used to process the annotated documents using Jape rules.

In the second stage the documents which saved as XML format are grouped so are classified using Java-based classifier that assigns each document to the most relevant ontology concept. The classifier parses each document to access the delivered annotation and determine their category.

A series of experiments are performed and the obtained results show a high classification approach. As we explain next.

5.4. *Evaluation*

We are interested in the approach ability to correctly classify documents to their categories. There are different measures that we can use to evaluate the classification approach. The basic measures that we can use are: accuracy, precision, recall and F-measure. Computation of these measures are based on computing confusion matrix as shown in Table 5.1. A confusion matrix also known as a contingency table or an error matrix, which is a matrix where test cases are distributed as follows:

1. **True Positive** (*TP*): refers to positive instances that are correctly classified.
2. **False Negative** (*FN*): refers to positive instances incorrectly classified as negative.
3. **False Positive** (*FP*): refers to negative instances incorrectly classified as positive.
4. **True Negative** (*TN*): refers to negative instances that are correctly classified.

Table 5.1 Confusion matrix for two classes' classification problem

		Predicted Class	
		Positive	Negative
Actual Class	Positive	*TP*	*FN*
	Negative	*FP*	*TN*

Since we have five classes classification problem, our confusion matrix will rebuilt as follows:

Table 5.2 Confusion matrix for five classes' classification problem

		Predicted Classes				
		A	B	C	D	E
Actual Classes	A	TP_A	E_{AB}	E_{AC}	E_{AD}	E_{AE}
	B	E_{BA}	TP_B	E_{BC}	E_{BD}	E_{BE}
	C	E_{CA}	E_{CB}	TP_C	E_{CD}	E_{CE}
	D	E_{DA}	E_{DB}	E_{DC}	TP_D	E_{DE}
	E	E_{EA}	E_{EB}	E_{EC}	E_{ED}	TP_E

Where E_{xy} refer to error classified class x as class y.

According to previous five confusion matrix, we can compute measures as follows:

Accuracy is a measure of the overall correctness of the approach, it's the number of documents that are correctly classified divided by sum of the total documents, and since Accuracy = (TP + TN) / (TP + FP + FN + TN) then:

$$Accuracy = \sum_{x=A}^{E} TPx \Big/ \left(\sum_{x=A}^{E} TPx + \sum_{x=A}^{D} \sum_{y=x+1}^{E} Exy + \sum_{x=A}^{D} \sum_{y=x+1}^{E} Eyx\right) \dots \dots (5.1)$$

Precision is the percentage of predicted documents that are correctly classified:

Since Precision = TP / (TP + FP) then:

$$Precision\ X = TPx \Big/ \left(TPx + \sum_{y=A, y\neq x}^{E} yx\right) \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (5.2)$$

Recall is the percentage of the total documents that are correctly classified:

Since Recall = TP / (TP + FN) then:

$$Recall\ X\ =\ TPx \Big/ \left(TPx + \sum_{y=A,y\neq x}^{E} xy\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5.3)$$

F-measure combines precision and recall. We use the F-measure to evaluate the performance of the classifier:

$$F\text{-}measure\ x\ =\ 2\,\frac{Precision\ x\ \times\ Recall\ x}{Precision\ x\ +\ Recall\ x} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5.4)$$

So using accuracy, precision, recall and F-measure we can evaluate our approach and compare our results with other experiments.

Table 5.3 shows the confusion matrix which summarizes the results of testing the classification approach of Aljazeera corpus which contain five categories with 20 documents in each.

Table 5.3 Confusion matrix for the five classification classes

		Predicted Classes				
		سياسة Politics	اقتصاد Economy	رياضة Sport	صحة Health	علوم وتكنولوجيا
Actual Classes	سياسة Politics	20	0	0	0	0
	اقتصاد Economy	3	16	0	0	1
	رياضة Sport	1	0	19	0	0
	صحة Health	0	0	1	18	1
	علوم وتكنولوجيا Science & Tech.	0	1	0	0	19

Table 5.4 shows the calculated values of these measures for the five classification classes of the corpus.

Table 5.4 Precision, recall and F-measure results for the five classification classes

Measure ***Category***	*Precision*	*Recall*	*F-measure*
سياسة *Politics*	83.33 %	100.00 %	90.91 %
اقتصاد *Economy*	94.12 %	80.00 %	86.49 %
رياضة *Sport*	95.00 %	95.00 %	95.00 %
صحة *Health*	100.00 %	90.00 %	94.74 %
علوم وتكنولوجيا *Science & Tech.*	90.48 %	95.00 %	92.68 %
Average	**92.59 %**	**92.00 %**	**91.96 %**

Figure 5.4 depicts the results for precision, recall and F-measure for the five classification classes.

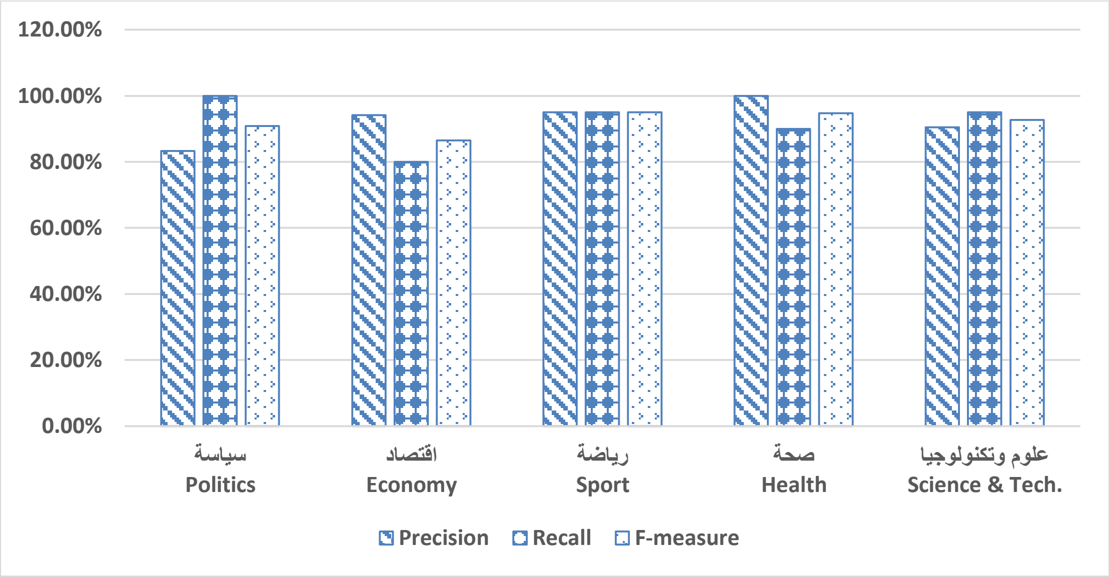


Figure 5.4 Precision, recall and F-measure results for the five classification classes

From the resulted confusion matrix and according to equation 5.1 we can compute the approach overall accuracy (the overall correctness of the approach) which is equal **92%**.

5.5. *Discussion*

The results in Table 5.3 show differences in the evaluation measures of the classification for the five mentioned categories. This is due to the following reasons:

- Some classes like *Politics* “سياسـة” have a low precision since there are some other documents which have been classified as politics. This because of the nature of some documents content which have mixed content, hence more than one category. We can classify documents to this category which is actually misclassified in Aljazeera website according to various considerations.
- Class *Economy* “اقتصـاد” have a low recall value since there is some economic files classified to other categories due to the same reason of mixed content.

From results we notice that there are different values depending on the size of the corpus. When we enlarge the corpus by adding new documents with new terms we should enumerate these new terms in our ontology to be taken into account in the process of annotation then classification.

Extending and enriching the ontology with more components which can be used in the process of document classification would give more accurate results. Because of building ontology takes a long term project, ontology enrichment needs more time to do.

We can say that the use of ontology contributes effectively in the process of Arabic documents classification.

5.6. Summary

In this chapter we have presented the stages of implementing our classification approach. Then we presented the tools, framework and programs used in our work. After that we talk about experiments and how to execute the approach. Then we evaluated the approach based on the different measures such as accuracy, precision, recall and F-measure. Finally we discussed the result of the classification approach.

Chapter 6: Conclusion and Future Work

This chapter concludes the thesis and its results and the future work.

6.1. Conclusion

We have developed an approach for ontology-based Arabic documents classification that facilitates the classification process. The approach achieves a significant improvement in the process of documents classification based on the different evaluation criteria. This ontology-based approach uses ontology components for annotating documents to be classified based on these annotations rather than depending on the traditional keyword based classifiers.

Our approach consists of several stages: preparing the corpus, constructing the domain ontology, documents annotation, and documents category assignment.

The results show that the proposed ontology-based approach achieves improvement in the process of documents classification based on the basic evaluation measures: accuracy, precision, recall and F-measure.

The main contribution of this research is that using the ontology to support the process of documents classification. Using the proposed approach, we overcome the problem of the traditional way used in the process of documents classification. This means saving time and returns better results.

Our contribution in this work includes the following:

- Building an ontology-based Arabic documents classification approach used in the process of documents classification.
- Building a domain specific ontology.
- Adaptation of GATE to work with Arabic documents.

6.2. *Future Work*

According to the results of experiments and the limitations that we faced in our thesis, this work can be improved in multiple directions:

- Extending the ontology by adding the other News categories, and enrich the ontology with more data and semantic information.
- Adding other domains to Arabic ontology (a top-level ontology), which enhance to build a generic classifier that is not limited to a specific domain.
- Contribution of building the Arabic language ontology by constructing the news domain ontology, which may be used for other researchers in the process of ontology building and reuse.
- Adopting the approach to deal with large documents corpus.
- Since only a prototype of the proposed approach is implemented, we look forward to implement a complete independent system that is make documents annotation independent of GATE.
- Extending our approach to work on the web, which will help to use in various applications such as documents search and retrieval.

Bibliography

[1] Aggarwal, C., “**Data Classification: Algorithms and Applications**”, Chapman and Hall/CRC, 2014.

[2] Alatrash, E., “**Using Web Tools for Constructing an Ontology of Different Natural Languages**”, Doctoral dissertation, University of Belgrade, 2013.

[3] Aliane, H., Alimazighi, Z., & Cherif, M., “**Al-Khalil: The Arabic Linguistic Ontology Project**”. Language Resources and Evaluation Conference LREC, May 2010.

[4] Aljazeera Website, http://www.aljazeera.net/portal, 2015, January, 12.

[5] Al-Safadi, L., Al-Badrani, M., & Al-Junidey, M., “**Developing ontology for Arabic blogs retrieval**”, International Journal of Computer Applications, vol.19, no.4, pp.41, 46, April 2011.

[6] Al-Shammari, E.T., "**Improving Arabic document categorization: Introducing local stem**", 2010 10th International Conference on Intelligent Systems Design and Applications (ISDA), pp.385-390, 29 November 2010 - 1 December 2010.

[7] Antoniou, G.; and Hormelen F., “**A Semantic Web primer**”, The MIT Press Cambridge, Massachusetts London, England, 2008.

[8] Apache JENA project, http://jena.apache.org/, 2015, January, 05.

[9] Arabic ontology project of Sina institute, http://sina.birzeit.edu/ArabicOntology/, January, 20015.

[10] Aurangzeb K., Baharum B., Lam H., Khairullah k., "**A Review of Machine Learning Algorithms for Text-Documents Classification**", Journal of Advances in Information Technology, vol. 1, no. 1, February 2010.

[11] Aurangzeb Khan; Baharum B. Bahurdin; Khairullah Khan, “**An Overview of E-Documents Classification**”, International Conference on Machine Learning and Computing IPCSIT, vol.3, 2001.

[12] Balakumar, M.; Vaidehi, V., "**Ontology based classification and categorization of email**", International Conference on Signal Processing, Communications and Networking, ICSCN '08, pp.199-202, 4-6 January 2008.

[13] Bhaskar, K., Savita, S., “**A Comparative Study of Ontology building Tools in Semantic Web Applications**”, International journal of Web & Semantic Technology (IJWesT) Vol.1, No.3, July 2010.

[14] Bouiadjra A.; Benslimane, S., “**FOEval: Full ontology evaluation**”, 7th International Conference on Natural Language Processing and Knowledge Engineering (NLP-KE), pp.464, 468, 27-29 November, 2011.

[15] Che-Yu Yang; Hua-Yi Lin, "**An automated semantic annotation based-on Wordnet ontology**", 2010 Sixth International Conference on Networked Computing and Advanced Information Management (NCM), pp.682,687, 16-18 August 2010.

[16] Corcho, O., Fernández-López, M., & Gómez-Pérez, A., “**Methodologies, tools and languages for building ontologies. Where is their meeting point?**”, Data & knowledge engineering 46, pp. 41-64, 2003.

[17] Cunningham H., et al., “**Developing Language Processing Components with GATE Version 8 :(a User Guide)**”, University of Sheffield, 2014.

[18] Dalloul, Y., “**An Ontology-Based Approach to Support the Process of Judging Hadith Isnad**”, Master's thesis, Islamic University of Gaza, 2013.

[19] Davies, J., Fensel, D., & Van Harmelen, F. (Eds.), “**Towards the semantic web: ontology-driven knowledge management**”, John Wiley & Sons, 2003.

[20] De Luca EW, Nrnberger A., “**Ontology-based semantic online classification of documents: Supporting users in searching the web**”, Proc European Symp on Intelligent Technologies, Aachen, 2004.

[21] De Luca EW, Nrnberger A.: “**Using clustering methods to improve ontology-based query term disambiguation**”, International Journal of Intelligent Systems, 21:7, pp. 693-709, 2006.

[22] Durga; A., Govardhan, A., “**Ontology Based Text Categorization - Telugu Documents**”, International Journal of Scientific & Engineering Research Volume 2, Issue 9, September 2011.

[23] El-Halees A., “**Arabic Text Classification Using Maximum Entropy**”, The Islamic University Journal (Series of Natural Studies and Engineering), vol. 15(1), pp. 157-167, 2007.

[24] El-Monsef, ME Abd, et al. "**Arabic Document Classification: A Comparative Study**" Journal of computing, volume 3, issue 4, April 2011.

[25] Elberrichi, Zakaria, and Karima Abidi. "**Arabic text categorization: A comparative study of different representation modes**" The International Arab Journal Information Technology, volume 9, No.5 September 2012.

[26] Falquet, G., Claudine M., Jacques T., Christopher T., “**Ontologies in urban development projects**”, Vol. 1, Springer Science & Business Media, 2011.

[27] Fang J.; Guo L.; Wang X.; Yang N., “**Ontology-Based Automatic Classification and Ranking for Web Documents**”, Fourth International Conference on Fuzzy Systems and Knowledge Discovery, 2007. FSKD 2007, vol.3, pp.627, 631, 24-27 August 2007.

[28] Fensel, D.; Hendler, J.; Lieberman, H.; Wahlster, W. (eds.). “**Spinning the Semantic Web**”. Cambridge, Mass. MIT Press, pp. 1 – 25, 2003.

[29] Fernández-López, M., Gómez-Pérez, A., Euzenat, J., Gangemi, A., Kalfoglou, Y., Pisanelli, D., & Sure, Y. “**A survey on methodologies for developing, maintaining, integrating, evaluating and reengineering ontologies**”, OntoWeb deliverable D, 1, 2002.

[30] GATE Software, https://gate.ac.uk/, 2014, November, 18.

[31] Hadi, W., “**EMCAR: Expert Multi Class Based on Association Rule**”, International Journal of Modern Education and Computer Science (IJMECS), pp.33, 34, 5 March 2013.

[32] Hadni, M.; Lachkar, A.; Ouatik, S.A.;, “**A new and efficient stemming technique for Arabic Text Categorization**”, 2012 International Conference on Multimedia Computing and Systems (ICMCS), pp.791-796, 10-12 May 2012.

[33] Han, Jiawei, and Micheline Kamber. "**Data mining: concepts and techniques**." (2001).

[34] Hani M. O. Iwidat, Zhou Yi-ming, “**Automatic Arabic Document Classification via KNN**”, School of Computer Science and Engineering, Vol.18 No.2, 2008.

[35] Harrag, F.; El-Qawasmah, E., "**Neural Network for Arabic text classification**", Second International Conference on the Applications of Digital Information and Web Technologies, ICADIWT '09, 2009, pp.778-783, 4-6 August 2009.

[36] Harrag, F.; El-Qawasmah, E.; Al-Salman, A.M.S., "**Stemming as a feature reduction technique for Arabic Text Categorization**", 10th International Symposium on Programming and Systems (ISPS), pp.128-133, 25-27 April 2011.

[37] Hwang M.; Kong H.; Kim P., “**The Design of the Ontology Retrieval System on the Web**”, 8th International Conference Advanced Communication Technology, 2006. ICACT 2006, vol.3, pp.1815, 1818, 20-22 February 2006.

[38] Jarrar M., “**Building a Formal Arabic Ontology**”, In proceedings of the Experts Meeting on Arabic Ontologies and Semantic Networks. Alecso, Arab League. Tunis, April 26-28, 2011.

[39] Jellouli, I.; El Mohajir, M., “**Towards automatic semantic annotation of data rich Web pages**”, Third International Conference on Research Challenges in Information Science, 2009, RCIS 2009, pp.139,142, 22-24 April 2009.

[40] Kaloub, A., “**Automatic Ontology-Based Document Annotation for Arabic Information Retrieval**”, Master's thesis, Islamic University of Gaza, 2013.

[41] Kayed, A., “**Ontology evaluation: Which test to use”**, 5th International Conference on Computer Science and Information Technology (CSIT), 2013, pp.45,48, 27-28 March 2013.

[42] Korde, V.; Mahender, C., “**Text Classification and Classifiers: A Survey**”, International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.2, pp. 85,99, March 2012.

[43] Lam, W.; Ruiz, M.; Srinivasan, P.; “**Automatic text categorization and its application to text retrieval**”, IEEE Transactions on Knowledge and Data Engineering, vol.11, no.6, pp.865-879, November/December 1999.

[44] Lin Y.; Jiang J.; Lee Sh., “**A Similarity Measure for Text Classification and Clustering**”; IEEE Transactions on Knowledge and Data Engineering, vol.26, no.7, pp.1575,1590, July 2014.

[45] Liu, J.N.K.; Yu-Lin He; Lim, E.H.Y.; Xi-Zhao Wang; , “**A New Method for Knowledge and Information Management Domain Ontology Graph Model”**, IEEE Transactions on Systems, Man, and Cybernetics, vol.43, no.1, pp.115-127, January 2013.

[46] Ma J.; Xu W.; Sun Y.; Turban, E.; Wang Sh.; Liu O., “**An Ontology-Based Text-Mining Method to Cluster Proposals for Research Project Selection**” IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, , vol.42, no.3, pp.784,790, May 2012.

[47] Marina, Litvak, et al. "**Improving classification of multi-lingual web documents using domain ontologies**" The Second International Workshop on Knowledge Discovery and Ontologies. Porto, Portugal October 2005.

[48] Ming-Syan Chen; Jiawei Han; Yu, P.S., "**Data mining: an overview from a database perspective**," in Knowledge and Data Engineering, IEEE Transactions on , vol.8, no.6, pp.866-883, Dec 1996

[49] Mu-Hee Song; Soo-Yeon Lim; Dong-Jin Kang; Sang-Jo Lee; , "**Automatic classification of Web pages based on the concept of domain ontology**," 12th Asia-Pacific Software Engineering Conference, APSEC '05., 15-17 December 2005.

[50] Ning H.; Shihan D., "**Structure-Based Ontology Evaluation**" IEEE International Conference on e-Business Engineering, 2006. ICEBE '06, pp.132, 137, Oct. 2006.

[51] Noaman, H.M.; Elmougy, S.; Ghoneim, A.; Hamza, T., "**Naive Bayes Classifier based Arabic document categorization**", The 7th International Conference on Informatics and Systems (INFOS), pp.1,5, 28-30 March 2010.

[52] Noy, N., & McGuinness, D. "**Ontology development 101: A Guide to Creating Your First Ontology**", Knowledge Systems Laboratory, Stanford University, 2001.

[53] Oliveira, P.; Rocha, J., "**Semantic annotation tools survey**", 2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM), pp.301,307, 16-19 April 2013.

[54] Prabowo, R., Jackson, M., Burden, P., Knoell, H.-D.: "**Ontology-Based Automatic Classification for the Web Pages: Design, Implementation and Evaluation**". In Proceedings of the 3rd International Conference on Web Information Systems Engineering, WISE 2002 (2002).

[55] Protege Software, http://protege.stanford.edu/, 2015, January, 15.

[56] Ranganathan, G.; Biletskiy, Y.; Kaltchenko, A., "**Semantic annotation of semi-structured documents**", Canadian Conference on Electrical and Computer Engineering, 2008. CCECE 2008, pp. 919, 922, 4-7 May 2008.

[57] S. Al-Harbi, A. Almuhareb, A. Al-Thubaity, M. S. Khorsheed, and A. Al-Rajeh, "**Automatic Arabic Text Classification**", In Proceedings of 9th International Conference on the Statistical Analysis of Textual Data, JADT'08, France, pp. 77–83, 2008.

[58] Saad M., Ashour,W.; “**Arabic text classification using decision trees**”, 12th international workshop on computer science and information technologies In Proceedings of the CSIT, Moscow–Saint-Petersburg, Russia, 2010.

[59] Seremeti, L.; Kameas, A., "**A Task-Based Ontology Engineering Approach for Novice Ontology Developers**", Fourth Balkan Conference in Informatics, 2009. BCI '09, pp.85, 89, 17-19 September 2009.

[60] Shian-Hua Lin; Meng Chang Chen; Jan-Ming Ho; Yueh-Ming Huang;, “**ACIRD: intelligent Internet document organization and retrieval**”, IEEE Transactions on Knowledge and Data Engineering, vol.14, no.3, pp.599-614, May/Jun 2002.

[61] Sure, Y., Staab, S., & Studer, R. “**Handbook on ontologies**”, Springer Berlin Heidelberg, pp. 135,152, 2009.

[62] Taswell, C., "**DOORS to the Semantic Web and Grid With a PORTAL for Biomedical Computing**", IEEE Transactions on Information Technology in Biomedicine, vol.12, no.2, pp.191,204, March 2008.

[63] Thabtah, F.; Gharaibeh, O.; Abdeljaber, H., “**Comparison of rule based classification techniques for the Arabic textual data**”, 2011 Fourth International Symposium on Innovation in Information & Communication Technology (ISIICT), pp.105,111, 29 November 2011-1 December 2011.

[64] Witten, Ian H., and Eibe Frank. "**Data Mining: Practical machine learning tools and techniques**". Morgan Kaufmann, 2005.

[65] Xiaogang, P.; Choi., B. “**Document Classifications based on Word Semantic Hierarchies**”, Artificial Intelligence and Applications. Vol. 5. 2005.

[66] Yahya, A.H.; Salhi, A.Y., "**Enhancement tools for Arabic web search**", International Conference on Innovations in Information Technology (IIT), pp.71-76, 25-27 April 2011.

[67] Yang Liu; Zhiqing Shao, "**A framework for semantic Web Services annotation and discovery based on ontology**", 2010 IEEE International Conference on Progress in Informatics and Computing (PIC), vol.2, pp.1034,1039, 10-12 December 2010.

[68] Yip Chi Kiong; Palaniappan, S.; Yahaya, N.A., "**Health ontology system**", 2011 7th International Conference on Information Technology in Asia (CITA 11), vol., no., pp.1, 4, 12-13 July 2011.

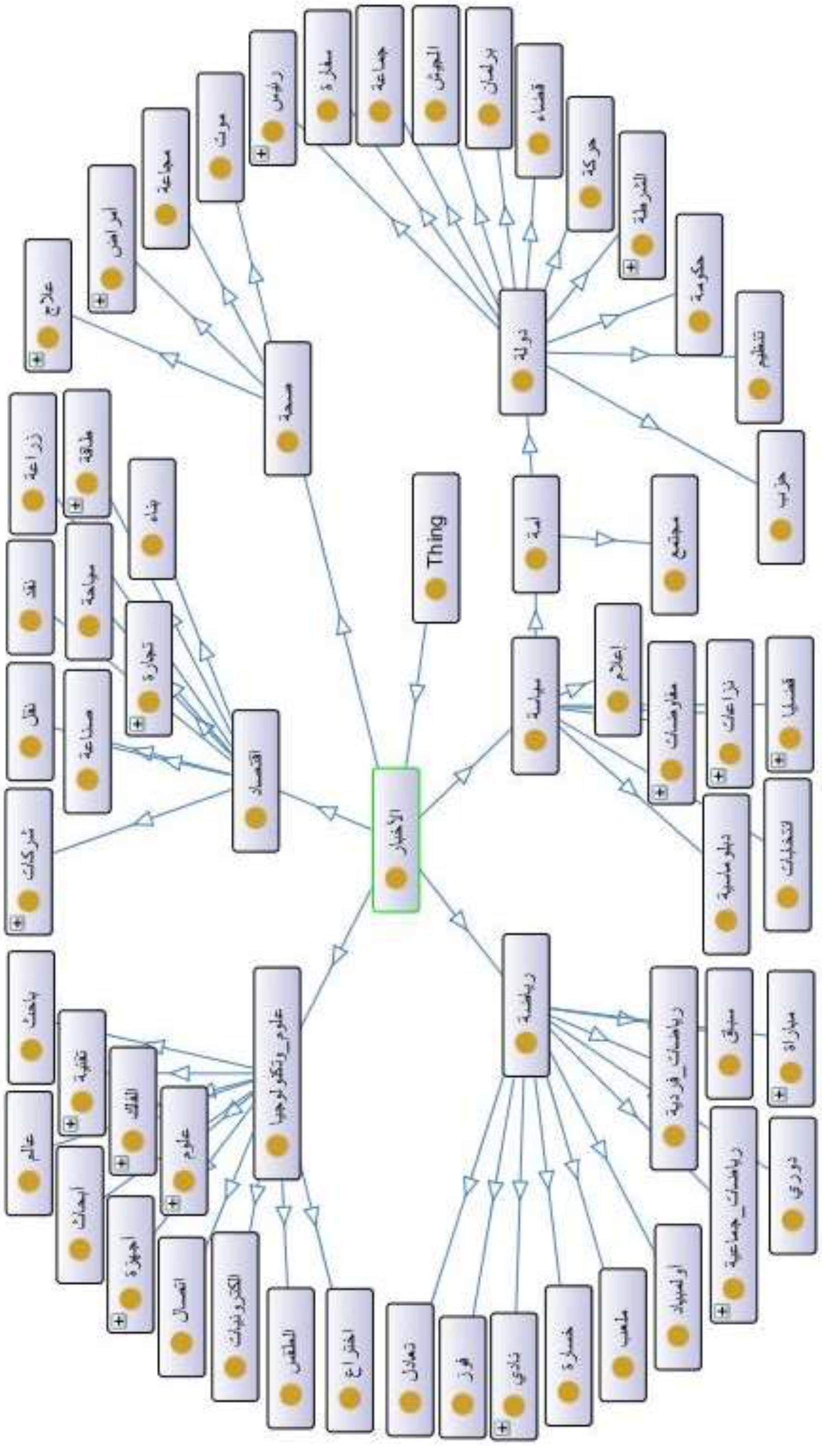
[69] Zaidi, S., Laskri, M. T., & Bechkoum, K. "**A cross-language information retrieval based on an Arabic ontology in the legal domain**", In Proceedings of the International Conference on Signal-Image Technology and Internet-Based Systems, SITIS'05, pp.86, 91, November 2005.

[70] Zakaria E., Amel B., Malika T.: "**Medical Documents Classification Based on the Domain Ontology MeSH"**, International Arab Journal e-Technology, 2(4): pp. 210-215, 2012.

[71] Zhang, J., "**Ontology and the Semantic Web**", Proceedings of the North American Symposium on Knowledge Organization. Vol. 1, 2007.

Appendices

Appendix A: The News Ontology



Appendix B: Part of the News Ontology Source Code in Owl

```
<?xml version="1.0"?>

<!DOCTYPE rdf:RDF [
    <!ENTITY owl "http://www.w3.org/2002/07/owl#" >
    <!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
    <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
    <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
]>

<rdf:RDF xmlns="http://www.semanticweb.org/mohammed/ontologies/news-ontology#"
     xml:base="http://www.semanticweb.org/mohammed/ontologies/news-ontology"
     xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
     xmlns:owl="http://www.w3.org/2002/07/owl#"
     xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
     xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
    <owl:Ontology rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology"/>

    <!--
    ////////////////////////////////////////////////////////////////////////////////
    //
    // Annotation properties
    //
    /////////////////////////////////////////////////////////////////////////////-->

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#Synonyms -->

    <owl:AnnotationProperty
rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#Synonyms"/>

    <!--
    ////////////////////////////////////////////////////////////////////////////////
    //
    // Object Properties
    //
    ////////////////////////////////////////////////////////////////////////////-->


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#إلى_تحتاج -->

    <owl:ObjectProperty rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#إلى_تحتاج">
        <rdfs:domain rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#أمراض"/>
        <rdfs:range rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#علاج"/>
    </owl:ObjectProperty>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#بواسطة_تحل -->

    <owl:ObjectProperty rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#بواسطة_تحل">
        <rdfs:range rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#مفاوضات"/>
        <rdfs:domain rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#نزاعات"/>
    </owl:ObjectProperty>
    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#بواسطة_تستخدم -->

    <owl:ObjectProperty rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#بواسطة_تستخدم">
        <rdfs:range rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#أجهزة"/>
        <rdfs:domain rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#تقنية"/>
    </owl:ObjectProperty>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#في_يلعب -->

    <owl:ObjectProperty rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#في_يلعب">
        <rdfs:range rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#فريق"/>
        <rdfs:domain rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#لاعب"/>
    </owl:ObjectProperty>

    <!--
    ///////////////////////////////////////////////////////////////////////////////////////
    //
    // Classes
    //
    ///////////////////////////////////////////////////////////////////////////////////////-->

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#أبحاث -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#أبحاث">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#وتكنولوجيا_علوم"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#أجهزة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#أجهزة">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#وتكنولوجيا_علوم"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#أحياء -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#أحياء">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#علوم"/>
        <Synonyms>الأحياء</Synonyms>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#أمة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#أمة">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#سياسة"/>
        <rdfs:comment>الأمم</rdfs:comment>
    </owl:Class>
    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#أمراض -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#أمراض">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#صحة"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#أولمبياد -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#أولمبياد">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#رياضة"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#إرهاب -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#إرهاب">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#قضايا"/>
        <rdfs:comment>الإرهاب</rdfs:comment>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#إعلام -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#إعلام">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#سياسة"/>
        <rdfs:comment>الإعلام</rdfs:comment>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#إنترنت -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#إنترنت">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#تقنية"/>
        <Synonyms>الإنترنت</Synonyms>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#اتصال -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#اتصال">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#علوم_وتكنولوجيا"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#احتلال -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#احتلال">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#قضايا"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#اقتصاد -->
    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#اقتصاد">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#الأخبار"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الأخبار -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#الأخبار"/>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الجيش -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#الجيش">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#دولة"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الحرس -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#الحرس">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#الشرطة"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#السوق -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#السوق">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#تجارة"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الشرطة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#الشرطة">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#دولة"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الشمس -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#الشمس">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#الفلك"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الطقس -->
    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#الطقس">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#وتكنولوجيا_علوم"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الفضاء -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#الفضاء">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#الفلك"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الفلك -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#الفلك">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#وتكنولوجيا_علوم"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#القمر -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#القمر">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#الفلك"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الكترونيات -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#الكترونيات">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#وتكنولوجيا_علوم"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#انتخابات -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#انتخابات">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#سياسة"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#باحث -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#باحث">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#وتكنولوجيا_علوم"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#برامج -->
    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#برامج">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#تقنية"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#برلمان -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#برلمان">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#دولة"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#بريد -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#بريد">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#تقنية"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#بناء -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#بناء">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#اقتصاد"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#بنك -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#بنك">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#شركات"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#بيولوجيا -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#بيولوجيا">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#علوم"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#تجارة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#تجارة">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#اقتصاد"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#تعادل -->
    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#تعادل">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#رياضة"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#تقنية -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#تقنية">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#علوم_وتكنولوجيا"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#تنظيم -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#تنظيم">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#دولة"/>
        <rdfs:comment>التنظيم</rdfs:comment>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#جماعة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#جماعة">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#دولة"/>
        <rdfs:comment>الجماعة</rdfs:comment>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#حارس -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#حارس">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#لاعب"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#حاسوب -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#حاسوب">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#أجهزة "/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#حرب -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#حرب">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#قضايا"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#حركة -->
    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#حركة">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#دولة"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#حزب -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#حزب">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#دولة"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#حكومة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#حكومة">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#دولة"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#خسارة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#خسارة">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#رياضة"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#دبلوماسية -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#دبلوماسية">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#سياسة"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#دوري -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#دوري">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#رياضة"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#دولة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#دولة">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#أمة"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#رئيس -->
    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#رئيس">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#دولة"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#جماعية_رياضات -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#جماعية_رياضات">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#رياضة"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#رياضة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#رياضة">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#الأخبار"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#زراعة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#زراعة">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#اقتصاد"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#سباق -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#سباق">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#رياضة"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#سفارة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#سفارة">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#دولة"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#سلام -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#سلام">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#قضايا"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#سياحة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#سياحة">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#اقتصاد"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#سياسة -->
    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#سياسة">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#الأخبار"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#صيدلة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#صيدلة">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#علوم"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#طاقة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#طاقة">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#اقتصاد"/>
    </owl:Class>



    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#طب -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#طب">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#علوم"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#عالم -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#عالم">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#علوم_وتكنولوجيا"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#علاج -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#علاج">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#صحة"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#علوم -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#علوم">
        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#علوم_وتكنولوجيا"/>
    </owl:Class>


    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#علوم_وتكنولوجيا -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-ontology#علوم_وتكنولوجيا">
```

```
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#الأخبار"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#فريق -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#فريق">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#نادي"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#فوز -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#فوز">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#رياضة"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#فيزياء -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#فيزياء">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#علوم"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#كرة_السلة -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#كرة_السلة">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#رياضات_جماعية"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#كرة_القدم -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#كرة_القدم">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#رياضات_جماعية"/>
    </owl:Class>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#كيمياء -->

    <owl:Class rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#كيمياء">
        <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-ontology#علوم"/>
    </owl:Class>


    <!--/////////////////////////// Individuals///////////////////////////-->

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الأنيميا -->

    <owl:NamedIndividual rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#الأنيميا">
        <rdf:type rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#أمراض"/>
    </owl:NamedIndividual>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#البترول -->
    <owl:NamedIndividual rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#البترول">
        <rdf:type rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#مصادر"/>
        <Synonyms>بترول</Synonyms>
    </owl:NamedIndividual>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#التيفوئيد -->

    <owl:NamedIndividual rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#التيفوئيد">
        <rdf:type rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#أمراض"/>
    </owl:NamedIndividual>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الجزائر -->

    <owl:NamedIndividual rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#الجزائر">
        <rdf:type rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#دولة"/>
    </owl:NamedIndividual>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الجفاف -->

    <owl:NamedIndividual rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#الجفاف">
        <rdf:type rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#أمراض"/>
        <Synonyms>جفاف</Synonyms>
    </owl:NamedIndividual>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الحصبة -->

    <owl:NamedIndividual rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#الحصبة">
        <rdf:type rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#أمراض"/>
    </owl:NamedIndividual>

    <!-- http://www.semanticweb.org/mohammed/ontologies/news-ontology#الحمى -->

    <owl:NamedIndividual rdf:about="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#الحمى">
        <rdf:type rdf:resource="http://www.semanticweb.org/mohammed/ontologies/news-
ontology#أمراض"/>
        <Synonyms>حمى</Synonyms>
    </owl:NamedIndividual>

<!-- Generated by the OWL API (version 3.5.0) http://owlapi.sourceforge.net -->
```