بسم الله الرحمن الرحيم

جامعة السودان للعلوم والتكنولوجيا

كلية علوم الحاسوب وتقانة المعلومات

Building Of University Ontology

بناء إنطولوجيا الجامعة

بحث مقدم كمتطلب جزئي لنيل درجة بكلاريوس هندسة البرمجيات

إعداد الطلاب:
قبس محمد عثمان
منال الغفاري إبراهيم
ميادة كمال عبد القادر

إشراف:
د.عفاف مدني

نوفمبر 2020

الآية

قال تعالي :(الذين ضل سعيهم في الحياة الدنيا و هم يحسبون أنهم يحسنون صنعا)

[سورة الكهف الآية: 104]

شكر وتقدير

يسرني أن أشكر كل من درسني أو ساهم في تدريسي من دكاترة جامعة وكل الأساتذة الذين يرجع لهم الفضل بعد الله عز وجل ، كما اقدم الشكر والتقدير للأساتذة المشرفين على هذا البحث المتواضع، الذي اسال الله تعالى أن يضيف قيمة إلى هذا العلم، وشكر موجه كذلك لإدارة جامعة السودان للعلوم و التكنولوجيا لحسن توفيرهم وتسهيلهم الخدمات للطلاب ومساعدتهم في كل الأمور التي من شانها أن تخول لهم فضاءً مريحا للدراسة وطلب العلم في أمان ونظام.

إهداء

إلى أمي الحنونة لا أجد كلمات يمكن أن تمنحها حقها، فهي ملحمة الحب وفرحة العمر، ومثال التفاني والعطاء.

إلى أبي العطوف قدوتي، ومثلي الأعلى في الحياة؛ فهو من علَّمني كيف أعيش بكرامة وشموخ.

إلى إخوتي سندي وعضدي ومشاطري أفراحي وأحزاني.

إلى جميع الأخلاء، أهدي إليكم بحثي العلمي.

Abstract

The current web is based on HTML which is not able to be exploited by information retrieval techniques and hence processing of information on web is mostly restricted to manual keyword searches which results in irrelevant information retrieval. This limitation may be overcome by a new web architecture known as semantic web which is an intelligent and meaningful web proposed by Sir Tim Berner's Lee. In his roadmap for semantic web, Ontology plays a pivotal role in information exchange, use and re-uses knowledge, shared and common understanding of a domain that can be communicated between people and across application systems which is the goal of semantic web. Ontology is used to capture knowledge about any domain of interest with the objective of incorporating the machine understandable data on the current human-readable web. Web Ontology Language (OWL) is a semantic markup language for sharing ontologies on the web and is designed for use by software agents to empower them to comprehend the meaning of web documents. Ontology is a broad term including a wide range of activities, complexities and issues in which Ontology Development is one of the most fundamental and significant concern. There may be various methodologies or tools for ontology development .we consider the university domain and demonstrate the development of a University Ontology using TopBraid compose .we have various aspects like super class and sub class hierarchy, creating a sub class, instances for classes illustration, query retrieval process, Graph corresponding to a sub-class,, and the Sudan University of Science and Technology was taken as a case study.

المستخلص

يعتمد الويب الحالي على HTML الذي لا يمكن استغلاله من خلال تقنيات استرجاع المعلومات ، وبالتالي فإن معالجة المعلومات على الويب تقتصر في الغالب على عمليات البحث اليدوية بالكلمات الرئيسية التي تؤدي إلى استرجاع معلومات غير ذات صلة. يمكن التغلب على هذا القيد من خلال بنية ويب جديدة تُعرف باسم الويب الدلالي وهي شبكة ذكية وذات مغزى اقترحها لي السير تيم بيرنر. في خارطة الطريق الخاصة به للويب الدلالي ، يلعب علم الوجود دورًا محوريًا في تبادل المعلومات ، واستخدام وإعادة استخدام المعرفة ، والفهم المشترك والمشترك للمجال الذي يمكن توصيله بين الناس وعبر أنظمة التطبيق وهو هدف الويب الدلالي. يتم استخدام علم الوجود للحصول على المعرفة حول أي مجال من مجالات الاهتمام بهدف دمج البيانات المفهومة آليًا على شبكة الويب الحالية التي يمكن للبشر قراءتها. لغة Web Ontology (OWL)هي لغة ترميز دلالية لمشاركة الأنطولوجيا على الويب وهي مصممة للاستخدام بواسطة وكلاء البرامج لتمكينهم من فهم معنى مستندات الويب. علم الوجود مصطلح واسع يشمل مجموعة واسعة من الأنشطة والتعقيدات والقضايا التي يعتبر فيها تطوير علم الوجود أحد أهم وأهم الاهتمامات. قد تكون هناك منهجيات أو أدوات مختلفة لتطوير الأنطولوجيا ، فنحن نأخذ في الاعتبار مجال الجامعة ونثبت تطوير علم الوجود الجامعي باستخدام مؤلف TopBraid. لدينا جوانب مختلفة مثل التسلسل الهرمي للفئة الفائقة والفئة الفرعية ، وإنشاء فئة فرعية ، وحالات لتوضيح الفصول ، عملية استرجاع الاستعلام ، والرسم البياني المقابل لفئة فرعية ، وجامعة السودان للعلوم والتكنولوجيا كدراسة حالة.

Table of Figures

Contents

Chapter One

Introduction

1.1 Background:

Several tools and methods have been developed to build ontologies. Rather than focusing all the attention to information, it was all in the core concepts in using ontology and its relationships. The most well-known and widespread tool for editing and developing ontologies is Protégé. Definition of the University concept is clarified through the university ontology. Creating a university ontology using TopBraid is the object of this research. Sudan University of Science and Technology is taken as a case study for the development of the ontology and several phases are outlined e.g., superclass, hierarchy of subclasses, creating subclasses instances, retrieving queries, graphs and visualization views. The case study is limited to few departments and courses, as an example. This implies that since the model works for one university, it will work for other universities, with minor changes.

1.2 The problem of Statement:

Lack of university ontology, No complete structure contains all entities “and their relationship between them” in the university and No standardization of terms.

1.3 The Objectives:

1) Building Ontology for university:
2) Specify the terms, entities and concepts.
3) Sampling the relation between the entities.
4) Building a high level of knowledge that has benefits and meaning.
5) Organizing and classify the knowledge to integrate with database and other software, hence retrieval can done easily and quickly.

1.4 Scope:

Create a complete structure “Ontology”, which contain all the entities and terms that form the university and the relation between them for Sudan University for Science and Technology as a case study.

1.5 Methodology:

Using the waterfall model for the methodology in several steps:

- Collect information.
- Analyse information.
- Determine entities and linking them together.
- Enter data related to the university.
- Using Sparql query to retrieve specific information.

1.6 Thesis Structure:

This chapter described the problem, objectives, scope and the methodology.
The later chapters will show the previous studies and literature, how to create the ontology and what are the tools to create it, the results and recommendation.

Chapter Two

Literature review and previous studies

2.1 Introduction:

In this chapter we talk about the ontology, its definition, types, tools and techniques used, their applications, previous studies, and how to benefit from these studies

2.2 Ontology

Is the branch of philosophy that studies concepts such as existence, being, becoming, and reality, It includes the questions of how entities are grouped into basic categories and which of these entities exist on the most fundamental level. Ontology is traditionally listed as a part of the major branch of philosophy known as metaphysics.

2.3 Examples of applications:

- Enterprise applications. A more concrete example is SAPPHIRE (Health care) or *Situational Awareness and Preparedness for Public Health Incidences and Reasoning Engines* which is a semantics-based health information system capable of tracking and evaluating situations and occurrences that may affect public health.
- Geographic information systems bring together data from different sources and benefit therefore from ontological metadata which helps to connect the semantics of the data.
- Domain-specific ontologies are extremely important in biomedical research, which requires named entity disambiguation of various biomedical terms and abbreviations that have the same string of characters but represent different biomedical concepts. For example, CSF can represent Colony Stimulating

Factor or Cerebral Spinal Fluid, both of which are represented by the same term, CSF, in biomedical literature. This is why a large number of public ontologies are related to the life sciences. Life science data science tools that fail to implement these types of biomedical ontologies will not be able to accurately determine causal relationships between concepts.

2.4 Tools and techniques:

2.4.1 Protégé:

Protégé is an ontology and knowledge base editor produced by Stanford University. Protégé is a tool that enables the construction of domain ontologies, customized data entry forms to enter data. Protégé allows the definition of classes, class hierarchies, variables, variable value restrictions, and the relationships between classes and the properties of these relationships. Protégé is free and can be downloaded from the internet. Protégé comes with visualization packages such as OntoViz; all of these help the user visualize ontologies with the help of diagrams. The main strong point of Protégé is that it supports at the same time tool builders, knowledge engineers and do main specialists. This is the main difference with existing tools, which are typically targeted at the knowledge engineer and lack flexibility for meta-modeling. This latter feature makes it easier to adapt Protégé to new requirements and/or changes in the model structure.

2.4.2 TopBraid:

TopBraid Composer is a visual modelling environment from industry experts for creating and managing domain models and ontologies in the Semantic Web standards RDF, RDFS and OWL. Composer is an ontology editor and

knowledge-based framework that provides visual editing support as well as interoperability with UML, XML Schema and databases.

It offers a convenient drag-and-drop, form-based user interface with the ability to view and edit ontologies in a variety of serialization formats. Testing, consistency checking and debugging is supported by built-in OWL Inference engine, SPARQL query engine and Rules engine.

2.5 Previous studies:

- There are many applications for ontology in many domains like:
- A process for building domain ontology: An experience in developing a government budgetary ontology
- Developing an University Ontology in Education Domain using Protégé for Semantic Web
- The Computer Science Ontology: A Large-Scale Taxonomy of Research Areas
- Case Study University of Palestine
- Case Study Ahlia University" AL-Bahrain ".

In the domain of universities there are two examples: Ahlia University and University of Palestine:
Ahlia University Ontology was a reuse of the University Ontology, describes all the departments under the University structure and the relationships that exist between them, they have modified the OWL version of the University Ontology and added more classes and restrictions based on the University organization chart of Ahlia University to get the final OWL of the Ontology .
The owl file of the ontology is imported into TopBraid Composer for more powerful data retrieval software, to get the data needed from the ontology easily with short SPARQL queries.

In University of Palestine they selected the UNIFIED process for Building the ontology (UPON) collect the requirement from the university, which consist from the employees and the stockholders.
Ontology was built from scratch, after finishing the design phase the implementation of building ontology, using protégé software was started. Tested the ontology using the protégé software either from graph to emphases that their work was correct or from SPARQL query.

2.6 Summary:

There was a clear difference in terms of terminology in our university and other universities and some difference in the main structure of the university as the structure was built in general in terms of the main entities. As for the structure of our university, we have mentioned in detail the main entities and their sub-entities, as explained in chapter three.

Chapter There

Methodology

3.1 Introduction:

This chapter describe the methodology steps that we used in building the ontology, and we will mention in detail how to do each step.

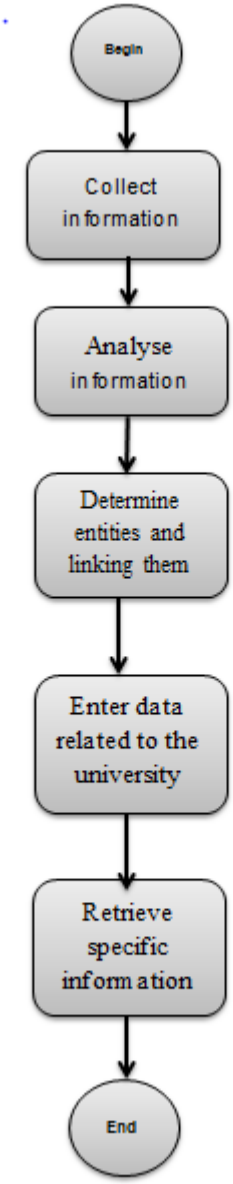

Figure 3.1: The Methodology Steps

3.2 The first step:

Collecting data and information from the interviews and referring to the university's website and the university law book, and an example of the interviews. We went back to some professors and the director of the information centre, and the university's website, so we took a lot of information from it, such as all the colleges in the university and the departments of these colleges and the deanships of the university.

Figure 3.2 Show the main page of the university website

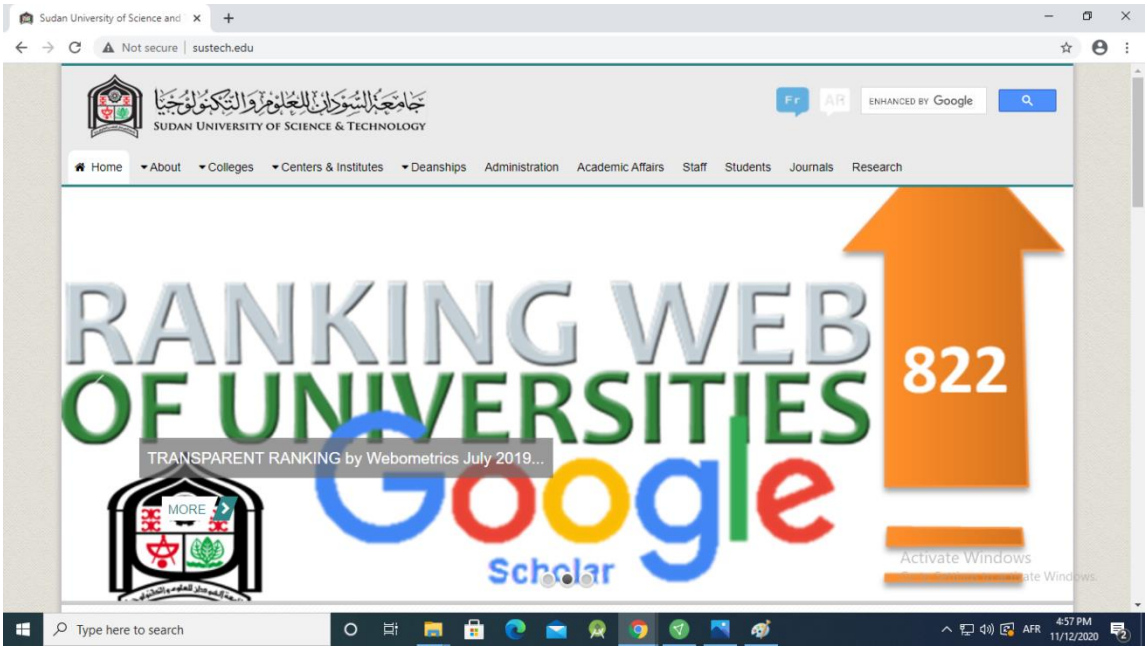

Figure 3.2: University's website

Figure 3.3 show all the colleges in the Sudan University of Science and Technology

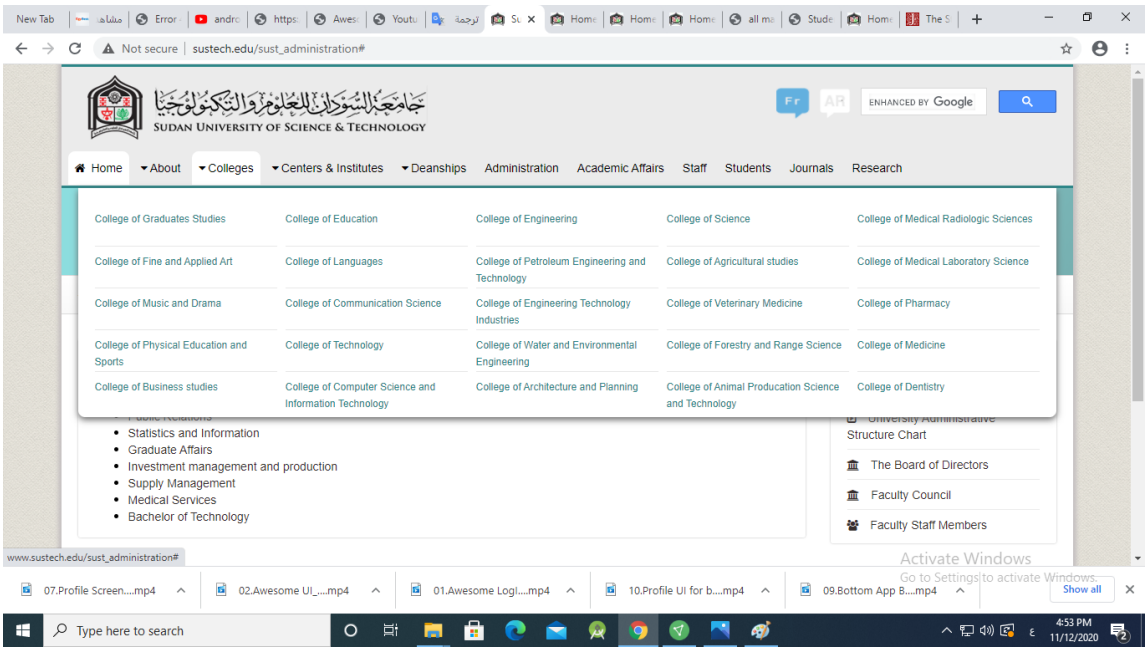

Figure 3.3: University's colleges

Figure3.4 show all the deanships in the Sudan University of Science and Technology.

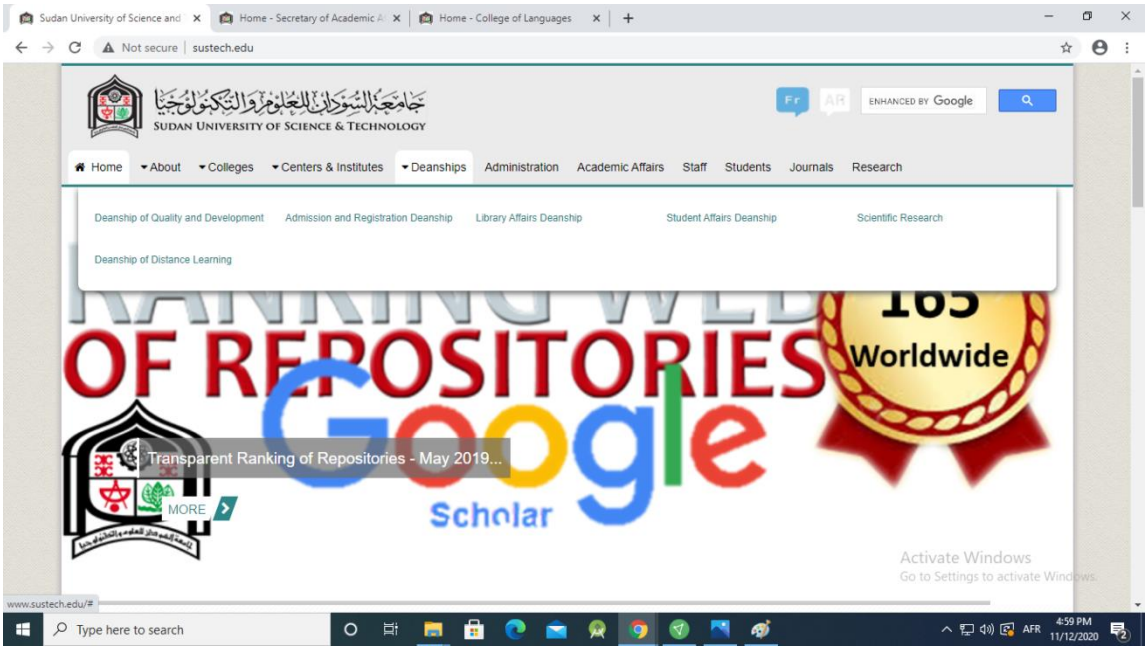

Figure 3.4: University's deanships

Figure 3.5 One example show the departments of the colleges on the university's website and the departments of the College of Computer Science and Information Technology.

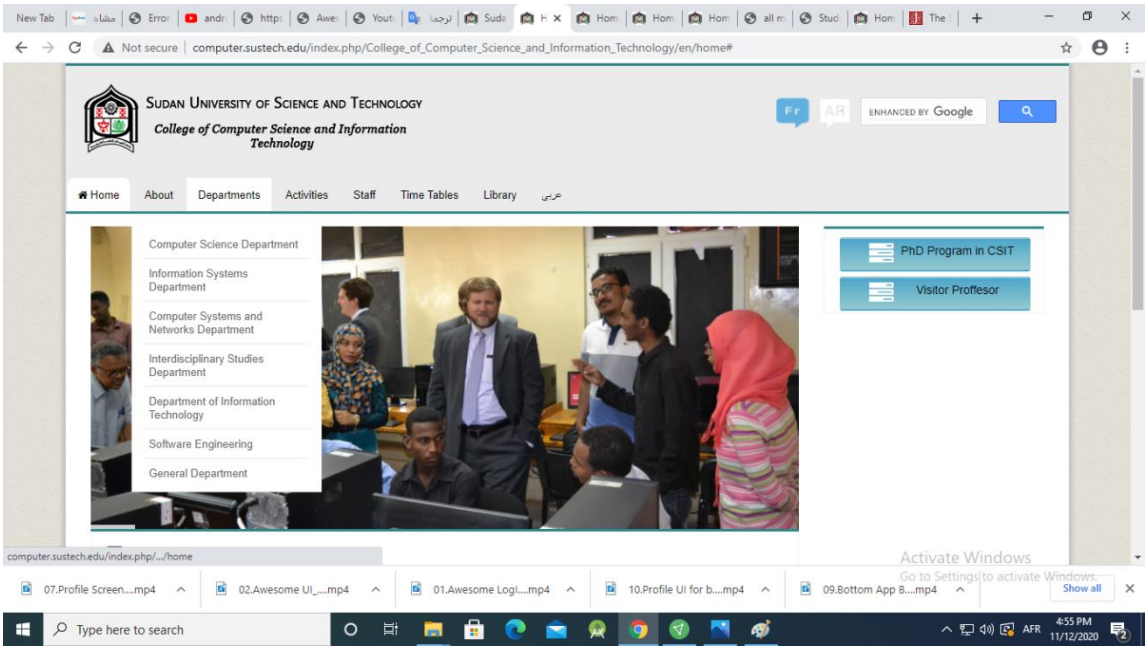

Figure 3.5: Departments of the College of Computer Science and Information Technology

The university law book also contains a lot of information that we have benefited from, such as the types of councils that exist at the university and its members, their relationship with departments, the relationship of the teaching staff with students, and the administrative and executive structure of the university.

3.3 The second step:

Analysis is Examination and audit of data, combing it to be more accurate and reconfiguring it to obtain information on the basis of which decisions can be made and determined , and appropriate relationships identified and and entities defined based on this information.

3.4 **The third step:**

We used Topbraid Composer to build an ontology for the university, explaining the main components of the university and the relationships between them.

The following are the figures illustrated

Show main component of topbraid :

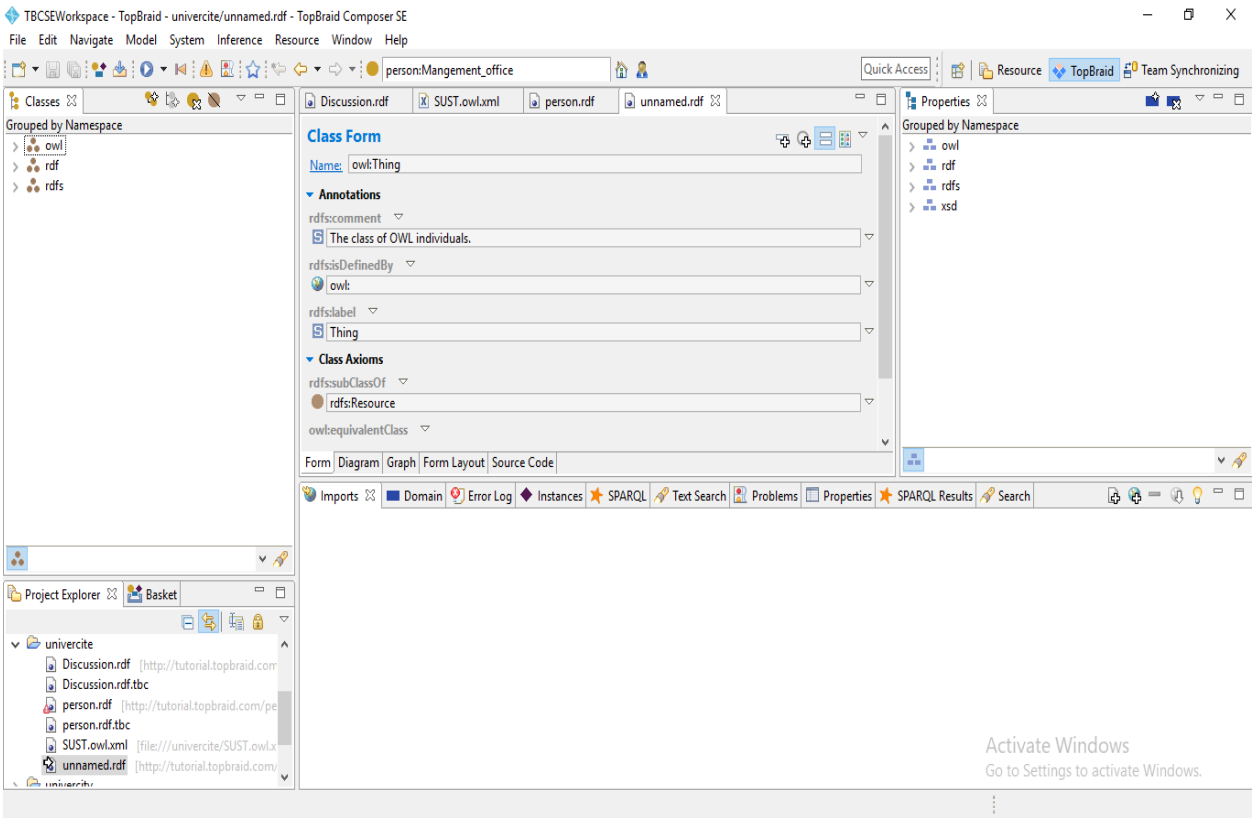

Figure 3.6: Main component of TopBraid

Figure 3.7 Show create new class in the topbraid.

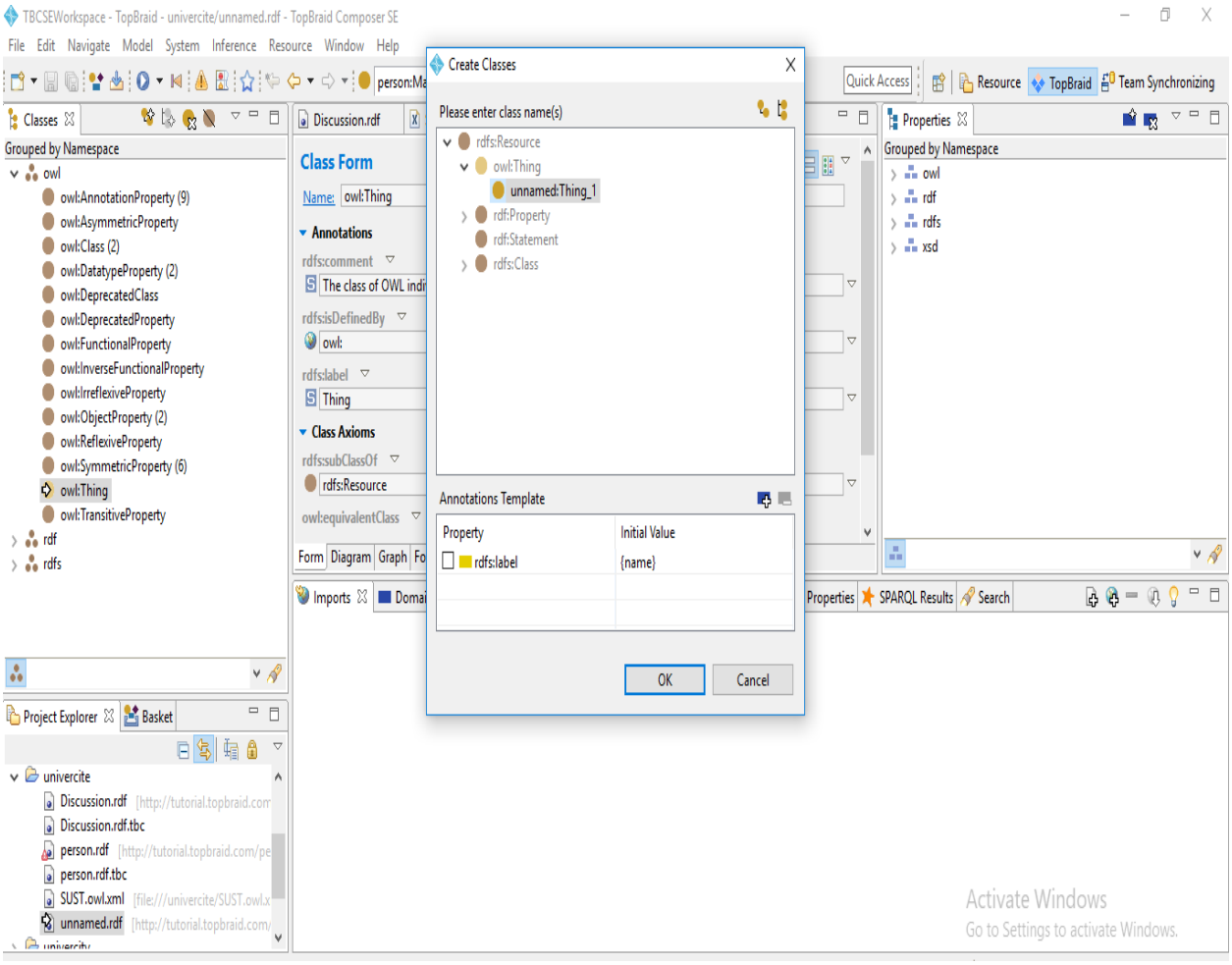

Figure 3.7: create class

Figure 3.8 show form to specific class.

Figure 3.8: Form class

Figure 3.9 show Diagram to specific class:

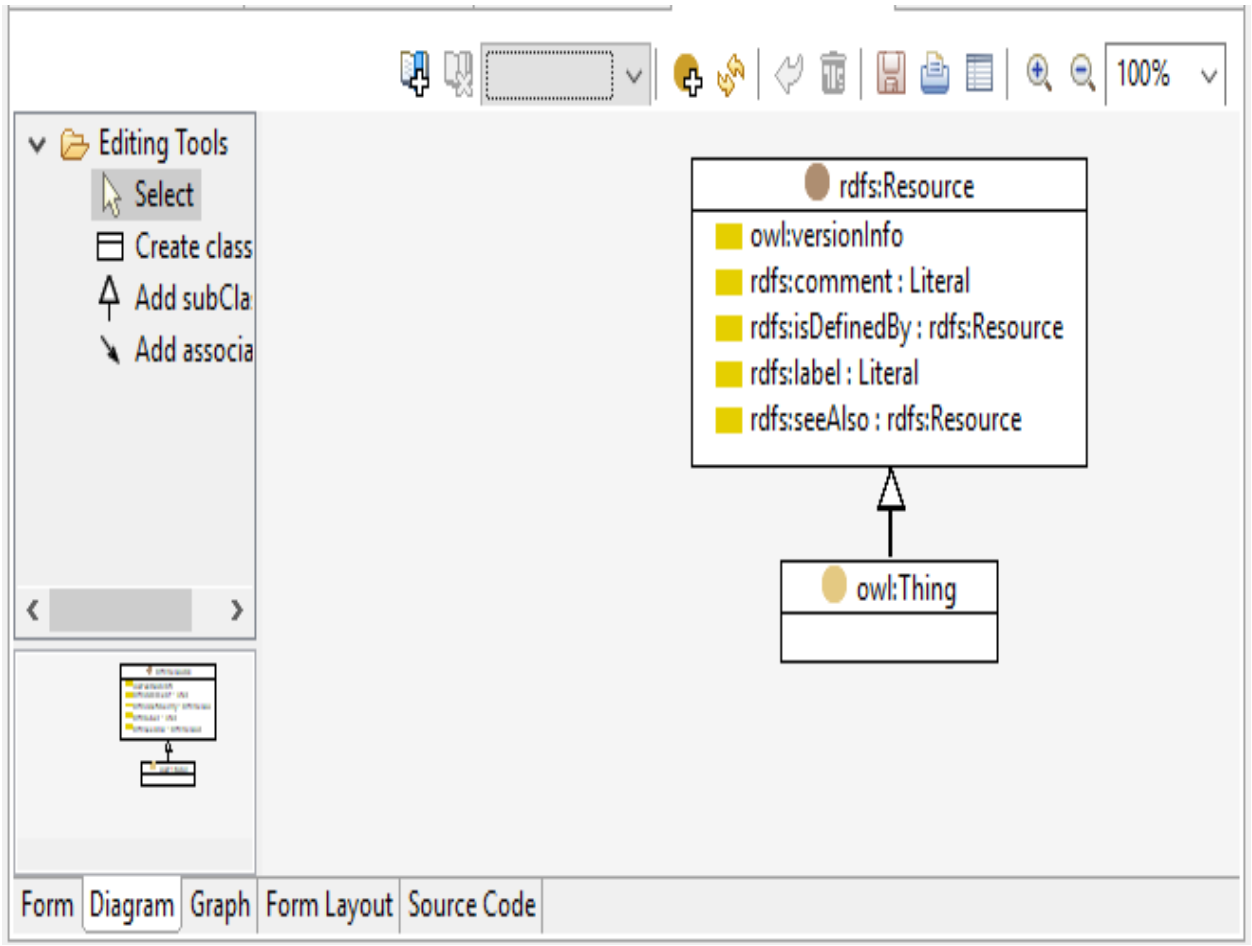

Figure 3.9: specific diagram

Figure 3.10 show create object property that specify relationship between entities.

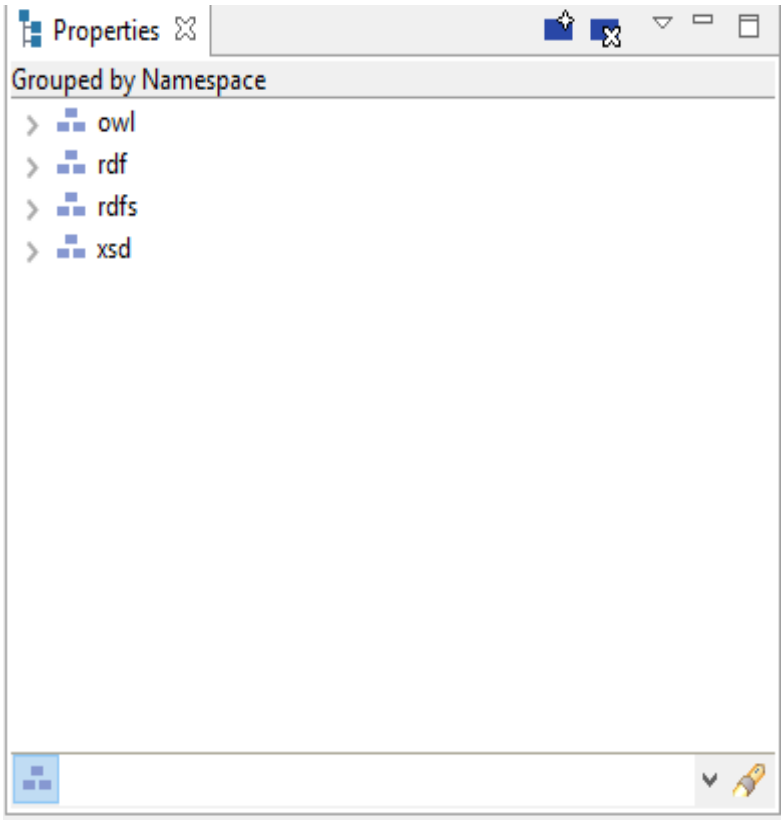

Figure 3.10: specific property

- We identified a number of major entities of university collage, person, administration office, publishing house, etc….. .
- And identified sub-entities: student, teacher, dean, departments, agent, etc…
- There are a number of relationships that bind these entities. Examples:
- The student studies in the classes and the labs and registers in the administration office, the professor publishes in the publishing house, professor taught student in specific class-room, teacher-assiest taught student in specific lab.

2.5 The four step:

Entered data related the Sudan University of Science and Technology.

We entered all collage in the university, All department to all collages, All councils in the university, etc.

Figure 3.11 Show All collage in the Sudan University of Science and Technology

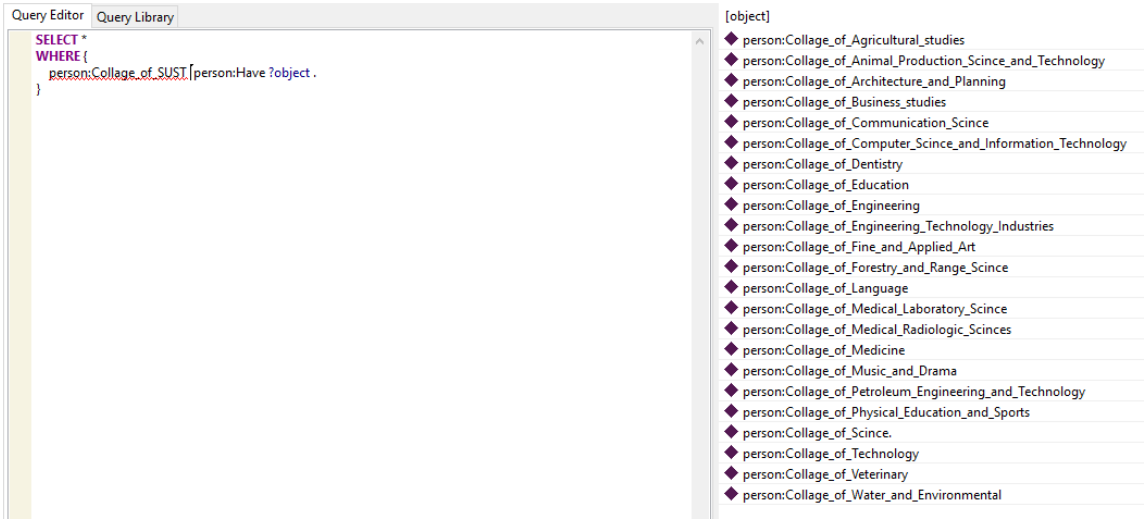

Figure 3.11: collages of Sudan

Figure 3.12 show collage of computer since departments

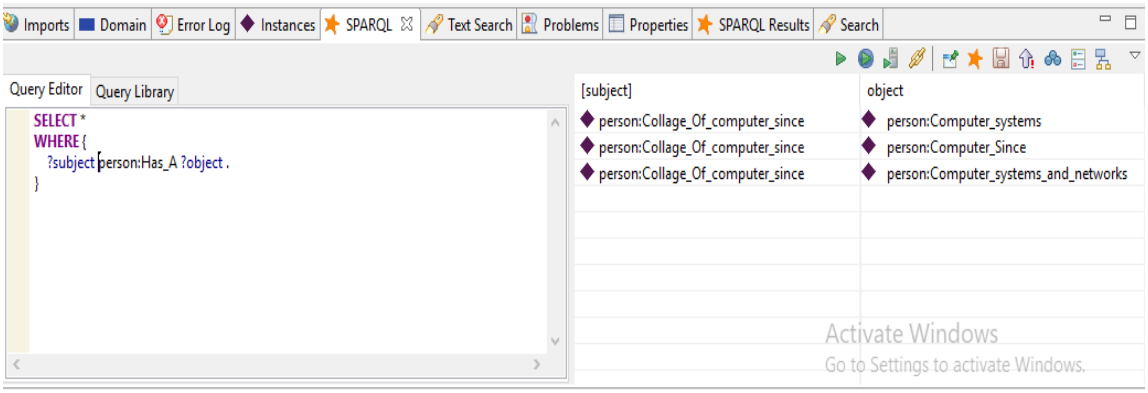

Figure 3.12: departments of one collage

3.6 The five step:

Using sparql query to retrieve specific information.

Figure 3.13 retrieve the Information for specific student (odai):

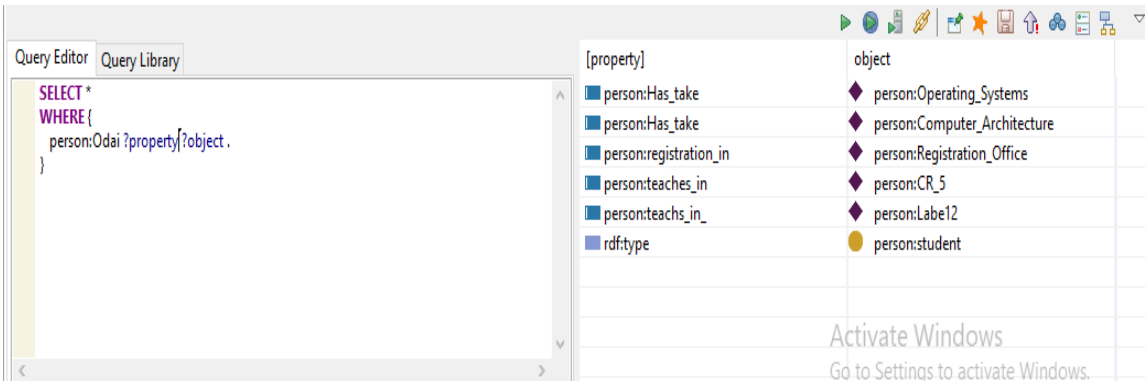

Figure 3.13: Information odai

3.7 Summary:

How to collect the information, to build the ontology by using the TopBraid tool, to obtain the specific results.

Chapter Four
The Results

In figure 4.1, Show all the components of the university, which are the main entities that make up the university.
In the university there have many entities, we show the most importance entities: Administration, classes-room, councils, collages, councils of other university institutions, management-office, person, university-publication-house.

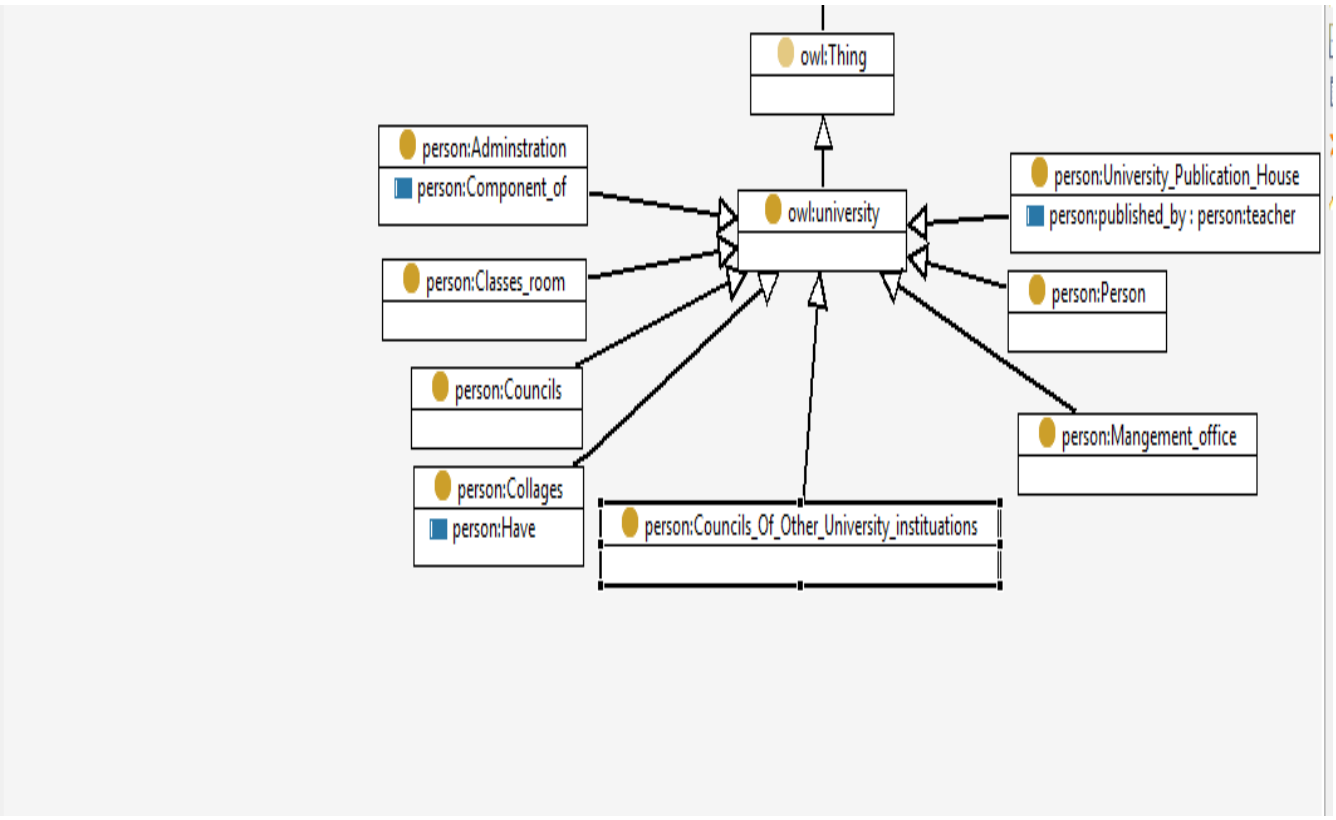

Figure 4.1: University components

In figure 4.2, Show some of the colleges in the university.
In the university there have many collage :collage of communication scince, collage of Education etc….

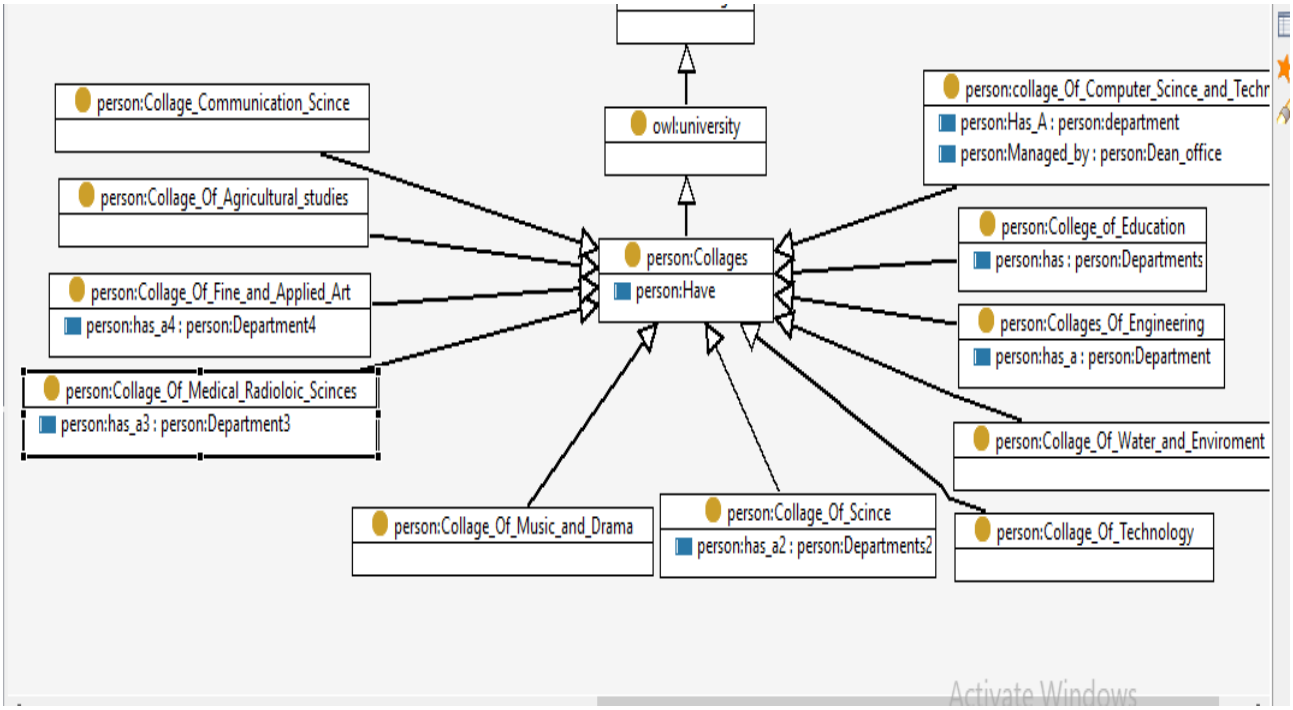

Figure 4.2: University colleges

This figure 4.3 show all the councils that exist in the university.

University have councils : senate-council, student-affair-council,university-council,university-hospitals-council.

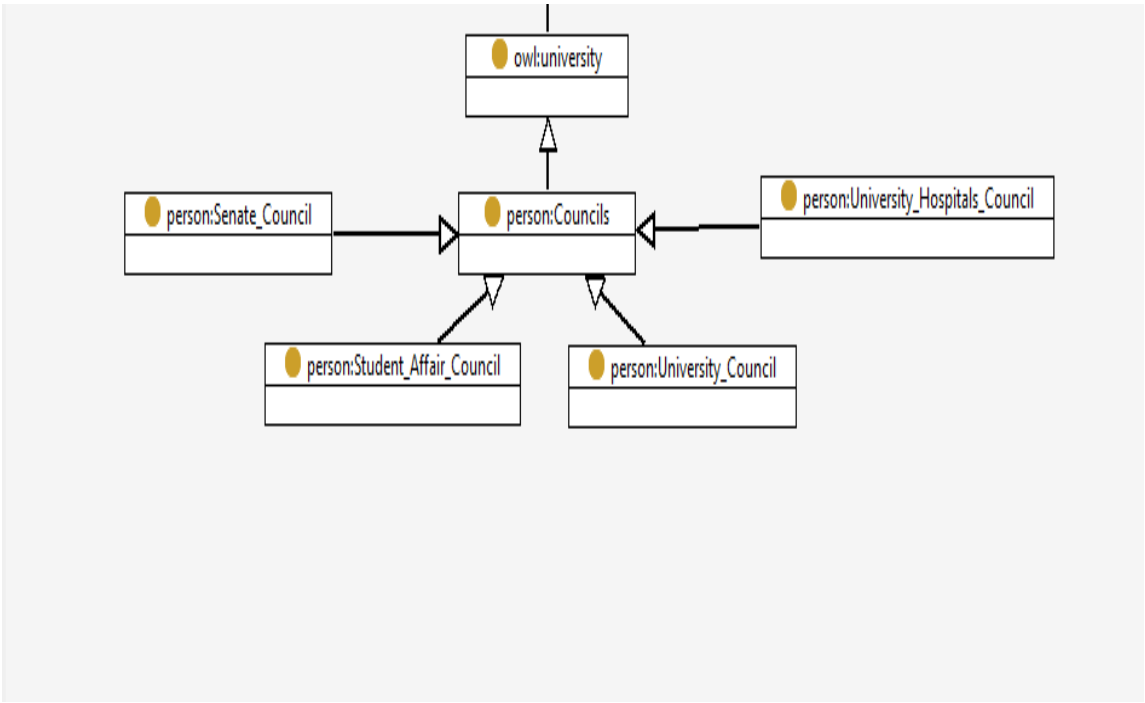

Figure 4.3: University councils

This figure 4.4 show the person entity and all its sub-entities
The person in university has student, teacher, employee.
Student relationship: Has-take, registration-in, teaches-in.
Teacher relationship: publish.
Employee relationship: managed-by.

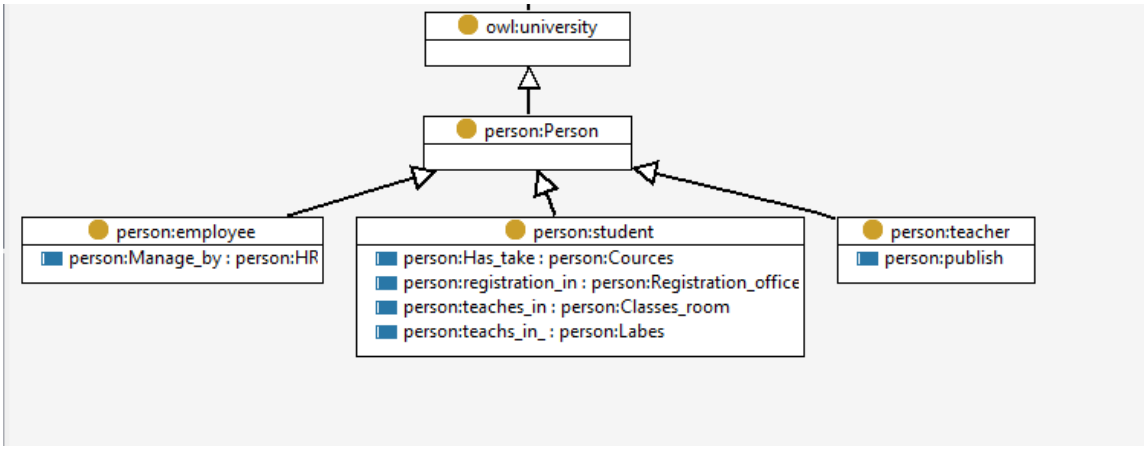

Figure 4.4: person entity

This figure 4.5 show the sub-entity, the teacher, its sub-entities, and their relationship to other entities

Teacher has subclass professor and Assistance-teacher.

Professor teaches in class-room, Assistance-teacher teaches in labs

Teacher has publish in university-publication-house.

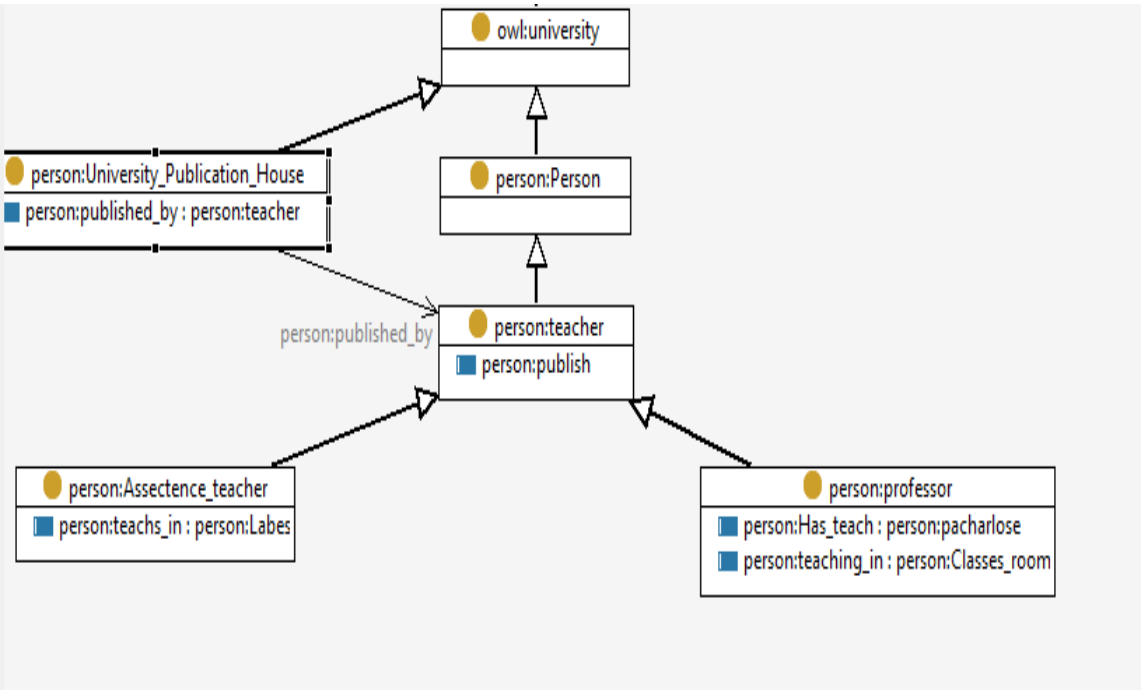

Figure 4.5: Teacher entity

This figure 4.6 shows the student entity and its relationship with other entities. Student has relationship Has-take, registration-in, teaches-in.

Student Has-take courses, registration in registration-office, teaches-in class-room and labs.

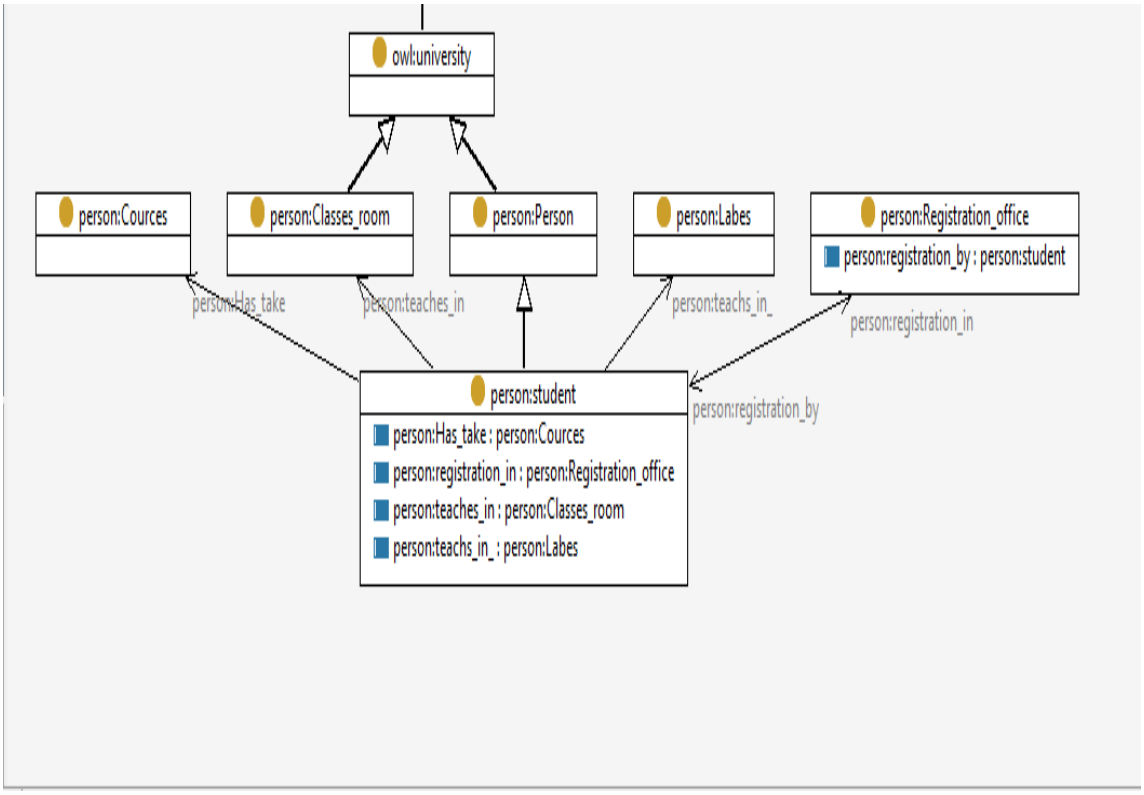

Figure 4.6: student entity

Figure 4.**7** Show All Collage in the Sudan University of Science and Technology:

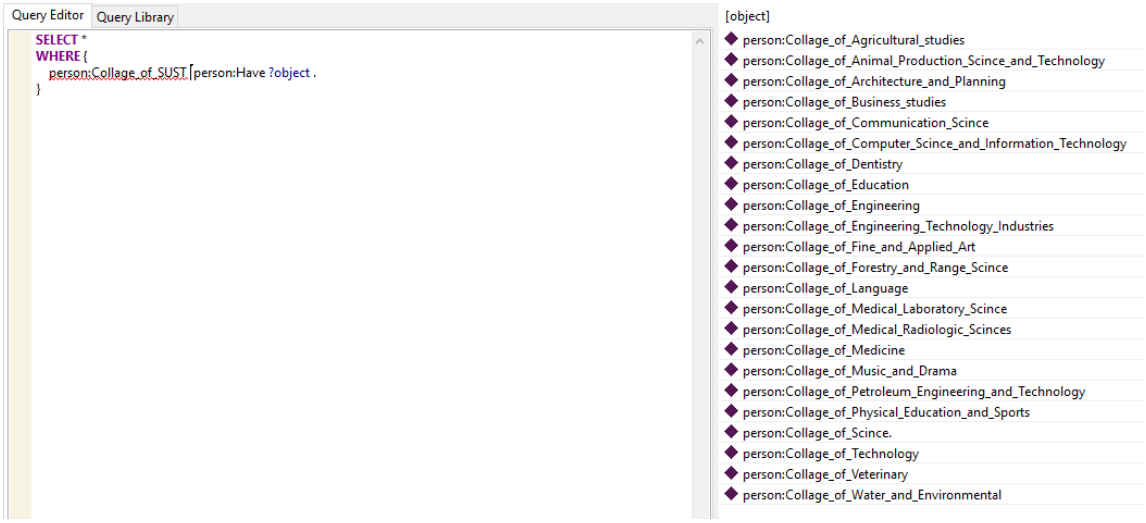

Figure 4.7: Sudan Collages

Figure 4.**8** Show All Department in the Collage of Engineering :

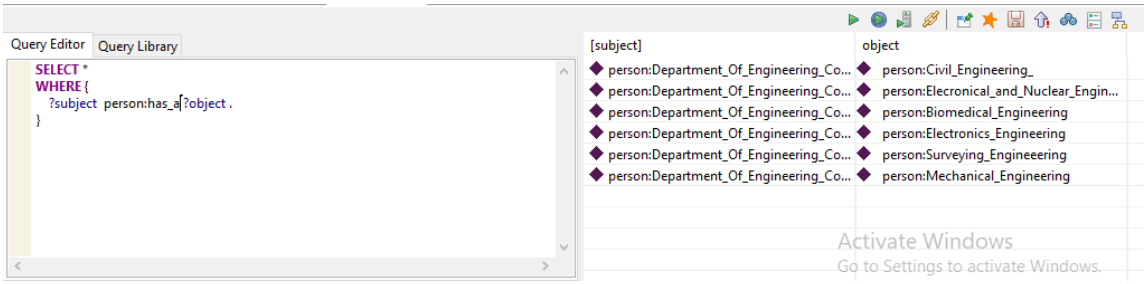

Figure 4.8: Collage of Engineering Department

Figure 4.**9** Show the Administration in the University:

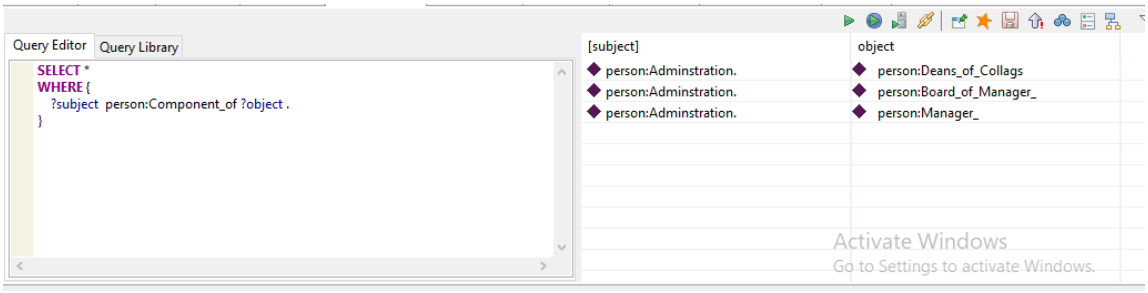

Figure 4.**9:** Administration

Figure 4.10 Show all courses has targeted by all teacher

Relationship: Has-teach.

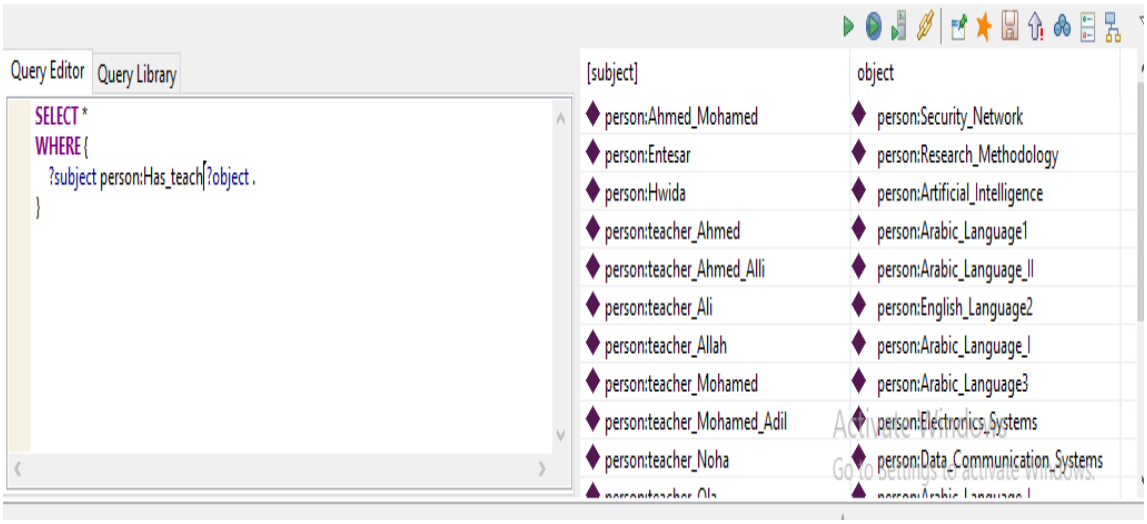

Figure 4.10: courses

Chapter Five
Conclusion and Recommendation

5.1 Recommendations:

- Make website for the university's ontology.
- Using ontology browsing.

5.2 Conclusion:

The need for clear and complete structure is the main reasons to build the University ontology, first describe all the departments under the University structure and the relationships that exist between them from university website and from the low book which describe all concepts and super and sub classes that form the university.

After that the ontology was expressed in OWL starting from creating classes and subclasses to properties, restrictions and instances. Then the owl file of the ontology is imported into TopBraid Composer for more powerful data retrieval software, to get the data needed from the ontology easily with short SPARQL queries.

5.3 References:

1) Gruber T.R, A translation approach to portable ontology specifications, Knowledge Acquisition, 5 (2), pp. 199-220, (1993).
2) Noy N. F. and McGuinness D. L. *Ontology development 101: A guide to creating your first ontology;* Stanford Knowledge System Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, (2001).
3) Powers, David (1991). Preface: Goals, Issues and Directions in Machine Learning of Natural Language and Ontology. AAAI Spring Symposium on Machine Learning of Natural Language and Ontology.
4) Powers, David (1991). Preface: Goals, Issues and Directions in Machine Learning of Natural Language and Ontology. AAAI Spring Symposium on Machine Learning of Natural Language and Ontology.
5) G Budin (2005), *"Ontology-driven translation management"*, in Helle V. Dam (ed.), Knowledge Systems and Translation, Jan Engberg, Heidrun Gerzymisch-Arbogast, Walter de Gruyter.