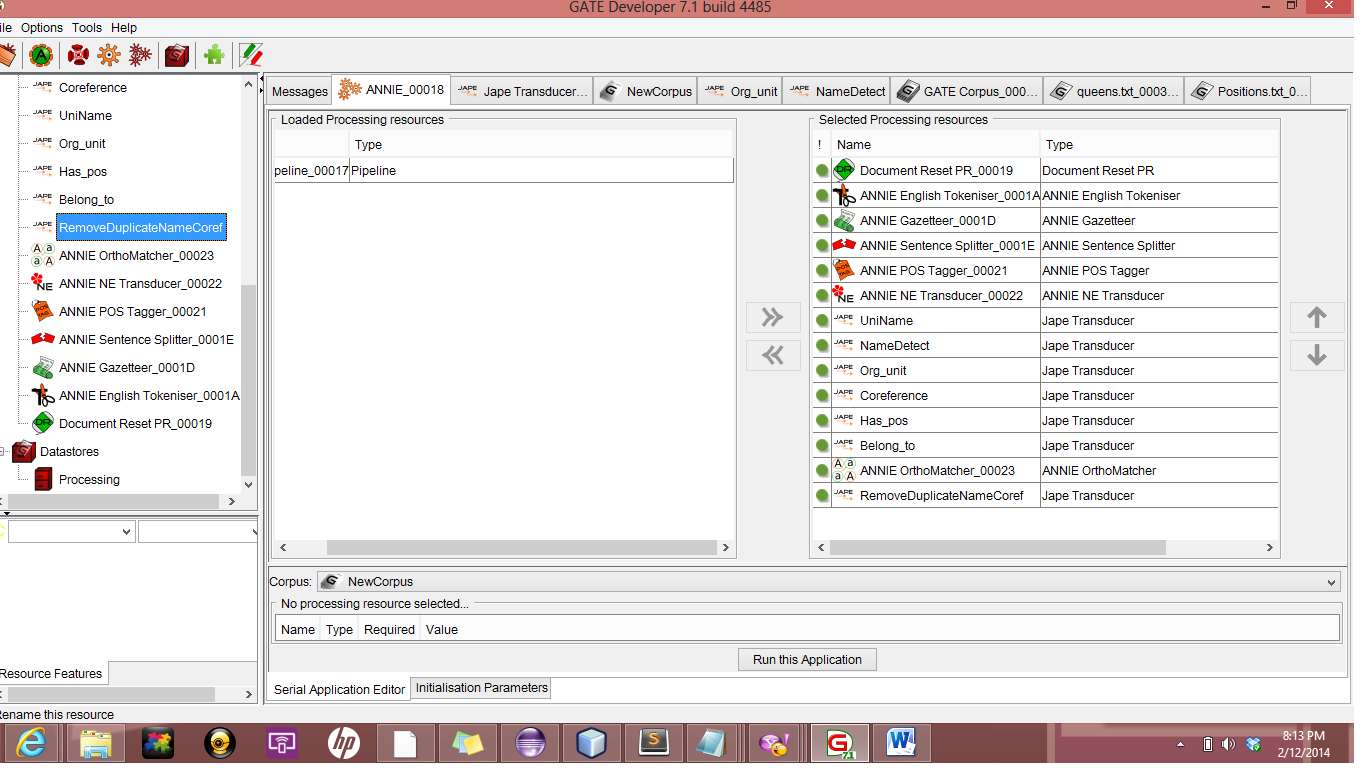
1. **Pipeline:**



* **UniName -**

This JAPE rule has been used for Detecting the University name. This JAPE rule also take a help from Gazetteer list to detect the universities name. For, detecting university name , it takes token as an input.

* **Org\_Unit-**

This JAPE rule used for finding the organizational units from the given documents. Mostly it finds the organizational units start from Department/Division/School of .

* **NameDetect –**

This Jape rule is helping to find the named entities. This jape rule take the help of Person look up to enhance the capability to detect the names.

* **Coreference –**

I made customized Coreference processing resource. It helps to find the relations between the detected entities.

* **Has\_Position –**

This JAPE rule is used for finding the relation between name entity and it’s position in the university. It take the Coreference PR and NameDetect PR as an input to extract the relations between Person’s position and Person.

* **Belong\_to –**

This JAPE rule used for find the the relationship between Person and University. Basically, it gives the output of person is related to which university. It takes Coreference PR and NameDetect PR as input to extract the relation.

1. **Gazetteer List**

I have used two customized and one default Gazetteer list. First Gazetteer list is University name. It help to find the universities name in the text. Second Gazetteer list is used for the find the belonging position of particular person.

First Gazetteer list has size of 500.

Second Gazetteer list has size of 66.

1. **Name Entity Recognization**

* **Person Name Recognization**

In this jape rule, first try to find the name start with title e.g. Mr., Dr., Prof., because It has more possibility to be a name of person. In addition to, in university domain mostly all the name start from some title. But, I also consider the exceptions. To tackle the exceptions, I use the person lookup. So, when I combine the two methods to recognize the entity. It was giving the better output.

* **University Name Recognizer**

In this JAPE rule, first I try to find the university’s name start from University of or Noun and University. But, problem with this method is that it is not considering all the universities names. Therefore, I am using gazetteer list of universities. So, by combing both the techniques, it correctly detects most of the Universities names.

* **Organization Unit Detection**

In this JAPE rule, I try to find the string which starts from Department of/school of and then words. It basically correctly detects most of the departments and divisions and faculties.

1. **Other Components.**

* **Has Position Relationship**

First I run the coreference Jape made by me which detects all the coreference. After that, when I found out the position through gazetteer list, first I will find for near by coreference and then get named entity which it pointed to. Otherwise, I used the nearest named entity as a relation.

* **Belong to** **Relationship**

This jape rule is quite similar to has position jape rule. It also take the input from the coreference which helps to find the named entity. When I detect the university name, then I search for nearby coreference or Named entity to find the belonging relationship.

* **Remove Duplicate Names**

This JAPE rule is used after Orthomatcher. Basically, orthomatcher find all the related entities and put id’s as a feature in person annotation. So, I took all the id’s an remove all the duplicated ones and export only non- duplicated entities.

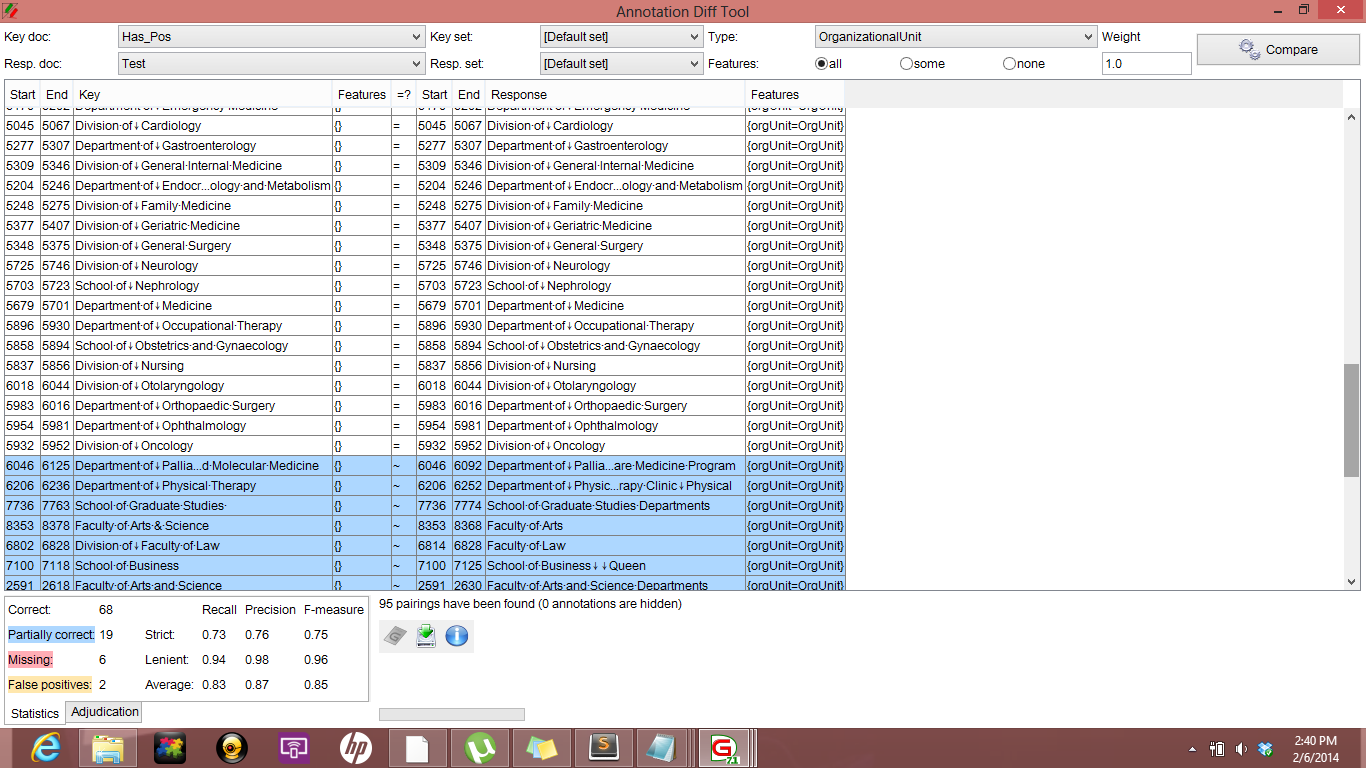
* **Export Result**

I am exporting the result in five different files. All files are different type of detected entities. When the rule has been fired at that time, on the RHS part of JAPE grammar, I am writing the detected entity on the HTML file. Therefore, five files for all five different types of entities.

1. **Evaluation**

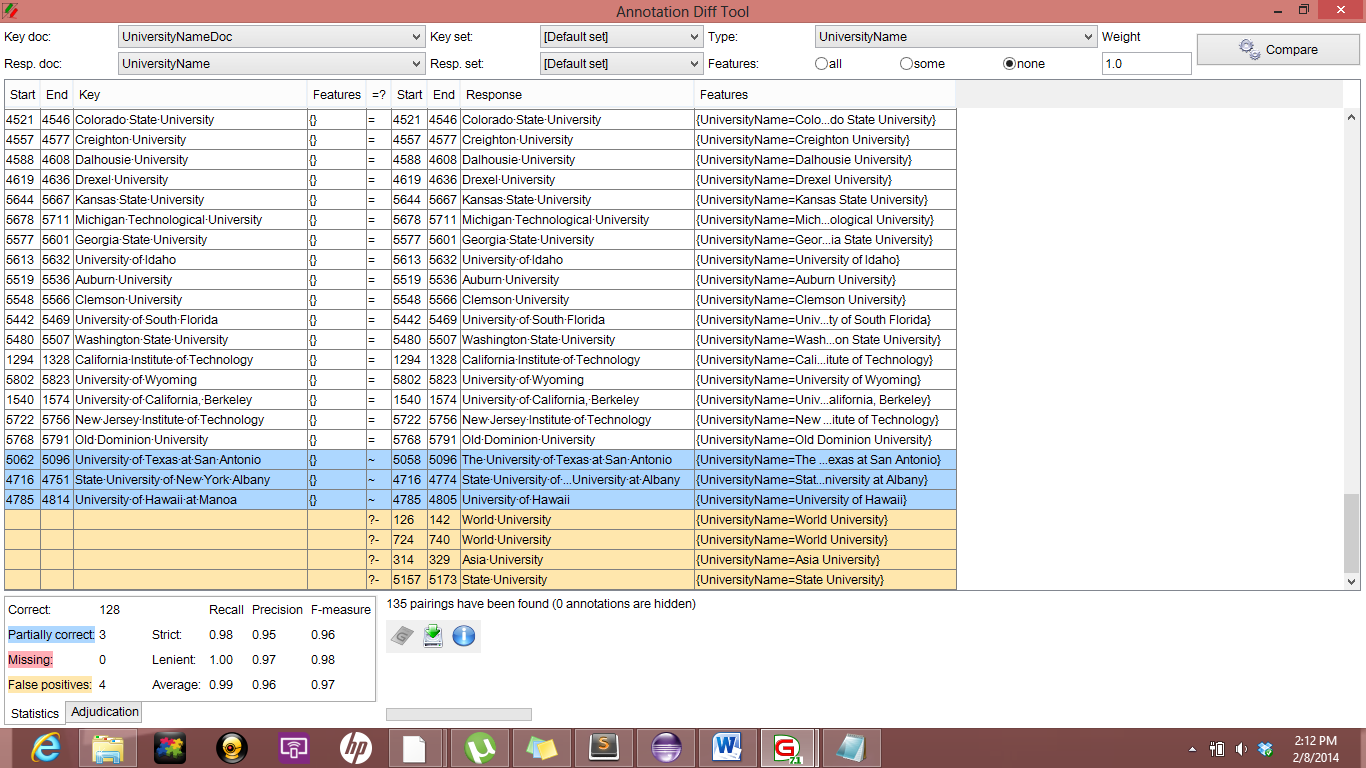
* **OrganizationUnit Detection**:-

I used the document containing 1138 words for Organization unit detection. I compare the result of generated results and manually annotated ones.

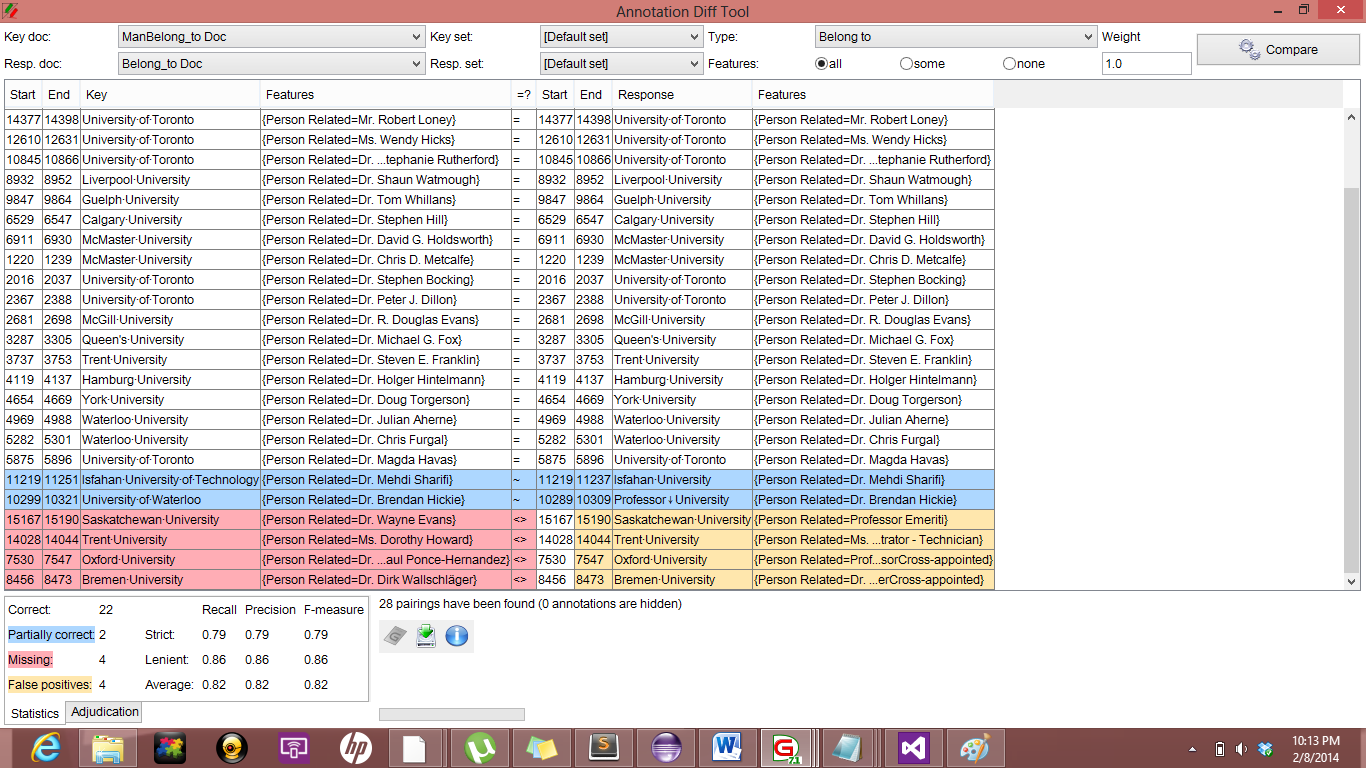


* **University Name Detection**

For University name detection , I have used the document containing 886 words. After I compare the JAPE rule generated annotation with manual annotations.



* **Belong to Relationship**

For finding the belong to relationship, I have used 2302 words document for the experiment

* **Has\_ Position Relationship**

For finding has position relationship, I have used the document of 2322 words. Then, I compare the JAPE rule generated result with manually annotated document.

