



**11756 Storage and Data Recovery**  
**Master's Degree in Intelligent Systems**  
**Final Project – Option A**

The Search Engine Marketing Campaigns are oriented to improve the web positioning in the Search Engine Results Page when the Search Engine users looks for some searching terms.

The Search Engine Positioning can be improved by SEO Strategies (Search Engine Optimization) that consist in improving some Positioning Factors, named SEO Factors or SEOF.

There are several SEOF, but one of the most important is the content of the HTML Title Tag. That is the content between the <title ...> and the </title> tags. This content should include the keywords of the organization.

In our project we want to analyze the title tag content keywords used by a set of organizations that operates in an specific economic sector.

The Final Project consists in:

- Choose an specific economic sector (insurances, hotels, rent a car, newspapers), in English or Spanish. After that you should search about 10-15 website URL related with the chosen sector. This set of URL should be inserted in a database table.
- Using the previous table, using web scraping techniques, retrieve the title tag content of every URL and insert it in the database table.
- Process every title considering the next functionalities:
  - Split the title in individual words, called keywords.
  - Delete superfluous keywords.
  - Clean data, changing similar keywords.
- Show a bar chart report (or similar) with the number of appearances of every keyword.
- Export the previous report to the next formats: XML, JSON and Excel.

The documentation to be uploaded consists in:

- Database structure
- Programs code
- A 5-10 pages document (or slides) with an explanation of the work, including screen copies of all developed functionality.
- A video of 5-10 minutes with a demo.