

# Information Retrieval - Report - Project: Art For Sale

**Morales Mariciano Jeferson, Piloni Filippo**

Universita' della Svizzera italiana, Faculty of Informatics, Lugano, Switzerland

## 1 Obtained results

Github code repository: <https://github.com/JekxDevil/IR-ArtForSale>

For the *Art-For-Sale* project we selected the following websites to collect the data from:

1. [www.artsy.net](http://www.artsy.net)
2. [www.saatchiart.com](http://www.saatchiart.com)
3. [www.artfinder.com](http://www.artfinder.com)

For each website, a spider to crawl the site was created. Website specific pre-fixed tags have been used to choose categories to explore, searching content for each of the tag and retrieve categorized grouped results.

The crawler explores the pages of the single artworks and gathers from each one the following information:

- Name of the artwork and author
- Url of the image of the artwork
- Price
- Link to the page
- Tags associated to the artwork
- Description of the artwork (if present)

Results are saved in json files and we are starting the creation of an inverted index combining all the results obtained.

In total, we obtained around  $\approx 9000$  documents to be used in our project

## 2 Planned results

The next step will be implementing the **inverted index** and the **frontend** and **backend** part.

For the inverse index, documents are indexed based on all the text information, providing a quick access based on tags, price and characteristics listed above.

We then decided to implement a simple yet highly customizable interface in Vuejs that will allow the user to filter the results based on the tags associated to the artworks and showcase all the important information about the artworks obtained thanks to the retrieval model.

Our plan is to implement at least the following features:

- Result presentation
- Automatic recommendation

With these, it would be possible to accomplish our goal, so to create a retrieve system with a good interface that returns results based on the user's search history.