IMIA

IMIA is an application, created as a project for Information Retrieval course at TU Delft (the Netherlands) by the following students:

- Anelia Dimitrova
- Ioana Leontiuc
- Mengmeng Ye
- Ilse Verdiesen

Task: To create an information retrieval system where user votes for terms regarding a specific company. User first answers a simple math question as a captcha service, then gets to a match application where they vote for one out of 4 terms showing on the screen for a random company.

Method: The way the system is realized is as follows:

1. A process model for the components of the application is created (Ilse Verdiesen)
2. Use of Twitter API to collect tweets for specific companies: Collected are two types of tweets: tweets that companies post themselves and tweets of users, mentioning the companies. Java is used for this part of the application (Mengmeng Ye)
3. We then extract the hash tags and use them as “description” for the company. We create a live listener for Twitter where each new tweet is fetched and tags are added to the database. We use Java and Python for this task (Ioana Leontiuc).
4. The collected terms are added in a matrix and have an equal initial rank, i.e. all of them are equally likely to be true.
5. A web application is developed with PHP, MySQL, HTML and CSS (Anelia Dimitrova) where the user answers a math question and then matches the term that corresponds most accurately to their opinion about a specific company.
6. A vote for one term, increases the rank of that term and decreases the rank of the other terms as per the ELO algorithm.

UPDATE: The project was further expanded by Anelia Dimitrova and Ioana Leontiuc to meet criteria of real user registration and login and to get actual data from users. Real users were asked to fill in information about 15 companies: 5 fast food, 5 clothing and 5 car companies, users were also asked to provide personal information such as age, gender, nationality, education, job status. The original database information was narrowed down to terms, seen at least 3 times which gave much better results and then after the crowdsourcing information was gathered, different clustering algorithms were applied based on the data and on the images used so that statistics and results could be provided to companies about their user’s perception of the company and about the status of users, using their products.

DEMO at http://worldofinspiration.net/tests/mmsr/

Notes:

1. Login and Register forms are dummy, so go to **Log in > NEXT** and you will be at the math problem page as if you are logged in.
2. If you answer the math problem correct, but the system says it is wrong, please refresh. (part of the first version of the project only: IRDB)
3. At the match application: make sure you submit the answer in order to refresh the terms and to be able to vote for another company. If you do not refresh, the terms will stay the same.

P.S. There are still things like these to be improved, but they probably won't be improved for this specific project as it is a prototype, not a real application.