



Faculty of Engineering and Architecture
Computer Science Engineering

Advanced Multimedia Applications

Apps4Ghent

Progress Report

Jeroen ARENS

Enver BRAL

Harald DE BONDT

Damian TARNACKI

Academic year 2014–2015

Contents

1	Introduction	2
2	General project decisions	3
2.1	Role decisions	3
2.2	Infrastructure decisions	3
3	Application	4
4	Overview of meetings	5
4.1	Meeting with Team 13/02	5
4.2	Meeting with Client 25/02	5
4.3	Meeting with Team 25/02	6
4.4	Meeting with Prof. Mannens 26/02	6
4.5	Meeting with Team 26/02	7
4.6	Meeting with journalists 03/03	7
4.7	Short meeting with people from the library 03/03	8
4.8	Short meeting about linked data 09/03	8
5	Overview of events	10
5.1	Apps for Ghent data-dive 09/03	10

Introduction

The library of Ghent has a massive considerable amount of data regarding its collection of books and the people it lends books to. The library is considering to provide most of this as open data that is accessible by everyone. The purpose of this project is to think of and develop a creative tool/app/website that uses this open data, which will be used in the library of Ghent itself and will also be used to promote the university of Ghent as a multimedia university. The project is executed in regard of the Apps4Ghent edition of 2015, which tries to stimulate the use of open data and the development of application that use the data that has been opened up.

To accomplish this project, a multi-disciplinary team will be assembled consisting of students of engineering, communication science, data analysts and economists. Together a well-thought and good looking/functioning concept will be rolled out.

General project decisions

2.1 Role decisions

- Jeroen will execute the role of Project Leader. He will take care of the communication between the client and the different groups of students. (13/02)
- Jeroen will be the scheduler for this project. (25/02)
- Harald will be the continuous integration manager. (25/02)
- Damian will take care of the tasks of the system administrator. (25/02)

2.2 Infrastructure decisions

- A private git repository is set up via Github on which the project development will take place.
- To keep the communication clear and well structured, Slack is used to handle the communication between the computer science students. Communication with other sources will be done by email.
- Slack will also be used for stand-up meetings with the group. If a call is needed, Google hangouts (integrated into Slack) will be used.

Application

Overview of meetings

4.1 Meeting with Team 13/02

present: Everyone of the team

purpose: Assign team leader who will contact the client to schedule a meeting.

result: Jeroen is assigned as team leader, he will make an appointment with our client during the 3rd week of the semester.

4.2 Meeting with Client 25/02

present: Everyone of the team, Professor Mannens, Koen from Digipolis, Professor Van den Poel, Pieter Colpaert from REC, Pieter Blomme, 2 spokes people from the library and Leenke (journalist student)

purpose: First introduction to every party and getting to know the scope of the project.

result:

- Each party discussed their involvement in the project and what they would like to gain from it. This information will be used to define the actual output of the project. We proposed a website that visualizes borrowed books on a map of Ghent. All parties agreed this would (in some way) fit in with their individual goals.
- The library of Ghent has two goals in mind. The first one is to make the distribution of books between the different libraries at Ghent more efficient. The idea is to find correlations based on geographical data to predict what kind of books need to be available at the different library sites.
The second suggestion is to define new (hidden) genres. These can be used to recommend books towards the readers themselves.
- The journalists want to find a story which can be interesting for the people of Ghent. This needs to be presented in a creative manner.
- The master students of marketing want to use the data to do analysis upon. Based on this analysis, they can develop a predictive model in regard with the books and users at the library.
- REC wants to have an inspiring project and wants to profile the UGent as a multimedia university.

- A lot of suggestions were given at the meeting and have been documented in a separate document so that they may be used to figure out what can be done within the limited time that is provided to us, as well as to give us inspiration for future suggestions.

4.3 Meeting with Team 25/02

present: Everyone of the team

purpose: Define an initial scope to work upon during the semester

- For the front-end of the website, first a user interface will be developed on which queries can be executed by the journalists as aid for them to find a story behind the data. Afterwards the visualization of this data will be made more creative. E.g. an interactive map with the data displayed, several graphs, ...
- To store the data at the back-end a NoSQL database tool will be used. The tool we will use is MongoDB. With MongoDB we can store the data and already make a basis semantical network. For the data-analysis, a basic model can be employed which will just look at obvious correlations. Later on the students of marketing can analyse the data to enhance this model.

4.4 Meeting with Prof. Mannens 26/02

present: Prof Mannens, assistant, everyone of the team

purpose: Discuss previous meetings and decide on what needs to be done next, get some feedback about the current progress and our current approach

- We should focus our attention on the two most important clients, being the library of Ghent and the city of Ghent. We should then try to find a way to incorporate the journalists students into the project.
- Just forget about the marketing students, since we can't satisfy their needs.
- We will focus on using **open data** instead of the raw data.
- The most promising project we will focus on is a *visualisation of the distribution of the books across Ghent over time*.

- If possible it might be nice to incorporate the concept of *Goodreads*, which we could suggest on the *data dive*.

4.5 Meeting with Team 26/02

present: Everyone of the team

purpose: Discuss feedback, decide on architecture

- We worked out the basic of the architecture and identified the biggest separate blocks as *Fronted*, *Backend*, *Database* and *Syncing*.
- We decided to use Django on the backend.
- On the frontend (obviously) HTML5/CSS/Javascript will be used. A visualisation framework for drawing maps and charts still need to be decided on before next meeting.
- The syncing part will be written in Python as to minimise the number of different languages used and can probably be integrated into Django.
- A REST API will need to be decided to provide communication between frontend and backend.
- As soon as the data for the library has been put online, we can start defining the Rest API between the *Syncing* block and the *TheDataTank* endpoint.

4.6 Meeting with journalists 03/03

present: Pieter Blomme, Pieter Colpaert, both journalist students, everyone of the team

purpose: Discuss what we decided on and what the journalists want from us

- We explained what was discussed during our meeting with prof. Mannens and told them the project we decided on and its scope.
- The journalists need to find a story using our visualisation. They will their ideas to Pieter Blomme, who will then help to select the final story by the main AppsforGhent event on 21/03.

- Leenke would like to be involved in every aspect of the project. As such, she would like to be notified of any meetings concerning the technical parts of the project and, when possible, attend them.
- We asked Pieter Colpaert for a server on which we could host the application, which he would provide as soon as possible.
- Damian will setup a sample Django project with MongoDB in a production environment, as soon as we have access to the server.
- Harald will focus on the frontend. Highcharts was suggested for drawing charts and maps and he will look into the usability of the framework or come up with a better alternative.
- Enver will work on the syncing between data provided by *TheDataTank* and *MongoDB*. He will also work on the REST API provided to the frontend by the backend.
- Jeroen will focus on the backend and will help out with all problems that are Django related (including setting up the production and development environments).

4.7 Short meeting with people from the library 03/03

present: Pieter Blomme, Leenke Dedonder, everyone of the team, person from the library

purpose: Give an update on what decided and inform the library on the next steps

- We explained the scope of the project and it was decided that it fits in to what they would like to gain from this project.

4.8 Short meeting about linked data 09/03

present: Pieter Colpaert, Anastasia Dimou, Leenke Dedonder, everyone of the team

purpose: Inform us about a project Pieter's team will be working on for AppsForGent.

- Pieter informed that a team would be working on a project that involves linked data. As a first step they would need to transform the library data (provided as CSV) to linked data.
- They would like it if we could make our data available as linked data as well, as such we have decided on using the same mapping from CSV to JSON-LD.

- We will also need to provide an API that returns JSON-LD data.

Overview of events

5.1 Apps for Ghent data-dive 09/03

present: Pieter Colpaert, Leenke Dedonder, Everyone of the team except Damian

purpose: Attend presentations about open data that is provided/available concerning the libraries of Ghent and their collections

results:

- It became clear that the data that is provided to us lacks a lot of metadata about the books (e.g: authors, genre, number of pages, ...), but does provide general identifiers that could be used to hook into other services (e.g. ISBN).
- The Bibnet catalogus API could be used to get the additional metadata that is lacking from the library data set.
- A lot of data on the collection of books in the UGent are also provided as open data and appears to be uniquely identified by URLs. This might be useful in the context of linked data.