

Project Proposal - Project and Seminary 2019/2020

Title

I-On Integration

Background

In order to carry out their activities, education institutions need to publish information regarding term calendars, timetables, evaluation schedule and others. Traditionally this information was posted on boards on campus at the beginning of each academic year or in the beginning of the semester. After these institutions established their online presence, it became available on school websites.

To consult their class schedules, students and teachers in some institutions are required to download the full class timetable for the entire course, despite only having interest in a limited subset of the information. In addition to this, it is not easily accessible and it is often times presented in a heterogeneous way.

I-On is an academic information aggregation and distribution system.

The main goal of the project is to facilitate the access to academic information according to the user's preferences. The user will then be able to consult their own timetable and evaluation schedule, ignoring excess information.

The I-On Core sub-project acts as the repository providing a Web API for access. I-On Integration is responsible for collecting all the relevant data from external sources and uploading it to the I-On Core sub-project. The data will then be made available to end users via web and mobile applications.

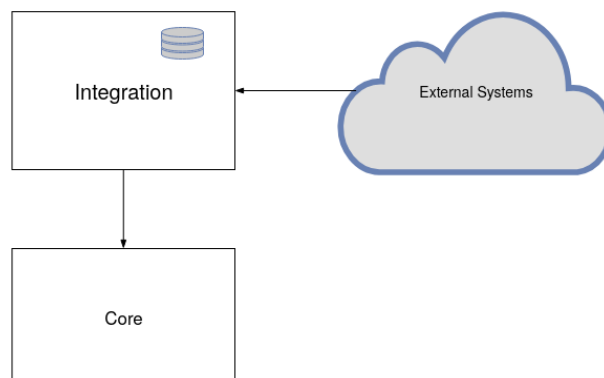


Figura 1: System diagram

Objectives

Development Infrastructure

- Configure *Continuous Integration pipeline* providing an automated way of building, packaging and testing, while guaranteeing confidence in implemented code;
- Configure *Continuous Deployment pipeline*, making the latest version of the project available;
- Configure *log* analysis platform;

Integration Sub-System

- Provide configurable *batch jobs* to obtain information from external sources and upload it to I-On Core sub-project via *HTTP API*;
- Enable *batch jobs* to run periodically or to be triggered by events;
- Create solid and well-documented project following industry best practices, enabling future improvements;

Additional Objectives

Depending on time availability, the functionality of collecting and uploading administrative information (e.g. term calendar) to the I-On Core will be available through the submission of a job that receives information directly through a file or presents a form. Ideally this functionality should support a more general use case, i.e. dynamically generated forms.

Justification

The project will enable us to explore programming techniques learned throughout the course and apply them to a domain that is close to our recent experience as students.

Furthermore, it will allow us to deepen our understanding of batch processing systems.

The use of *Continuous Integration*, *Continuous Deployment* will give us the opportunity to reproduce a real-world development job.

Scope

We will use ISEL as a test case given our familiarity with the information and close association. At a later stage, the system will also become available to other academic institutions.

Approach and Resources

In order to create a well-developed project we will use *Git* as the version control system and apply *Git Flow* as our development style.

The technology stack will be:

- *Kotlin* - Main development language
- *Spring Batch* - Batch processing framework
- *iText*, *Tabula* - PDF manipulation libraries
- Automation server - e.g. *Jenkins*
- Source analysis, visualization tool - e.g. *Grafana*
- *Web Hosting* - e.g. *AWS*, *Azure*, *Heroku*

Additionally for storage we can use *PostgreSQL* for *RDBMS* and *Elastic search* for search engine.

Major milestones

Our initial plan contemplates the following milestones:

- Requirements catalog Mar 18
- Technical design Mar 31
- Infrastructure set up Apr 21
- Initial report May 3
- Beta version May 27
- Poster Jun 1
- Integration with I-On Core Jun 8
- Final version Jul 25

Constraints and assumptions

Due to our lack of experience with some of the technologies to be used, we have to account for time to increase proficiency.

We depend on the I-On core sub-project to consume the data we produce, namely providing an *HTTP API endpoint*.

As our use case is ISEL, we assume that the data we want to obtain will continue to be available and not change its schema/format.

Finally there are hard deadlines that must be met so that the project can succeed.

Some software resources are cost-free for students. For hosting solutions, the school will cover the licensing costs.

Risks

The main risk we identify originates from the fact that we don't control the information we depend on. The project should be able to identify these issues and alert the responsible end-user. Our limited experience with the *Kotlin* programming language, *batch processing* frameworks and *Continuous Integration/Continuous Deployment pipelines* is also a risk worth noting. However there are many quality sources of information we can refer to and also we are counting of the expertise of our supervisors.

Project Organization

Participants

Student No.	Name	E-mail	Telephone
43314	Miguel Teixeira	miguelbarbosateixeira@gmail.com	960013992
43552	Samuel Costa	sampaicosta.samuel@gmail.com	936696494

Supervisors

Name	E-mail
Pedro Félix	pedrofelix@cc.isel.pt
João Trindade	jtrindade@cc.isel.pt

Clients

I-On Core and Supervisors.

Project Plan

teamgantt
Created with Free Edition

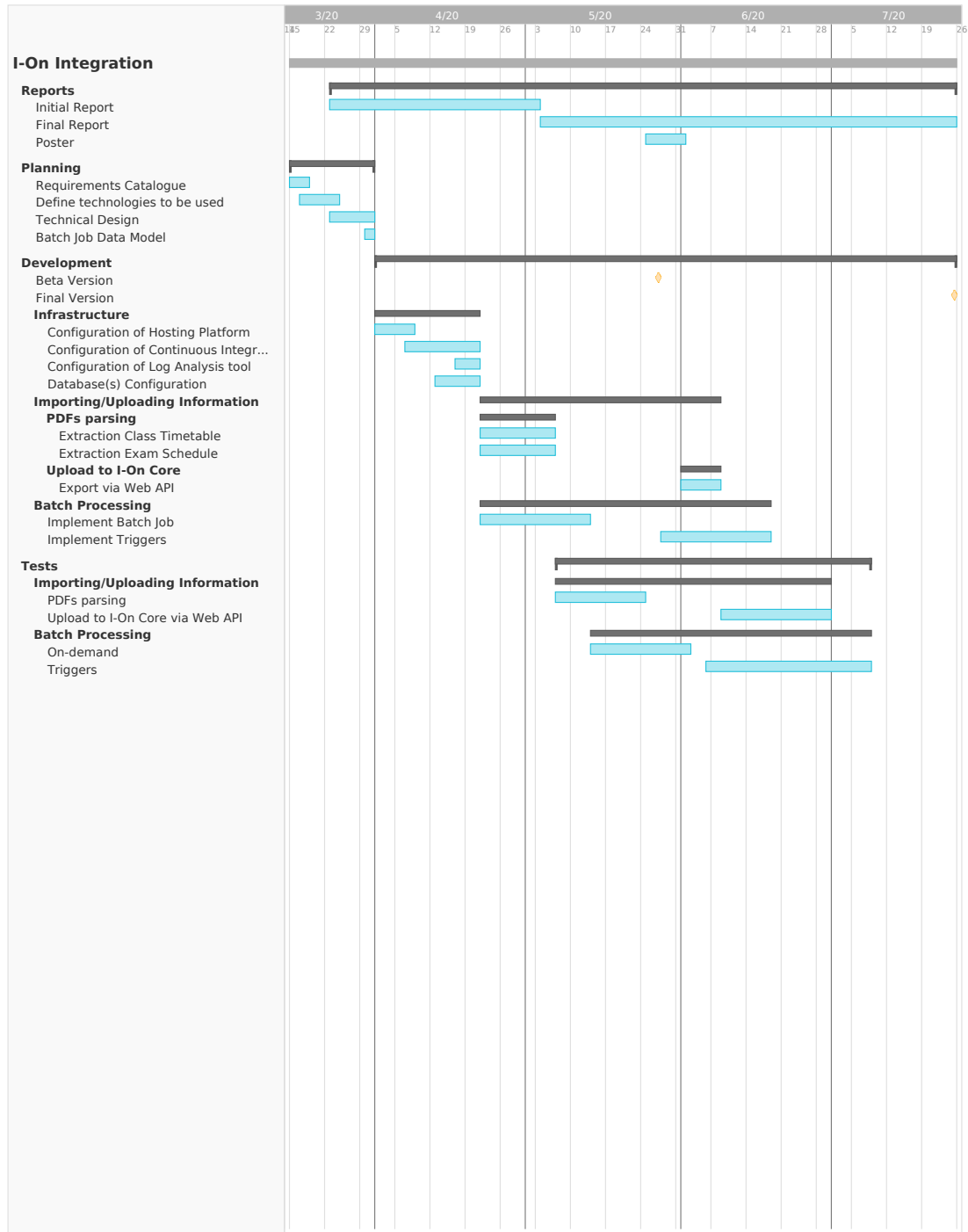


Figura 2: Gantt Diagram