ELECTRONIC PUBLISHING
AND DIGITAL STORYTELLING
Lesson 1

COURSE INTRODUCTION

TABLE OF CONTENTS

01

INTRODUCTION

course overview

02

BACKGROUND

what you need to learn by your own, what you learn in this course

03

PROJECT

Topic, groups, documentation

04

EVALUATION

Assignments and final exam

ELECTRONIC PUBLISHING
AND DIGITAL STORYTELLING
Lesson 1

INTRODUCTION

IN NUTSHELL

Overview of methods to manage and query data, e.g. statistical analysis

DATA ANALYSIS

Features and techniques to visualise data for multiple purposes

DATA VISUALISATION

How to present data to support exploration, discovery, and decision making

DIGITAL COMMUNICATION

WHAT YOU'LL LEARN

DATA ANALYSIS

Manipulate data with python.
Access and query RDF data on the web.

DATA VISUALISATION

Choose appropriate visualisations according to your data.

DIGITAL COMMUNICATION

Present your results in an engaging way on the web.

BACKGROUND

WHAT I WISH YOU ALREADY KNEW...

Computational Thinking and Programming (1) (LM)
Usability and User Experience (1) (LM)

Basic-intermediate skills in python programming: install python libraries, usage of notebooks (Jupyter)

PYTHON

Basic knowledge of Javascript for IU (modify the DOM, interactivity).

JAVASCRIPT

WHAT I WISH YOU ALREADY KNEW...

Computational Thinking and Programming (1) (LM)

Create, access and manipulate CSV files with Python and Javascript

CSV

Create, access and manipulate JSON files with Python and Javascript

JSON

WHAT I WISH YOU ALREADY KNEW...

Computational Thinking and Programming (1) (LM)
Introductory seminars
Usability and User Experience (1) (LM)

Good skills in web development (static web pages).

HTML / CSS

Good skills in usage of version-control interfaces (e.g. sourcetree)

GITHUB

WHAT I WISH YOU ALREADY KNEW...

Library, Archive, Museum and Information Science (I.C.) (LM)
Information Technologies and Knowledge Management (I.C.) (LM)

Basics of RDF, syntaxes (.ttl, RDF/XML), SPARQL, and OWL.

LINKED OPEN DATA

How to organise knowledge, how to query/explore data according to an ontology.

ONTOLOGIES

WHAT I WILL INTRODUCE YOU

Py Libraries for data exploration.
Jupyter to document your work.

DATA-RELATED

Js Libraries for data visualisation.
Github pages to publish a website.

WEB-RELATED

WHAT I WILL INTRODUCE YOU

Py Libraries for creating RDF data and querying SPARQL endpoints. Short tutorial on SPARQL.

SEMANTIC WEB-RELATED

How to frame your work (research questions, exploration, discovery). **Digital storytelling** strategies.

COMMUNICATION-RELATED

WHAT YOU WILL HAVE TO LEARN BY YOUR OWN

GITHUB

[A short introduction to git and github]
[Github guides]
[sourcetree GUI to Git]

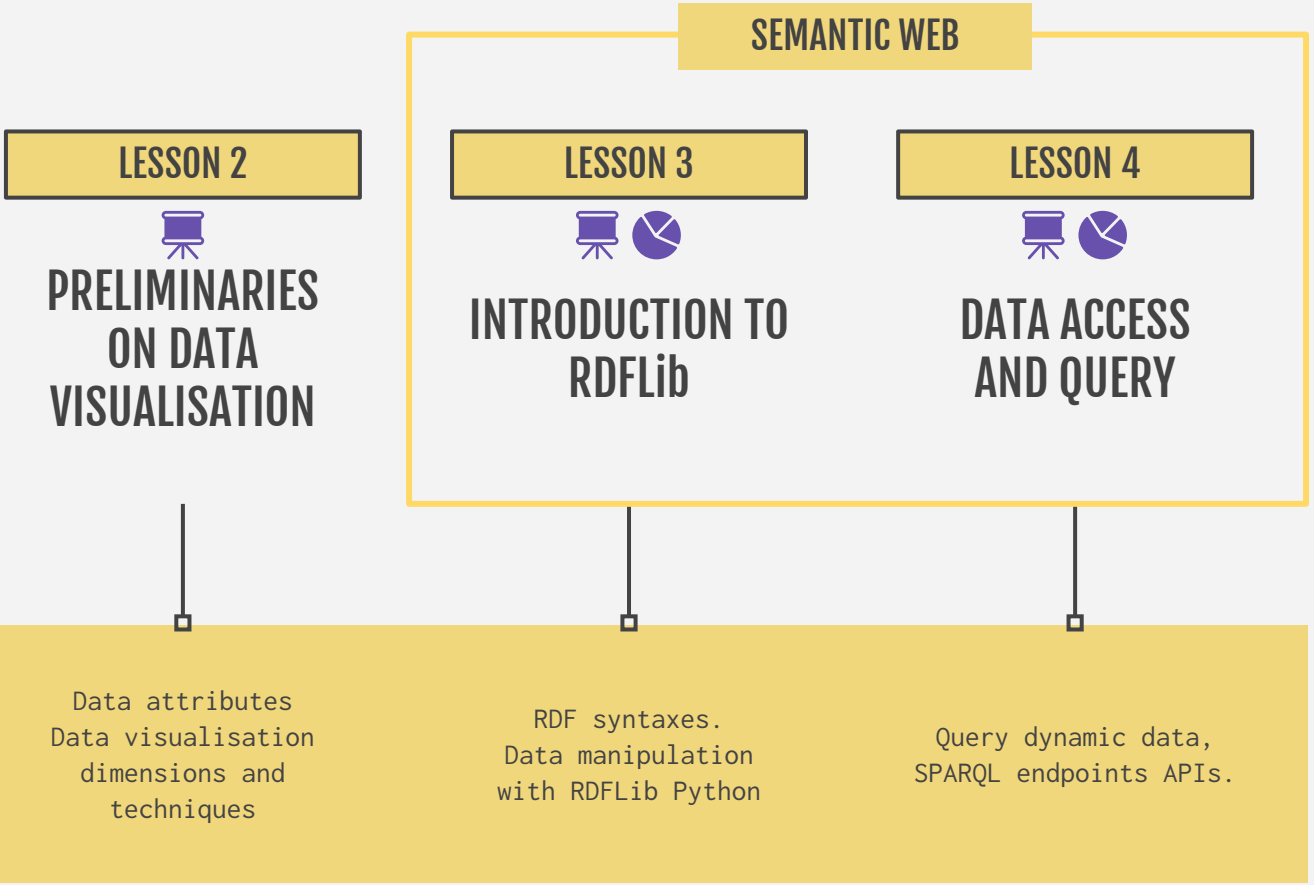
Read online documentation on how version-control systems work.

Learn how to use github (via shell or interfaces, e.g. sourcetree)

PYTHON BASICS

How to install python and libraries.
How to manipulate data structures (lists and dictionaries).
How to read / write files.

CLASSES OVERVIEW



lecture

hands-on

CLASSES OVERVIEW

LESSON 5

DATA SENSE MAKING

Formulate questions and answers with data visualisation.
Introduction to Jupyter.

LESSON 6

DATA ANALYSIS WITH PYTHON

Data wrangling and analysis

LESSON 7

DATA VISUALISATION WITH JS

Web development.
JS libraries for data visualisation.

lecture

hands-on

CLASSES OVERVIEW

LESSON 8

DIGITAL STORYTELLING

Communication strategies and best practices

LESSON 9

SEMINAR

Designing for data-driven experiences and research. Calibro
Lecture and tutorial

LESSON 10

WEB PUBLICATION AND WRAP-UP

Binder
Github pages

lecture

hands-on

ELECTRONIC PUBLISHING
AND DIGITAL STORYTELLING
Lesson 1

PROJECT

A DATA VISUALISATION PROJECT

Given a topic, find relevant research questions that can be answered by visualisations (or that can support domain experts to understand certain phenomena)

WHAT YOU NEED

THE TOPIC

THE DATA

A GROUP?

THE TOPIC

Art historians and history of art history

Art historians' relations
E.g. Which countries are historians from and where did they work?

The debate on research topics (artists, periods, movements)
E.g. In which periods the research focussed on Mannerism? (i.e. the archival collections)

Resources
E.g. which types of resources in archival collections represent research topics?

For whom?

In this course we address a topic that is relevant to scholars, cataloguers, market dealers, museum curators, etc.

OUR FOCUS
Research questions to be answered by data visualisation must be tailored to what is relevant to scholars, i.e. **art historians.**

How do I know it’s the correct topic?

Once you defined your research questions, drop an email to marilena.daquino2@unibo.it for feedback.

SPOILER ALERT This will ensure your work is not going to be trashed and can be useful to somebody in the near future (see at the end of this presentation for clarifications)

ARTchives Linked Open Data

Use primarily data from **ARTchives (mandatory)**
http://artchives.fondazionezeri.unibo.it/

Data in ARTchives include newly generated data created by cataloguers of 6 institutions and reuse data harvested from Wikidata.

You can either use only ARTchives data or (recommended) integrate ARTchives data with other relevant sources, both Linked Data or not (e.g. Wikidata, DBpedia, artistorians.info)

GROUPS vs LONELY WOLVES

3 people max.
You need to justify your contribution to the project. Grades are individual.

GROUPS

You can work alone, but you need to achieve the same results.
No discounts :)

WOLVES

ELECTRONIC PUBLISHING
AND DIGITAL STORYTELLING
Lesson 1

EVALUATION

WHAT HAPPENS THE DAY OF THE EXAM

it's not an exam on python or javascript!

Likely to be remote. Everybody can attend by requesting the link to Teams meeting.

PRESENT YOUR PROJECT

While you talk I'll have a look at your source code.

DIG INTO THE CODE

General questions on how you handled the problem and on the **soundness** of your project.

ANSWER QUESTIONS

WHAT YOU NEED TO HAVE BY THE DAY BEFORE THE EXAM

A Jupyter notebook **published on the web** including documented workflow to create and analyse your data

JUPYTER NOTEBOOK

A static web page or a web application **published on the web** showing results of your data analysis

WEB PAGE

A presentation of your work according to the template I give you. With slides or not.

15min PRESENTATION

PREPARE YOUR MATERIALS: THE JUPYTER NOTEBOOK

1. Install **Jupyter** [1] on your laptop (individual)
2. Create your <projectname> notebook (individual)
3. Describe your project aims in a short Abstract section.
4. Prepare (extract, integrate, clean) the data you want to visualise and save them into a **CSV or JSON file**.
5. For every function you define and for each step in the data analysis, add an explanation in the notebook in **Markdown** [2] syntax.
6. [If you are a group] Merge your notebooks into a **single notebook**.

[1] https://jupyter.org/
[2] https://www.markdownguide.org/

PREPARE YOUR MATERIALS: THE JUPYTER NOTEBOOK

1. Create an account on **GitHub** (individual)
2. Create a repository for the project (collective or individual)
3. Load the jupyter notebook and the CSV/JSON data on the repository.
4. Connect your repository to **Binder** [1] to make the notebook interactive on the web.
5. Include the binder link into your README file in the repository.

DONE!

[1] https://mybinder.org/

PREPARE YOUR MATERIALS: THE WEB PAGE

1. Create a **index.html** page (collective or individual)
2. Create a **CSS** stylesheet for the webpage (collective or individual)
3. Create a **Javascript** file (collective or individual)
4. Access the data previously created (CSV/JSON) from the JS file
5. Add style to your data (CSS/JS).
6. Upload the website files on the repository.
7. Use **github pages** to publish your website.
8. Open the repository of the course [1]
9. Open or comment on an **issue** called “Exam DD/MM/YYYY” (or similar) with:
    a. Project title
    b. Website URL
    c. Repository URL (if not github)
    d. People involved

DONE!

[1] https://github.com/marilenadaquino/epds

PREPARE YOUR MATERIALS: LINK YOUR PROJECT

1. Open the repository of the course [1]
2. Open or comment on an **issue** called “Exam DD/MM/YYYY” (or similar) with:
   a. Project title
   b. Website URL
   c. Repository URL (if not github)
   d. People involved

DONE!

[1] https://github.com/marilenadaquino/epds

PREPARE YOUR MATERIALS: RESPONSIBILITY AND LICENSE

IF COLLABORATIVE PROJECT
Please MAKE CLEAR who was responsible for what aspect of the project (e.g. data access and cleaning, data analysis, data visualisation, web development, graphics, communication strategy and so on) during the presentation

BOTH INDIVIDUAL/COLLABORATIVE PROJECTS
Specify your names and tasks:

- in the README file of the repository
- on the webpage (e.g. in the footer, in a dedicated page called *Credits*)

LICENSE FOR THE REUSE OF DATA
In the README file of the repository specify that:

- ARTchives data are available under CC0 license
- Other reused data (look at their websites)
- Your derivative data (please, use either CC0 or CC-BY licenses to allow future reuse)

PREPARE YOUR MATERIALS: DO NOT PANIC

no need to do everything from scratch!

*You can either create your own static web page with your styles or you can **reuse existing projects or tools** (e.g. CMS, HTML templates)*

*You don't need to style from scratch the visualisation, you can (definitely) **reuse existing libraries**.*

no need to use github to publish your website!

Github is a free-of-charge solution to publish a project website and to handle the code in the same environment. Feel free to use other solutions for publishing the website! However, it is mandatory for publishing the notebook and the data.

but especially, no need to panic!

We will set up all the pieces of your project (install Jupyter, use libraries, create an account on github, create and publish a webpage) in dedicated hands-on classes :)

PREPARE YOUR MATERIALS: THE PRESENTATION

If slide presentation, max 10 slides.
No need to share it anyway before the exam day.

Use the following **template** to prepare the speech.

Title

1. Background
2. Goals
3. Research questions
4. Data preparation and data analysis
5. Data visualisations selected and reasons
6. Data communication strategies
7. Summary of results

AN EXAMPLE:
THE PRESENTATION

Title: Trends in the study of artistic periods

Background. Artistic periods are differently studied by art historians over time. An artistic period may be overlooked at a certain time, possibly due to a general interest (market, research discoveries, exhibitions) towards some artist or genre.

Goals. Discover trends of interest towards artistic periods by looking at the dates of activity of art historians (and their archival collections) that studied a certain period.

Research questions.

- When there is a research interest towards a certain artistic period?
- How the interest evolves over time?
- Which artistic periods show a significant trend over time?

AN EXAMPLE: THE PRESENTATION

Data preparation and data analysis.
We studied ARTchives and Wikidata data models.
Data about archival collections, art historians' activity dates, and related artistic periods are collected from ARTchives. Descriptions and dates of artistic periods are collected from Wikidata.
We query ARTchives and Wikidata SPARQL endpoints, we reconcile the data, we prune duplicate concepts (e.g. "Baroque" and "Baroque art").
We perform some preliminary analysis to understand the distribution of periods over the archival collections.
We analyse the trend of artistic periods as subject of art historians' collections over time.

Data visualisations selected and reasons. We show trends in a line chart having on the x axis the dates of art historians activities (corresponding to the time range of their collection) and on the y axis the artistic periods. So doing a user can see all trends at the same time and figure if there are correlations in trends.

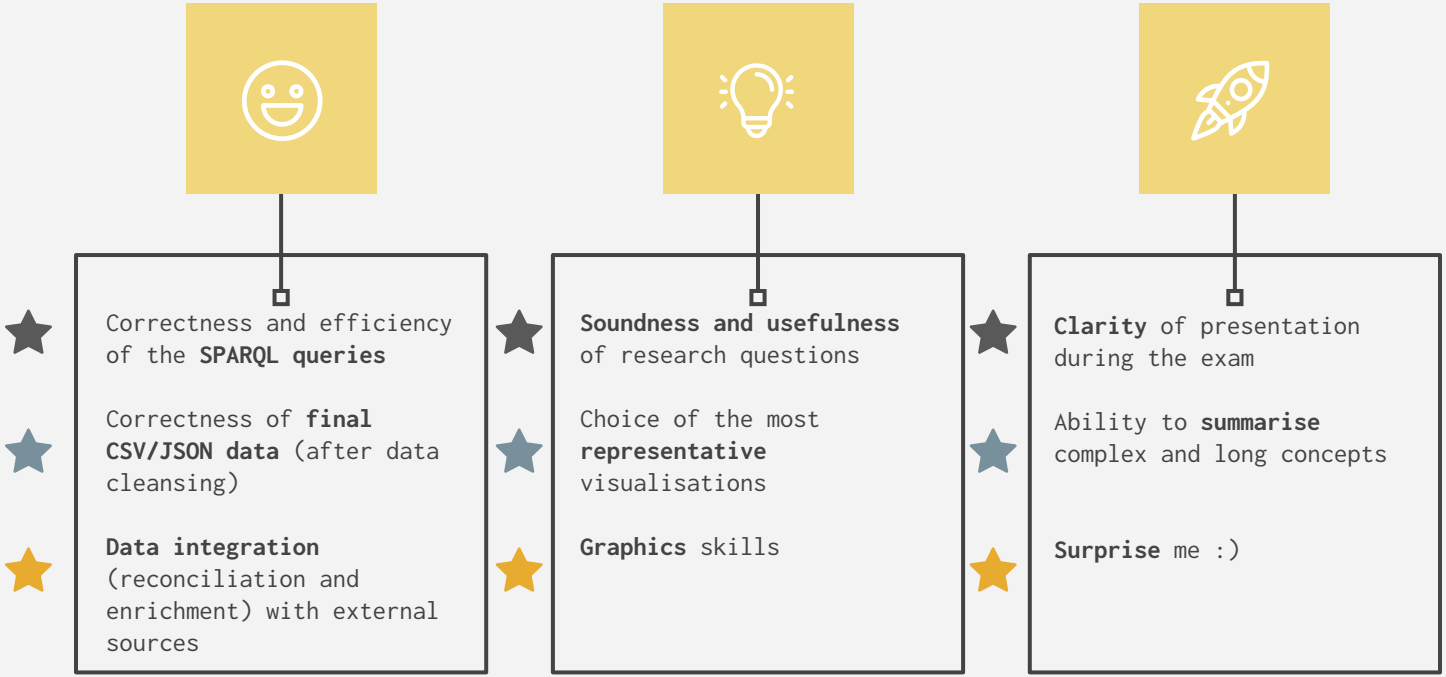
AN EXAMPLE: THE PRESENTATION

Data communication strategies. We first show preliminary exploratory visualisations about the distribution of periods as subjects addressed by art historians so as to demonstrate the validity and representativity of results. We provide brief descriptions of artistic periods for whom may not be acquainted with art history. We show results of our investigation as an interactive line chart where users can select/remove the period to be shown in the graph.

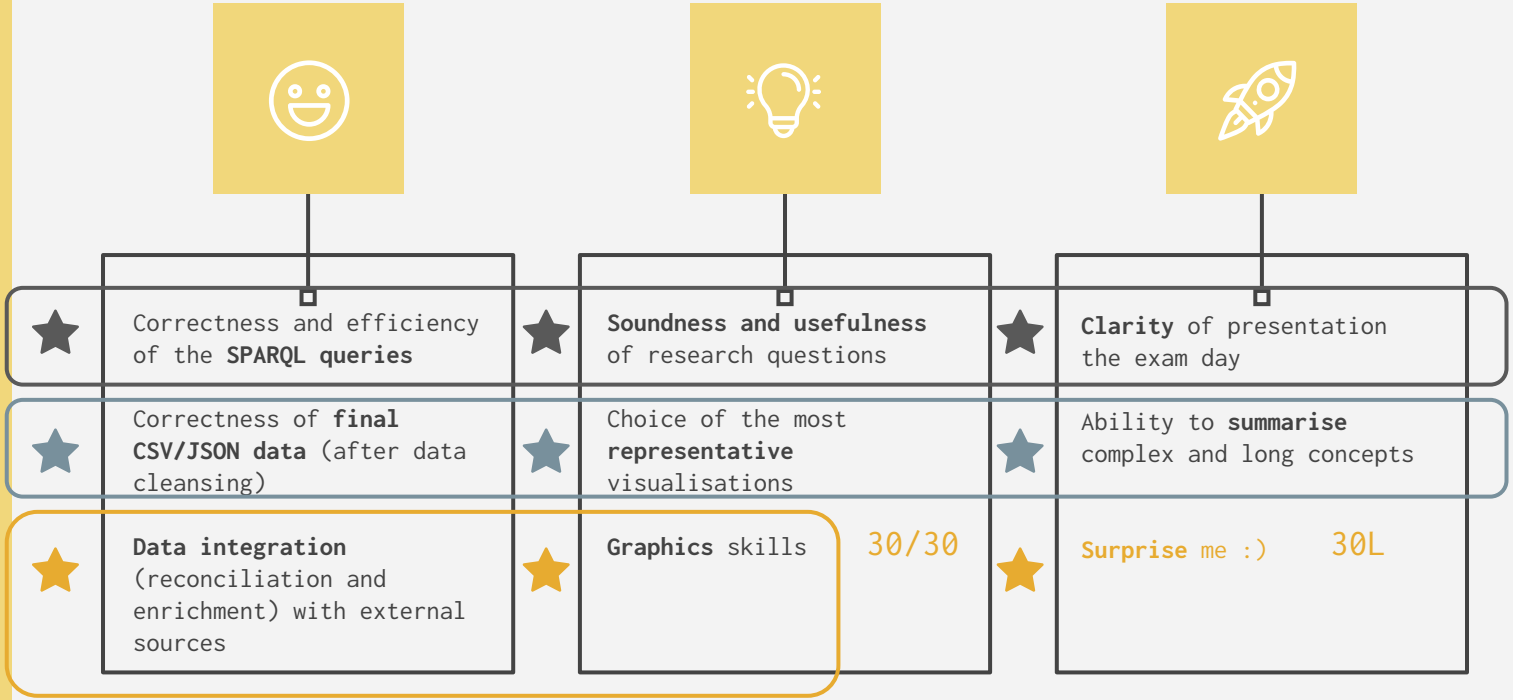
Summary of results. Most significant results show that artistic period XX received lots of attention in the 19XXs, while period YY shows a significant loss of attention in the same period. Notably, period XX and ZZ have similar trend over time.

[TO EXPLAIN WHY THIS HAPPENS IS NOT UP TO YOU - if you are not an art historian]

What I value most



What I value most



Explore ARTchives

ELECTRONIC PUBLISHING
AND DIGITAL STORYTELLING
Lesson 1

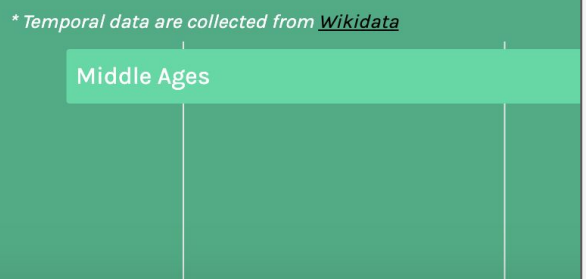
Explore art historians' archival collections th
Discover archival collections that are relate
historical periods, artists, art historians
organisations.

A SMALL REWARD

The most insightful and well-designed visualisations will be published/linked to the website.

ARTISTIC PERIODS

Explore archival collections through



ELECTRONIC PUBLISHING
AND DIGITAL STORYTELLING
Lesson 1

A BETTER (?) REWARD

Co-authoring a scholarly publication dedicated to quantitative art history (in a conference or journal)

LET'S TRY TO GET INTO THE SPIRIT (1)

1. Read this article (20min) on Medium

Daniel Bourke, A gentle introduction to Exploratory Data Analysis. Medium 2019.

Take away message (If you run out of time)

The preliminary questions when exploring a new dataset:

1. What question(s) are you trying to solve (or prove wrong)?
2. What kind of data do you have and how do you treat different types?
3. What's missing from the data and how do you deal with it?
4. Where are the outliers and why should you care about them?
5. How can you add, change or remove features to get more out of your data?

LET'S TRY TO GET INTO THE SPIRIT (2)

2. Explore ARTchives (30 min) and try to (roughly) answer the questions in the prior slide
Understand what types of data are there by navigating through some cataloguing records.

HINT 1: Use the question shown in the presentation example
Trends in the study of artistic periods

HINT 2: Here a preliminary visualisation that helps you
http://artchives.fondazionezeri.unibo.it/contents

3. Fill in the questionnaire (**by the next class**) with your answers
https://forms.gle/gJ1NFrBRzpuDbssg9

WHY A QUESTIONNAIRE?

A few questionnaires will be given to you over time. These are primarily meant to give the teacher **an estimate of your general understanding**.

Results of the tests will not contribute to the final grade, hence these are not mandatory, you should not be afraid to answer wrong (this won't be used against you), and you may even decide to use a pseudonym if you don't want me to know your identity (just be consistent please!).

WHY A QUESTIONNAIRE?

However...

Remember that the final presentation lasts 15 minutes, and it's the first moment I get to know you closely (which can be either good or bad).

If I knew you had **good** results in the questionnaires and you gave a **good** presentation this may encourage me to give you the maximum grade.	If I knew you had **good** results in the questionnaires but you gave a **bad** presentation, I'd take into account your constant effort and you'd not be disadvantaged.
If I knew you had **bad** results in the questionnaires but you gave a **good** presentation this would highly impress me and I'd tend to be more generous.	If I knew you had **bad** results in the questionnaires and you gave a **bad** presentation, well...at least you tried!

Take your chance!

THANKS

Does anyone have any questions?

marilena.daquino2@unibo.it
github