FROM WRITTEN MANUSCRIPTS TO BIG HUMANITIES DATA

Marco Büchler, Greta Franzini, Emily Franzini & Maria Moritz

eTRAP
Electronic Text Reuse Acquisition Project

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

TABLE OF CONTENTS

WHO AM I?

- 2001-2002: Head of Quality Assurance department in a software company;
- 2006: Diploma in Computer Science on big scale co-occurrence analysis;
- 2007: Consultant for several SMEs in IT sector;
- 2008: Technical project management of the eAQUA project;
- 2011: PI and project manager of the eTRACES project;
- 2013: PhD in Digital Humanities on Text Reuse;
- 2014: Head of Early Career Research Group eTRAP at the University of Göttingen.

Electronic Text Reuse Acquisition Project (eTRAP)

Interdisciplinary Early Career Research Group funded by the German Ministry of Education & Research (BMBF).

Budget: €1.6M.
Duration: March 2015 - February 2019. Research since October 2015.
Team: 4 core staff; 5-9 research & student assistants; Bachelor, Masters and PhD thesis students.

- Interdisciplinary: Classics, Computer Science, German Literature, Mathematics, Philosophy, Cognitive Psychology and Literature Studies.

- International: Currently from eight nationalities.

MODULE/COURSES & OBJECTIVES

Ulrike Rieß (*Big Data bestimmt die IT-Welt*):

- Large amounts of data that can't be processed and analysed manually;
- Less structured data, e.g. in comparison to databases and data warehouse systems;
- Heterogeneous and distributed data across resources.

Information overload = large amounts of data (Big Data).
Information poverty = noisy, fragmentary (Humanities Data).

MODULE OBJECTIVES & LEARNING OUTCOMES

- Designed for both students of Computer Science and of the Humanities.
- Working in groups of up to four people and solving problems as a team.
- Involved in the entire process of transforming assets of our cultural heritage into digital data (Digital Transformation).
- Work with the transcriptions of manuscripts, by analysing digitally available texts with text mining and information retrieval techniques.
- Gain knowledge and experience with the problems that arise because of information overload and information poverty.
- Learn how to deal with uncertain data.

Two courses in the module:

- The letters and tales of the brothers Grimm (seminar);
- Cultural Heritage Programming (practical course).

MIXED CLASSES

MIXED CLASSES

Mixed classes are classes with students of different backgrounds such as students from Humanities and Computer Science.
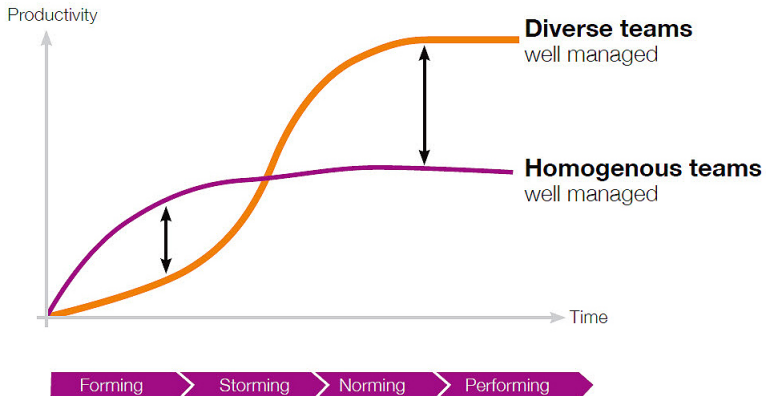
Some facts:

- To cover and improve both scholarly and digital skills.
- Often have no overlapping skills but each bring unique skills to the team.
- Mixed classes take easily twice as much time as "normal" classes for preparation and postprocessing.
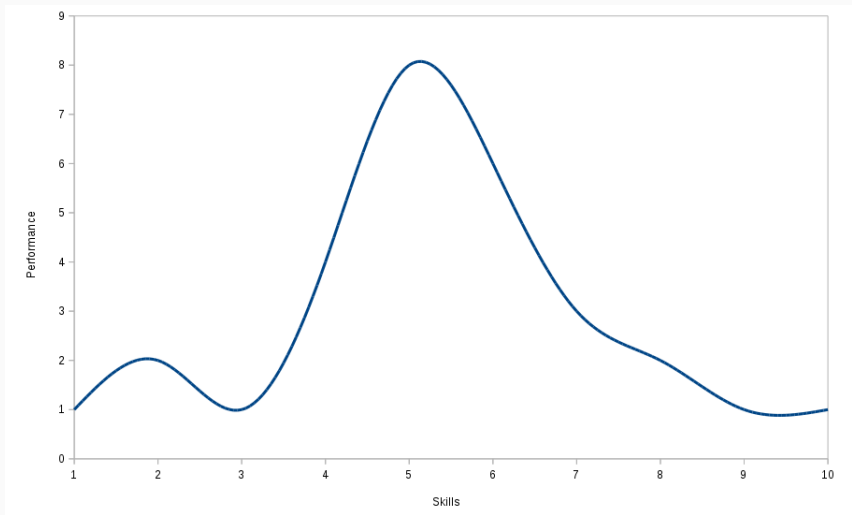
STRUCTURE OF COURSES

- "From teaching to learning."

- "See one, do one, teach one."

- Teaching:
  - **Task**: Sharing as much information as possible to the students ("frontal courses" or a "teacher perspective")
  - **Pro**: Easy to prepare for teachers; interaction with students not necessary as content is on slides or script.
  - **Con**: Students often only gain knowledge but can't apply the knowledge by themselves, i.e. knowledge is not transformed into skills.
- Learning:
  - **Task**: Everybody learns differently and starts from a different background. Focus on training of skills. ("student perspective").
  - **Pro**: It improves proactively skills of students (that are needed for jobs).
  - **Con**: It takes much more time to prepare.

DO ONE, SEE ONE, TEACH ONE.

- **See one**: Teacher provides necessary knowledge and shows how to do something.
- **Do one**: Student practises a skill under supervision of the teacher.
- **Teach one**: Student trains the gained skill to another person (or, ideally, create a new elearning module).

Implication: 2/3 of the content has to be removed from course plan.

STRUCTURE OF COURSES

The Course plan is developed at the beginning of the semester together with the students, and everybody defines skills to be obtained.

Three course sections:

- **1/3: Acquiring knowledge**:
  - Usage of digital teaching material to balance the skills of the students (different materials are provided to different students.)
  - No usage of MOOCs (too high failure rate).
  - Usage of specific elearning material (such as vidoes)
- **1/3: Practising Knowledge**:
  - We invite experts to hold online classes, to share their expertise and inspire with cutting-edge research or send students to the experts.
  - We review submitted papers with the students.
- **1/3: Sharing knowledge**:
  - Creation of elearning modules or even publish results in papers.

Two feedback loops between the three sections by a one minute paper each.

eTRAP's mission is to train skills to students in a way as we would need them to hire staff.

"Train people well enough so they can leave. Treat them well enough so they don't want to."
Sir Richard Branson

Visit us

🌐 http://www.etrap.eu

✉ contact@etrap.eu

Stealing from one is plagiarism, stealing from many is research
(Wilson Mitzner, 1876-1933)

SPONSORED BY THE

eTRAP

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Federal Ministry
of Education
and Research

The theme this presentation is based on is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. Changes to the theme are the work of eTRAP.