

# ➤ Web Information Retrieval

## Module Overview

Frank Hopfgartner, Stefania Zourlidou  
Institute for Web Science and Technologies

# About the WIR module

- **Course ID**
  - 04IN2028
- **Aim**
  - understanding the fundamentals and basics of information retrieval and web search
- **For whom it is intended**
  - Master students in: 1) Web Science, 2) Computer Science, 3) Computer Visualistics, 4) Mathematical Modelling, 5) etc.
    - Elective
  - (some) Bachelor students (please check your programme regulations)
    - Elective

# About the WIR course – Learning Materials

## Material and further information on OLAT

Web Information Retrieval

Module Overview

Lecture Materials

Exercise Materials

Forum

Web Information Retrieval SoSe 2023

The lecture will give an introduction in established retrieval models for text based documents, models that exploit the graph structure of the WWW, the topic of evaluating the performance of retrieval systems and related tasks like classification and clustering of web documents. The concepts communicated in the lecture will be applied in practical exercises and tutorials.

In order to complete the module, students first have to achieve a score of at least 60% in the weekly exercises and complete at least 80% of all assignments, which then qualifies them to sit the final exam. The threshold to pass the exam is 50%. Exam qualifications from previous years are *not* carried over.

Module Overview

In this folder, you can find general information about the module.

Lecture Materials

In this folder, you can find lecture slides and further reading materials. We will release new learning materials each week.

Exercise Materials

As entry requirement for sitting the final exam, students are required to score 60% or higher in the weekly exercises and complete at least 80% of all assignments. The solutions are discussed in the following week in our tutorial sessions. We will release new assignments each week.

Forum

You can use the forum to ask questions about the module content. We will also use it to announce issues relevant to the module.

<https://olat.vcrp.de/url/RepositoryEntry/4071063853>

# About the WIR course - Lectures

- **Where**
  - M001 and some (potentially) online on BigBlueButton
- **When**
  - Wednesdays at 04:00pm (except public holidays and semester breaks)
  - See the schedule on KLIPS
- **Lecturers**
  - Stefania Zourlidou, Frank Hopfgartner
- **Consultation**
  - Only via OLAT Forum
- **Klips Link**
  - <https://klips.uni-koblenz-landau.de/v/123564>
- **Credit points**
  - 6 ECTS

# About the WIR course - Tutorials

- **Where and When?**
  - Thursdays in M201 or online (at 14:00)
  - Starting from: 26 April
- **Given by**
  - Stefania Zourlidou, Frank Hopfgartner
- **How to contact us?**
  - Use OLAT Forum
  - **Only** for emergency cases, send an email

- **The next assignment sheet**
  - Will be published weekly on Thursdays after the tutorial
  - Must be submitted
    - before 23:59 on Wednesdays
    - on OLAT
- **Requirements**
  - The assignment sheet must be solved individually
  - The programming language (if applicable) is Python

- **Where**
  - 08 August 2023
- **When**
  - D 028
- **How**
  - Written
- **Requirements**
  - Achieving 60% of practice sheet points in tutorials
  - Submitting 80% of all assignment sheets before their deadlines
  - Registering for the exam before the deadline

# Can I pass the course?

- Yes, if .....
- Attend both lecture and tutorials
- Ask questions
- Take notes
- Study the recommended material
- Solve the exercise sheets by yourself
- Practice programming
- Never cheat nor plagiarize



# Prerequisites

- Programming skills in general (Python in particular)
- Basic understanding of maths, algebra, probability theory, etc

# Books

- *Introduction to Information Retrieval.*  
Manning, Raghavan, Schütze  
Cambridge University Press, 2009.
- *Web Data Mining.*  
Liu.  
Springer, 2007.
- *Modern Information Retrieval.*  
Baeza-Yates, Ribeiro-Neto  
ACM Press, 2012.



# What is plagiarism?

- Turning in someone else's work as your own
- Copying words or ideas from someone else without giving credit
- Failing to put a quotation in quotation marks
- Giving incorrect information about the source of a quotation
- Changing words but copying the sentence structure of a source without giving credit
- Copying so many words or ideas from a source that make up the majority of your work, whether you give credit or not (see the section on "fair use" rules)
- Acknowledging that the material has been borrowed and providing your audience with the necessary information to find the source is usually enough to avoid plagiarism

# Important to know

« Any sharing, copying or plagiarism of any exercise or assignment will result in a repeat of the course the following year for both students.»

**We take this very seriously!**

# Credit for these slides

These slides have been adapted from

- Web IR (Zeyd Boukhers-WeST)