Course memo

Marcel Bollmann

Introduction

This document contains general information about the course *Language Technology*. The course is co-taught as `729G86` and `TDP030`; unless a course code is specifically mentioned in the text, the information here applies to both course codes equally.

Contents

1 Course content

Language technology – technology for the analysis and interpretation of natural language – forms a key component of smart search engines, personal digital assistants, and many other innovative applications. The goal of this course is to give an introduction to language technology as an application area, as well as to its basic methods. The course focuses on methods that process text.

The course covers:

- basic methods and techniques for the analysis and interpretation of words, sentences, and texts
- language technology systems
- validation methods
- tools, software libraries, and data

The course content is structured into *concepts* and *procedures*. By *concepts* we mean terms and models that you should be able to explain and apply. By *procedures* we mean standard tasks that you should be able to perform. If a concept or procedure is classified as *advanced*, it is beyond what is expected from you for a passing grade.

1.1 Intended learning outcomes

On completion of the course, you should be able to:

1. explain basic methods for the analysis and interpretation of words, sentences, and texts
2. practically apply language technology methods and systems to texts and text collections
3. evaluate language technology components and systems using standard validation methods
4. judge the difficulty and the feasibility of language technology applications

For each intended learning outcome, there is a set of more specific knowledge requirements that express what you need to demonstrate in order to attain a particular grade. These knowledge requirements are listed under each graded module below.

1.2 Teaching and working methods

The course is taught in the form of lectures, tutored lab sessions, and supervision in connection with a project. You are also expected to study independently, both individually and in groups. When you plan your time for the course, you should calculate approximately:

- **53 hours** to prepare for, attend, and follow up on the **lectures**.
- **53 hours** to prepare for, carry out, and follow up on the **labs**.
- **53 hours** to plan, carry out, and reflect on the **project**.

1.3 Course literature

The reading for this course consists of individual sections from the following book, which is available to read for free online:

- Dan Jurafsky and James H. Martin. *Speech and Language Processing*. Draft chapters of 3rd edition, January 2023.

For follow-up and in-depth reading, we recommend the following books:

- Emily M. Bender. *Linguistic Fundamentals for Natural Language Processing: 100 Essentials from Morphology and Syntax*. Springer, 2013.
- Emily M. Bender. *Linguistic Fundamentals for Natural Language Processing II: 100 Essentials from Semantics and Pragmatics*. Springer, 2020.
- Markus Dickinson, Chris Brew, and Detmar Meurers. *Language and Computers*. Wiley–Blackwell, 2012.

2 Examination

The course has the following examination modules:

DIT1	Digital written examination	2 credits
LAB1	Practical assignments	2 credits
UPG1	Project assignments	2 credits

Each examination module is graded on the full grading scale depending on your course code (`729G86` *Fx, E, D, C, B, A*; `TDP030` *U, 3, 4, 5*).

> 🛈
>
> Grade conversion for `TDP030` Throughout this document, we use the ECTS scale to describe knowledge requirements and grade requirements. If your course code is TDP030, we convert them to your grading scale as follows:
>
> | `729G86` | Fx | E | D | C | B | A |
> |---|---|---|---|---|---|---|
> | `TDP030` | U | 3 | 3 | 4 | 4 | 5 |

- To **pass the course**, you must have a passing grade on all three modules; in other words, you cannot pass the course if you have a failing grade (*Fx/U*) on any of the modules.
- Your **final grade** is the *median* of the grades for the three modules. (To compute the median of three grades, you sort them and pick the middle one.)

2.1 Examination dates

The formal examination date for the DIT1 module is the date of the digital written exam, which you will find in the Student Portal. The formal examination date for the other modules (LAB1 & UPG1) is the final examination date for the course session, which is **2025-03-28**.

Additional examination For each module, there are **two** additional examination opportunities during the year following the course session:

- during the re-exam period for VT1 (last day: **2025-06-12**)
- during the general re-exam period after the summer break (last day: **2025-08-30**)

After this, you can still be examined in connection with the next course session. Note however that the next session may feature different content, different assignments and different examination requirements.

2.2 Cheating and plagiarism

Each piece of work you present for examination must be entirely your own. You are not allowed to give or receive aid on an assignment unless such collaboration is explicitly permitted. The use of prohibited aids is cheating.

The following, in particular, is not allowed:

- copying code from other lab groups, or letting other lab groups copy your code
- making lab solutions available via public channels, such as Github

When using external sources (such as text or code) in work that you present for examination, you must appropriately acknowledge these sources. This rule also applies to materials obtained from the internet. Failure to acknowledge your sources is plagiarism.

Use of generative AI   For using generative AI tools including (but not limited to) GitHub Copilot or ChatGPT, the following rules apply:

- For the examination in **DIT1**, you **may** *not* use generative AI tools in any form.
- For the examination in **LAB1**, you **may** *not* use generative AI tools for any of the free-text answers. We also discourage the use of generative AI tools for the programming exercises, since these exercises already provide a lot of guidance and consist of small, focused tasks; additionally, lab assistants offer you personal guidance and feedback. The use of generative AI tools would therefore not aid, but rather undermine, the intended learning outcomes.
- For the examination in **UPG1**, generative AI tools *may* be used for providing inspiration, assisting you with producing code you may need to implement your project, and helping you to prepare your presentation. They **may** *not* be used to produce entirely new text for your post-project paper, regardless of whether you copy the text verbatim or edit/rephrase it; the post-project paper is intended to be *your own* personal reflection.

ℹ

Consequences of cheating and plagiarism

We **must** report suspected cheating and plagiarism cases to the University Disciplinary Board. No exceptions exist to this rule, no matter how slight the potential offence.

2.3 Commonly asked questions

When do I get my grade? For each attempted module, the examiner will decide about your grade within 15 working days after the module's formal examination date. Your final grade will be set as soon as you have passed all modules.

When does a module count as attempted? A module counts as attempted when you take an examination or submit an assignment that belongs to that module, e.g. one lab in a lab series.

What happens if I fail to pass a module? If you attempt but do not pass a module, you will receive a failing grade for this module (*Fx/U*). You have two additional examination attempts for each module during the year following the course; see Sec. 2.1.

Can I still get a higher grade if I submit for one of the additional examinations? Yes. There are no grade penalties for late submissions or re-submissions; the grade requirements are exactly the same for each examination date.

3 DIT1: Digital written examination

The digital written exam tests how well you are able to explain basic methods for the analysis and interpretation of words, sentences, and texts (learning outcome 1). This includes relevant validation methods.

Knowledge requirements

Grade E You demonstrate a *good* understanding of the basic methods that are covered in the course. You assess the applicability, requirements, and limitations of methods with *simple* judgements. You use relevant terminology and notation with *some certainty*.

Grade C You satisfy all of the requirements for grade E and some of the requirements for grade A.

Grade A You demonstrate a *thorough* understanding of *all* the methods that are covered in the course. You assess the applicability, requirements, and limitations of methods with *well-developed* judgements. You use relevant terminology and notation with *certainty*.

Form of the examination & grade requirements   The digital written exam will be conducted through the Inspera platform. Items on the exam will award points; point requirements for different grades will be prominently stated on the exam.

(This document may be updated later with more information on the exact format and point requirements for the exam.)

Registration   In order to take the digital written exam, you have to register for it in the Student Portal. The last day to register is 10 days before the exam. This also applies to re-exams.

Feedback   To get feedback about how well you meet the knowledge requirements for this module, you can answer practice questions that will be given during the lectures and/or made available during the course. In addition, you can always get personalized feedback from the examiner.

4 LAB1: Practical assignments

The practical assignments test how well you are able to (a) practically apply language technology methods and systems to texts and text collections (learning outcome 2), and (b) evaluate language technology components and systems using standard validation methods (learning outcome 3).

Knowledge requirements

Grade E You correctly apply the basic methods that are covered in the course according to *given instructions*. You make *minor* modifications to existing systems and implement *simple* functions. You interpret results and assess the applicability, requirements, and limitations of methods with simple judgements.

Grade C You correctly apply the basic methods that are covered in the course, *as well as several of the more advanced methods*, and use them to *independently solve practical problems*. You make *extensive* modifications to existing systems and implement *complex* functions. You interpret results and assess the applicability, requirements, and limitations of methods with simple judgements.

Grade A You correctly apply the basic methods that are covered in the course, *as well as most of the more advanced methods*, and use them to *independently solve practical problems*. You make *extensive* modifications to existing systems and *implement your own systems from scratch*. You interpret results and assess the applicability, requirements, and limitations of methods with simple judgements.

Form of the examination   The assignments of this component consist of lab assignments, or "labs," which are done in pairs. The labs come at two levels:

Basic labs (six in total) test your command of the **basic** methods that are covered in the course. These labs come with step-by-step instructions and ready-to-run systems. The instructions ask you to make minor modifications to these systems, evaluate the systems on text data, and interpret your findings.

Advanced (X) labs (five in total) test your command of the more **advanced** methods that are covered in the course. These labs are formulated as problem specifications and either (a) come with ready-to-run systems to which you will have to make *significant modifications*, or (b) come *with little or no code* and require that you implement your own (limited-size) systems. In either case, you will have to evaluate the final systems on text data and interpret your findings.

Grade requirements

Grade E Pass all the basic labs.

Grade D Pass all the basic labs and *one* of the advanced labs.

Grade C Pass all the basic labs and *two* of the advanced labs.

Grade B Pass all the basic labs and *three* of the advanced labs.

Grade A Pass all the basic labs and *at least four* of the advanced labs.

Registration Before submitting your first lab, you and your lab partner must sign up in Webreg. Links to the sign-up page are provided on the course website. If you cannot find a lab partner, let us know *at the latest* during the first lab session, and we will pair you up with a random student.

Each lab assignment has two due dates: The first due date is stated in the instructions for the assignment on Lisam. The second due date is the final examination date for the course session, **2025-03-28**.

> ℹ️
>
> Try to meet the first due date! If you meet the first due date, you will get formative feedback on your assignment and the chance to revise it before its final assessment after the examination deadline. If you do *not* submit by the first due date, we will not be able to assess your submission until after the second due date.

Feedback To get feedback about how well you meet the knowledge requirements for this module, you can attend the tutored lab sessions and make sure that you submit your assignments in time for the first due date. In addition, you can always get personalized feedback from the examiner.

5 UPG1: Project assignments

The project assignments primarily test how well you are able to judge the difficulty and the feasibility of language technology applications (learning outcome 4). They also test how well you are able to (a) practically apply language technology methods and systems to texts and text collections (learning outcome 2), and (b) evaluate language technology components and systems using standard validation methods (learning outcome 3).

The assignments are centered around a **group project** in which you investigate a language technology application of your choice.

Knowledge requirements

Grade E In collaboration with the other members of your group, you identify, plan, and carry out a *simple* project, and present the project in a way that makes it *clear* what application of language technology you investigated, what results you obtained, how you interpret these results, and what conclusions you draw regarding the difficulty and feasibility of the chosen application. You make *simple* reflections on your project and your own learning.

Grade C You meet all of the criteria for grade E and most of the criteria for grade A.

Grade A In collaboration with the other members of your group, you identify, plan, and carry out a *complex* project, and present the project in a way that makes it *very clear* what application of language technology you investigated, what results you obtained, how you interpret these results, and what conclusions you draw regarding the difficulty and feasibility of the chosen application. You make *well-developed* reflections on your project and your own learning.

Form of the examination　The project should be done in groups of approximately six students, and is examined based on hand-in assignments and an oral project presentation. Some of these deliverables are group work, some are individual. Your grade is based on the following two parts:

Part 1 is the project's complexity and presentation (group work).

Part 2 is your post-project paper (individual work).

🛈

> **More details in a separate document!**　We provide more details about the project work in a separate document that you can find under "Project" on the course website. This includes the assessment criteria for the two examined parts, an overall timeline of the project, guidelines for the post-project paper, and other tips for successfully completing your project.

Grade requirements　Your overall grade for the project module takes into account both the group part and the individual part. For more information, please refer to the separate "project" document on the course website.

Additional examinations　If you do not fulfill all of the passing requirements for the group part of the project, you need to consult with the examiner at least 10 days before an examination date. If you fulfill the passing requirements for the group part, but not the individual part, you can simply submit (or re-submit) your post-project paper for another examination date.

6 Other course policies

6.1 Feedback policy

What you can expect from us   We try our best to give you prompt, constructive, and meaningful feedback on how well you meet the knowledge requirements set out for the course. Our focus is on *formative feedback*, which you can use to improve your learning (and we can use to improve our teaching!) while the course is ongoing.

What we expect from you   We expect you to familiarise yourself with the knowledge requirements set out for the course and to actively seek our feedback on how well you meet these requirements. We also expect you to reflect on the feedback we provide and grasp opportunities to put it to good use.

6.2 Communication policy

What you can expect from us   When you contact us via email or Teams chat, you can expect an answer during standard working hours, 9–17. (Do not expect us to respond to email/chat in the evening or on the weekend.) For more personal contact, you can talk to the examiner in class or book an appointment.

What we expect from you   The course website and the documents linked there are the primary sources of information about the course, and we expect you to keep yourself up-to-date with what we publish there. We also send out information via the University's email list for the course, and we expect you to check for and read this information while the course is ongoing.

6.3 Special needs

Accessibility   If there is any portion of the course that is not accessible to you due to challenges with technology or the course format, please let the examiner know so we can make appropriate accommodations.

Students with disabilities   If you have a documented disability, you should contact the examiner as soon as possible regarding accommodations.

Credits

This document is in large parts based on material developed by Marco Kuhlmann.