

MODELLO D (inglese)	
General Information	
Academic subject	Intelligent Information Access and Natural Language Processing
Degree course	Master of Science (MSc) in Computer Science (LM18)
Curriculum	Knowledge Engineering and Machine Intelligence
ECTS credits	6
Compulsory attendance	No, but attendance is strongly recommended
Language	English

Subject teacher	Name Surname	Mail address	SSD
	Giovanni Semeraro	giovanni.semeraro@uniba.it	INF/01
Office and Office Hours	Office 757 Dipartimento di Informatica Campus Universitario Via E. Orabona 4 70126 Bari	Tuesday: 4:00pm-7:00pm	

ECTS credits details			
Basic teaching activities	Lectures	4 T1 credits	
	Laboratory and guided exercises	1 T2 credit	
	Practical project work	1 T3 credit	

Class schedule	
Period	I semester
Year	Second
Type of class	Lectures - laboratory and guided exercises - practical project work

Time management	
Hours	150
Hours of lectures	47
Tutorials and lab	103

Academic calendar	
Class begins	28th September 2020
Class ends	13th January 2021

Syllabus	
Prerequisites/requirements	Formal prerequisites: none Cultural prerequisites: knowledge of a programming language (necessary), discrete mathematics (relevant), probability and statistics (important), first year courses (important).
Expected learning outcomes (according to Dublin Descriptors) (it is recommended that they are congruent with the learning outcomes contained in A4a, A4b, A4c tables of the SUA-CdS)	The learning goals, depicted according to Dublin Descriptors, are the following: <i>Knowledge and understanding</i> The learner will acquire knowledge about fundamentals of natural language processing and intelligent information access. In particular, the learner will be able to know and understand:

- the theoretical, methodological and operational aspects of natural language processing with particular reference to the main levels of linguistic analysis;
- the techniques and the main open source platforms for natural language processing;
- the theoretical, methodological and operational aspects of intelligent information access and semantic information indexing and sharing;
- the theoretical, methodological and operational aspects of information filtering systems and recommender systems;
- the techniques and major open source platforms for the design of recommender systems.

Applying knowledge and understanding

The learner will be able to:

- apply knowledge on natural language processing and the main levels of linguistic analysis to address problems related to the management and integration of unstructured data (text) and intelligent information access, even in new, non-familiar or interdisciplinary contexts;
- exploit the techniques and the main platforms for natural language processing and the design of recommender systems in order to develop and evaluate complex computer systems that require the management and integration of unstructured data (text) and the filtering of information;
- apply his/her own competences to identify effective solutions to complex problems related to the processing of unstructured data and information filtering, and to justify, support and argue his/her choices.

Making informed judgements and choices

The learner will be able to:

- make his/her own assessment and define a critical judgment, and support it within the working group with which the case study has been developed;
- integrate necessary knowledge in an autonomous way as well as manage complexity resulting from the limited or incomplete information available;
- make decisions and identify solutions in the areas of intelligent information access and natural language processing taking into account the social and ethical implications and the professional responsibilities that they imply.

Communicating knowledge and understanding

The learner will be able to:

- choose the appropriate form and means of communication for the interlocutors, both specialists and non-specialists;
- communicate effectively information, ideas, problems and solutions related to intelligent information access and natural language processing.

Lifelong learning skills capacities to continue learning

The learner will acquire the ability to:

- develop a high level of autonomy in learning the disciplines of natural language processing and intelligent information access;

	<ul style="list-style-type: none"> - keep up-to-date with the evolution of the disciplines of natural language processing and intelligent information access by drawing on bibliographic sources in both Italian and English; - pursue his/her own training course by undertaking high-level studies.
<p>Contents</p>	<ol style="list-style-type: none"> 1. Introduction to natural language processing. 2. Elements of computational linguistics. 3. Lexical level. 4. Part-of-Speech (PoS) tagging. 5. Syntactic level. 6. Semantic level: basics. 7. Semantic level: advanced techniques. 8. Open source frameworks for natural language processing. 9. Introduction to intelligent information access. 10. Open data. 11. Recommender systems. 12. Collaborative filtering and Content-based filtering. 13. Other recommender systems paradigms. 14. Evaluation of recommender systems and advanced techniques. 15. Open source frameworks for the design of recommender systems.
<p>Course program</p>	
<p>Bibliography</p>	<ol style="list-style-type: none"> 1) D. Jurafsky and J. Martin, Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice-Hall Series in Artificial Intelligence, 2000. ISBN: 0130950696. 2) C. Manning and H. Schutze, Foundations of Statistical Natural Language Processing. MIT press, 2000. 3) Julia Hirschberg, Christopher D. Manning. Advances in natural language processing. Science, Vol. 349 no. 6245, pp. 261-266, 2015. DOI: 10.1126/science.aaa8685 3) Mark Stevenson, Word sense disambiguation: the case for combinations of knowledge sources. Stanford, CA: CSLI Publications, 2003. ISBN: 1575863898. 4) Eneko Agirre and Philip Edmonds, Word sense disambiguation: Algorithms and Applications. Springer Text, Speech and Language Technology, Vol. 33, 2007. ISBN: 1402048084.

	<p>5) Dominic Widdows, <i>Geometry and Meaning</i>. CSLI Publications, Stanford CA, 2004. ISBN: 1575864487.</p> <p>6) Dietmar Jannach, Markus Zanker, Alexander Felfernig, Gerhard Friedrich, <i>Recommender Systems: An Introduction</i>. Cambridge University Press, 2010. ISBN: 0521493366.</p> <p>7) F. Ricci, L. Rokach, B. Shapira (Eds.), <i>Recommender Systems Handbook</i>. 2nd Edition, Springer, 2015. ISBN: 9781489976369.</p> <p>8) Armano, G.; de Gemmis, M.; Semeraro, G.; Vargiu, E. (Eds.), <i>Intelligent Information Access. Studies in Computational Intelligence</i>, vol.301, Springer, 2010. DOI: http://dx.doi.org/10.1007/978-3-642-14000-6. ISBN: 3642139994.</p> <p>9) P. Lops, C. Musto, F. Narducci, G. Semeraro, <i>Semantics in Adaptive and Personalised Systems: Methods, Tools and Applications</i>, Springer Nature, 2019. ISBN: 978-3-030-05618-6.</p> <p>10) AA.VV. <i>Open Data: come rendere aperti i dati delle PA - Linee guida per i siti web della PA</i>. Vademecum. Formez PA, Ottobre 2011.</p> <p>11) F. Di Donato, <i>Lo stato trasparente: Linked open data e cittadinanza attiva</i>. Edizioni ETS, 2011 ISBN: 9788846728876 (versione con licenza Creative Commons Attribuzione-Non commerciale 2.5 Italia: http://www.linkedopendata.it/wp-content/uploads/statotrasparente.pdf).</p>
Notes	<p>The teacher will provide the students with supplemental material, which will be available on the e-learning platform: https://elearning.di.uniba.it/</p>
Teaching methods	<p>Lectures (32 hours), exercises in the classroom (15 hours)</p>
Assessment methods (indicate at least the type written, oral, other)	<p>Exam sessions (at the end of the course). The exam consists of an oral test concerning the presentation and discussion of a case study or project work of experimental/applicative nature, chosen among those proposed during the lectures. The discussion of the test is individual, while the development of the case study or project work can be carried out in groups of maximum three students. The duration of the test varies depending on the type of case study or project work; however, it does not exceed 45 minutes.</p>
Evaluation criteria (Explain for each expected learning outcome what a student has to know, or is able to do, and how many levels of achievement there are)	<p>In the oral test the student will have to demonstrate:</p> <ul style="list-style-type: none"> - to have acquired knowledge and have thoroughly understood the theoretical, methodological and operational aspects of the intelligent access to information and the processing of natural language; - how to properly apply the knowledge on natural language processing and intelligent information access in order to identify effective solutions to problems encountered in the development of the case study or project work (even integrating techniques and components made available by the open source platforms object of study and exercises); - to be able to adequately justify the project choices, supporting them through critical arguments and demonstrating awareness of the social and ethical implications as well as the professional responsibilities they entail; - to know how to communicate clearly and comprehensively;

	<p>- to have developed a high level of autonomy, through the enucleation of his/her own contribution when the case study or project work has been developed by a team.</p> <p>The evaluation of the test is expressed over a 30-point scale. The test is passed with a minimum score of 18/30. The determination of the score takes into account the following aspects:</p> <ol style="list-style-type: none"> 1) correctness of the solutions proposed during the development of the selected case study or project work; 2) completeness of the solutions proposed during the development of the selected case study or project work; 3) logic followed by the student when proposing the solutions; 4) use of an appropriate formalism for the description of the solutions proposed during the development of the selected case study or project; 5) degree of innovation of the solutions proposed during the development of the selected case study or project. <p>To pass the exam, the student should be able to propose a solution that satisfies at least aspect 1). Students able to develop the case study or the chosen project work in order to also satisfy aspects 2) -5) get higher scores.</p>
Further information	