

中国社会科学院大学“大数据与计算传播”课程大纲

课程基本信息 (Course Information)					
课程编号 (Course ID)	102073020167	*学时 (Credit Hours)	32	*学分 (Credits)	2
*课程名称 (Course Name)	大数据与计算传播				
	Big Data and Computational Communication				
先修课程 (Prerequisite Courses)	大学计算机				
*课程简介 (Description)	本课程是新闻学专业和广播电视学专业的专业选修课，是一门理论和实践结合的实务课程，其前置课程是“大学计算机”。 本课程内容丰富，涉及面广，在教学中帮助学生培养社会主义核心价值观和创新精神，深刻理解习近平新时代中国特色社会主义思想。坚持正确的舆论导向，弘扬主流价值观，中华民族自豪感。本课程目标是要培养学生的问题意识和计算思维，以及以数据为视角的问题分析能力。理解计算传播学研究中从数据到模式,再从模式到机制的研究思路。学生能够以大数据，文本和情感分析、建模等为主要分析工具。（以非介入地方式）大规模地收集并分析人类传播行为数据，分析人类传播行为背后的模式和法则，分析模式背后的生成机制与基本原理。 本课程采取跨学科的研究方法和研究视角，以经典案例剖析为路径，介绍探索人类传播行为的模式和规律的前沿方法和思路。主要教学内容包括大数据基础，计算传播的起源、概念和应用。数据获取，数据清洗，统计思维，文本分析，情感分析和语义建模。难点是文本分析和情感分析。 主要教学方法包括讲授法、讨论法、案例教学法。 基本要求：遵守课堂纪律，积极参与课堂讨论，按时出勤并完成平时作业。				
*课程简介 (Description)	This course is a professional elective course for journalism and radio and television majors. It is a practical course combining theory and practice. Its prerequisite course is "University Computer". This course is rich in content and covers a wide range of subjects. In teaching, it helps students cultivate socialist core values and Marxist news concept, and deeply understand Xi Jinping's socialist thought with Chinese characteristics in the new era. Adhere to the correct guidance of public opinion, promote mainstream values, Chinese national pride and awareness of cybersecurity. The goal of this course is to cultivate students' problem awareness and computational thinking, as well as their ability to analyze problems from a data perspective. Understand the research approach in computational communication studies from data to patterns, and then from patterns to mechanisms. Students can use big data, text and sentiment analysis, modeling, and other analysis tools as their main analysis tools. Collect and analyze human communication behavior data on a large scale (in a non-invasive manner), analyze the patterns and rules behind human communication behavior, and analyze the generation mechanisms and basic principles behind the patterns.				

	This course adopts interdisciplinary research methods and perspectives, using classic case analysis as the path to introduce cutting-edge methods and ideas for exploring the patterns and patterns of human communication behavior. The main teaching content includes the basics of big data, the origins, concepts, and applications of computational propagation. Data acquisition, data cleaning, statistical thinking, text analysis, sentiment analysis, and semantic modeling. The difficulty lies in text analysis and emotional analysis. The main teaching methods include lecture method, discussion method and case teaching method. Basic requirements: Obey classroom discipline, actively participate in classroom discussions, attend work on time and complete homework.		
*教材 (Textbooks)	《计算传播学导论》，张伦等编著，北京师范大学出版社，2018 年，第一版，ISBN：9787303241200。		
参考资料 (Other References)	1. 《计算传播》，刘峰，王虎编著，华中科技大学出版社，2022 年，第一版，ISBN：9787568080620 2. 《python 计算传播学实战》，隆广庆主编，科学出版社，2021 年，第一版，ISBN：9787030698940		
*课程类别 (Course Category)	☐ 公共基础课/全校公共必修课 ☐ 通识教育课 ☐ 专业基础课 ☐ 专业核心课/专业必修课 ☒ 专业拓展课/专业选修课 ☐ 其他_____		
*授课对象 (Target Students)	新闻学专业、广播电视学专业	*授课模式 (Mode of Instruction)	☐ 线上，教学平台_____ ☒ 线下 ☐ 混合式 ☐ 其他 ☐ 实践类（70%以上学时深入基层）_
*开课院系 (School)	新闻传播学院	*授课语言 (Language of Instruction)	☒ 中文 ☐ 全外语_____ ☐ 双语：中文+_____（外语讲授不低于 50%）
*授课教师信息 (Teacher Information)	课程负责人姓名及简介	课程负责人：刘英华 中国社会科学院大学新闻传播学院副教授，硕士生导师，博士。中国科技新闻学会数据新闻专委会常务理事。长期从事新技术与数字传播、数据新闻与可视化传播、计算传播等方面的研究。已发表学术论文 30 余篇，出版学术专著 1 部，译著 1 部，教材 5 部，主持并参与国家自然科学基金面上资助项目 2 项、中国青少年中心和北京市课题 1 项，北京市“双一流”建设经费项目 1 项，教育部产学研合作协同育人项目 2 项。	
	团队成员姓名及简介	无	
学习目标 (Learning Outcomes)	1. 形成社会主义核心价值观和马克思主义新闻观，深刻理解习近平新时代中国特色社会主义思想； 2. 使学生全面了解大数据和计算传播的发展历程； 3. 掌握数据科学的编程工具并熟练完成数据收集和数据清洗； 4. 训练学生统计思维，熟练掌握文本分析、情感分析和语义建模方法。		

*考核方式 (Grading)		本课程的总成绩由平时成绩和期末成绩两部分构成： 平时成绩 30% ： 出勤、平时作业等； 课下作品 70% ： 小组完成。					
*课程教学计划（Teaching Plan）							
周次	周学时	其中					教学内容摘要 (必含章节名称、讲述的内容提要、实验的名称、教学方法、课堂讨论的题目、阅读文献参考书目及作业等)
		理论讲授	实践讲授	习题课	课程讨论	其他环节	
第一周	3	3	0	0	0	0	第 1 章 导论 大数据简介。 传统传播学简介。 【教学方法】讲授法、讨论法、案例教学法。 【阅读文献】 [1] http://ghostweather.slides.com/lynncherny/what-is-big-data-any-way [2] 祝建华, 彭泰权, 梁海, 王成军, 秦洁, & 陈鹤鑫. (2014). 计算社会科学在新闻传播研究中的应用. 科研信息化技术与应用, 5(2), 3-13. 【讨论题目】 大数据中“大”的具体含义？
第二周	3	1	2	0	0	0	第 2 章 数据科学的编程工具 编程工具的特点和类库。 案例：英国卫报数据博客“2009 年英国国会议员开支丑闻”的幕后花絮。 案例：澎湃新闻作品“转角遇到爱”背后的数据 案例：CGTV 作品“Who runs China？”背后的数据 【教学方法】讲授法、案例教学法。
第三周	3	1	2	0	0	0	第 3 章 数据收集 数据收集的基本原理。 数据收集的流程。 案例：抓取当下的某类热点新闻。 【教学方法】讲授法、讨论法、案例教学法。 【讨论题目】 数据收集过程中需要注意哪些问题？

第四周	3	2	1	0	0	0	第4章 数据清洗 数据清洗的目的。 按行或块的方式对预处理大数据。 数据清洗流程。包括：处理缺失数据、添加默认值、删除不完整的行和列、规范化数据类型、必要的转换、变量重命名、保存结果。 案例：对抓取的数据进行数据清洗。 【教学方法】 讲授法、案例教学法。
第五周	3	2	1	0	0	0	第5章 统计思维 基于案例理解“幸福感”的概念、缺失值的处理、人口统计变量、认知水平和抑郁水平的测量、大五人格、描述性分析、价值观的测量。 案例：中国家庭追踪调查。 【教学方法】 讲授法、讨论法、案例教学法。 【作业】 分析 kaggle 大赛 Titanic 数据集。
第六周	3	2	1	0	0	0	第6章 文本分析 文本分析研究现状。文本分析与传播学研究。文本分析的基本步骤。 【教学方法】 讲授法、讨论法、案例教学法。
第七周	3	2	1	0	0	0	第6章 文本分析 文本分析的常用库和基本模型。包括 NLTK、Textmining、LDA、Gensim 和 jieba 等及其案例。 【教学方法】 讲授法、讨论法、案例教学法。 【阅读文献】 http://www.nltk.org http://pypi.org/project/textmining/ http://pypi.org/project/lda11/ http://pypi.org/project/gensim/
第八周	3	1	2	0	0	0	第6章 文本分析 案例：基于政府工作报告的文本分析。 【教学方法】 讲授法、案例教学法。
第九周	3	2	1	0	0	0	第7章 情感分析 情感分析简介。情感分析与传播学研究。使用 Python 进行情感分析案例。 【教学方法】 讲授法、讨论法、案例教学法。

第十周	3	2	1	0	0	0	第7章 情感分析 情感分析的基本算法。语词的情感倾向判断。句子的情感分析。 案例：对抓取的某类新闻进行情感分析。 【教学方法】 讲授法、案例教学法。 【阅读文献】 http://sentiment.nrc.ca/lexicons-for-research/ http://ir.dlut.edu.cn/
第十一周	2	1	1	0	0	0	第8章 语义建模 语义建模与传播学研究。LDA 主题生成模型基本原理。语义模型的 Python 实现案例。 【教学方法】 讲授法、案例教学法。
总计	32	19	13	0	0	0	
备注（Notes）							