"Studying Digital News with Computational Methods"

Information about instructor:

Dr. Valerie Hase

Department of Media and Communication, LMU Munich

valerie.hase@ifkw.lmu.de

Information about seminar:

Part of the undergraduate program "Communication studies"

takes place weekly (Fridays, 10:00 am – 2:00 pm, Fall semester 2024/2025)

in the CIP-Pool (room 002), Akademistraße 7 (entrance to the left of the main building)

Link to course in Moodle

Link to R Tutorial

Summary of Seminar

In an increasingly fragmented digital information environment, audiences can retrieve online news across various platforms, such as news websites and social media platforms. Furthermore, information often spreads across these and is updated by the minute, something often discussed as digital news flow. In this seminar, we will learn how to use computational methods - in particular, how to scrape websites to track the flow of information online.

The seminar is held in English (planned as an in-person meeting, but we may switch to virtual if needed). Students will have to give presentations and write a paper on an empirical project they conducted for the final assessment.

Important: First, please be aware that this is a research seminar employing computational methods, meaning you will have to learn how to program with R. Previous knowledge of R is not a requirement, but you will have to be interested and willing to learn R via digital tutorials, which we will discuss in class. Second, please don't feel anxious about visiting a seminar in English. I am not a native speaker, and I assume most students won't be either. See this seminar as a chance to test and improve your English in a constructive environment - something that will prove useful for your future studies and/or the job market.

The seminar is held in English (planned as an in-person meeting, but we may switch to virtual if needed). Students will have to give presentations and write a paper on an empirical project they conducted for final assessment.

Learning Goals

- Students will learn basic knowledge of theories and empirical studies on news flow, especially in digital information environments.
- Students will be able to conduct their own analyses, using their knowledge of R, web scraping for data collection and manual content analysis for data analysis.
- Students will be able to apply this knowledge to independently conduct a research project on the flow of information in digital information environments. This means that they formulate research questions/hypotheses, develop a suitable operationalization, conduct descriptive analyses/statistical tests to answer these questions/test hypotheses, and critically interpret results.
- Students will be able to critically discuss current research on the topics of digital information flow, summarize its strengths and weaknesses, and evaluate it.
- Students will be able to engage in factual and competent discussions. They are able to communicate their acquired knowledge alone and in groups in a comprehensible and structured manner, both orally and in writing. They are also able to work in groups, resolve potential conflicts, and organize themselves.

Workload & Assessment

This is a 12 ECTS seminar which equals 360 hours of work (different regulations may apply for incoming students). The final grade is based on three different assessments:

- Assessment 1: Preregistration document (20%)
- Assessment 2: Task in R (20%)
- Assessment 3: A term paper (60%)

Please know that the time spent in in-person meetings in class is only a fraction of the actual workload needed to pass the class. The bulk of the required work needs to be done between sessions. The workload includes attendance in seminar sessions, preparing for these via readings, working through R tutorials and related tasks, handing in your preregistration and the graded task in R, collecting and analyzing data for your research project, and a final seminar paper.

Assessment 1: Preregistration document

The first graded assignment is a *preregistration document.* In short, it outlines which research questions and hypotheses you plan to test as well as the dependent variable you will focus on. You will register the preregistration via OSF and submit the resulting PDF file. For further instructions, see slides from session "S4 Research Design II". Preregistrations can be delivered individually or in groups with which you conducted your research projects. The preregistration makes up **20%** of the final mark. The preregistration document is due by Wednesday, November 20th, 2024 (deadline: 23:59 pm). Please upload your solutions via Moodle (see folder Session 5).

Assessment 2: Task in R

The second graded assignment is a *Task in R*. In the bookdown tutorial, you will find instructions for the "Graded task in R" at the very end of the tutorial. You can work on this task either in groups or individually. The Task in R makes up **20%** of the final mark. It is due by Wednesday, December 4th, 2024 (deadline: 23:59 pm). Please upload your solutions via Moodle (see folder Session 7). In the script, please specify the name of all students who worked on the corresponding task. You can also use the "Template Solution Script" for handing in your solution (voluntary), which provides some structure.

Assessment 3: Term Paper

Term papers can be delivered individually (35.000 characters) or in groups (45.000 characters for groups of two, 55.000 characters for groups of three, 65.000 characters for groups of four). This includes everything (i.e., from to introduction to references, including blank spaces). It makes up **60%** of the final mark and is due on February 21st, 2025 (via Moodle, see folder at the end of the course). The paper should be formatted according to the following guidelines:

- Times New Roman, 12 pt., double spacing; please use template for the title page (Moodle)
- Please cite according to APA7
- Please use the provided title page
- Please add the affirmation on independent work. If you hand in your work as a group, everyone has to sign the affirmation (this can be done on the same document or by each handing in a separate affirmation).
- Please make sure to include all necessary Appendices (e.g., codebook, code sheet, R code, R working environment).

You will find information on how these presentations will be graded in the document "*Grading Scheme_Term Paper*" on Moodle.

"Studying Digital News via Computational Methods" (Fall 2024/2025)

Date	Phase	Session	Tasks & Deadlines before each class
18.10.2024	**Conceptualization of Research Design**	Session 1: *Introduction*	
25.10.2024		Session 2: *Theory Hackathon*	Before class: Read Thorson & Wells (2016); Wallace (2018)
01.11.2024		No in-person meeting: Public holiday!	
08.11.2024		Session 3: *Research Design I*	Before class: Read texts for your research area
15.11.2024		Session 4: *Research Design II*	Before class: Upload screenshot of whiteboard; read Dienlin et al (2019)
22.11.2024	**Development of Research Methods & Data Collection**	Session 5: *Introduction to R I*	Before class: **Submit Graded Preregistration** (by Wednesday, 20th); Work through R Tutorials 1–2
29.11.2024		Session 6: *Introduction to R II*	Before class: Work through R Tutorials 3–5
06.12.2024		Session 7: *Introduction to Web Scraping*	Before class: **Submit Graded Task in R** (by Wednesday, Dec 4th); Work through R Tutorials 6–7
13.12.2024		Session 8: *Introduction to Web Scraping II*	
20.12.2024		Session 9: *Data Collection*	No in-person meeting: Group work on data collection with R
10.01.2025	**Data Analysis**	Session 10: *Data Analysis I*	Before class: Finalization of data collection
17.01.2025		Session 11: *Data Analysis II*	
24.01.2025		Session 12: *Data Analysis III*	Before class: Finalization of reliability tests
31.01.2025		Session 13: *Term paper Q&A*	
07.02.2025		Session 14: No in-person meeting: Individual office hours for questions on term papers. Please book office hour until Thursday, February 6th, via Moodle.	

References

Carter, M. C., Cingel, D. P., Ruiz, J. B., & Wartella, E. (2024). Social media use in the context of the Personal Social Media Ecosystem Framework. *Journal of Communication*, *73*(1), 25–37. https://doi.org/10.1093/joc/jqac038

Chadwick, A. (2017). *The hybrid media system: Politics and power* (2nd edition). Oxford University press.

Lacy, S., Miller, S., & Lovejoy, J. (2024). Improving the Creation of Social Science Theory in Journalism and Mass Communication Scholarship. *Journalism & Communication Monographs*, *25*(2), 117–133. https://doi.org/10.1177/15226379231167125

Harder, R. A., Sevenans, J., & Van Aelst, P. (2017). Intermedia Agenda Setting in the Social Media Age: How Traditional Players Dominate the News Agenda in Election Times. *The International Journal of Press/Politics*, *22*(3), 275–293. https://doi.org/10.1177/1940161217704969

Paulussen, S., & Harder, R. A. (2014). Social Media References in Newspapers: Facebook, Twitter and YouTube as sources in newspaper journalism. *Journalism Practice*, *8*(5), 542–551. https://doi.org/10.1080/17512786.2014.894327

Thorson, K., & Wells, C. (2016). Curated Flows: A Framework for Mapping Media Exposure in the Digital Age. *Communication Theory*, *26*(3), 309–328. https://doi.org/10.1111/comt.12087

Vonbun, R., Königslöw, K. K., & Schoenbach, K. (2015). Intermedia agenda-setting in a multimedia news environment. *Journalism*, *17*(8), 1054–1073. https://doi.org/10.1177/1464884915595475

Von Nordheim, G., Boczek, K., & Koppers, L. (2018). Sourcing the Sources: An analysis of the use of Twitter and Facebook as a journalistic source over 10 years in *The New York Times* , *The Guardian* , and *Süddeutsche Zeitung*. *Digital Journalism*, *6*(7), 807–828. https://doi.org/10.1080/21670811.2018.1490658

Wallace, J. (2018). Modelling Contemporary Gatekeeping: The rise of individuals, algorithms and platforms in digital news dissemination. *Digital Journalism*, *6*(3), 274–293. https://doi.org/10.1080/21670811.2017.1343648

Wendelin, M., Engelmann, I., & Neubarth, J. (2017). User Rankings and Journalistic News Selection: Comparing news values and topics. *Journalism Studies*, *18*(2), 135–153. https://doi.org/10.1080/1461670X.2015.1040892

Zamith, R. (2018). On Metrics-Driven Homepages: Assessing the relationship between popularity and prominence. *Journalism Studies*, *19*(8), 1116–1137. https://doi.org/10.1080/1461670X.2016.1262215