

CDMM White Paper

Teaching Digital Methods Experiences on the BA Level in Media Studies at Aarhus University

Christoph Raetzsch & Niels Brügger

reviewed by Janne Nielsen

Department of Media and Journalism Studies, Aarhus University, Denmark

Introduction

The growing popular and academic interest in digital methods for social science and humanities (SSH) research requires adapting humanities education to new skill demands and topic areas. In this white paper, we would like to share our experiences of developing and running an introduction to digital methods in a BA Media Studies programme over five years (2022-2026). Such academic 'background work' of developing curricula and didactic approaches are rarely present in academic discussions but forms the basis of future development in the social sciences and humanities disciplines. The experience of introducing humanities students to rather technical skills and competences comes with a range of difficulties that teachers need to be prepared for. We hope this contribution informs a wider debate about the conjunction and (future) requirements of teaching technical skills in the humanities without giving up core tenets that are constitutive of such research traditions.

Digital methods are a case in point here: We have seen a significant diversification and vast new range of methods proliferating within only the last ten years (Hine et al. 2016; Madsen and Munk 2026). New domains of research have been established and we need to develop new approaches to teaching digital methods as well. What originally emerged as an experimental approach to studying communication on the Web, the Internet, on social media or in online communities has become fairly institutionalised in new sophisticated approaches that broadly fall under categories such as the "Computational

Analysis of Communication" (van Atteveldt et al. 2022), "Computational Communication Science" (Haim 2023), or digital humanities (Brown 2020), web (archive) studies (Aasman et al. 2025), and social media research more generally (Sloan and Quan-Haase 2018).

Within this broader turn to studying the digital as constitutive of social communication practices and infrastructures, digital methods were initially defined by Richard Rogers as an approach to 'repurposing' the "methods of the medium, or the methods that are in some sense built into the web" for social science research (Rogers and Lewthwaite 2019: 14; cf. Rogers 2013). In this respect, digital methods were developed based on the new possibilities that technically standardised digital communication environments offered for data collection, analysis and investigations of social phenomena. Instead of porting existing methods to a new (digital) context, e.g. by doing a survey online, a hallmark of the digital methods approach is to develop methods through data and media exploration, generating new questions about the materiality and meaning of communication in highly technical environments, alongside the development of new kinds of research tools (Rogers 2014). By using large amounts of data that reflect and (to a certain extent) represent entire populations of users, digital methods embed the possibility to overcome the qualitative-quantitative divide of social science disciplines that has emerged over the course of the 20th century. Because data of social interactions is now ample, yet often restricted by the conditions of digital media platforms and their governance and commercial models, digital methods allow to "unfold the origami of collective existence" with 'thick descriptions' of vast populations and their interactions in digital media (Venturini, Jensen and Latour 2015: n.p.).

It is important to underline that the study of digital data and the development of research sensitivities are co-constitutive for the digital methods research approach. Data analysis techniques may be borrowed from other disciplines, yet the purpose of digital methods is to both understand digital objects in their social and cultural dimension and in their digital materiality.

"Digital methods are techniques for the study of societal change and cultural condition with online data. They make use of available digital objects such as hyperlink, tag, timestamp, like, share and retweet [and] endeavour to repurpose the online methods and services with a social research outlook." (Richard Rogers, *Doing Digital Methods*, 2019: 3)

Digital methods place an emphasis on understanding the circulation of digital objects in their role as 'epistemic objects' that have different meanings in relation to the communities that create and circulate them (Raetzsch and Bødker 2016). They target media environments and infrastructures that make this circulation possible in the first place (Plantin and Punathambekar 2018).

Teaching such a dual perspective on cultural objects and their technical environment is not an easy task. Teaching methods that presuppose a minimum of data literacy and certain technical skills goes against what many humanities students expect to achieve and master. Methods education in this domain also often relies on case studies that are published in international academic journals. Such studies document very specific conditions in terms of their technical setup, data collection possibilities or competence level of the research team but are difficult to recreate for a teaching environment and cannot be readily adopted by students. We recognise some notable exceptions for actionable and insightful digital methods pedagogies, e.g. the recent introduction for students by Venturini and Rogers (2025) or the first and second edition of "The Data Journalism Handbook" (Bounegru et al. 2012; Bounegru and Gray 2025). Such works exemplify the need for concurrent and continuous adjustment of methods to the platforms, media and digital contexts under investigation, which embody the digital methods approach.

From a didactic perspective, we need to acknowledge that competences in digital methods are only acquired over time with a propensity for technical abstraction. From a curricular perspective, students should develop technical competences in relation to humanistic research designs. At the same time, many students feel challenged to translate individual research interests into sound research designs. Demonstrating digital analytic skills in relation to data and methods requires a high degree of abstraction from the empirical research object, e.g. a specific online discourse, an online community, or a meme traversing platforms. There are to date very few contributions that outline and demonstrate how digital methods are taught in a regular university curriculum. We assume that syllabi and teaching materials are often not deemed relevant for publication to a wider community, and remain mostly functional documents within a course context. They often have little relevance outside the respective institutional settings where courses take place. At the same time, scholarly communication in academic journals tends to prefer the ground-breaking new approaches, the methodological innovation, and new data skills in a short-term perspective while sidelining the infrastructural and institutional work required

to implement new didactic approaches for new topic areas in media and communication studies or digital humanities more broadly.

Our white paper seeks to remedy this gap in scholarly communication.

How to teach Digital Methods?

In this white paper, we would like to present our approach to teaching a basic introduction to digital methods for Bachelor students in the Media Studies programme at Aarhus University. The course has run for five years now (2022-2026) and has been continuously developed from year to year based on feedback from students (see section 'What Students are Saying') and the instructors didactic approach. Despite the annual modest innovation, our basic didactic concept has remained fairly stable. We start from the students' interest in media and data exploration to gradually refine their skills in handling and analysing digital data. The course takes students through different types of media, methods and technical skills to understand mediated communication through the "natively digital objects" and underlying "methods of the medium" that are characteristic for digital methods research (Rogers 2009: 29; Rogers 2013). Along the way, the course embeds an applied approach with a strong potential to teach transferable skills in working with digital data that are relevant beyond media studies (See Appendix 1 for the 2026 course plan).

We usually start with the students' interests in media, present the current academic research informing analyses of media and data, and offer practical training in working with digital data in an analytic sense. This combination of skills, reflection and training has been consistently appreciated by students over the years. In the beginning, we encountered difficulties in relating the academic demands and purposes of the course to students. Yet, the continuous development of the course's approach has been positively reflected in student evaluations and feedback since then, signalling that there is a demanding but rewarding learning curve for many students.

When students encounter the course during the second semester, many are initially reserved, probably cautiously curious about the possibilities. Most of them are quite competent at the end of the course to implement basic data collection and analysis routines, argue critically about the benefits and shortcomings of their approaches, and discuss possible alternatives if their study did not work out as expected. We do acknowledge that some students at the end of the course and after passing the exam conclude that digital methods are not for them and they honestly would prefer not to work with

them again. That is also a part of their learning. Yet, we would like to emphasise that such a conclusion is based on students actually working with data and analysis routines, with methods and software as well as their own critical reflections in relation to a subject they find (or found) interesting. Even if they never use digital methods again, they have learned how such approaches work, and they can assess the usefulness of digital methods for specific research interests or challenges in future work environments.

Every year, we also see students, who approached the course with reservations and no significant background knowledge, who excel in surprising proficiency of employing or even developing methods, creating insightful and mature projects with relevance, scope and sophistication. This experience has encouraged us to summarise our didactic approach in this white paper. Our goal is simply to share best practices to inform pedagogical development in social sciences and humanities for the adoption and critical evaluation of digital methods. The topics of this course can shift from year to year, but the overall integration of elements makes the didactics of teaching digital methods robust to ensure that a majority of students without technical backgrounds are enabled in their development of digital literacy and practical analytic skills.

In the following, we present the purpose, structure and organisation of the course, followed by an explanation of our didactic approach. We reflect the core stages of the digital methods research process across data collection, cleaning, analysis and communication. We also include exemplary student feedback from the last five years. Our main contribution is a list of insights in respect to managing the learning experience, foregrounding learners' interests in topics and methods, fostering a knowledge culture, and training in specific, practical as well as generalisable skills for data-based research and analysis. The graphs we include in this paper are also used in our communication with students and summarise our key didactic considerations.

Purpose, Structure and Organisation of the Course

The course was developed in 2021, following a reform of the BA study regulations for Media Studies at Aarhus University which placed more emphasis on digital analytic skills in relation to subject domains of media studies like social media, public broadcasting, online discourses and communities. The core purpose of the course is to "enable students to understand, explain, evaluate and apply digitally supported methods within media studies". Another important aspect is that students should be able to

“account for some of the opportunities and challenges that characterise data collection, data cleaning, analysis and dissemination as well as law, research ethics, and data management in a digital media environment” (our translation from Danish study regulations). The course started in 2022 and has run for five years with a structure that has only been slightly modified over the years following student feedback. Usually the course is attended by around 100 BA students in their second semester in Media Studies. Both authors of this white paper have been the main teachers in this course over the whole period, and have been assisted by changing student assistants (usually six), recruited from previous cohorts and the department.

The course runs for 13 weeks in the Spring term and is concluded with a short written synopsis on the students’ own research project (written in groups or individually) and an oral exam. The teaching sessions include lectures, seminars with the teachers and practical training sessions with student instructors. The sessions are typically spread over the whole week and each teacher covers 3-4 weeks of lectures and seminars in a row as well as the coordination of their team of three student instructors. This division allows students to become familiar with the teacher’s content, style of teaching, and specific topic area. Quite importantly, it also allows the teachers to specialise in their respective areas of expertise and update content and exercises from year to year.

Alongside the lectures, which scope a particular analytic approach, method or theoretical concept, the teachers demonstrate practical approaches to implementing certain types of analysis during the seminar sessions. Students first get a demonstration of a particular analytic procedure and can then try it out for themselves. We also created video tutorials for certain introductory data cleaning and conversion routines, which allow the students to learn at their own pace and relieve the teacher of having to attend to 15 different groups at the same time. The seminars can pass quite quickly, and students will face obstacles in implementing a data analysis routine immediately, since there are a lot of details to be observed to make these routines work. Occasionally, the students’ own computers, system settings or connectivity problems can cause problems. During the last teaching session of the week, students have hands-on exercises with student instructors, who offer repetition and bring in their own perspectives and experiences. These tutorial sessions are usually designed in close coordination with the respective lecturer in each week. The weekly structure of lecture (scoping), seminar (applying) and student instructor sessions (practicing) supports students to learn basic procedures and use cases

over the week, so that they are able to apply them with some proficiency by the end of a week and develop them beyond the teaching itself.

Since every cohort of students faces slightly different problems from year to year, it is essential to have continuous exchange between teachers and students instructors (ie. TA's) and to adjust the course content as the semester progresses. The students have to work in groups on two separate case studies, which are presented in class. These presentations serve as a preparation for the exam, where each teacher is accompanied by a second examiner recruited from colleagues at the department, who are familiar with digital methods against different disciplinary backgrounds. The presentation of the case studies in class is accompanied by feedback from the teachers and fellow students. The students' assessment of the course at the end is discussed again in class during the last lecture, which allows teachers and students to exchange experiences to inform the next year's design of the course.

Didactic Approach

Most of the students in this course have no prior experience with digital methods, data collection or computational analyses of communication. Students are interested in media as organisations, in audiences, publics and above all, in media content and its dedicated modes of delivery (mobile, television, print, podcast, stream, website). Our approach addresses media and content that students interact with on a daily basis in relation to its underlying technical infrastructures, digital encodings of content or metadata that make this content available to them as end users and audiences. We deliberately refrain from making the technical or computational part of the course too prominent at the start, seeking to create interest in digital methods through very approachable off-the-shelf, open source software and browser extensions, which put a very low threshold on students' engagement with data.

We foreground the *experience of approachable data collection routines* as something doable, and then gradually build up technical skills through such interactions with data. It is usually from the moment that students "hit a button" to 'collect' data that many become seriously interested in the possibilities of digital methods. This low threshold of engaging with data should not be underestimated for fostering students' dedication to the objectives of the course and their own willingness to invest in their own learning. In contrast to typical data science approaches, which start with basic programming skills in languages such as R or Python, we start with the *experience of media* reflected

in an exploration of data and processing routines which students can relate to. Starting with data means that we encourage students to develop their interests by acquiring new technical skills that are productively applicable from the start.

In the beginning of the course, we introduce existing collections of data like data repositories, or web and media archives, as possible starting points for research. Then we move on to develop approaches to data collection that reflect students' own interests e.g. in streaming content, online communities or social media. It is usually enough to generate interest in digital methods by slightly defamiliarising the 'seamless' interface of a web browser that enables the consumer experience of media. One way to do this is to disassemble a website into its constitutive parts (scripts, data types, graphic elements) and understand how they work together to enable the experience in the first place. When students see how a browser fetches each element from a server to build the HTML website (e.g. by showing the *Network* feature in the browser's developer tools), they realise how the structure of a website is standardised and how certain elements are processed when they interact with them. Another way is to split up a website into types of data (image, link, text, interactive features) and target each type as a potential source of data for research.

On the one hand, we show students a particular way of studying different types of media content (images, video, written texts such as news articles, social media posts and comments), their encoding, their relevant metadata and delivery mechanisms. On the other hand we introduce basic methodologies for structuring, cleaning, systematising digital data in an analysis process without making the data science aspect too dominant in the teaching. We introduce students to some basic routines of working with, understanding and using digital data, which they can further develop and apply in other contexts. Our usual approach relies on standardised exports of data from websites, tools or social media. Comma-separated value files (csv) are here the most flexible and universal file format for data exchange. Such files can contain a lot of data while being easy to handle from a computational point of view. As text files they can be imported to a range of different tools and programmes. The benefit of csv files is that they rely on the distinction between categories and content, or types of data and the actual data associated with a distinct category such as user names, date-time, or links. Imported to a standard spreadsheet programme, they can teach students the difference between data and metadata quite easily, and working with such structured data becomes more of an exploration of available data in rows and columns than a theoretically-driven

quest for quick results. Applying formulas to a whole dataset, rows or whole columns, shows students the powerful effect of transforming data based on standardised scripts and formats. Once students realise the benefit of such structured data, we introduce further processing routines that rely on uniform inputs such as seed link collections or uniform text inputs for e.g. sentiment analysis or topic modeling for further exploration. The size of data sets we prefer is usually between a few hundred to a few thousand data points (e.g. comments, links or social media posts). The size should be too large to analyse by, for example, close reading, yet large enough to justify semi-automated routines for data processing. Data in our context then becomes the object through which we enable new ways of thinking through issues of broader social, political or cultural concern. Yet, the data that students collect remain attached to the contexts where they are generated, a platform, a forum, or a website, allowing them to approach media theory in digital research practice with new digital competences as they explore a subject they find interesting.

The course relies on students having their own computer at hand to perform dedicated data collection and analysis tasks. This condition is beneficial for students since they know (to a certain extent) how their computer is organised and operated. Once the course starts, we ask students to install a limited amount of open-source software for specific analytic tasks but initially start with software they already know such as a web browser or the Microsoft M365 and Apple Office packages. This approach has the benefit that students are not confronted with software environments and interfaces, which are completely new and unknown (such as programming environments or Jupyter notebooks), lowering the threshold to engage with simple data handling routines significantly. We also observe that most students use Google Chrome or Safari for their private browsing and that open source alternatives like Firefox are largely unknown. This makes it somewhat easier to suggest Firefox as their dedicated research browser with a separate researcher profile, sensitising them to the differences between private online communication and data collection for research. In **Figure 1** we have summarised the levels of technical and analytic competences developed over this course in relation to different software(s) and digital research environments.

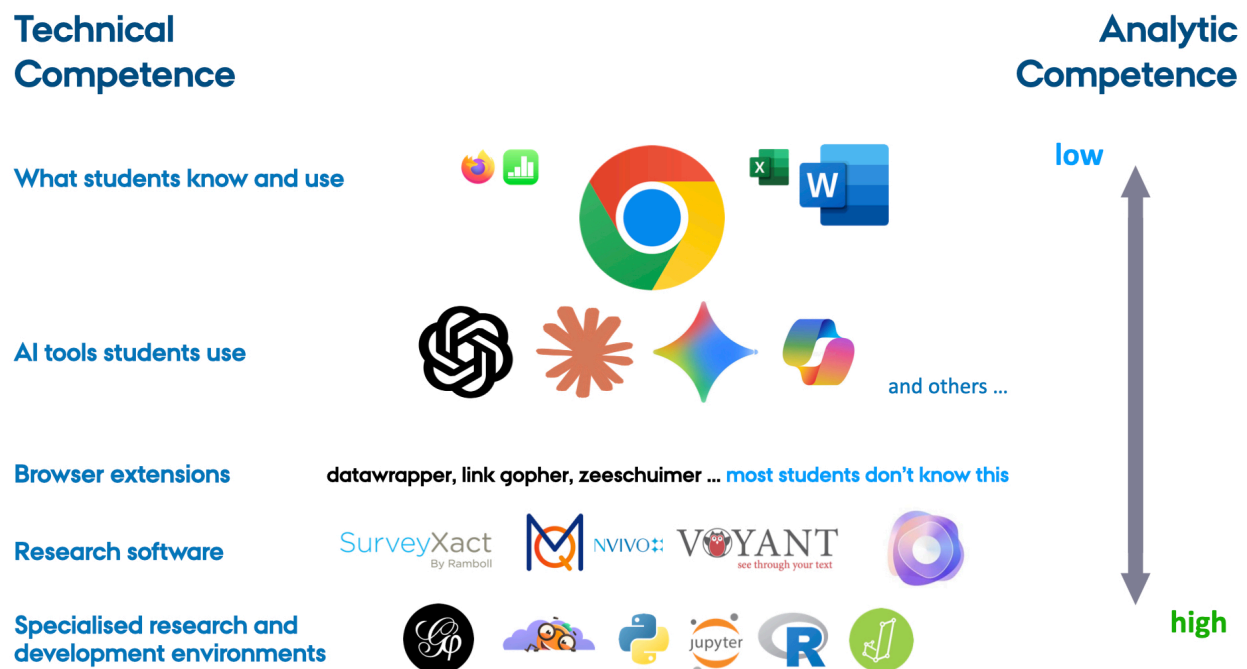


Figure 1: Overview of software that students know with some technical competence (top) requiring low analytic competence contrasted with specialised research software (bottom) requiring higher analytic competence. (Graph: authors' own)

Using software that students already know has the advantage that we can gradually build up new competences without disrupting the familiarity with interfaces and environments. It is significantly easier to teach some basic data transformation routines in Excel (e.g. by using formulas and generic functions) than to teach Python scripts for the same purpose in a completely new (and less intuitive) environment. The same can be said for browser extensions: a web browser is a known and familiar software, and just adding an extension or plugin lets students experience the same software in a new way and for a new purpose. Again, the same routines could be taught based on Python or R, yet didactically the introduction of extensions that have graphic interfaces feels a lot more doable and intuitive for students.

We only require students to install these research softwares:

Orange Data Mining (<https://orangedatamining.com/>) is based on Python and is a visual programming, open-source environment developed for the purpose of teaching data science. It is organised around the logic of workflows for data analysis routines which are built from widgets, small programmes which are connected to enable replicable data analysis and transformation routines. Widgets are organised into libraries for specific application fields, such as mapping, text analysis, text mining or image

analysis. In contrast to Excel, Orange does not affect the original data set and allows to export subsets as new files. The graphic representation of the processing chain also allows to understand what happens to data in each stage and control these stages through experimenting with different settings. Experiments with data can fail without impact on the original data collection, which is reassuring for students.

The Orange environment is very didactic because the tool reveals how data is exchanged between widgets. When data is transformed or processed, students understand which part of their data is used. Contextual error messages help to understand if and why a particular workflow is not possible or data is not correctly processed. We provide a basic workspace for students with several templates for typical analysis approaches to prepare the sessions of the course (see Figure 2). Some widgets are already 'wired' to allow immediate basic analyses; others are placed as optional to experiment with and understand differences between processing routines.

Typical Stages of Orange Workflow Sections (Course Template)

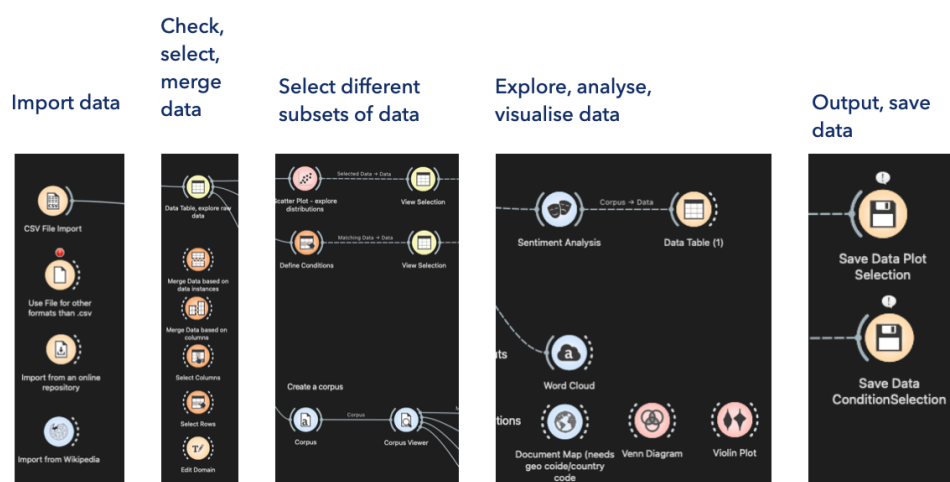


Figure 2: Overview of stages in our Orange workflow template. Widgets linked with a line are set to process imported data automatically or step by step. Widgets not currently connected are optional for each stage of a workflow. See Appendix 3 for our full Orange workspace with several templates for different research approaches (using the workflow requires Orange Data Mining Version 3.40.)

These workflows embed the same logic like a Python script: they let students understand what happens at each stage of a workflow to 'their data' and thereby gradually make them aware of how the conditions they set for a computational routine have consequences for the outcome of an analysis. Additional Python scripts can be included in workflows (e.g. by AI-generated inputs) through a dedicated widget when needed in an analysis.

- **Gephi** (<https://gephi.org/>) is the state of the art, open source network analysis tool widely used in academia and beyond. Its interface and design likewise encourage exploration, critical reflexivity and technical competence in the creation of network graphs and analyses. As van Geenen et al. argue, network analyses and network graphs can be pedagogical tools that shape crucial critical sensitivity to understanding “how networks become meaningful and are invested with particular kinds of analytical, interpretive and narrative capacities” (van Geenen et al. 2022: 11). When using Gephi, most students initially encounter the ‘hairball’ after they import node-edge tables, a meaningless twirl of connections between data points. Once they understand the different elements and layers, they also see how weights, data preprocessing, and visualisation settings affect the usefulness and appropriateness of a network graph to the issue they are researching.

Both Orange and Gephi are introduced and applied in the context of the course for 3-4 weeks in a row, allowing to build technical competences from week to week and showing the versatility of the respective software for different kinds of research scenarios. Aside from these standalone environments, we mostly use open source tools for data collection offered by the Amsterdam Digital Methods Initiative (<https://wiki.digitalmethods.net/Dmi/ToolDatabase>), and the tool collections and tutorials developed in-house by the Centre for Digital Methods and Media at Aarhus University (<https://cc.au.dk/en/cdmm/tools-and-tutorials>).¹

The data collection process is a special challenge for teachers and students alike. From year to year we review open-source applications that currently work without cumbersome applications for developer accounts from platforms directly. Platforms often change their access policies or terms of service (as happened to Twitter, now X), or tools can become outdated or replaced by others. The challenge of teaching in a dynamic media environment requires us to be flexible without overhauling the entire curriculum from year to year.

In our didactic approach, we distinguish a) between underlying methodologies of investigating social relations, content or users, and b) between digital methods and tools developed for these research purposes. **Figure 3** shows how the didactic approach and the practical implementation of the course are linked in the course structure and content.

¹ The Centre was founded by Niels Brügger in 2020, and Christoph Raetzsch currently serves as its Director.

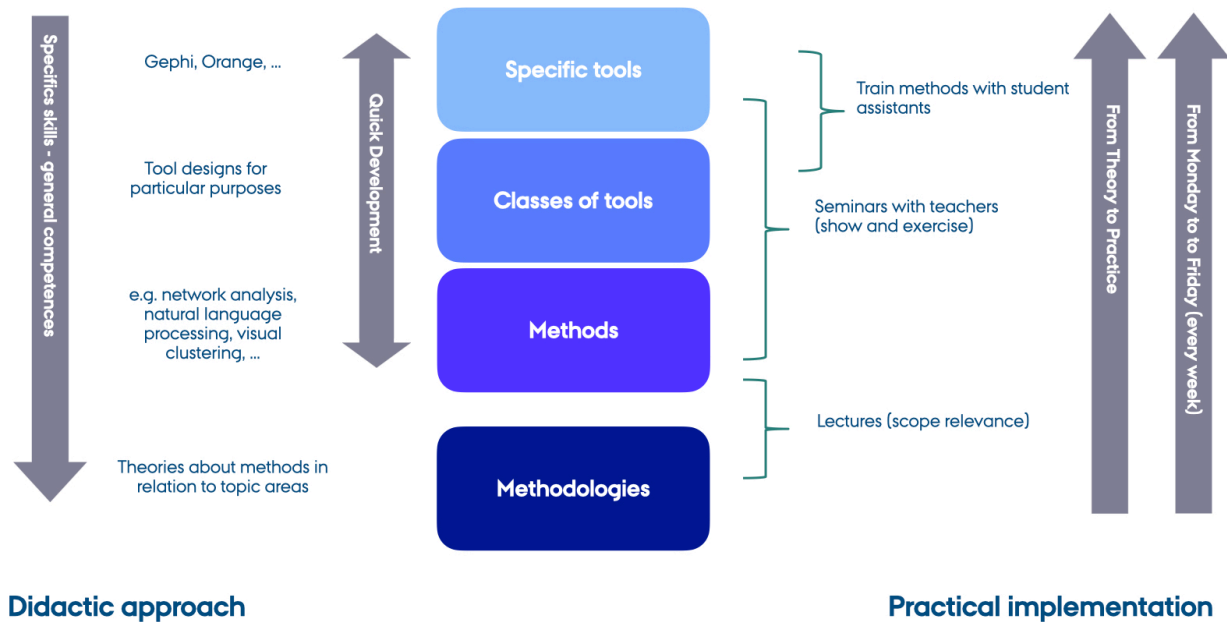


Figure 3: Overview of the didactic approach and its relation to the practical implementation. Tools can change quickly, but underlying methodologies of digital methods remain rather stable. Specific skills that students learn in relation to a software or routine need to be grounded in general competences and underlying methodologies. The weekly structure of the course reflects this move from theory to practice, teaching specific skills as well as helping to develop general competences through exercises and case studies. (Graph: authors' own)

Acknowledging that tools can change quickly, we emphasise that there are always classes of tools for similar purposes, which can complement or even replace each other. On the level of methods, we make students aware that tools also embed certain methods of handling data in their technical designs, yet remain rather stable in their underlying methodologies: Network analyses can be done entirely without software; visual or linguistic analyses target particular attributes of cultural objects and can likewise be done without any computation. It is important to stress these differences between tools, methods and methodologies so that the specific skills that students acquire with a given routine or software are not mistaken as equivalents of competences in an analytic perspective, ie. the methodology. Through the structure of the course, the didactic approach is mirrored in the weekly sequence of teaching sessions:

from lectures on theories and current research to seminars for practicing new methods, and training sessions with student instructors to develop skills for long-term knowledge.

The Digital Methods Research Process

In contrast to traditional research designs in social sciences and humanities, we encourage the exploration of topics that students are already interested in from their daily lived (media) experiences. A literature review can be helpful, but for many new or emerging topics that students are interested in, academic literature will not be the most useful starting point.

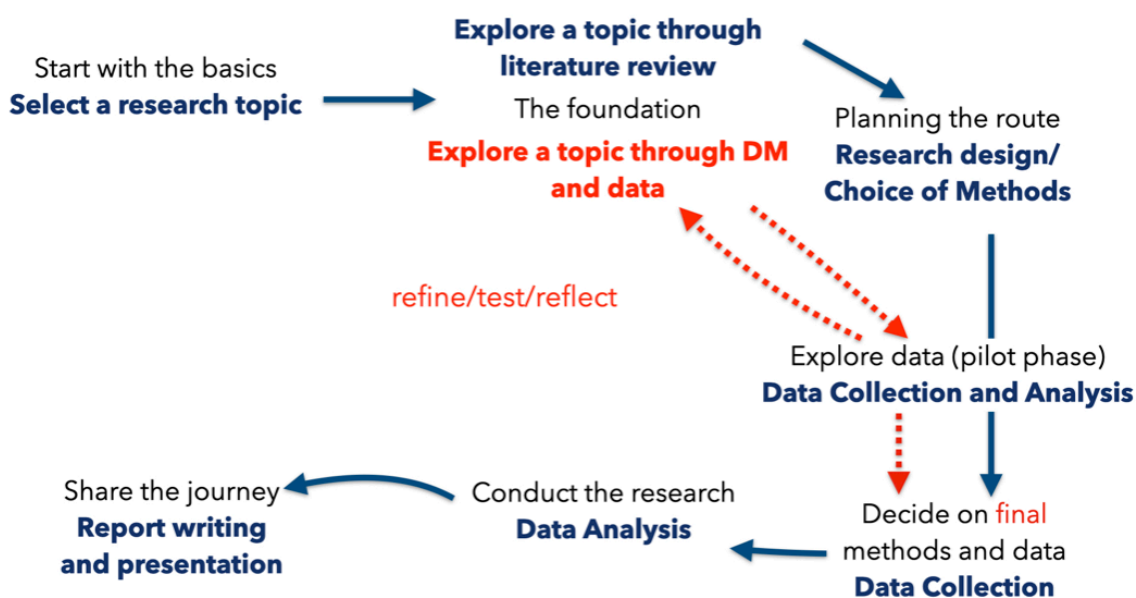


Figure 4: Overview of a (prototypical) research process with digital methods. The **blue elements** reflect typical steps in a traditional research design, where theory and existing knowledge inform a new methodological approach or choice of data collection before the analysis starts and the findings are communicated. The **red elements** foreground how a digital methods research design will typically include exploration of a topic through (accessible) data and tools (ie. the pilot phase), often together with learning new digital skills. This is where exploring the ‘methods of the medium’ is most crucial for the final research design. (Graph: authors’ own with contributions from Cecilia Arregui Olivera).

In a digital methods research process, scoping available data sources (platforms, channels, user profiles, web spaces) is often the first part of an exploration: Where does this phenomenon happen which I want to study? Who are the principal actors? What types of discourses can I find? These are questions that students can start working with easily and the course structure provides ways in which such topics can be explored. Using tools and data are

an essential part of a piloting phase before a research design and research question are set up. This approach requires students to invest in their own skill development to develop research designs that rely on digital methods proper.

This way of approaching the research process may in some ways run counter to social science handbook recommendations for research designs. In our experience, though, the formulation of a relevant research question that is feasible to answer within the students' competences and the course's time frame requires to become familiar with different data collection and analysis routines *before* a suitable research design is found and implemented. The research process is thus broken up into different stages:

1. the **data collection** itself, relying on existing collections or developing a routine to combine different tools for different purposes
2. the effort required for **data cleaning**, or pre-processing, available data, for the purpose of the analysis. This stage will involve some data structuring routines and reflected choices of which data to include in an analysis.
3. conducting the **data analysis** for the chosen topic and data sources often involving new combinations of methods (in pilot and main research phase)
4. **communicating** the research outcome, including critical reflections of the data collection, tool criticism, the research design and visualisations.

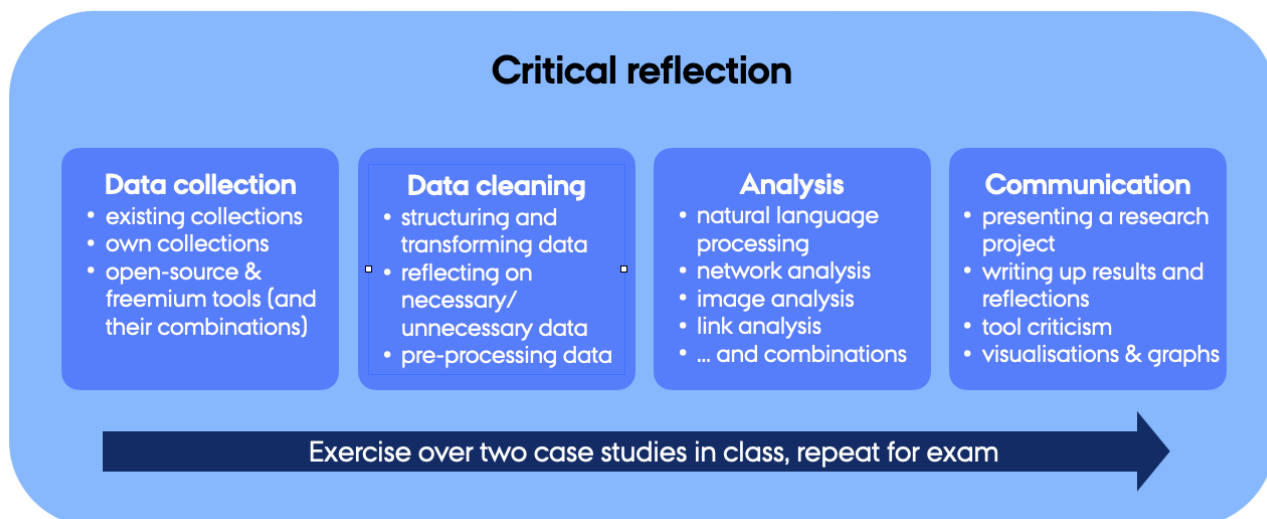


Figure 5: The four stages of a DM research process - from data collection to cleaning/preprocessing data to analysis stages and the communication of results. Students conduct two case studies over the course of the semester with different types of data and analysis methods before they repeat the process for their exam project, written up in a synopsis and then examined in an oral presentation and discussion. The structure of the course prepares for the oral exam, which is very similar to the case study work and presentation in class. (Graph: authors' own)

We emphasise that each of these stages needs to be informed by critical reflection. We encourage students to document their experiences and reflections in a personal logbook that they can keep for themselves. Over the years, we have observed that student groups often start a joint logbook, to capture their learnings and share insights. Based on the study regulations, students need to demonstrate critical reflexivity across every stage of the research process. A core approach for us is to introduce 'tool criticism' across all stages, a very useful and applicable concept developed by van Es, Schäfer and Wieringa (2021). Tool criticism is defined by the authors as

"the critical inquiry into digital tools and how they are used for various purposes within the research process. It reviews the qualities of the tool in light of, for instance, research activities, and it reflects on how the tool (e.g., its data source, working mechanisms, anticipated use, interface, and embedded assumptions) affects the user, the research process and output, and its reliance on the user's training." (van Es et al. 2021: 52)

Tools can embed routines that are opaque to their users or can have settings which affect the quality and scope of a data collection. Blackboxing such routines can pose a significant problem for assessing the quality and outcome of an analysis. Yet, tools such as Gephi can also inspire new perspectives on a given data set, encouraging exploration and challenging preconceived knowledge. Our emphasis on tool criticism support students in their critical perspectives as well as justifying the eventual outcome of an experimental approach. Tool criticism is both about tool use itself and a given tool's usefulness for a specific research purpose. It helps students to interrogate the outcome of their analysis in self-critical terms: Does a given analysis reveal new knowledge or just confirm previous conceptions? Is the analysed data sufficient for the claims made in the results? Does the tool (mis-)represent the research object? Are we using the right data? Are the visualisations and graphs convincing or misleading? Tool criticism across the four stages is essential for moving from the specific tool skills to the more general competences in digital methods. It can be frustrating, admittedly, to be critical of one's own work and be reminded of that throughout the process. Especially when an analysis required extensive data collection and pre-processing routines, students understandably want to defend the outcome of their analysis. But it is important to develop a sensitivity to the limitations and inherent design choices of tools to justify the outcome of an analysis or research design.

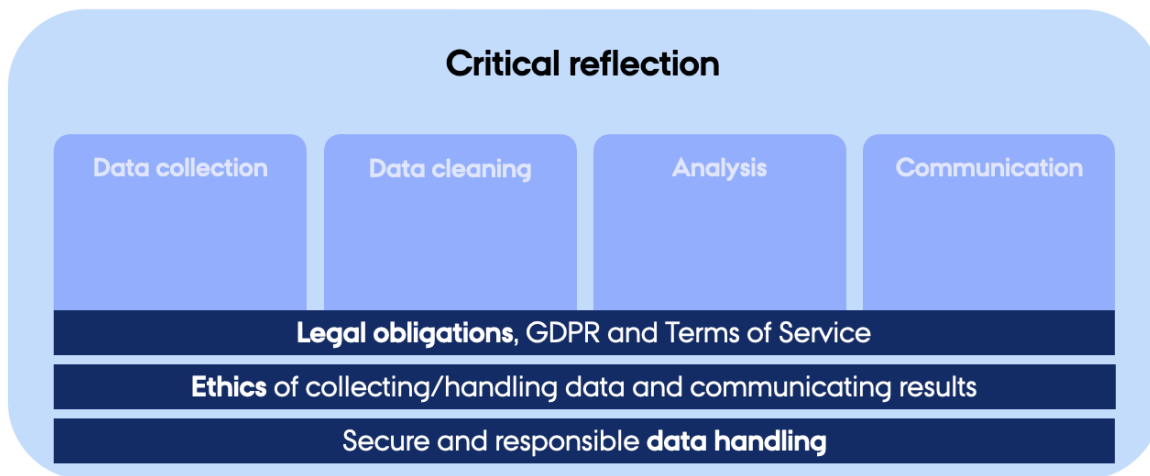


Figure 6: The four stages of the research process in relation to the core considerations of law, ethics, and secure and responsible data handling.

A second element of critical reflection are the dimensions of legal obligations (e.g. GDPR or platforms' Terms of Service), ethical considerations of collecting and handling data, communicating results, and best practices for secure and responsible data handling. These dimensions stretch across all stages of the research process – with different emphases across the semester. They are important to consider across the DM process to build up general data and digital methods competences that remain relevant beyond the course.

In conclusion, the didactic approach centers on translating research topics and approaches in digital methods to the research ideas and projects students have in mind based on their own experience of media. The approach relies on the introduction of basic but useful data collection and processing routines to enable students in defining research projects, where digital methods are central. Across the four stages of the research process, we encourage tool criticism and critical self-reflection in relation to case studies which students conduct in groups. By using largely familiar software environments at the start, we avoid an overly technical conceptualisation of the course and foreground the quick learning outcomes in terms of skills, which motivate students to learn and acquire more generalisable competences in the long run.

What Students are Saying

We reviewed our course evaluations from the last five years and tried to find a representative selection of positive and critical comments from each year. Throughout the years, students have especially appreciated sessions with student instructors and exercises where they could get their 'hands on data' to experiment with possible research designs. A lot of students have also invested their time in trying to manually code large corpora of data. *We would like to thank all of these students for their dedication and honest interest in making digital methods work.* Here are few original comments by students in course evaluations over the last five years (translated from Danish by authors).

2022

- ➡ "Det tog mig godt og vel 4-5 ugers tid at forstå kurset og dets indhold. Det var først halvvejs inde i kurset, jeg virkelig forstod Digitale Metoder, og kunne koble emnerne sammen." [It took me about 4-5 weeks to understand the course and its content. It wasn't until halfway through the course that I really understood Digital Methods and was able to connect the topics.]
- ➡ "Synes det er gået for stærkt når vi har gennemgået forskellige programmer, og jeg synes vi er gået for hurtigt igennem de forskellige programmer. Måske har der været for mange programmer vi skulle lære." [I think it went too fast when we went through different programs, and I think we went through the different programs too quickly. Maybe there were too many programs we had to learn.]
- ➡ "Jeg har lært de forskellige programmer, og dem har jeg styr på - men jeg ved ikke, hvordan jeg skal bruge dem til en problemstilling eller til den slags opgave, som eksamensformen er." [I have learned the different programs, and I have them under control - but I don't know how to use them for a problem or for the kind of assignment that the exam format is.]

2023

- ➡ "Instruktør har været godt til at øve sig i programmerne. Men holdundervisningen, hvor man arbejder i studiegrupper hvor undervisere går rundt mellem grupperne, fungerer også rigtig godt. Massere af god øvelse." [The instructor [sessions] have been good for practicing the programs. But the group teaching, where you work in study groups with teachers moving between the groups, also works really well. Lots of good exercises.]
- ➡ "Det kunne være fint at få et overblik over hvordan de forskellige metoder og programmer kan arbejde sammen i en eksempelvis eksamen." [It would be nice to get an overview of how the different methods and programs can work together in an example exam.]

2024

- ➡ "Det har fungeret rigtig godt, når Christoph og Niels har gennemgået et workflow på tavlen og vi så har skulle gøre det på vores egne computere imens!" [It has worked really well when Christoph and Niels have gone

through a workflow on the board while we have had to follow along on our own computers!”]

- ➡ “Føler ofte læsningen har været lidt irrelevant fordi det først giver mening når man arbejder med det, men øvelserne har været brugbare.” [I often feel that the reading has been a bit irrelevant because it only makes sense when you work with it, but the exercises have been useful.]

2025*

- ➡ “I starten var det så udfordrende for mig, at jeg ikke syntes det var spændende. Men efterhånden som jeg blev mere fortrolig med min computer, blev det meget interessant.” [At first, it was so challenging for me that I didn't find it exciting. But as I became more familiar with my computer, it became very interesting.]
- ➡ “Jeg har fået meget ud af når vi har haft samtaler om, at der ikke findes én rigtig måde at gøre ting på. Det er også dejligt når undervisningen er meget i øjenhøjde.” [I have learned a lot from our conversations about the fact that there is no one right way to do things. It is also nice when the teaching is very much at eye level.]
- ➡ “Instruktor har været meget givende, ift. at lave nogle lidt anderledes øvelser, og samtidig har der været tid og plads til at spørge og undre sig undervejs.” [The [student instructors] have been very rewarding in terms of doing some slightly different exercises, and at the same time there has been time and space to ask questions and wonder along the way.]

2026

- ➡ “Det er i dette fag, hvor eventuelle jobmuligheder faktisk åbner sig.” [It's in this course that possible job opportunities actually open up.]
- ➡ “Instruktortimer og konkrete opgaver i holdundervisning hjalp med at få de digitale værktøj ind under huden. Caseopgaverne har også været god eksamensforberedelse.” [Sessions with student instructors and concrete assignments in class helped to get the digital tools under my skin. The case studies have also been [a] good preparation for the exam.]

*In 2025, the course was evaluated independently on the level of the three seminar groups, while in all other years, it was evaluated by the whole cohort in one survey. We include one statement from each seminar group for 2025.

Based on the feedback of students we have changed elements in the curriculum from year to year. In the beginning, we introduced a lot of different tools which was confusing for students. The solution was to use **Orange** as a main data analysis environment and collect student-developed workflows in a dedicated archive where students can get inspiration from previous cohorts. We also introduced the first case study based on Orange as an additional presentation opportunity and research project. Two case studies before the actual exam give students a stronger sense of having space for experimentation and explore methods without the pressure of having to achieve a grade. We also now share examples of exams and actively coach

students on how they can expand from their case studies to the final exam. For data collection, we have relied on **Youtube Data Tools** (<https://ytdt.digitalmethods.net/>) because the whole suite is versatile and Youtube still allows data access. Data from Youtube is perfectly sufficient for our exercises, even if it is a special form of social media environment. For data collection from platforms like TikTok or LinkedIn, students have successfully employed **Zeeschuimer** (<https://www.digitalmethods.net/Dmi/ToolZeeschuimer>) together with a custom-made manual (Harlung 2025). Of course, we review and update the syllabus and exercises every year responding to student feedback and also integrate new approaches coming from developments in the field, from instructor sessions or even exam projects.

Teaching Digital Methods: Our Key Insights

The experience of teaching this introduction to digital methods at the BA level has revealed a couple of key insights over the years which we would like to share. Some are related to the subject matter itself, which is very dynamic and requires continuous updates. Some insights are more consistent with the general didactic approach at our department, which is widely shared by colleagues across different educations. Our course addresses the specific challenge of bridging the technical, the theoretical and the methodological development of students' general competences in digital methods through hands-on exercises with a clear academic purpose for their whole study trajectory. We hope it is therefore helpful and instructive for other teachers in the humanities to apply and adopt our approach to their specific needs.

1) Managing the Learning Experience

- Communicate the structure and goals of teaching clearly and continuously throughout the course, reminding students of the purpose of a given exercise and underline the progression in their learning
- Differentiate types of instruction by different activities: Lectures for scoping a field or approach, seminars with group or individual exercises for learning new skills, instructor sessions for skill training and application to new cases
- Distinguish between the abstract and the specific elements of training and how they are serving the learning goals, incl. the expectations for using AI
- Keep discussions about exams to the end of the course, focusing on the learning, exercises and development of skills before the exam is prepared
- Where possible, build bridges between current and previous cohorts to show continuity and reflexivity, e.g. by hiring student assistants or sharing examples of exercises and exams from previous cohorts

2) Fostering the Learners' Interests and Skill Development

- Start with the experience of media and interests that students bring to class
- Demystify data and technical procedures: Show approachable ways to collect, explore and question data as just *one element* in a research process
- Respect and respond to the different skill levels & learning styles in a class
- Support responsive troubleshooting in class and beyond, and encourage students to find solutions themselves, e.g. with the help of AI or online fora
- Refer to future or non-academic application fields of methods to enable students in transferring skills and knowledge beyond their case or the course objectives
- Create space for critical self-assessment that is not interpreted as "failure" to comply with exam expectations and "getting a good grade"

- Embed individualised AI-agents as sparring partners to learn technical terminology, explain concepts or help with data processing questions
- Students feel more in control and encouraged to experiment when they can use a familiar graphic interface and can see all 'their' data being transformed by a script or formula that is understandable in all its parts

3) Enabling a Joint Knowledge Culture

- Emphasise that teachers, student assistants and students *are all part of a joint knowledge culture* during a course. *The teachers are also learners.* It does not need to be emphasised that teachers are experts but if they can approach students as fellow learners, they can make a huge difference
- Continuously develop didactic elements e.g. new exercises, tools, or research scenarios through inputs by students and student assistants
- Share knowledge across cohorts, e.g. by creating libraries of workflows and case studies created by students (if your course runs several years)
- Encourage students to help each other, e.g. by having students explain solutions that worked to the whole class. With only one teacher in the room, not every issue can be solved immediately
- Foster the use of AI for solving complicated data collection or pre-processing challenges, observing that data is shared responsibly
- Create space for experimentation, avoid competition for a "best solution"
- Allow students to showcase what they have learned (and be proud of it!)

4) From Specific Skills to General Competences

- Make data collection and analysis relatable for non-expert users, e.g. by starting with general software that is already known to students, and progressing gradually to more specialised applications
- Be agnostic about tools and stay focused on the methods: As long as a tool works, stay with it to teach more general competences that can be applied in a different software as well
- Acknowledge the spectrum of complexity involved in digital methods research: For some users, acquiring basic skills is already enough; others may want to learn the basics first before developing advanced skills independently later
- Prioritise understanding of *data-based problem solving* over one-off routines for a specific problem
- Unless your students are already familiar with programming, *avoid code and the blinking cursor on a black screen.* When students are familiar with the approaches in digital methods, they can graduate to use programming in Python or R (also with the help of AI). At the start of a course, coding in the humanities feels alien to what most of the students are interested in. Respect this feeling and work creatively with accessible possibilities.

The course establishes basic competences for the entire BA programme in Media Studies. Its design is embedded in an ongoing exchange with teachers in the same programme and the same semester. The ambition is to make digital methods applicable to other research contexts at later stages of the education, e.g. understanding media audiences and their interactions, or investigating strategic communication approaches by media organisations and businesses. This coordinated approach requires to communicate the course's setup and core ideas to teachers across the department. At the same time, the course feeds into general strategic interests of the department and contributes to dedicated knowledge exchange activities and workshops with fellow researchers and teachers for joint competence development.

Over the last few years, we have also engaged in knowledge exchange with colleagues in the Humanities at Aarhus University through workshops and courses. The main lesson from these activities is that our setup is flexible, understandable and proven in practice. It starts with the respective disciplinary needs, interests and challenges, to which digital methods are added as a new way of doing research, e.g. in history or anthropology. The main difference to dedicated trainings in data science or computational social science is that we start with the *emic* disciplinary context where such methods are needed. We do not try to implant data science methods into domains, which do not understand their research object as being represented in data. This distinction is quite important because it foregrounds the learner's as well as the researcher's own goals and ambitions by rephrasing their research approach in terms of experimental, data-oriented methods that can be implemented without arduous technical training. These experimental methods enable further specialisation (when needed) to include advanced data science methods or interdisciplinary collaborations with various domain experts.

Teaching Digital Methods in the Future: Two Notes on Institutional Agreements and IT Infrastructures for Teaching

With the broad integration of AI in teaching digital skills in the humanities, we will very likely see fundamental shifts in the next years as to what is possible for students to learn and how teachers need to respond to more advanced demands on skills and competences. Yet, the main part of this white paper has foregrounded the learning trajectory for technical skills in a humanities setting, and we are confident that new topics and new methods can be embedded in this framework. The didactic approach is transferable to other domains and seeks to balance both the technical demands on skills as well as the critical and reflective competences of working with digital methods in the humanities. We feel encouraged that the implementation of critical AI literacy will strengthen rather than replace the existing didactic approach without shedding the educational ambition behind digital methods training in the humanities. As a recent report on the impact of AI underlines, critical digital competences are needed more than ever, demanding a “slowing of the charge towards innovation and instead focus on what the technology is for and whether it is in fact the best tool for the job at hand.” In a practice-based learning process the combination of technical and socio-cultural research objectives will help to address “a broader range of human values than simply innovation” (European Commission; Luger and Stavridou 2026: 27).

In this last section, we would like to reflect on two aspects that institutionally enable our approach, and that warrant some more attention for the years to come. With questions of digital sovereignty and critiques of digital monopolies increasingly informing public debates about digitisation, we would like to highlight these institutional arrangements and support infrastructures which significantly make such a course possible in the first place.

First, we rely on a **course-specific legal agreement for data processing** with Aarhus University, which puts students under the same legal protections for collecting and analysing data as full-time researchers, in Danish *Forskerhjemmel* (or ‘research permission’, see Appendix 2). Students are informed about this agreement and its conditions at the beginning of the course, which sensitises them to the legal basis governing data use, storage and processing. The agreement was needed because the tasks of collecting, storing and analysing digital data from the Web or social media was written

into the study regulations, thereby requiring the university to ensure a sound legal basis for the exam. The importance of this agreement has to be emphasised, since ordinarily students can of course collect data for exam projects without explicit permission by the university, leaving compliance risks for personal data (e.g. in respect to GDPR) largely on the students' shoulders. Data from social media or the Web can be scraped *at scale* with off the shelf, open source, and free (or freemium) tools. *This is a reality now*. There are endless tutorials, software solutions, and tools available on the Web for data collection and processing purposes. Many of these tools come with insecure ways of data processing, dubious business models or technical dependencies that raise severe legal risks (e.g. in an EU context).

In this course, we make students responsible for safe data processing while protecting the liberty of research. Under the same conditions as full-time staff, our students can collect and process data as long as data is stored securely, e.g. on an encrypted drive, not shared publicly, and deleted once the purpose of the collection and analysis has expired (ie., once the exam is passed and the legal complaint period is over). Given that many courses in social sciences, humanities and beyond nowadays make use of data from publicly available sources such as websites and social media, we pose our arrangement as a best practice to be emulated across departments. This would significantly lower the burden for individual teachers to negotiate course-specific agreements with their Faculties or legal departments. We also see significant compliance risks emerge for the university as a whole if there is a strong divergence in terms of how data collection and processing in a teaching context is legally interpreted and institutionally supported. Outsourcing such risks to students and leaving legal arrangements to the teacher's own good will as they perform their contractual duties severely obstructs the development of advanced digital methods education in the humanities in the long run.

Second, our course concept relies on data being somehow available for research purposes (through scraping, crawling, API access, etc.) and that students have capable (and up-to-date!) private computers that they can use in class. We do not really have an **IT infrastructure for teaching** to use in classrooms. Although our course was a pioneer in this respect, the approach still relies heavily on private equipment, available open source tools and services, as well as teachers' and assistants' good will and flexibility. When we started, our course became a pilot project for the Danish supercomputer (UCloud), a federated national research IT infrastructure, to be used in a

teaching context. Over the years, UCloud has been widely adopted and offers a broad range of advanced professional research and computing services. Yet, there is still a *strong division between IT infrastructures for research and IT infrastructures for teaching*. The latter does not need to be as advanced in its functions but rather needs to be well defined and well designed (!) in its interfaces so that students can intuitively understand and employ tools from the start. Further, outsourcing compute time and equipment costs to students creates a skewed balance in terms of the actual institutional costs of running such a course.

We also rely on dedicated IT support at the department for teaching, where we appreciate the work by our colleague Ulrich Have for managing the access to UCloud, setting up virtual environments for required software, and helping students with advanced technical challenges. The second component of our institutional support comes from Asger Harlung, who is responsible for reviewing and updating the Tools & Tutorials section of the Centre for Digital Methods and Media website (apart from other tasks at the department). These tutorials are an indispensable resource for teachers in other courses as well. They allow us to be flexible in our teaching and also be visible outside Aarhus University as such. Both IT infrastructure and complementary teaching resources account for a flexible, safe and up-to-date course environment that requires considerable resources beyond the course budget of instruction hours.

Lastly, we want to raise the question how much students can be expected to provide a private laptop at the university for rather specialised, computational tasks. Not all laptops are created equal. Not all private hardware is capable of running the software required in this course, is able to store data in larger amounts, or is secure enough to ensure data protection (apart from encryption). Teachers cannot know what else students are doing on their hardware, which processes are running in the background, or how students get distracted by technical troubleshooting. In the interest of security on the one hand, and equity in teaching on the other hand, it would be a significant step forward to create well-designed virtual environments for flexible use across the humanities to teach digital methods. The private laptop should be merely an interface to a controlled and centrally hosted application run by teachers and the university. For a better integration of open source software, data storage and federated, secure access, such virtual environments like the one available on UCloud would need to be adapted for use in teaching. This would lower the

dependency on private laptops for safe collection and processing of data. But most importantly, we would create infrastructures to *ensure that essential learning outcomes do not depend on private financial resources*. We are aware that such a change would need high initial investment and an institutional agreement about long-term support and update. Ideally, such an interdisciplinary solution (ie. virtual environments) could be set up for teaching across departments in the humanities with a strong potential for international collaboration and funding schemes to moderate high development costs.²

In the future, the methods and topic areas of this course are likely to change. We believe that our didactic approach supports humanistic research in a digital communication environment. We hope that this white paper contributes to enabling a broader academic debate about the hidden but equally important institutional work of developing teaching formats for this future.

Contact the Authors

Christoph Raetzsch (craetzsch@cc.au.dk)

Niels Brügger (nb@cc.au.dk)

Cite as: Christoph Raetzsch, Niels Brügger (2026). "CDMM White Paper: Teaching Digital Methods. Experiences on the BA Level in Media Studies at Aarhus University" Aarhus: Centre for Digital Methods and Media.

Shared CC-BY-4.0 <https://creativecommons.org/licenses/by/4.0/deed.en>

Appendices

- **Appendix 1:** Syllabus for the 2026 edition of the course (in Danish/English)
- **Appendix 2:** Instruction on Data Processing regulation and GDPR risk assessment template for personal data (in Danish) - *forskerhjemmel*
- **Appendix 3:** Template Orange Workspace with different workflow suggestions and options (requires Orange Data Mining version 3.40, <https://orangedatamining.com/download/>, and the optional widget libraries *Educational*, *Geo*, *Image Analytics*, *Network*, *Text*, *Textable*, *Timeseries*, *World Happiness*)

² Aarhus University offers access to UCloud through the Centre for Humanities Computing (CHC), whose main purpose is to support research but less teaching.

References

- Aasman, Susan; Ben-David, Anat; Brügger, Niels (eds.) (2025). *The Routledge Companion to Transnational Web Archive Studies*. Abingdon: Routledge.
- van Attevelde, Wouter; Trilling, Damian & Arcila, Carlos (2022.) *Computational Analysis of Communication*. Wiley Blackwell. <https://cssbook.net/>
- Bounegru, Liliana; Chalmers, Lucy; Gray, Jonathan (eds.) (2012). *The Data Journalism Handbook 1*. <https://datajournalism.com/read/handbook/one>
- Bounegru, Liliana; Gray, Jonathan (eds.) (2025.) *The Data Journalism Handbook 2. Towards A Critical Data Practice*. London: Routledge. <https://datajournalism.com/read/handbook/two>
- Brown, Kathryn (ed.) (2020). *The Routledge Companion to Digital Humanities and Art History*. Abingdon: Routledge.
- van Es, Karin; Schäfer, Mirko Tobias & Wieringa, Maranke. (2021). Tool Criticism and the Computational Turn. A "Methodological Moment" in *Media and Communication Studies*. *Medien & Kommunikationswissenschaft*, 69(1), 46-64. <https://doi.org/10.5771/1615-634x-2021-1-46>
- European, Commission: Directorate-General for Research and Innovation; Luger, Ewa; Stavridou, Marianthe (eds.) (2026). *The Impact of the Artificial Intelligence on Our Societies*. Brussels: Publications Office of the European Union. <https://dx.doi.org/10.2777/2377840>
- van Geenen, Daniela; Gray, Jonathan; Bounegru, Liliana; Venturini, Tommaso; Jacomy, Mathieu & Meunier, Axel (2022). Staying With the Trouble of Networks. *Frontiers in Big Data*, 5, 510310. <https://doi.org/10.3389/fdata.2022.510310>
- Haim, Mario (2023.) *Computational Communication Science. Eine Einführung. [Computational Communication Science. An Introduction]*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-40171-9>
- Harlung, Asger (2025). *CDMM Tutorial Zeeschuimer*. Aarhus: CDMM Publications. https://cc.au.dk/fileadmin/dac/CDMM/CDMM_Zeeschuimer_Tutorial.pdf
- Hine, Christine; Watson, Hayley; Snee, Helene; Roberts, Steven; Morey, Yvette (eds.) (2016). *Digital Methods for Social Science. An Interdisciplinary Guide to Research Innovation*. Basingstoke: Palgrave Macmillan.
- Madsen, Anders Koed; Munk, Anders Kristian (eds.) (2026). *Handbook of Digital and Computational Research Methods in the Social Sciences and Humanities*. Cheltenham (UK): Edward Elgar Publishing. <https://dx.doi.org/10.4337/9781802208993>
- Plantin, Jean-Christophe; Punathambekar, Aswin (2018). Digital Media Infrastructures: Pipes, Platforms, and Politics. *Media, Culture & Society*, 41(2), 163-174. <https://doi.org/10.1177/0163443718818376>
- Raetzsch, Christoph; Bødker, Henrik (2016). Journalism and the Circulation of Communicative Objects. *Tecnoscienza*, 7(1), 129-148. <http://www.tecnoscienza.net/index.php/tsj/article/view/256/170>
- Rogers, Richard; Lewthwaite, Sarah (2019). Teaching Digital Methods: Interview With Richard Rogers. Interviewer: S. Lewthwaite. *Revista Diseña*, 14), 12-37. <https://doi.org/10.7764/disen.14.12-37>
- Rogers, Richard (2019.) *Doing Digital Methods*. London: SAGE.
- Rogers, Richard (2014). Political Research in the Digital Age. *International Public Policy Review*, 8(1), 73-87.
- Rogers, Richard (2013.) *Digital Methods*. Cambridge MA.: MIT Press.
- Rogers, Richard. (2009.) The End of the Virtual. Digital Methods. Inaugural Lecture. University of Amsterdam, Vossiuspers. <https://doi.org/10.7551/mitpress/8718.001.0001>
- Sloan, Luke; Quan-Haase, Anabel (eds.) (2018). *The Sage Handbook of Social Media Research Methods*. Thousand Oaks: Sage.
- Venturini, Tommaso; Jensen, Pablo & Latour, Bruno (2015). Fill in the Gap: A New Alliance for Social and Natural Sciences. *Journal of Artificial Societies and Social Simulation*, 18(2), 37. <https://doi.org/10.18564/jasss.2729>
- Venturini, Tommaso; Rogers, Richard (2025.) *Digital Methods: A Short Introduction*. Hoboken: Polity Press.