Version Control of Code & Data with Git

An international teaching collaboration between the Universität Hamburg and Erasmus University Rotterdam (Application for the ILLF@home program)

Dr. Lennart Wittkuhn   Dr. Ana Martinovici

An online version of this application is available at

https://lennartwittkuhn.com/illf-application/.

Project Description

Course Content

The digital objects on our computers are in constant flux. Manuscripts, programming code, and research data undergo continuous and often collaborative changes over extended periods of time. The systematic documentation of these changes forms the basis for controlled and reproducible work on digital objects.

In a seminar with practical exercises, students learn "Version Control of Code and Data with Git." Version control is the notebook for a digital world, and Git is perhaps the most well-known version control system. It enables precise documentation of changes in digital objects, allowing one to trace who changed what, when, how, and why in a file. Changes can be revised, versions can be compared, and restoration is possible. Moreover, Git allows simultaneous work on the same file, with parallel versions systematically integrated. Git also facilitates effective collaboration, as platforms like GitHub or GitLab enable sharing, transparent viewing, utilization, and collaborative development of code and data with the world. Thus, version control contributes to making knowledge generated from data transparent, accessible, and verifiable. As an effective method for storing and processing code and data, version control represents a core competency of Data Literacy.

In their studies, students in many research domains receive a multifaceted education in research methodology, statistics, and principles of responsible scientific work. The current teaching concept complements practical skills in the effective management of programming code and research data. The acquired skills are particularly applicable in empirical dissertation projects.

Discussing concepts such as reproducibility, open code, and open data supports understanding of principles of good scientific practice. Due to their extensive methodological training, students are in demand not only for research but also for careers as data analysts or data scientists in the business sector. Since version control is a standard in professional software development, the employability and innovation capacity of graduates are enhanced.

In addition to these practical competencies, students are taught theoretical concepts in the areas of open science, open code, open data, and open source. This enables a reflective engagement with increasingly digital and data-driven economic and societal processes, promoting participatory self-empowered citizenship.

The continuation of the teaching concept at the University of Hamburg, especially in a newly established research-oriented master's program and a graduate school for doctoral students, is anticipated. Skills in version control are relevant in all fields where digital data is processed. The teaching content and practical exercises are also cross-disciplinary and transferable to other study programs. The publication of all teaching materials as Open Educational Resources allows independent reuse and development (see teaching concept).

Teaching Concept

The didactic teaching concept consists of a seminar (2 teaching hours per week) with lectures by the instructors (20%) and practical collaborative hands-on exercises (80%). The lectures provide theoretical foundations (e.g., reproducibility, open code, open data) and transition into interactive demonstrations of Git ("code-along"). These teaching formats are complemented by supervised exercises in small groups or pairs, which may continue outside of class time. The courses are accompanied by continuous evaluation and self-assessment quiz questions, implemented in formR and supervised by a student assistant. Evaluation data and quiz responses are stored anonymously and reused within the course as exemplary datasets. All software tools used are open-source, freely available, compatible with all operating systems (macOS, Linux, Windows), and mostly provided by the University of Hamburg and the partner university, Erasmus University Rotterdam.

As Git is primarily developed for command-line usage, the teaching concept also focuses on interaction with the command line. The command line is a part of a computer program that takes a text line as input and interprets it as a command in context. In the initial teaching units, basic command line commands are demonstrated and practiced in joint exercises. Graphical user interfaces (GUIs) are also available as an alternative to the command line. GitKraken or GitHub Desktop are used as the GUIs for Git, available in a free Pro version for students through the GitHub Student Developer Pack. Digital technologies such as Jupyter Notebooks and Quarto are primarily used for creating teaching materials (project website, presentations, practical exercises). To prevent potential issues with software installation, Jupyter Notebooks allow interaction within prepared computational environments provided through Jupyter-Hub at the University of Hamburg or Binder, accessible through a web browser.

All teaching materials are published as Open Educational Resources, including the project website, under a CC BY-SA 4.0 license. Discoverability and citability of teaching materials are ensured through persistent identifiers on the Zenodo platform. The courses are conducted in an online teaching-learning setting using Zoom. Technical issues can be addressed simultaneously in "Breakout-Rooms" through Zoom with the help of the tutors. Communication outside the classes can take place through the GitHub platform. Basic programming skills are advantageous for the course but are not mandatory. The introduction to all used software technologies is part of the course. The applicants have extensive experience in using all mentioned software tools in research and teaching.

Course Goals

This course aims to teach students the fundamentals of version control using Git, with a focus on precise documentation of changes in digital objects, collaborative work, and effective use of platforms like GitHub. It integrates practical exercises to enhance skills in code and data management, aligning with the broader goals of promoting transparency, accessibility, and verifiability of knowledge in science and society. Additionally, the course covers theoretical concepts in open science and reproducibility fostering a reflective engagement with digital and data-driven processes and enhancing students' employability in research, data analysis, and other sectors requiring digital competences.

Beyond these subject-related goals, the course aims to foster virtual exchange among students and researchers from the two partner universities. Through collaborative learning and practical exercises, students engage in shared projects, encouraging active participation and the exchange of diverse perspectives. Using online platforms like GitHub for collaborative development not only facilitates real-time interactions but also provides a space for continuous communication. Virtual discussions, joint problem-solving sessions, and collaborative coding exercises allow students to learn from each other's experiences, practices, and cultural backgrounds. This international teaching collaboration, therefore, not only enhances technical proficiency in version control using Git but also promotes a rich exchange of ideas among students and researchers from the partner universities.

Timetable

- **Project period:** 01.04.2024 to 31.03.2025
- **Course delivery:** 14.10.2024 to 01.02.2025 (winter term)

Step	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
Preparation of new learning resources	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan		
Adaptation of existing resources to virtual setting	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan		
Continuous documentation and publication of all learning resources as OER	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan		
Preparation of questions for evaluation and self-assessment (quiz questions) in formR					Aug	Sep	Oct	Nov	Dec	Jan		
Delivery of the courses							Oct	Nov	Dec	Jan		
Analysis of the evaluation data and self-assessment (quiz questions)							Oct	Nov	Dec	Jan		
Supervision of students in open office hours and via GitHub							Oct	Nov	Dec	Jan		
Final analysis of the continuous teaching evaluations											Feb	
Project evaluation and preparation of a results and experience report											Feb	Mar
Persistent archiving of all Open Educational Resources via e.g., Zenodo												Mar

Cost Justification

Overview

	Description	Category	Duration	Costs
1	Travel and accommodation for Lennart Wittkuhn for trip to Rotterdam to initiate project	Travel and accommodation	2 days	350.00
2	Travel and accommodation for the Student Research Assistant for trip to initiate project	Travel and accommodation	2 days	350.00
3	Student Research Assistant (MSc level) for 40 hours / month	HR costs	01.04.2024 to 31.03.2025	9428.34
4	Student Teaching Assistant 1 (MSc level) for 2 SWS / month	HR costs	14.10.24 to 01.02.25	2256.99
5	Student Teaching Assistant 2 (MSc level) for 2 SWS / month	HR costs	14.10.24 to 01.02.25	2256.99
6	License for Plausible.io	Partial costs	01.04.2024 to 31.03.2025	107.10

Sum of total costs: 14749.42 Euro ~ 14750 Euro.

Justification

Regarding 1 (Travel and accommodation for Lennart Wittkuhn for trip to Rotterdam to initiate project):

The travel costs for initiating the collaborative teaching project are justified as they enable essential face-to-face interactions with the collaborators at the partner university. This facilitates stronger interpersonal relationships, fostering effective communication and teamwork. In-person engagement during the initiation phase is crucial for efficient planning, alignment of teaching methodologies, and the establishment of a shared vision for the project. The travel expenses contribute to a cohesive collaboration, enhancing the likelihood of project success and the delivery of high-quality educational outcomes.

Regarding 2 (Travel and accommodation for the Student Research Assistant for trip to initiate project):

Including the student research assistant in the travel costs is justified as it enhances their effectiveness in the collaborative teaching project by providing valuable face-to-face interactions with collaborators during the initiation phase. This in-person engagement allows the student

research assistant to establish direct connections with team members, fostering effective communication and building a stronger working relationship. The opportunity for face-to-face interactions ensures a shared understanding of project goals and methodologies, enabling the student research assistant to work more efficiently and collaboratively in later stages. Additionally, participating in the initial meetings and discussions enhances the student research assistant's overall contribution, promoting a more integrated and cohesive team dynamic for the duration of the project.

Regarding 3 (Student Research Assistant (MSc level) for 40 hours / month):

The student research assistant will help to adapt the existing learning materials to an international virtual learning setting. In addition, they will enrich the curriculum by adding additional content that caters to diverse learning styles and the virtual setting. The preparation of surveys and quizzes by the assistant contributes to formative assessments, facilitating ongoing improvements to the course. The student research assistant will also analyze responses to online surveys, expected to provide valuable insights into the learning progress of students. The student research assistant's role in analyzing web traffic (see note about website analytics below) contributes to optimizing the online learning materials for accessibility and user experience. Finally, their involvement in preparing each teaching session ensures a seamless integration of the various learning resources, elevating the overall quality of the virtual learning experience for students.

Regarding 4 (Student Teaching Assistant 1 (MSc level) for 2 SWS / month):

In an online Git course, a teaching assistant (TA) is crucial for providing individualized support, timely feedback, and technical troubleshooting to students navigating the complexities of version control. With the absence of in-person interaction, a TA becomes instrumental in facilitating discussions, fostering collaboration, and adapting to diverse learning styles. They supplement instructional content, manage assessments, and ensure a consistent learning experience for students with varying levels of familiarity with Git concepts. Overall, a TA's presence enriches the online learning environment by offering personalized assistance, promoting collaborative engagement, and contributing to the successful learning of Git practices.

Regarding 5 (Student Teaching Assistant 2 (MSc level) for 2 SWS / month):

Introducing a second teaching assistant (TA) in an online Git course further enhances the learning experience by addressing the demands of the diverse course participants. The second TA can efficiently manage additional volume of inquiries, providing additional individualized support and timely feedback. This ensures that students receive prompt assistance, particularly during challenging learning sections. Moreover, having two TAs facilitates more comprehensive coverage of discussion forums and collaborative activities, fostering a richer learning environment. With varying learning paces and backgrounds, having two TAs allows for even

greater adaptability to diverse learning styles, offering students an enhanced and well-rounded educational experience.

Regarding 6 (License for Plausible.io):

All course materials will be shared as publicly accessible websites hosted via GitHub Pages. We will use website analytics to gain insights into reach and dissemination of the teaching materials. These insights will allow us to better understand how users engage with our online learning materials and derive opportunities for improvements. We will use Plausible, which provides "intuitive, lightweight and open source web analytics", does not use cookies and is fully compliant with GDPR, CCPA and PECR. A subscription plan with a monthly quota of up to 10k pageviews for 1 year will be selected.

Application Form

Short Project Description

Short version

Students from Universität Hamburg and Erasmus University Rotterdam will engage in a course on "Version Control of Code & Data" with Git, emphasizing its fundamental role in managing collaboration on digital objects in academic work. The course, originally in-person at Universität Hamburg (https://lennartwittkuhn.com/version-control-course-uhh-ws23/), will be adapted for an international and virtual learning environment.

Long version

In an international and virtual course with collaborative hands-on exercises, students from Universität Hamburg and Erasmus University Rotterdam will learn "Version Control of Code and Data" with Git. Version control is as a fundamental approach for managing the evolution of digital objects, with Git standing out as one of the most widely recognized version control systems. Its importance goes beyond its wide application in professional software development, proving valuable in academia by offering an organized and systematic approach to tracking and managing changes in scholarly work. Git and the online platform GitHub provide precise documentation of changes, support revision and comparison of versions, allow simultaneous work on files, and facilitate effective collaboration through transparent sharing, viewing, and interactive development of digital objects, in particular code and data. The course will be based on existing learning materials of a course, previously supported by the Digital & Data Literacy in Teaching Lab program and currently taught as an in-person seminar at Universität Hamburg (https://lennartwittkuhn.com/version-control-course-uhh-ws23/). With the support of the ILLF@home program, the course will be adapted for an international and virtual learning environment. The application documents can also be found on the project website at https://lennartwittkuhn.com/https://lennartwittkuhn.com/illf-application/

Project Goals

> *"In addition to imparting subject-related knowledge, virtual exchange among students and researchers from the partner university must be the primary goal."*

Short version

This course teaches version control using Git, emphasizing documentation, collaboration, code and data management. It aligns with goals of promoting transparency, and verifiability in

knowledge, covering concepts in open science and reproducibility. Beyond subject-related objectives, the course fosters exchange among partner universities through collaborative learning, collaborative practical exercises, and communication on platforms like GitHub.

Long version

This course aims to teach students the fundamentals of version control using Git, with a focus on precise documentation of changes in digital objects, collaborative work, and effective use of platforms like GitHub. It integrates practical exercises to enhance skills in code and data management, aligning with the broader goals of promoting transparency, accessibility, and verifiability of knowledge in science and society. Additionally, the course covers theoretical concepts in open science and reproducibility fostering a reflective engagement with digital and data-driven processes and enhancing students' employability in research, data analysis, and other sectors requiring digital competences. Beyond these subject-related goals, the course aims to foster virtual exchange among students and researchers from the two partner universities. Through collaborative learning and practical exercises, students engage in shared projects, encouraging active participation and the exchange of diverse perspectives. Using online platforms like GitHub for collaborative development not only facilitates real-time interactions but also provides a space for continuous communication. This emphasis on communication skills not only contributes to the success of joint projects but also prepares students for effective participation in digital, data-driven workplaces. Virtual discussions, joint problem-solving sessions, and collaborative coding exercises allow students to learn from each other's experiences, practices, and cultural backgrounds. This international teaching collaboration, therefore, not only enhances technical proficiency in version control using Git but also promotes a rich exchange of ideas among students and researchers from the partner universities.

Partner University

> *"Explain your selection of partner university/universities. If the selected university is not a strategic or profile partner, separately explain the importance of the partnership."*

Short version

Erasmus University Rotterdam, a profile partner of Universität Hamburg, is a leader in Open Science, evidenced, for example, by hosting the Netherlands's National Open Science Festival. Dr. Martinovici, an Assistant Professor at EUR, not only actively applies open science principles in her research but also conducts Git workshops, making both her and EUR exceptionally suitable partners for the proposed ILLF@home program.

Long version

The Erasmus University Rotterdam (EUR) is one of the profile partner universities of the Universität Hamburg. The student target group of the proposed course at EUR come from diverse study programs, including psychology, marketing, neuro-marketing and medicine. The EUR is at the forefront of the Open Science movement, promoting transparency, collaboration, and responsible research practices. For example, in August 2023, EUR hosted the Netherlands's National Open Science Festival, attracting 400 visitors. Git and GitHub are tools that allow students and researchers to put principles of open science into practice, and they are in high demand within EUR's research and student community. Dr. Martinovici is an Assistant Professor of Marketing at the Rotterdam School of Management at the EUR. Beyond implementing principles of open and reproducible science in her own research, she has taught several workshops on version control with Git and GitHub, for example at the renowned conference of the Society for the Improvement of Psychological Science (SIPS). Together, Dr. Martinovici and Erasmus University Rotterdam present an ideal partnership for the proposed project under the ILLF@home program.

Contributions

"What content-related or financial contribution will the partner make?"

Short version

Dr. Martinovici's expertise in teaching Git and GitHub will enhance course material development, addressing shortcomings and tailoring content for students from both partner universities. Her contributions may cover topics like version control for large datasets and data protection and privacy. Anticipated financial support from the partner university, once the course is established, could involve assistance from EUR's teaching assistants.

Long version

Leveraging Dr. Martinovici's extensive experience in teaching version control with Git and GitHub, she will contribute to the content-related development of the course materials. Her expertise will help address current shortcomings in the existing course materials and tailor the course to meet the needs of students from both partner universities. More specifically, content-related contributions may cover topics like version control for large datasets, automation using Makefiles, as well as data protection and privacy. Financial support from the partner university is anticipated in the future, once the course is established. This support could, for instance, involve assistance from teaching assistants at Erasmus University Rotterdam.

Intercultural experience

> *"How will you help students gain intercultural experience in addition to subject-specific knowledge through the courses on offer?"*

Short version

Git and GitHub are essential for collaboration on digital objects, fostering intercultural exchange among students. Course exercises involve contributing to shared GitHub repositories, encouraging collaboration between students from both universities in mixed teams. This approach enhances subject-specific knowledge and provides valuable intercultural experience through peer interactions.

Long version

Fundamentally, Git and GitHub serve not only as tools for version control but also as crucial elements for fostering effective collaboration on digital objects. While these tools are domain-agnostic, our approach involves using examples that encourage intercultural exchange among students and researchers from the partner universities. For example, as part of the course exercises, students will contribute to a code & data repository hosted on GitHub, which is shared by students from both universities. To enhance students' intercultural experience in addition to subject-specific knowledge, the courses will introduce collaborative projects that deliberately require students to work in mixed teams. This approach enables students not only to learn about version control and collaboration, but also provides them with the opportunity to interact with peers from the partner university. Through this interaction, students may exchange information about their field of study, student life in their respective cities, or other topics of mutual interest, thereby gaining valuable intercultural experience.

Embedding in the curriculum

> *"The ILLF@home program aims to establish international virtual learning and teaching formats for the long term. To what degree is the project embedded in the curriculum and how do you plan on ensuring the project being applied for will become a long-term aspect of the curriculum?"*

Short version

The course, introduced at Universität Hamburg in 2023, is set to become a regular offering for Psychology MSc and PhD students. At Erasmus University Rotterdam, the course may extend to PhD students at other research institutes, and promoted through the Open Science Community Rotterdam. The ILLF@home program will greatly facilitate this integration into the regular curricula of both partner universities.

Long version

At Universität Hamburg, the course was introduced for the first time as an in-person seminar during the winter term 2023/24 as part of the elective module in the Master's program in Psychology. It is planned to regularly incorporate the course into the curriculum and offer it each winter term. Discussions are underway regarding an additional offering of the course as part of the PhD graduate school at the Institute of Psychology at Universität Hamburg. Both course formats could consistently include international virtual learning experience, thereby establishing this course format for the long term. Moreover, to build flexibility into the course design and accommodate evolving educational trends, technological advancements, and changes in the academic landscape, we will actively seek feedback from students and other instructors through regular surveys and evaluations. This continuous feedback mechanism will allow us to make real-time adjustments, ensuring that the course remains responsive to emerging developments, thereby contributing to its long-term viability. Similarly, at Erasmus University Rotterdam, the course may be extended to ERIM PhD students. ERIM (Erasmus Research Institute of Management) is a collaborative research institute of Rotterdam School of Management, Erasmus University (RSM), and the Erasmus School of Economics (ESE). Through the Open Science Community Rotterdam, the course will also be promoted to PhD students at other EUR research institutes, ensuring a diverse audience spanning various research fields, including psychology and medicine. To complement these efforts, ongoing professional development opportunities will be provided for instructors involved in the course, and comprehensive documentation and best practices guidelines will be developed. This documentation will serve as a reference for future instructors, offering insights into successful strategies, challenges, and adjustments made over time, thereby contributing to the continuity and consistency in the delivery of the international virtual learning format. The ILLF@home program will provide the necessary support to establish an international virtual learning and teaching format that can be regularly integrated into the curricula of both partner universities.