VU VRIJE UNIVERSITEIT AMSTERDAM

Faculty of Social Sciences

Societal Analytics Lab

2025 Report

Abstract

For the Societal Analytics Lab accountability is a core value. In this report, the Lab summarises all what it has done and accomplished during the year 2025.

Dr. Sofia Gil-Clavel (Lab Manager/ Researcher)

b.s.gilclavel@vu.nl

SOCIETAL ANALYTICS LABORATORY TEAM IN 2025

Wouter van Atteveldt
(Director)

Kasper Welbers
(Co-Director)

Sofia Gil-Clavel
(Manager/ Researcher)

Gabriella Bollici Moreno
(Events Co-Organizer)

CONTENTS

TABLE OF FIGURES

THE SOCIETAL ANALYTICS LABORATORY

In the Societal Analytics Laboratory (here onwards referred to as the Lab), our main goal is to be the Faculty of Social Science's (VU FSW) first point of contact for those interested in applying computational methods (Figure 1) in their research on societal issues.

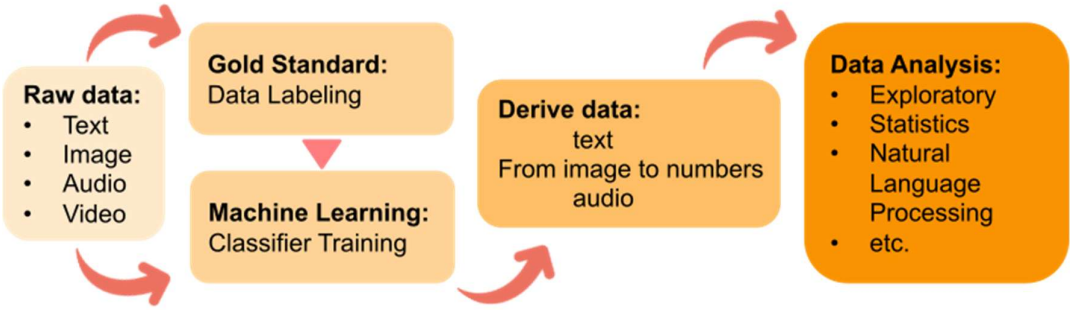


Figure 1: Computational methods pipeline

Specifically, we target two groups of scholars:

1. Those New to the Matter: those that are curious about how computational methods can benefit their research.
2. Those Already in the Know: those that are already familiar with computational methods but lack resources.

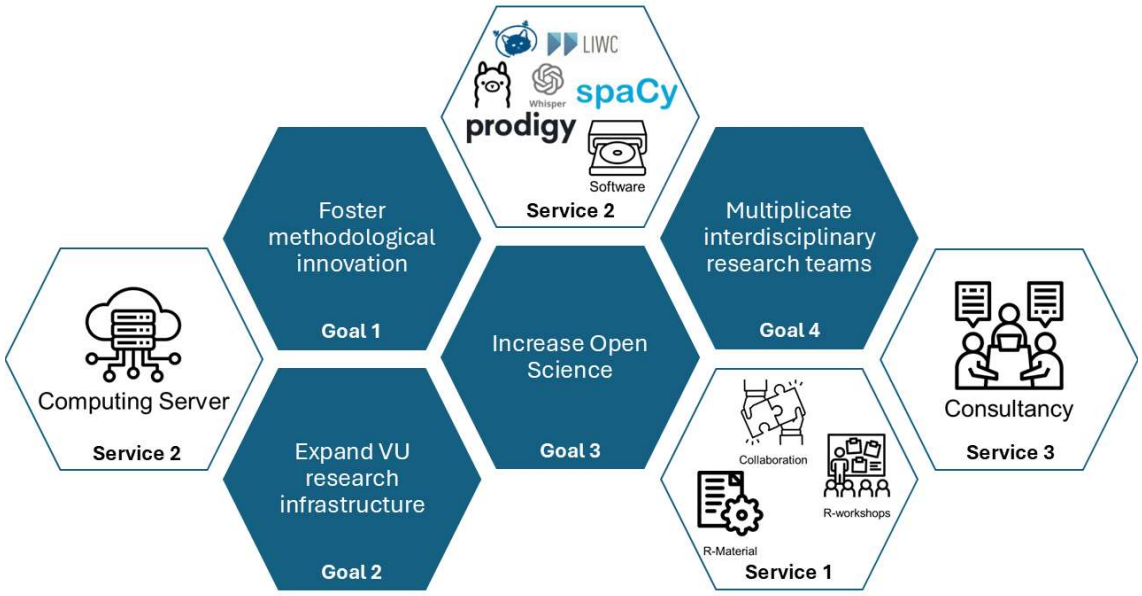


Figure 2: Lab's services and their relation with its goals.

For both groups, the Lab offers guidance, training, and essential tools to boost their career. We provide the necessary tools and infrastructure to facilitate VU FSW research. As such, the Lab

provides three main services: Collaboration and Learning, Computational Resources and Tools, and General Consultancy and support writing Grants and Research Proposals. With these services, we aim to foster methodological innovation, expand VU research infrastructure, increase Open Science through better reproducibility and transparency, and multiplicate interdisciplinary research teams. Figure 2 shows the relationship between the Lab's services and goals.

DISSEMINATION STRATEGY

In principle, any VU FSW scholar can become a member of the Lab by signing up via this form: https://forms.office.com/e/8Bgd2YsasJ. To keep our members informed and to reach out to our target groups of scholars. The Lab has four communication channels: Newsletters, webpage, BlueSky account, and physical posters and flyers.

First, in March 2025, the Lab started to communicate via Newsletters sent by email. Since then, we have sent one Newsletter every month. Second, the Lab launched its current webpage in April 2025. The webpage can be accessed here: https://societal-analytics.nl/. In the webpage, we inform members about our services, we also host our blogposts there, and we also advertise our workshops and events. Third, the Lab has its own BlueSky account (https://bsky.app/profile/societal-analytics.nl), where the Lab posts about any new events we have coming up. We also repost any other news that the followers may found interesting or useful. Finally, we also print physical posters and flyers. The posters are used during the vents and we have one permanently hanging at the Communication department (Figure 3). The flyers are distributed in all the departments any time we have an incoming workshop or event (check page 14 for an example). We use them to reach out to (new) Lab members that may not be that attentive to their email.

Using these advertising strategies, the Lab was able to recruit 25 new members. This means that by the end of 2025, the Lab moved from 40 to 65 VU FSW Lab members.

Figure 3: Lab poster

ACADEMICS COMPUTATIONAL NEEDS

As a way to tailor our services to the computational needs of the VU FSW researchers, the Lab ran a survey on November 2024. In this survey, the Lab asked VU FSW researchers about the type of research they are interested in performing, the type of programs they use in their everyday work, and the type of computational knowledge and resources they need. In this section, we will give a summary of the findings.

In total 29 academics participated, from there 24%, 43%, and 32% belong the categories "PhD Student/ PostDoc", "Lecturer/ Assistant Professor" and "Associate Professor/ Full Professor", respectively. The distribution by gender is in Figure 4.

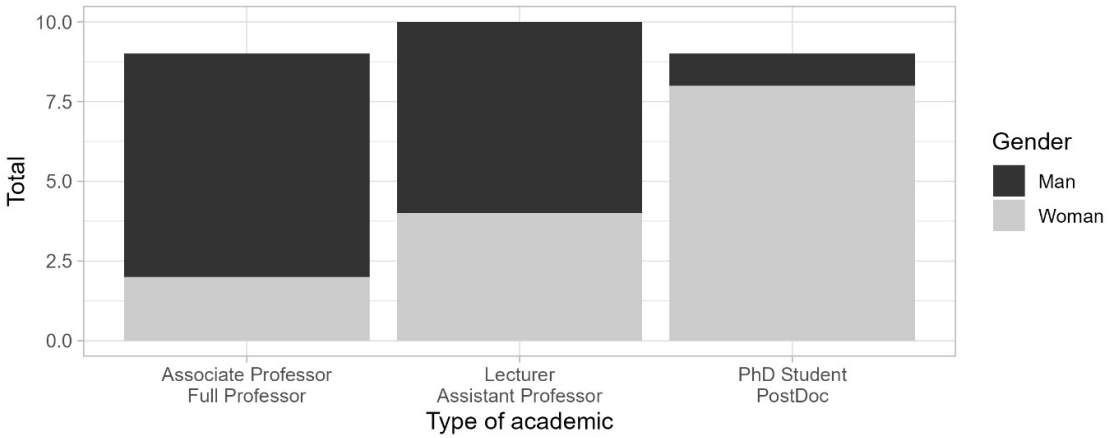


Figure 4: Number of academics that participated in the survey broken down by level and gender.

From here, academics mostly perform quantitative research (45%). When asked about the type of research they would like to perform in the future, we see that the percentage of computational research almost equals quantitative research (Figure 5).

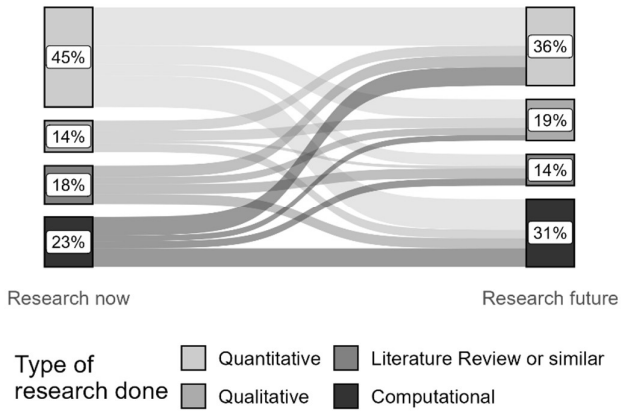


Figure 5: Distribution of the type of research academic are mostly performing now and the type they would like to perform in the future.

For the Operating System, we found that 95% of the researchers use Windows or Mac. In terms of the software, researchers mostly use R for data modeling, data handling, and data visualization, while for model training they prefer to use python (Figure 6).

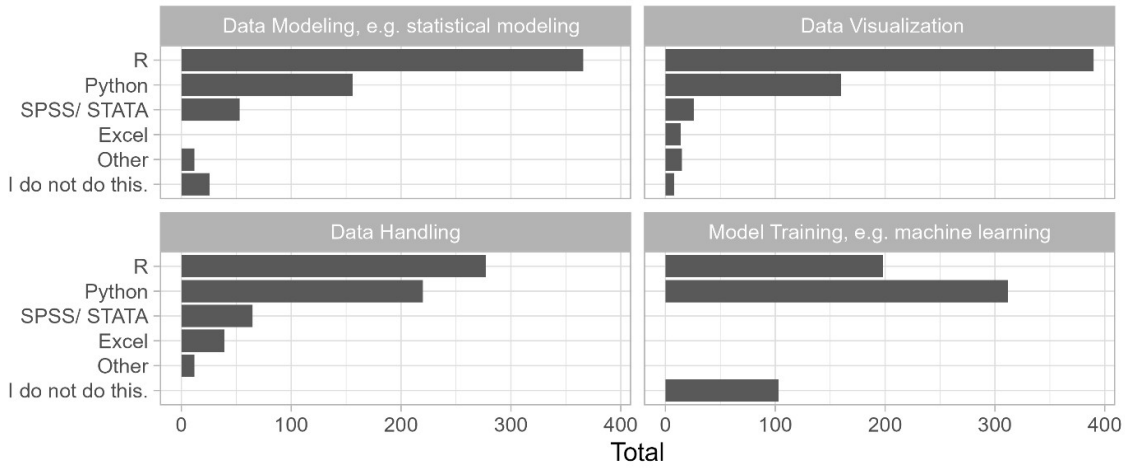


Figure 6: Most frequently used programs

When questioned about their computational constraints (Figure 7), 48% and 37.9% have experienced constraints related to knowledge and computing power, respectively.

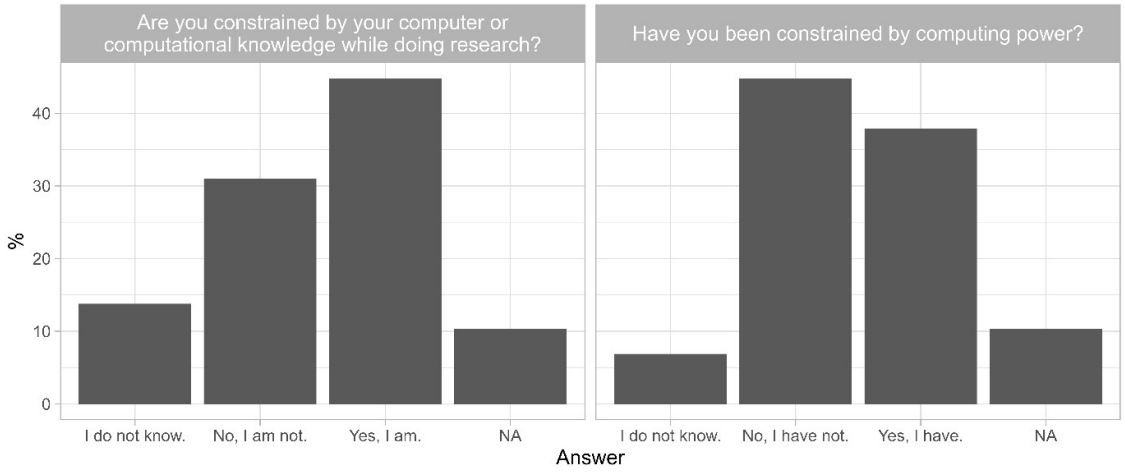


Figure 7: Computational Constraints

In terms of their awareness of VU IT hardware, around 57% knew about it, and 39% did not (Figure 8).

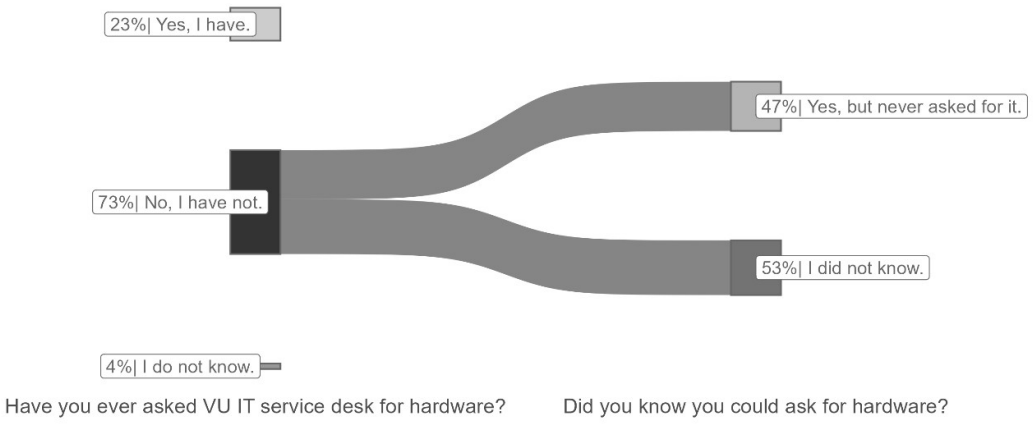


Figure 8: IT resources

To gauge the possible popularity of the Lab, we also asked respondents whether they would use our services (Figure 9). Figure 9 shows that more than 50% of respondents would use the services we provide.

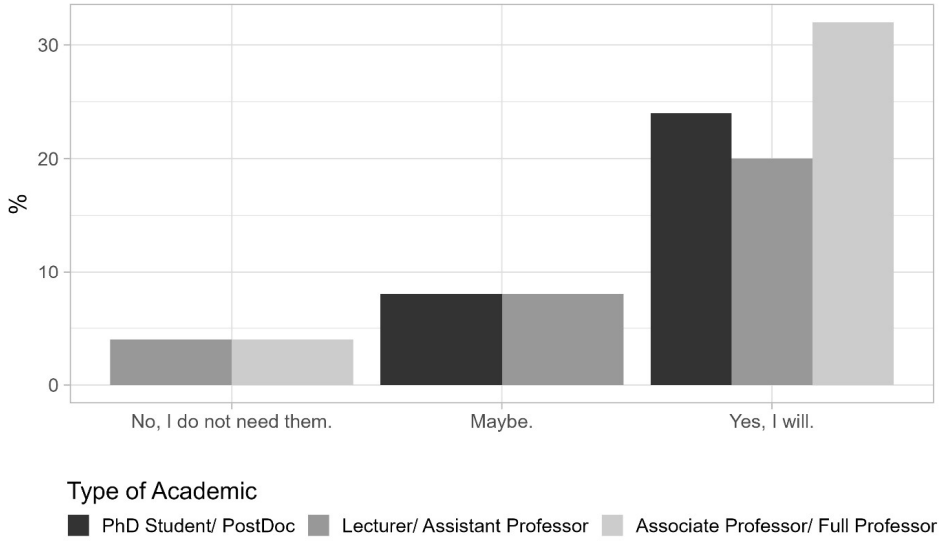


Figure 9: Possible popularity of the Lab

We use these insights to tailor our services. In the following sections, we will focus on describing what the Lab has done through the Lab main services (Learning and Collaboration, Tools and Resources, and Support writing Grants and Research Proposals) while relating them to the researchers computational needs.

SERVICE 1: COLLABORATION AND LEARNING

At the Lab, we aim to serve as a nexus for connecting various relevant skills in the area of computational analysis. For this, we promote computational knowledge exchange, and foster global connections. However, we know that in contemporary academia, scientists have little time to invest in learning something new. Therefore, we offer three options: Collaboration, Self-Study Material, and R Workshops.

COLLABORATION

The Lab aims to foster a collaboration network where we and other Lab members can speed-up science by exchanging ideas, knowledge, and skills. With this aim, we try to organize at least one main event per year, and we also founded the Film Club. To help organize these events, we hired the PhD candidate Gabriella Bollici as events co-organizer (2 hours per month).

ETHICS OF AI AND LAB'S TOOLS INTRODUCTION

Assistant Prof. Douglas Parry
(Communication Department)

"**The Ethics of AI event** hosted by the Societal Analytics lab **provided a clear overview of the Lab's tools and services**, along with a practical illustration of how these resources can be used in text-analysis research. **Dr Katleen Gabriels' keynote offered a thoughtful perspective on the ethical questions raised by generative AI**."

For the main event, we organized the "Ethics of AI and Lab's Tools introduction" on May 19th 2025. We had an attendance of 29 people. As main speaker, we invited Dr. Katleen Gabriels[1] to talk about the Ethics of AI. During her talk Dr. Gabriels focused on how Large Language Models (LLMs) are increasingly being used for very personal and intimate occasions. In the talk, she discussed the possible implications for society.

[1] Dr. Katleen Gabriels is a moral philosopher and a philosopher of technology, specialised in computer ethics (e.g., AI; monitoring technology; VR). She is an associate professor of philosophy at Maastricht University. She researches the co-shaping of morality and computer technologies. From August 2020 to September 2024, she was the programme director of the interdisciplinary BA Digital Society. In December 2024, the Dutch research council (NWO) awarded her an Aspasia grant to study the co-shaping of morality and monitoring technologies in families. More information about her and her work can be found https://www.maastrichtuniversity.nl/nl/k-gabriels. Accessed November 18th, 2026.

During the event, we also kicked-off of the Lab server together with some of the AI tools the Lab makes available to the researchers through it: AmCAT, Whsiper, LIWC, OpenWeb-UI + Ollama, and Prodigy. We also invited two Lab members (Johannes Gruber and Daniel Preciado) to share their experiences using the Lab's tools.

THE FILM CLUB

The Film Club aims to be an event that brings together researchers every two months to discuss research relevant films. This event aims to serve as an incentive to get to know people from other departments with similar research interests. During 2025, we organized two sessions.

- June 24th 2025: We watched "Picture a Scientists". 5 people attended.
- September 16th 2025: We watched "Feels Good Man". 10 people attended.

SELF-STUDY MATERIAL

R-MATERIAL

As R is the preferred statistical and programming software (Figure 6), we maintain a repository and a book with R course material. The topics range from data wrangling to text analysis and advanced statistics. These materials are openly available for anyone interested. If researchers have questions, they can always get in touch with the Lab. All the information is here:

- R course material: https://github.com/ccs-amsterdam/r-course-material/tree/master
- Communication Science R Book: https://vu-communication-science.github.io/R-canon/

The material was initially developed by the Lab directors, but, nowadays, it is also maintained by other Lab members. As such, the Lab started a collaboration with the VU Library, specifically

with Charles Greene[2] to continue updating the material and to help promoting its use across faculties.

BLOG POSTS

The Lab has also made available a blog post to help researchers access VU IT resources. This is necessary because VU FSW researchers do not use Linux, which is the main VU IT Operating System. These creates a challenge for the VU FSW researchers, as VU IT tends to write their manuals for Linux users. Therefore, the Lab translated the manuals for non-Linux users (Gil-Clavel, 2025b).

BOOKS

During this year, the Lab required three books about the societal implications of using Artificial Intelligence. The books costs were covered by the VU Library.

1. Benjamin, Ruha. (2019). *Race after technology: abolitionist tools for the new Jim code*. Polity.
2. Crawford, K., & Gallagher, L. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Unabridged. Tantor Media, Inc.
3. Olson, Parmy. (2024). *Supremacy: AI, ChatGPT, and the Race That Will Change the World*. St. Martin's Press.

R WORKSHOPS

We offer R-workshops specially designed for social scientists (Figure 10). They were developed by the Lab manager (Dr. Sofia Gil-Clavel), who has many years of experience teaching R to demographers and social scientists from different backgrounds (qualitative and quantitative). Figure 10 shows its build up. All the materials can be found here: https://github.com/SofiaG1l/R_Course. During its first edition (between February and June 2025), 15 people attended the workshops. Participants gave good rates to the workshops.

[2] Research Software Engineer at VU Library Team Research Support. https://vu.nl/en/about-vu/divisions/university-library/more-about/team-research-support. Accessed November 18th, 2026.

Martina Stanková, PhD candidate
(Political Science Department)

"I really enjoyed the workshops organized by the Societal Analytics Lab, led by Sofia. As a researcher with some pre-existing, though minimal, knowledge of R and computational text analysis, **I found these sessions extremely informative and engaging.** I learned about building and cleaning datasets, visualizing data and results, supervised and unsupervised learning, and the tools available to computational researchers at our faculty. **The way the sessions were structured, and how Sofia challenged me to problem-solve and find my own answers, stood out compared to other R-focused seminars I have attended.** On top of that, Sofia did a great job as a lecturer; she explained not only the main techniques in code but also the background information, helping me understand not only what to write but why to write it instead of an alternative. **By the end of these sessions, I had a much clearer understanding of what I would like to do in my analysis and the skills and tools I need to accomplish it.** I highly recommend this workshop to anyone interested not only in text analysis but in R more broadly, and I am looking forward to attending future workshops led by Sofia and the Societal Analytics Lab."

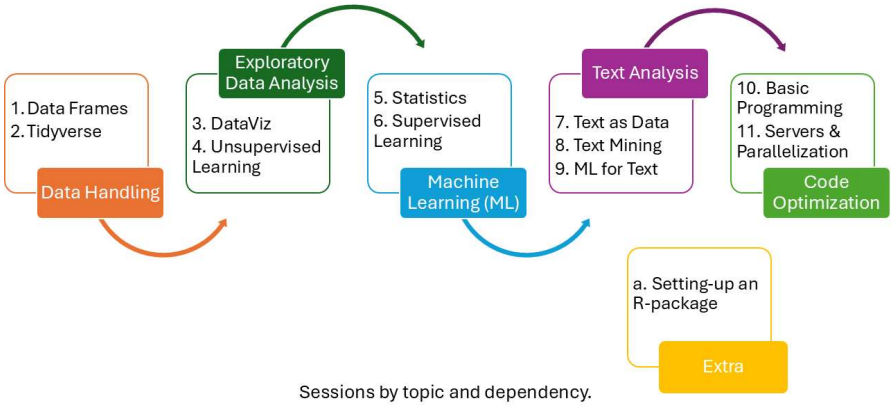


Figure 10: R-workshop for academics.

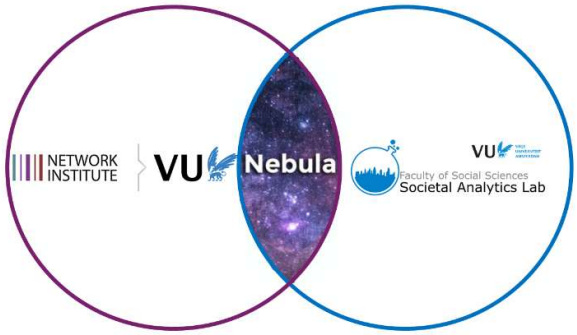
SERVICE 2: COMPUTATIONAL RESOURCES AND TOOLS

The Lab offers a range of computational resources for VU FSW researchers, and it is involved in the development and maintenance of various digital infrastructure projects. Our main goal is to facilitate the use of big multimodal (text, images, audio, and video) data and computational methods (Figure 1). For these goals, we acquired computational resources and tools, and we offer software as a service.

COMPUTATIONAL RESOURCES

In terms of resources, in 2024 the Lab invested on its own servers: one to host AmCAT and another to host Whisper and OpenWeb-UI + Ollama. Both servers continued to be maintained during 2025 by the Lab team. During 2025, the Lab also started a collaboration with the Network Institute in order to improve users' access to more powerful Large Language Models (LLMs).

NETWORK INSTITUTE COLLABORATION



To give researchers access to more powerful LLMs, the Societal Analytics Laboratory and the Network Institute have started a collaboration. Both groups invested equal amounts to buy more and more powerful GPUs. The groups also agreed that the Network Institute will take care of handling these resources to supercharge Nebula (Network Institute, 2025).

COMPUTATIONAL TOOLS

As a way to facilitate VU FSW researchers access to safe machine learning and artificial intelligence tools, we acquired and installed computational software in the aforementioned servers.

Prof. Damian Trilling
(Journalism Studies)

"There are many tools available nowadays that make a researcher's live easier. When I was a student, we were transcribing interviews by hand; now this can be automated. There are **commercial tools** for this, but **they can be expensive to use** and not everyone has always a budget that such costs could be charged to. Especially **with sensitive data, one** also **rather does not want them to leave the VU** for an online service. One can use AI-based tools like Whisper for this, but doing so on your own computer is not always easy, but in particular quite slow. **Thanks to the Social Analytics lab hosting such tools, I can just quickly get a transcript, without thinking too much about the technical details** (which I love doing as well, but don't always have the time or resources to do)."

The computational tools the Lab provides are:

- AmCAT: AmCAT makes it easy to upload, analyze, and sample big data. In it, users keep full control over their data, while still being able to share it, to keep access control, and to export it when needed. More info here: https://amcat4.labs.vu.nl/
- Whisper: Whisper is an automatic speech recognition system. It enables transcription in multiple languages, as well as translation from those languages into English. More info here: https://openai.com/index/whisper/
- OpenWeb-UI + Ollama: They are open-source projects that serve as powerful and user-friendly platforms to run Large Language Models. More info here: https://docs.openwebui.com/.

We also purchased one license of each of the following programs:

- Prodigy: It is an efficient annotation and training tool that enables entity recognition, detection, image classification, and more! More info here: https://prodi.gy/
- LIWC: It is the gold standard in software for analyzing word use. It can be used to study a single individual, groups of people over time, or all of social media. More info here: https://www.liwc.app/

Figure 11 shows how these tools are integrated with the computational methods pipeline of Figure 1.

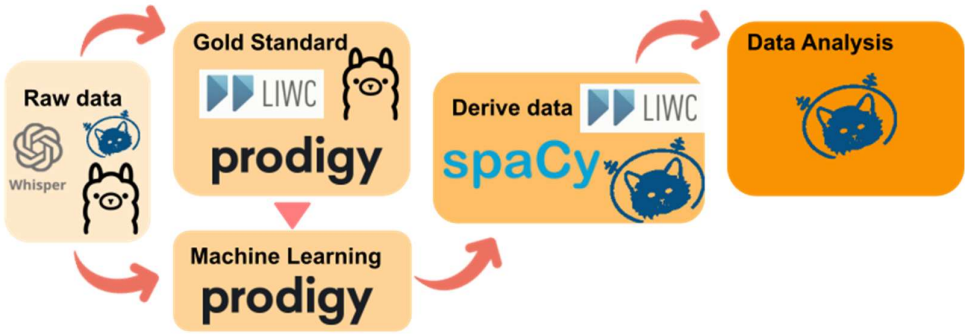


Figure 11: Computational tools integrated into the computational methods pipeline

In terms of users, Table 3 shows the number of people accessing the tools.

Table 1: Number of users per computational tool.

Computational Tool	Number of registered users
AmCAT	10
Whisper	5
OpenWeb-UI + Ollama	18
Prodigy	1
LIWC	0

SOFTWARE AS A SERVICE

As a way to continue contributing to the research community, we offer software as a service. The Lab is responsible for developing and maintaining AmCAT (Amsterdam Content Analysis Toolkit) (Gil-Clavel et al., 2025) and Annotinder (Welbers, 2022/2025).

During 2025, we continued to improve AmCAT by working on its front- and back-end developing, as well as on its users' experience. For the latter, we hosted a focus group to have a better understanding of what constraints users encounter when they work with AmCAT. We also tailored the AmCAT book (Gil-Clavel et al., 2025) for a social science audience and we wrote an easy to follow blog-post (Gil-Clavel, 2025a). The Lab also started to collaborate with the Monitoring Electoral Democracy[3] (MEDem) consortia to continue developing AmCAT, and to make it possible to integrate it with other software. This is done through our collaboration with the service center for text and multimedia data[4].

In the case of Annotinder, the Lab is working on integrating it into AmCAT. Towards this end, the Lab applied for an OpenScience grant[5] which we got (*Observatory for Political Texts, Images and MultimediA (OPTIMA)*, 2025).

[3] https://www.medem.eu/. Accessed November 18th, 2025.
[4] https://www.medem.eu/people/#service-center-for-text-and-multimedia-data. Accessed November 18th, 2025.
[5] https://www.openscience.nl/en/calls/open-science-infrastructure. Accessed November 18th, 2025.

SERVICE 3: GENERAL CONSULTANCY AND SUPPORT WRITING GRANTS AND RESEARCH PROPOSALS

The Lab offers VU FSW researchers support to write grant proposals with a computational component. Specifically, the Lab offers consultancy on what methods are appropriate and competitive, provides feedback on feasibility and formulation, and supports the researcher with resources and expertise. Depending on the amount of help provided, researchers may need to acknowledge the Lab and/or the Lab member that is helping them. This depends on the amount of time the Lab invests and it can go from just writing a line in the acknowledgments section to adding a Lab team member as co-author.

Assistant Prof. Ana I. Loureiro Lopes
(Communication Department)

"There are many things to appreciate about his lab, but **I would highlight the responsiveness and the quality and range of the offered activities**. Every time I've needed assistance I have received fast and helpful replies."

During 2025, the Lab gave 2 research grant consultancies, and answered 13 computational general questions (Figure 12).

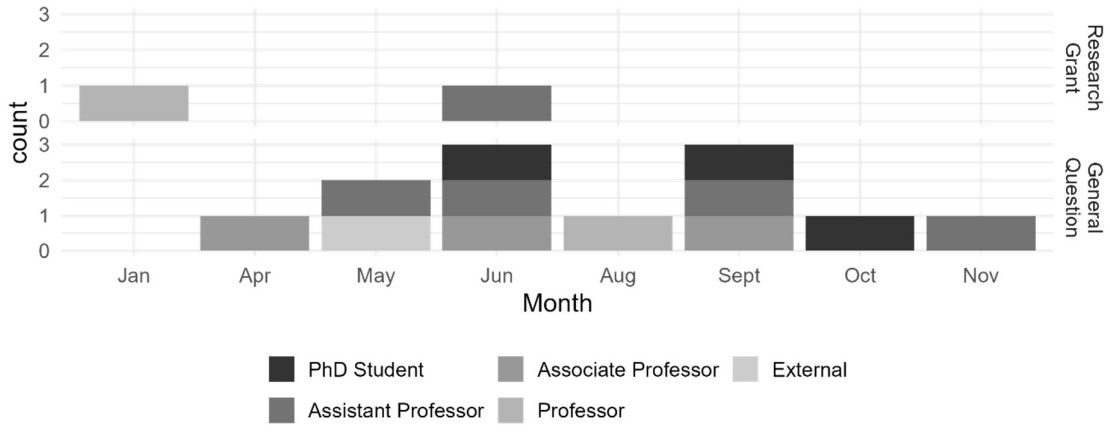


Figure 12: Number of unique consultancies given during 2025. Values are broken down by type of consultancy and academic.

INVESTMENT

For this service, the Lab did not incur in any costs.

A GLIMPSE INTO 2026

During 2026, the Lab will continue to work on its goals: foster methodological innovation, expand VU research infrastructure, increase Open Science through better reproducibility and transparency, and multiplicate interdisciplinary research teams. For this, the Lab will further promote and invest in its three core services: Collaboration and Learning, Computational Resources and Tools, and General Consultancy and support writing Grants and Research Proposals. So far, for 2026, we already have a couple of items prepared.

SERVICE 1: COLLABORATION AND LEARNING

LAB'S BEGINNING OF THE YEAR EVENT

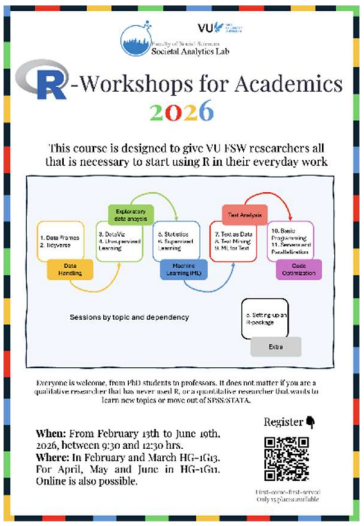
During the event, the Lab will present all what we have done and accomplished during the year 2025. But hey, a celebration is not complete without some fun activities! So, besides the presentation, the Lab will also screen the movie "Coded Bias" and finalize the event with drinks. More info here:

https://societal-analytics.nl/events/20251216_StartYear/

R WORKSHOPS

During the first semester of 2026, the Lab will continue to offer the R-Workshops for Social Scientists. As such, we have a full agenda prepared:

1. February 13th: Introduction to R, R-Studio, and data frames
2. February 27th: Introduction to R packages, Tidyverse, and R-Markdown
3. March 13th: Introduction to data visualization using ggplot2
4. March 27th: Introduction to unsupervised learning
5. April 10th: Introduction to statistical analysis with base R
6. April 24th: Introduction to supervised learning
7. May 8th: Introduction to text as data



8. May 22nd: Introduction to text mining with tidytext
9. June 5th: Introduction to machine learning for text with textrecipes
10. June 19th: Introduction to basic programming with R

MASSIVE OPEN ONLINE COURSES

Given the financial struggles universities are undergoing, the Lab will start working towards bringing a constant stream of money. For this, we will explore the possibility to offer paid (online) courses to people outside VU FSW.

SERVICE 2: COMPUTATIONAL RESOURCES AND TOOLS

FUNDING FOR BUYING OR ANNOTATING DATA

The Lab invited all the VU FSW researchers affiliated to the Lab to apply for funding for buying or annotating data. The Lab allocated 15,000EUR in grants that require up to 5000EUR. The proposals were examined by the two directors of the Lab and two independent members of VU:

- Mona Katharina Wallraff (Knowledge Valorisation Officer, VU)
- Charles Greene (Scientific Information Specialist, VU Library)

The main points to evaluate are:

- How well the project aligns with the Societal Analytics Lab objectives.
- How well the project applies computational methods in social science research.
- Impact and broader utility of the expected output
- Clarity of the proposal and research plan.
- Feasibility of the plan and timetable.

Applicants also needed to promise that the resulting data and/or annotated data will be archived as FAIR and as open as possible, and that its documentation will follow the "Datasheets for datasets" (Gebru et al., 2021). We believe that adding these requirements fosters methodological innovation, and increases Open Science through better reproducibility and transparency.

OBSERVATORY FOR POLITICAL TEXTS, IMAGES AND MULTIMEDIA (OPTIMA)

The Lab will use the grant of 250,000 EUR to work on the project "Observatory for Political Texts, Images and MultimediA (OPTIMA)" (*Observatory for Political Texts, Images and MultimediA (OPTIMA)*, 2025). This project is funded by the Open Science Infrastructure.

OPTIMA aims to advance open science by creating an innovative, collaborative infrastructure for securely storing, analyzing, enriching, and sharing multimedia data. In general, what this means is that we will improve and integrate AmCAT (Gil-Clavel et al., 2025) and Annotinder (Welbers, 2022/2025).

REFERENCES

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Iii, H. D., & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, *64*(12), 86–92. https://doi.org/10.1145/3458723

Gil-Clavel, S. (2025a). *Amsterdam Content Analysis Toolkit (AmCAT)*. https://societal-analytics.nl/blogs/20250101_amcat/

Gil-Clavel, S. (2025b, January). *Constrained by computing power? We can help you!* [Societal Analytics Laboratory]. https://societal-analytics.nl/blogs/20250201_computing-power/

Gil-Clavel, S., Gruber, J. B., Welbers, K., & Atteveldt, W. van. (2025). *AmCAT-Book* (Vol. 2). https://amcat.nl/book/

Network Institute. (2025). *Nebula: Empowering VU research with open-source generative AI | Network Institute*. https://networkinstitute.org/nebula/

Observatory for Political Texts, Images and MultimediA (OPTIMA). (2025, December). Vrije Universiteit Amsterdam. https://research.vu.nl/en/prizes/observatory-for-political-texts-images-and-multimedia-optima/

Welbers, K. (2025). *Ccs-amsterdam/annotinder-r* [R]. Computational Communication Science Amsterdam. https://github.com/ccs-amsterdam/annotinder-r (Original work published 2022)