CSSci

STRUCTURES: Applying Responsible AI to Reduce Inequality

Version 1: 15 January 2025

Note: This course manual may be updated during the course. Updates will be communicated via Canvas as well. Make sure you always use the most recent version of this manual. See final page of this manual for a list of updates.

Title of Course

STRUCTURES: Applying Responsible AI to Reduce Inequality

Course Catalogue Number

738200002Y

Credits

30 EC

Entry Requirements

The entry requirements for the course are specified in the Course Catalogue, available online at https://studiegids.uva.nl/xmlpages/page/2024-2025-en/search-course/course/122408.

Instruction Language

English.

Time Period(s)

Academic year 2024-2025, semester 2.

Location

Please see the UvA schedule webpage https://rooster.uva.nl/.

Video Recording of Lectures

None.

Examiners

Prof. dr. S. (Sennay) Ghebreab (s.ghebreab@uva.nl)
Prof. T. (Tobias) Blanke (t.blanke@uva.nl)

Core Lecturers

dr. Roland Adorjani
dr. Lukas Ansteeg
dr. Roberto Cerina

dr. Kunhe Li
dr. Kirils Makarovs
dr. Regina Nockerts (semester coordinator)
dr. Petter Tornberg

Course Content

You will study political science theories to assess structural issues, building on the knowledge that was acquired in the Foundation, Building Blocks, and Connections courses. You will further deepen your knowledge of qualitative and quantitative research methods and acquire skills in machine learning and prediction.

A major part of the semester is organised around a group project (theme: social inequality and algorithmic decision making). Students will work in small project groups to address a real-world problem identified by CSSci staff and using real-world data. In the course of this project, students will:

- Design a viable research question that could be approached by the means of qualitative and quantitative research methods.
- Work with a given dataset to create robust predictions.
- Assess that prediction from a critical perspective and implement techniques to reduce bias.
- Create professional presentations of their project and its deliverables.
- Present the intervention plan to the CSSci community.

Additionally, there will be five individual assignments to reinforce key aims of the course:

1. Essay on Algorithms and Social Inequality.
2. Future Scenario Planning.
3. Prototype Ethics Tool.
4. Coding Portfolio and Explainer Session.
5. Individual contribution to Group Project V.

Please note the week 40 of June 02 - 06 is reserved for collaborative group work. Although there may not be formal classes scheduled during this period, students shall be present, prepared, and participating in group work during this time. **Absence during non-designated vacation periods is incompatible with the program's essence** (with exception for extenuating circumstances, see the Force Majeure policy).

Teaching Methods/Learning Formats

Teaching and learning take place through a set of formats, including check-ins, lectures, workshops, practicals, and group work. You can find additional information on Canvas.

Course Evaluations & Adjustments of the Course

This course was last conducted in Semester 2 of the 2023-2024 academic year. A summary of the changes that have been made is as follows:

- All assignments have been redesigned.
- New datasets have been provided for the group project.
- Lectures and workshops on social science theory have been redesigned.
- Reduced focus on regression.
- ILOs updated to reflect all changes.

Professional Study Attitude

Computational Social Science is organised along a “just in time learning” approach, which means that the weekly content is directly applicable to the project work at that time. This

means that it is crucial to participate in all the sessions (check-in, workshops, lectures, and so on) and all scheduled meetings and events with external partners. Your well-prepared, on-site participation is expected for all learning activities and inherently reflected in the programme's exit qualifications. In the event of a learning activity that you have missed, it is your responsibility to make up for the content that was covered during your absence. In case of absence, you must always inform your Core Lecturer beforehand or as soon as possible.

Requirements

In the Computational Social Science programme, students work together, in student groups and with external partners. You must therefore demonstrate a professional study attitude (the four P's):

1. Present
2. Prepared
3. Participating
4. Proper and clear communication

After all, absence or low-quality participation from one team member will impair the ability of other group members to excel and perform well. Additionally, a student's own learning is inhibited by an insufficient professional study attitude.

Consequences

General lack of participation

If a student cannot take part in a learning activity, they should inform their Core Lecturer. It is the student's responsibility to catch up on the contents that were covered.

If the Core Lecturer notices patterns of absence or low-quality participation inhibiting a student's learning process, the Core Lecturer will request an explanation from the student. The Core Lecturer will schedule a meeting with the student to determine:

- If the student can still meet the participation requirement.
- Which pattern of professional study attitude the student needs to develop.
- In which timeframe this improvement in performance needs to be sufficiently demonstrated.

The Core Lecturer will make a reasonable attempt to schedule this meeting with the student; should the student be unresponsive, or a time and location cannot be agreed upon in a timely manner, the Core Lecture will instead issue a written warning covering the same points.

A summary of this meeting will be shared with the student via email and counts as a first written warning (if an in-person meeting was unable to be scheduled, the original written warning will serve and no additional summary will be sent). In some cases, the Core Lecturer may refer the student to the programme's Study Adviser to seek guidance.

In case the student has not sufficiently demonstrated improvement in performance during the set timeframe, a second written warning is issued. This email is also shared with the examiners of the semester course, the study adviser and the programme's Examinations Board.

Lack of immediate adequate further improvement of performance or accountability from the student will lead to an 'Unacceptable' result for the individual assignment 'Individual Contribution to Group Project'. As a consequence, the student will not be able to complete the semester course.

Group work

If, in the informed judgement of the Core Lecturer, the (lack of) participation of a student in a group project is such that the student has not contributed to the project at all stages of its development, that student alone may receive an 'Insufficient' result for the group assignment.

Failure to fulfil the participation requirements, failure to demonstrate professional study attitudes, or continuing and serious disruptive behaviour may also result in a student being removed from their project group or not assigned to new project groups. This student receives an 'Unacceptable' result for the individual assignment 'Individual Contribution to Group Project' and will not be able to complete the semester course.

Lateness Policy

Teaching and learning and students and staff participating in these activities, are disrupted and distracted when persons arrive after the starting time (late). In case a teaching and learning activity has commenced (e.g. if the door is closed and/or if a person is given instructions or a presentation), a student arriving late should not enter the room or activity. In this case the student will need to wait until the break before entering the room.

Force Majeure

When students are unable to fully participate because of reasons beyond their control, they should notify the Study Adviser and their Core Lecturer as soon as possible. Notification should be given ahead of time if the absence is foreseeable, or as soon as possible in case of unforeseen circumstances such as illness. The programme will make serious efforts to facilitate these students and prevent unnecessary delays. If students are not able to participate in graded assessments due to reasons beyond their control; or they wish to ask for a deadline extension or permission to participate in the resit/repair opportunity without having made Deadline Extension Request Procedure for Graded Assignments (see below).

Assessment

In the Computational Social Science programme, assessment is geared towards learning. Most assignments are formative, which means that the students receive a progress check report to further improve and learn, but do not receive a grade. In addition, there are summative and therefore graded assignments. These assignments receive a descriptive grade but - in principle - no feedback beyond an explanation of the grade.

Computational Social Science uses the UvA descriptive grading scheme: Excellent, Good, Satisfactory, Sufficient, Insufficient, and Unacceptable. See:
https://student.uva.nl/en/topics/assessment-of-your-grade-point-average

Grades will be determined in no longer than 15 working days.

The main part of each semester is the group work. In addition, each student is assessed on the four learning trajectories Social Science and Humanities Expertise, Changemaking Expertise, Research Expertise, and Digital Expertise. **All students are required to score at least sufficient for these individual assessments to pass the course**.

Students are also individually graded for their contribution to group projects. This assessment is based on roles and responsibilities as specified in a Team Charter, and the student's active and professional participation in group activities.

In case a group receives an insufficient final grade for the project(s), students can compensate this with their individual assignments. This is calculated using the grade points

attached to the UvA descriptive grades (see link above) and the weights mentioned in the table below.

All assignments need to be handed in through Canvas. Inspection of the graded work will be possible by appointment with the relevant Core Lecturer in the week following the release of grades. **All assignments must be submitted as static files**; links to online files/storage will only be accepted in extraordinary cases and with preapproval by the Core Lecturer.

Table 1: Graded Assessments
* Items marked with an asterisk are required components for a serious attempt at the assignment.

Graded Assessment	Date / Deadline	Weight (%)
Individual Assignments		
SSH: Essay on Algorithms and Social Inequality		10%
Final Submission*	Wednesday, May 28, 2025 (17:00)	
RE: Future Scenario Plan		10%
Final Submission*	Wednesday, May 28, 2025 (17:00)	
DE: Code Portfolio and Explainer Session		10%
Code Submission - Portfolio Notebook 1* - Portfolio Notebook 2* - Portfolio Notebook 3* - Portfolio Notebook 4* - Portfolio Notebook 5* - Portfolio Notebook 6* - Portfolio Notebook 7* - Portfolio Notebook 8*	Monday, March 31, 2025 (17:00)	
Code Explainer Session*	Monday, April 7th and Tuesday April 8th, 2025 Time by appointment.	
CME: Prototype Ethics Tool		10%
CME Sprint Days*	10 April 2025, Thursday (9:30 – 17:00) 11 April 2025, Friday (9:30 – 17:00)	
Prototype Testing Notes*	10 April 2025, Thursday (17:00) 11 April 2025, Friday (17:00)	
Prototype of Ethical Tool*	17 April 2025, Thursday (17:00)	
Individual Contribution to Group Project V		10%
Attendance at Opening Event*	Friday, Feb. 07, 2025 (15:00 – 17:00)	
Attendance at Camp Week*	February 3 – 7th, 2025 (09:00 – 17:00)	
Progress Check 1	Wednesday, March 05, 2025 (17:00)	
Progress Check 2	Wednesday, April 2, 2025 (17:00)	
Final Submission	Friday, June 06, 2025 (17:00)	
Attendance at Closing Event*	Tuesday, June 03, 2025 (10:00 – 14:00)	
Group Project V		
Code and Data Submission*	Thursday, June 5, 2025 (17:00)	50%
Algorithm Audit Report*	Thursday, June 5, 2025 (17:00)	

Rules Regarding Resits

In the case of an 'Insufficient' for a graded assessment following a serious attempt, one resit or repair possibility will be offered before the end of the semester. If no serious attempt has been made by a student, the grade will be registered as 'Unacceptable,' and no resit or repair opportunity will be possible (see the *Serious Attempt* section, below).

Note that there are two weeks reserved before the end of the semester for repairs or resits. The grade for all repair or resit assignments is capped at "Sufficient."

Table 2: Resit and Repair Deadlines

Eligible Graded Assessments	**Repair or Resit?**	**Date / Deadline**
DE: Code Portfolio and Explainer Session	Repair: Code Portfolio	May 19, 2025 (17:00)
	Resit: Explainer Session	June 6, 2025 (time by appointment)
SSH: Essay on Algorithms and Social Inequality	Repair	June 26, 2025 (17:00)
RE: Future Scenario Plan	Repair	June 26, 2025 (17:00)
CME: Prototype Ethics Tool	Repair	June 20, 2025 (17:00)
Individual Contribution to Group Project V	Resit	June 27, 2025 (17:00)

Serious Attempt

Students must make a "serious attempt" to complete every required component of a graded assignment in order to receive a grade for that assignment. A "serious attempt" is the deliberate and sustained attempt to fulfil the requirements of a task or project to the best of one's ability and in a timely manner. A "component" refers to a major aspect of the assignment, such one of the eight iPython notebooks for the DE coding portfolio or the "code and data submission" of the Group Project. (For example, a single assessment criterion within a larger written work or a single question within a code notebook would not, on its own, be considered a "component.") Required components are indicated in the "Graded Assessment" table in this Course Manual.

The teaching staff will determine if submitted work constitutes a serious attempt, looking at aspects such as (but not restricted to) the scale of the work, depth of work, alignment with stated requirements, utilization of available resources (including Progress Check opportunities), seriousness of revisions, and/or effort at self-reflection. If no serious attempt is made, the student will receive a grade of "unacceptable" on that assignment, be ineligible for repair of that assignment, thus ending their participation in the course.

Failure to submit ***any required component*** of an assignment is also considered to demonstrate a lack of serious attempt and the student will be given "NAP" (not submitted) for the assignment.

Description of Graded Assignments

Group Project V: The Unequal Machines Grand Challenge

The Unequal Machines Grand Challenge project is designed for you to engage simultaneously with computational social science and inequality research. You are encouraged to study the complex dynamics of inequality through the lens of rich datasets

and computational methodologies. This description aims to provide an understanding of the structure and requirements of this project.

The final project deliverable will consist of:

- Demonstration of a socially responsible predictive algorithm;
- Written audit report demonstrating potential biases in the analysis and the measures that were taken to address those biases;
- Code and data submission.

Individual Contribution to Group Project V

When participating in a group project, it is important for each individual to contribute effectively to the success of the project. This requires a professional attitude towards one's colleagues in the group, challenge partners and lecturers, as well as commitment to fulfilling the expectations outlined in the Team Charter.

The Team Charter is a document that describes the goals, responsibilities, and expectations of each group member, as well as defines the agreed mode of work in terms of communication and conflict resolution. The Team Charter will be revisited on a regular basis during the check-in meetings to evaluate the performance of the groups; the contents of the Team Charter could be updated during the Semester, if necessary. By behaving professionally and complying to the specifications of the Team Charter, the student will demonstrate their commitment to the success of the group project and help ensure that it is completed to a high standard.

Assessment will be determined by the Core Lecturer by evaluating student's performance over the project against the specifications of the Team Charter; the Core Lecturer will also take into account the feedback provided by other group members and own observations regarding the performance of the student.

The resit opportunity for these assignments includes three components and has a 2-week timeline:

1. Reparations: expected to take ~8 days for an average student to complete.
2. Reflection: expected to take ~1 day for an average student to complete.
3. Plan: expected to take ~1 day for an average student to complete.

Of particular importance is the first, Reparations, component. This is a substantial piece of work, the outlines of which must be agreed on between the student to take the resit and their Core Lecturer before the student begins working on it. It is the student's responsibility to reach out to their Core Lecturer to begin the planning process for this work as early as possible after the need for the resit becomes announced.

Individual Assignment SSH: Essay on Algorithms and Social Inequality

Students are asked to explore an issue of their choosing in which algorithmic decision-making has substantial impact on social inequality.

In Part A, you will first argue for an existing theoretical framework that justifies the current status quo in algorithmic decision-making. For instance, you may consider the Rational Actor Theory or similar frameworks. This part should take the form of an annotated bibliography that traces the origins and development of the current system or status quo.

In Part B, you will critique the theoretical framework from a critical theory perspective, such as feminist theory, critical economics, or post-colonialism. The critique should assess how algorithmic decision-making reproduces ideational and material structures and how stakeholders may confront these dynamics. Students may draw on concepts from lectures and workshops but are expected to conduct substantial independent literature reviews.

Individual Assignment CME: Prototype Ethics Tool

In this assignment you will prototype a web-based intervention to inform users about an ethical dilemma posed by an emerging technology. To do so, you will:

1. Select an emerging technology that you think poses a serious ethical dilemma.
2. Motivate your designated target audience to care about this ethical dilemma.
3. Identify the ethical framework(s) you will use to guide the development of your tool, and explain how that framework informs our thinking on the dilemma at hand.
4. Create a web-based tool which:
    a. Highlights the dilemma, using common ethical framework/s.
    b. Helps users clarify their thinking about the related emerging technology.
5. Test out your peers' prototypes and reflect on their effectiveness.

Appropriate web tools could include, for example, an interactive map or infographic, a game, a quiz or profile-builder, a calculator, or a simulation. This is not an exhaustive list; we encourage you to be creative.

Individual Assignment RE: Future Scenario Planning

Students will develop a framework for a long-range forecast of a complex societal phenomenon associated with social inequality. The predictive framework will be grounded in probabilistic reasoning and aimed at identifying the most likely future scenario. Students will conduct desk research to justify the framework's components and estimate the probabilities of future events.

The predictive framework will be accompanied by a cover note that explains the rationale for the prediction, justifies the components of the framework, and interprets the modelling results. The cover note should also reflect on the limitations of the model, including a critical evaluation of potential biases in the sources used.

Individual Assignment DE: Coding Portfolio and Explainer Session

After each DE practical from weeks 2 to 7, students will be provided with a coding assignment. Your task is twofold:

1. Complete, and keep a well-annotated record of, the code and generated results from these assignments using a different dataset;
2. Familiarise yourself with the code and the method enough such that if you are questioned on any part of the coding or modelling choices, you are able to defend your code.

You will submit a copy of your coding portfolio for assessment. Then you will undergo your individual oral Explainer Session with a CSSci Core Lecturer; a TA or other staff member may also be present. Your job will be to:

1. Answer a series of questions related to your code, and/or
2. Provide accurate justifications for the chosen methodology's compatibility with your data, and/or
3. Provide accurate justifications for the chosen methodology's compatibility with the goals of the analysis, and/or
4. Comment on the results of your DE effort.

Assignment Deadlines

Assignment deadlines are strict. **Canvas will not accept submissions after the deadline**. If you miss the deadline for a graded assignment, you will not receive a grade (NAP), you are

not eligible for a repair opportunity, and you cannot complete the course. Your Core Lecturer cannot change this. If you are unable to submit your assignment in a timely manner due to extenuating circumstances, contact the coordinating Core Lecturer and send the Core Lecturer your assignment by email immediately.

When submitting an assignment, please allow sufficient time to deal with potential minor technical issues (problems with your computer or internet connection, time required to upload large files, etc.) as they are not considered a valid reason for missing a deadline.

Please be aware that Canvas **does** allow multiple submissions of your non-quiz assignments. If you are concerned about the deadline, we recommend that you first submit your deliverable as soon as you have what you consider to be a serious attempt at the assignment. You may then later - but before the deadline! - resubmit the finished version of your deliverable. The most recent submission of the assignment that is made before the deadline will be graded.

Deadline Extension Request Procedure for Graded Assignments

Deadline extensions are given only in exceptional, unforeseen circumstances. To request an extension of the deadline from the Examinations Board due to extenuating circumstances, such as illness, **please submit your request as soon as you are experiencing the circumstance** (e.g., at the start of the illness). **Do not wait until after you have missed the deadline**.

If the request is sent to the Examinations Board in the five days leading up to the assignment deadline, the board cannot guarantee that your request will be processed before the deadline. If the outcome of your request is still pending on the day of the deadline, please make sure to upload your (incomplete) assignment to ensure that you are eligible for participation in the repair opportunity should your request be denied, and in case of an Insufficient result. In case of extenuating circumstances, it is crucial to communicate clearly and timely about your situation with both the study adviser and your Core Lecturer, and if possible, document your circumstances with the study adviser.

In case of sudden, short and unforeseen emergencies on the day of the deadline, or the day before it, that require an extension of a maximum of three days, students should directly contact their Core Lecturer and the coordinating Core Lecturer for an extension request without submitting a request to the Examinations Board. Students can be asked for documentation to substantiate their request. Again, please make sure to upload your (incomplete) assignment to ensure that you are eligible for participation in the repair opportunity should your request be denied, and in case of an Insufficient result.

Academic Integrity

Rules Regarding Fraud and Plagiarism

The provisions of the Regulations Governing Fraud and Plagiarism for UvA Students apply in full. Access this regulation at https://student.uva.nl/en/topics/plagiarism-and-fraud

Plagiarism can take many forms, including (but not limited to!):

- Making use of or reproducing another person's texts, data or ideas without complete and correct acknowledgement of the sources.
- Presenting the structure or central body of ideas taken from third-party sources as one's own work or ideas, even if a reference to other authors is included.
- Submitting a text that has previously been submitted, or which is similar to a text that has previously been submitted, in the context of assignments for other courses.
- Reproducing the work of fellow students and passing it off as one's own.

If you have not read these regulations before, please do so! Note that plagiarism and fraud is extremely serious. In case of alleged plagiarism and/or fraud your supervisor is obliged to inform the Examinations Board immediately.

It is important to realize that plagiarism applies to all of your academic output, **including programming code**.

In general: DON'T EVEN THINK ABOUT IT.

Rules Regarding the Use of Artificial Intelligence Software

As we continue to advance in the digital era, we wholeheartedly embrace the positive contributions that artificial intelligence offers to your learning journey. An example of this is Large Language Models (LLMs), like ChatGPT. These can help adequately phrase ideas that you feed them, summarise texts they are explicitly fed, or help you when you encounter clearly defined issues in coding.

However, we must emphasize that LLMs are no substitutes for your intellectual efforts and critical thinking. Prudentially, they do not rely on any external criterion for truth - this means they may present made-up information as real (even inventing sources), remain vague in their answers and neglect key aspects, or produce buggy code. Ethically, they raise issues of claiming authorship for ideas that are not your own and may undermine responsibility for your own output. Hence, the development of original work remains an essential part of our educational mandate. Remember that LLMs are here to assist, not to replace your creative process.

Please also get acquainted with the UvA general information on the use of ChatGPT and AI tools in your studies: https://student.uva.nl/en/topics/chatgpt-and-your-studies.

If a student has made use of ChatGPT or a similar tool for an assignment, this should always be mentioned by the student and the tool should be cited as a source in the bibliography.

Open Research, Data Integrity and Storage

CSSci supports open research. Unless otherwise specified, the results of research will be shared with those partners and may be further distributed. CSSci may choose to share student results, reports, data visualizations, or other products for research or educational purposes. If a student objects, they must contact their Core Lecturer as soon as possible to explain why the research should not be made available.

It is therefore of the highest importance that research by students should be conducted responsibly, neatly, and archived for a period of two years. Both during ongoing research and for archiving purposes, all data should be stored with UvA storage facilities (OneDrive or another option approved by your Core Lecturer) and appropriately secured. You may not use 3rd party options (Google Drive or Dropbox, for example) to store any data which may contain personal information as per the GDPR.

For the rules, see the "Ethical Guidelines for Students" on the GSSS page: https://assets-eu-01.kc-usercontent.com/2a7f7854-b38e-015d-c63a-62b2c8fa2017/102d0011-670c-4d57-af37-2d613c2a1821/gsss-ethical-guidelines-for-students-version-2021-22.pdf

Personal Safety and Security

We want students to be assured they have a positive environment and a safe basis for their studies. We hope you never experience an unsafe situation or undesirable behaviour at the UvA but should this occur there are different individuals and institutions you can turn to (such as the study adviser and the programme director). Please see https://student.uva.nl/en/topics/safe-study-environment

Literature and Materials
Readings and additional materials are announced through Canvas.

Date of Final Grade
No later than 15 working days after the last graded assessment.

Programme / Weekly Planning and Deadlines
Weekly schedules can be found in Canvas and on rooster.uva.nl.

Learning Objectives
The activities of this semester together work towards the following learning trajectory objectives.

Social Sciences and Humanities Expertise
- Use relevant theories to recognize and explain structural inequality, bias and stratification.
- Apply relevant theories in order to identify sources of bias which constrain the development of socially responsible AI solutions.
- Analyse how design choices in AI reproduce ideational and material structures and how stakeholders may confront this.
- Evaluate the ethical, legal and societal impact of AI/algorithms.

Changemaking Expertise
- Explain socially responsible AI as a change-making method.
- Reflect on effective collaboration and communication with all kinds of stakeholders.
- Suggest scenarios on how new digital tools may influence individuals and society with probable desirable and undesirable effects.
- Evaluate the ethical aspects of AI/algorithm-based interventions, defending your position.
- Design and implement projects that are able to take power and values into account.

Research Expertise
- Explain the social issues related to predictive (recommendation) systems
- Apply predictive statistics to make predictions about future scenarios.
- Apply qualitative data to make predictions about future scenarios.
- Draw connections between qualitative and quantitative data in order to answer a research question.
- Evaluate the reliability of predictions using the adequate techniques for the relevant method.
- Design a research project which balances the ethical, legal and social aspects of the collection of digital behaviour and/or communication data, given their underlying assumptions and biases

Digital Expertise
- Understand how social biases manifest in digital data and predictive algorithms
- Demonstrate the application of machine learning and other digital techniques for prediction.
- Apply basic mathematics underpinning vector representations (PCA, factor analysis).
- Apply algorithm audits to identify potential data-driven fairness issues.
- Implement appropriate social bias mitigation techniques.
- Analyse digital data and predictive algorithms in order to detect social biases.

- Evaluate the technical and ethical implications of the choices made when implementing digital analyses, including bias mitigation.
- Select and justify a database management structure fitting the project at hand.
- Devise a database management structure suited for the project at hand.
- Design a bias mitigation strategy appropriate to the project at hand.

Updates to Course Manual

Table 3: Version History

Version	Date	Changes
1.0	15 Jan 2025	None

= END OF DOCUMENT =