

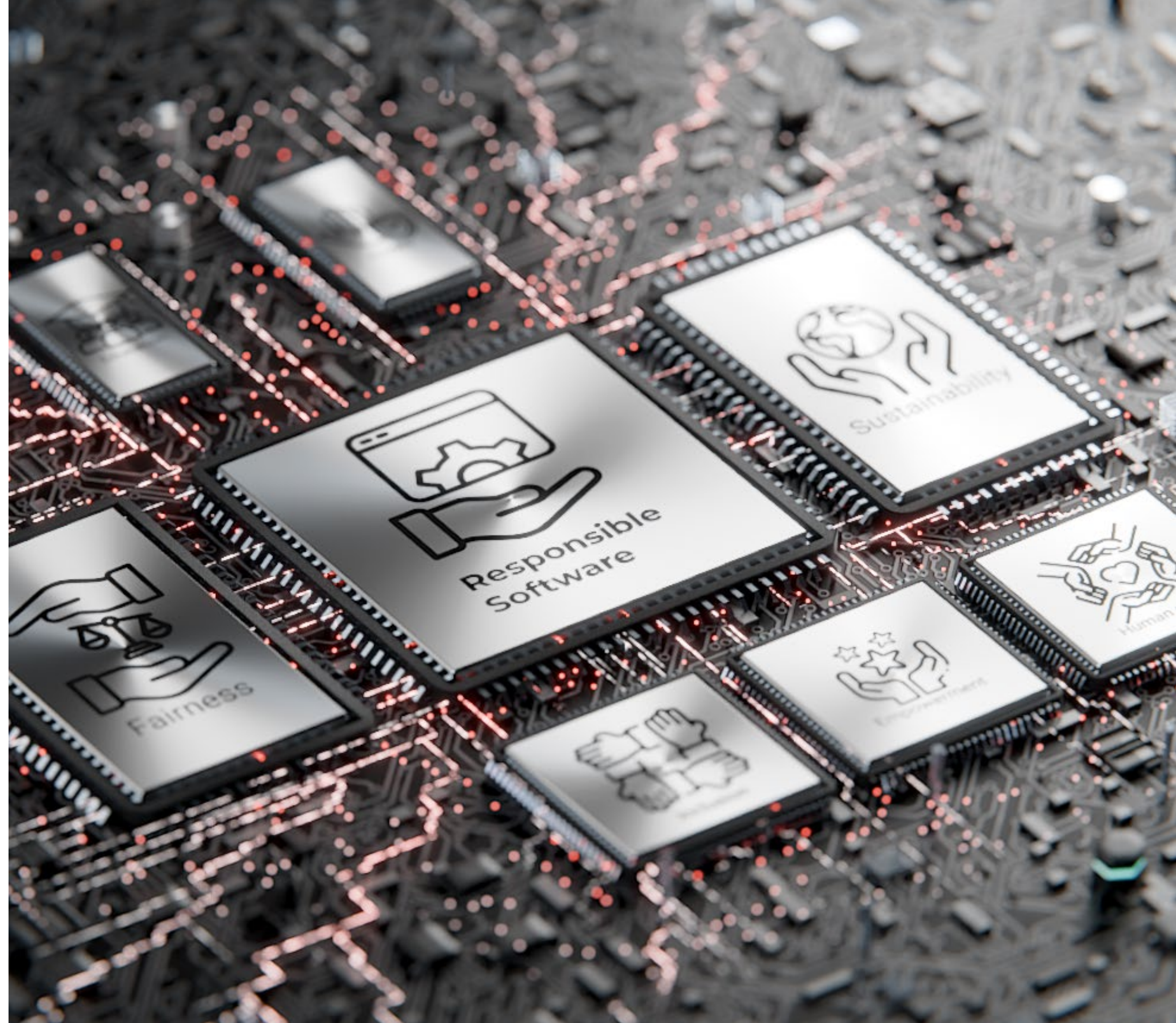
EPFL

Fairness 1 Review & Case studies

7 oct.

Cécile Hardebolle

**Responsible
Software**



Agenda for today

1. Graded assignment 1:

- a) Submission process
- b) Important advice

2. Interactive review questions on Fairness 1

3. Case studies:

- a) Analyzing values
- b) Inclusive design

Graded assignment 1 [reminder]

	Dates	Grading	Topics
Release	October 10	8% of total grade	Safety 1 & 2
Submission	October 14 at <u>23h59</u>	Coding questions + reflection questions	Fairness 1 & 2

■ Reminder of the rules:

- No GenAI
- No group work
- Use noto

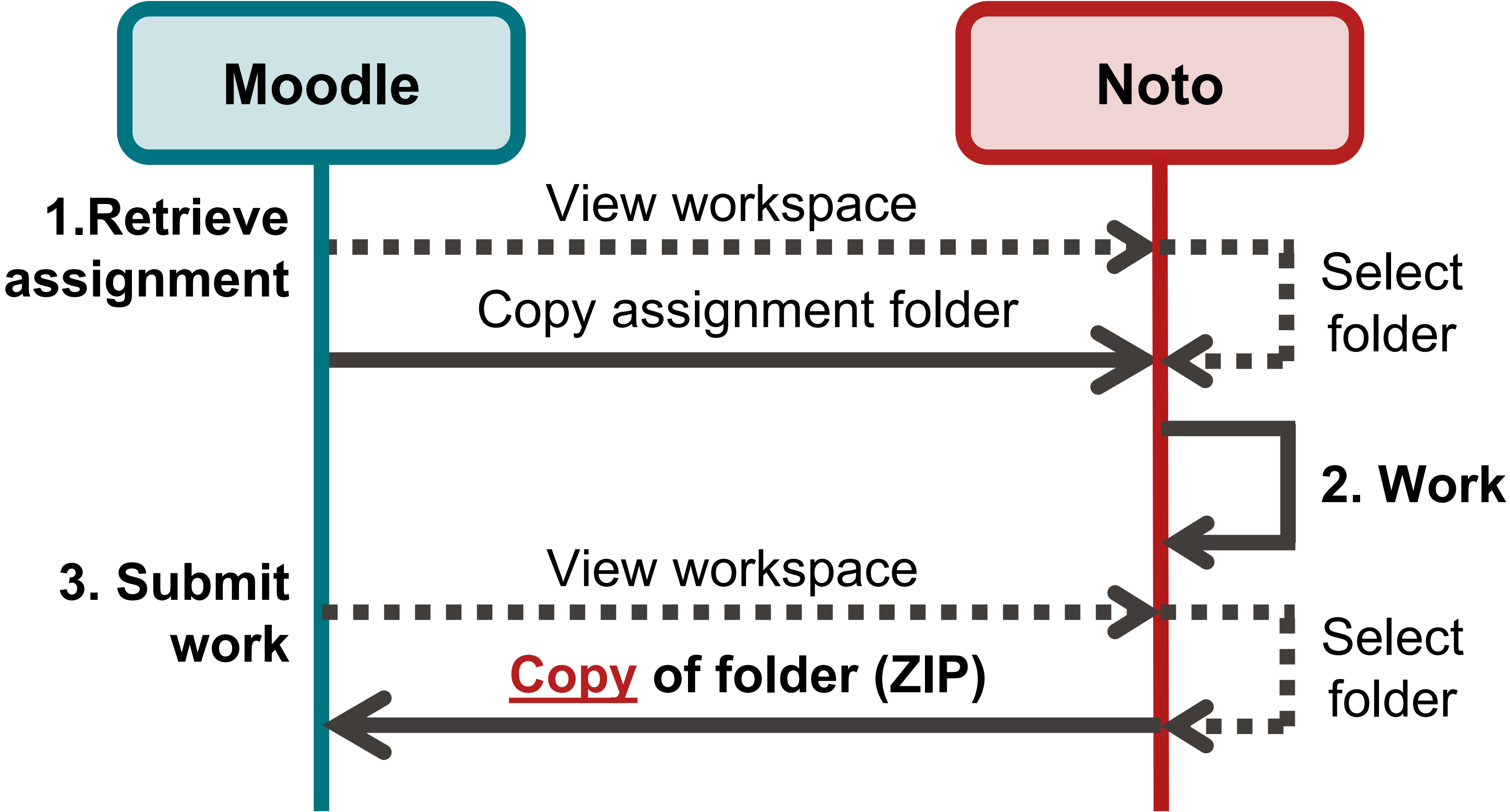
■ Support session with assistants on Tuesday, October 14, 10h-12h

👉 They do not have the solution!

They can help you debug or help with tech issues

They can help you submit your work on moodle

Assignment: process on moodle & noto



Submission process to test!

1. Get the assignment folder

Go to moodle:

- Find the assignment “Test the submission process” (Fairness 1 section)
- Retrieve the assignment folder -> to your noto workspace

2. Work on the assignment

Go to noto:

- Modify the notebook

3. Submit your work

Go to moodle:

- Add a submission
- Select the folder test-assignment from noto -> to moodle

Important advice – Submitting your work

- The content available on moodle is the **reference for grading**
- After submission, **changes on noto** are NOT reflected on moodle **unless you resubmit**
- **Do NOT wait until the last moment to submit!!!!**
 - You can re-submit as many times as you want until the deadline
 - Each re-submission will erase and replace the previous

Important advice – Working in the notebook

■ For your notebook to be correctly graded:

- Do not **delete** any cell!
- Remove all **calls to print()** from your code/functions (e.g., for debug)
- Regularly **restart** the kernel and run all cells
- Cite your **numerical results** from your code in your answers to open questions (we don't have code outputs when we grade open questions)

■ To obtain the best performance on noto:

- Do not open multiple notebooks in browser tabs!
- Open maximum 3 notebooks at the same time in noto (in panels)
- Regularly kill the kernels you do not use

Actions on the kernel

The screenshot displays the JupyterLab interface. The 'Kernel' menu is open, showing options such as 'Interrupt Kernel', 'Restart Kernel...', 'Restart Kernel and Clear Outputs of All Cells...', 'Restart Kernel and Run up to Selected Cell...', 'Restart Kernel and Run All Cells...', 'Restart Kernel and Debug...', 'Reconnect to Kernel', 'Shut Down Kernel', 'Shut Down All Kernels...', and 'Change Kernel...'. A red arrow points to the 'Restart Kernel...' option. A red callout box contains the text: 'Restart the kernel regularly to reset the execution state!'. The background shows a notebook with Python code and its output.

```
[2]: print(scenario)
print("The type of the scenario variable is", type(scenario))
```

A self-driving car with brake failure is heading towards five pedestrians crossing the street. The car can swerve to the other lane, hitting one pedestrian instead. What should the autopilot do?
The type of the scenario variable is <class 'str'>

Saving your work

The image shows a JupyterLab notebook interface. A red callout box with the text "Ctrl + S" points to the save icon (a floppy disk) in the notebook's toolbar. The notebook contains two code cells. The first cell, labeled [2]:, contains the following Python code:

```
scenario = "A self-driving car with brake failure is heading towards five ped |
```

The second cell, labeled [3]:, contains the following Python code:

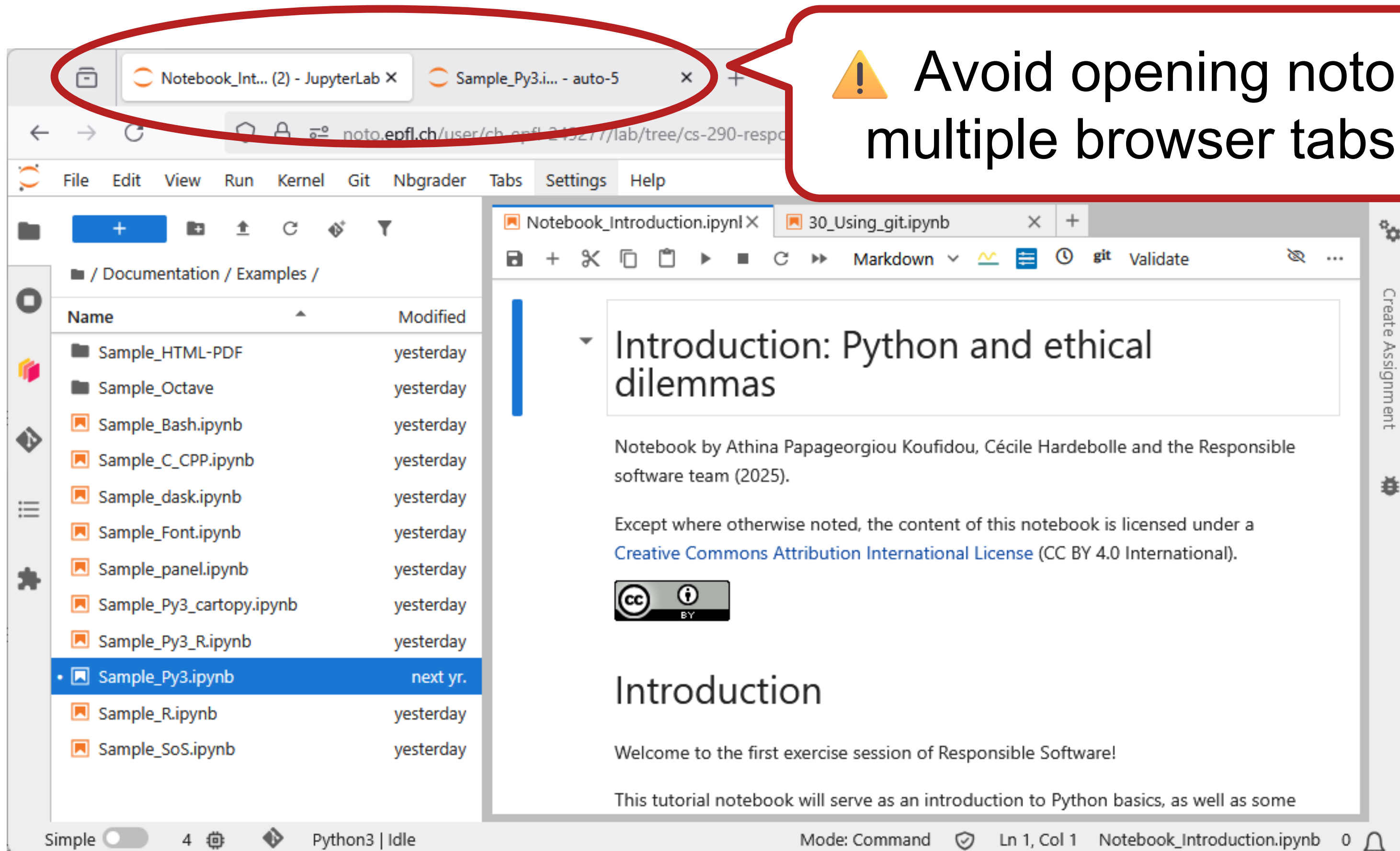
```
print(scenario)
print("The type of the scenario variable is", type(scenario))
```

The output of the second cell is:

```
A self-driving car with brake failure is heading towards five pedestrians cro
ssing the street. The car can swerve to the other lane, hitting one pedestria
n instead. What should the autopilot do?
The type of the scenario variable is <class 'str'>
```

The interface also shows a file browser on the left with a file named "Notebook_Introduction.ipynb" selected. The status bar at the bottom indicates "Mode: Command" and "Ln 1, Col 1".

Opening multiple notebooks



The screenshot shows the JupyterLab web interface. At the top, the browser's tab bar contains two tabs: "Notebook_Int... (2) - JupyterLab" and "Sample_Py3.i... - auto-5". A red oval highlights these tabs. A red callout bubble with a warning icon and the text "Avoid opening noto in multiple browser tabs!" points to the tabs. The main interface shows a file browser on the left with a list of files and folders, including "Sample_Py3.ipynb" which is highlighted. The main content area displays the content of "Notebook_Introduction.ipynb", which includes a title "Introduction: Python and ethical dilemmas" and a Creative Commons Attribution International License (CC BY 4.0 International) logo. The bottom status bar shows "Mode: Command" and "Ln 1, Col 1".

⚠️ Avoid opening noto in multiple browser tabs!

Name	Modified
Sample_HTML-PDF	yesterday
Sample_Octave	yesterday
Sample_Bash.ipynb	yesterday
Sample_C_CPP.ipynb	yesterday
Sample_dask.ipynb	yesterday
Sample_Font.ipynb	yesterday
Sample_panel.ipynb	yesterday
Sample_Py3_cartopy.ipynb	yesterday
Sample_Py3_R.ipynb	yesterday
• Sample_Py3.ipynb	next yr.
Sample_R.ipynb	yesterday
Sample_SoS.ipynb	yesterday

Introduction: Python and ethical dilemmas

Notebook by Athina Papageorgiou Koufidou, Cécile Hardebolle and the Responsible software team (2025).

Except where otherwise noted, the content of this notebook is licensed under a [Creative Commons Attribution International License \(CC BY 4.0 International\)](#).

CC BY

Introduction

Welcome to the first exercise session of Responsible Software!

This tutorial notebook will serve as an introduction to Python basics, as well as some

Opening multiple notebooks

The screenshot shows the JupyterLab interface. The top bar displays three open notebooks: Notebook_Introduction.ipyn, 30_Using_git.ipynb, and Sample_Py3.ipynb. A red circle highlights this top bar. A callout box with a red border points to the text "Instead, open the notebooks inside noto". The main content area shows the "Python3 Notebook" with the following text:

First example

This first example will plot a simple

$$e^{-x} = \sum_{i=0}^{\infty} \frac{(-1)^i x^i}{i!}$$

Well... not really. This was to demo *L^AT_EX* equations, courtesy of `MathJax`.

The equation we'll plot is the following:

$$y = \sin(x) + \cos(x)$$

```
[ ]: import numpy
      from matplotlib import pyplot
```

At the bottom, the status bar shows "Mode: Command", "Ln 1, Col 1", and "Sample_Py3.ipynb 0".

Opening multiple notebooks

The screenshot shows the JupyterLab web interface. At the top, there's a browser window with the URL `noto.epfl.ch/user/ch-epfl-249277/lab/tree/Documentation/Examples/Samp`. Below the browser is the JupyterLab menu bar with options: File, Edit, View, Run, Kernel, Git, Nbgrader, Tabs, Settings, Help. The main workspace contains two notebook panels side-by-side. The left panel is titled 'Notebook_Introduction.ipynl' and shows the start of an 'Introduction: Python and ...' notebook. The right panel is titled 'Sample_Py3.ipynb' and shows a 'Python3 Notebook' with a 'First example' section. The status bar at the bottom indicates 'Simple' mode, '3' tabs, 'main' environment, 'Python3 | Idle', 'Mode: Command', 'Ln 1, Col 1', and the active file 'Sample_Py3.ipynb'.

You can collapse the side bar

You can arrange panels side by side! (or any other config you want)

Where is the kernel?

The image shows a screenshot of the JupyterLab web interface. The browser address bar shows the URL `noto.epfl.ch/user/ch-epfl-249277/lab/tree/git_Noto/responsible-soft`. The JupyterLab menu bar includes File, Edit, View, Run, Kernel, Git, Nbgrader, Tabs, Settings, and Help. On the left sidebar, the 'Kernels' section is circled in red, and a red arrow points to the 'Notebook_Introduction.ipynb (d0...)' entry under the 'Python3' language server. A red speech bubble with the text 'List of active kernels' points to this section. The main notebook area shows a code cell with the following content:

```
[1]: scenario = "A self-driving car with brake failure is heading towards five ped
```

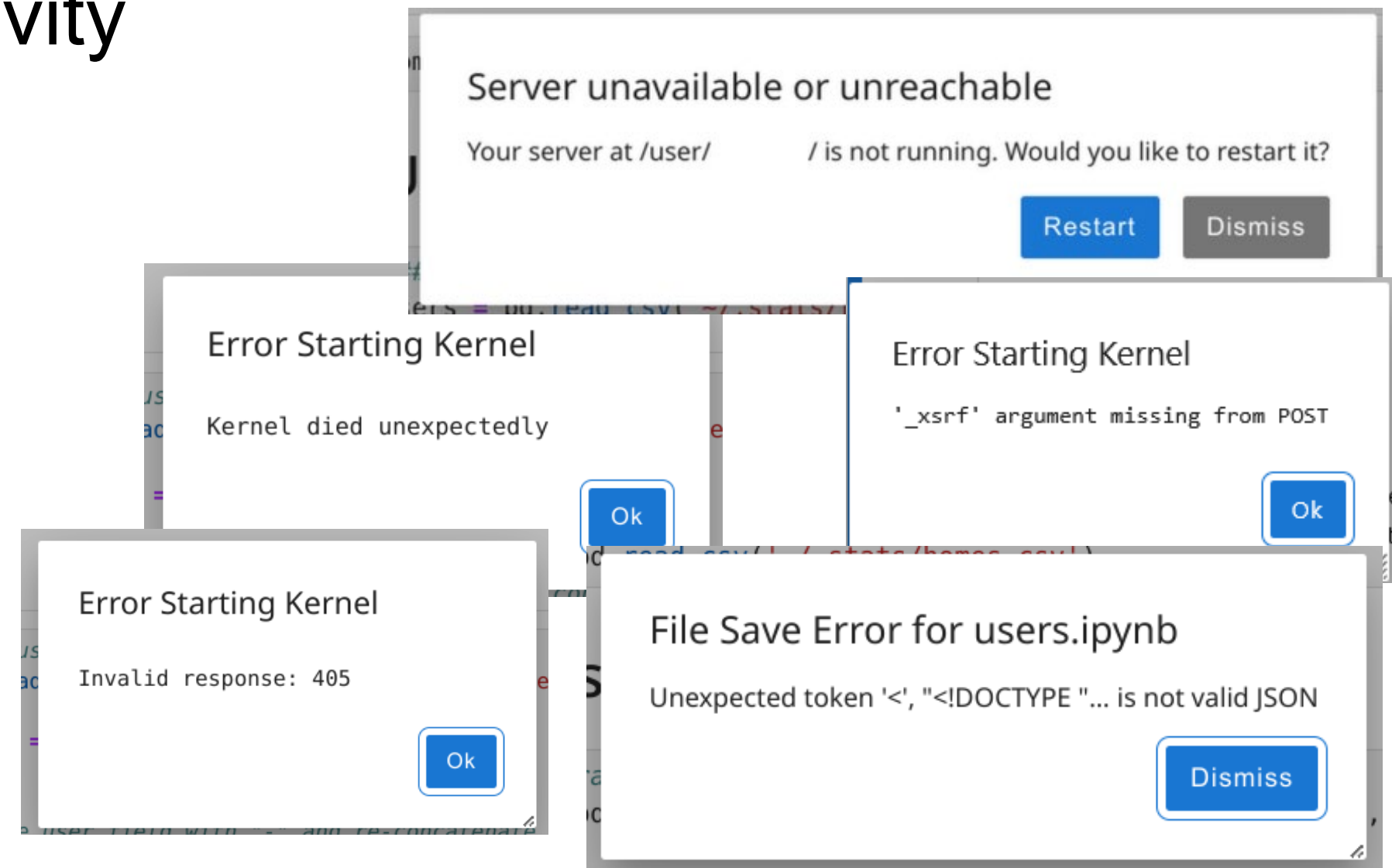
Below the code cell, there is an 'Instructions' box with the text: 'Execute the cell below to see the output.' Below the instructions, there is another code cell with the following content:

```
print(scenario)
print("The type of the scenario variable is", type(scenario))
```

The output of the second code cell is: 'A self-driving car with brake failure is heading towards five pedestrians crossing the street. The car can swerve to the other lane, hitting one pedestrian instead. What should the autopilot do? The type of the scenario variable is <class 'str'>'

Error messages

- If you start getting **popups with error messages**, it means that
 - ⚠ your personal server is no longer available ⚠
 - Timeout after \approx 30 minutes of inactivity
 - Too many resources used
- Do not simply dismiss or click ok!
The interface looks like it works but it does NOT!
- **Log out** and then log in again:
 - File > Logout
 - <https://noto.epfl.ch/hub/logout>



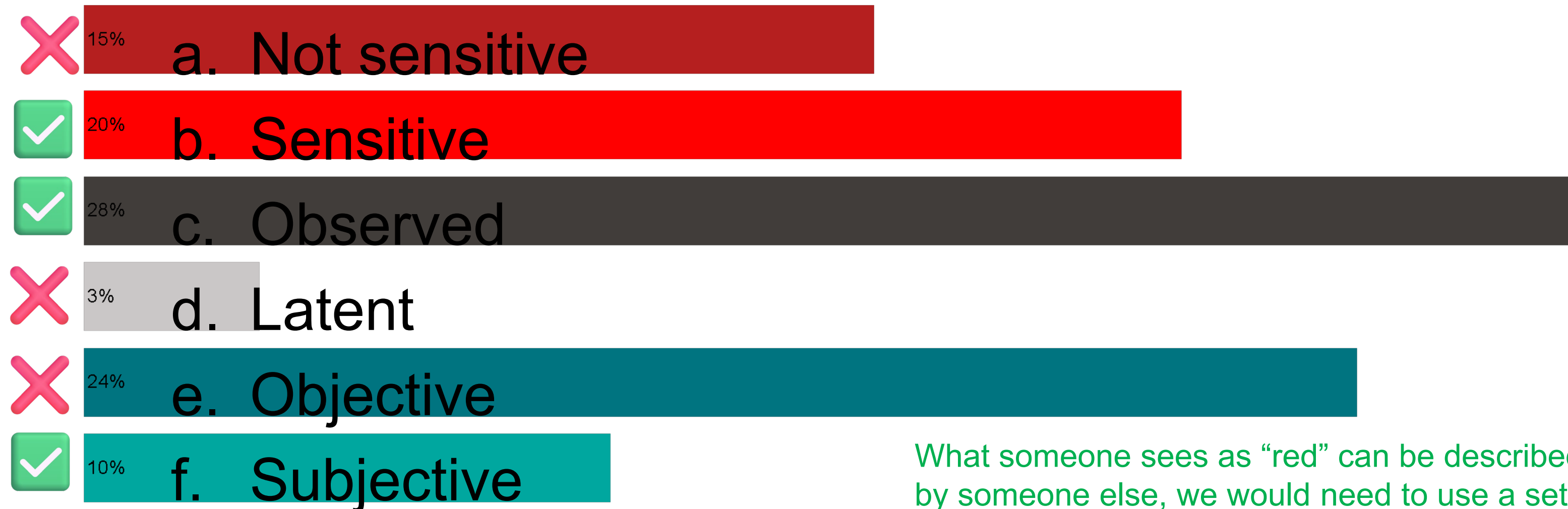
Review questions

Fairness 1

Attributes - 1

URL: ttpoll.eu
Session ID: cs290

What are the characteristics of hair color as an attribute to represent people? (select all that apply)



What someone sees as “red” can be described as “auburn” by someone else, we would need to use a set of predefined categories, and some cases would be difficult to fit in (there is the same issue with skin color by the way)
-> the only way to make it “objective” would be to measure the color with colorimetry

Attributes - 2

URL: ttpoll.eu
Session ID: cs290

Let's imagine a software that relies on SAT scores (standardized test for university admission in the US) to make recommendations of when to approve study loans.

What are the characteristics of the SAT score?

- 9% a. Not sensitive
- 27% b. Sensitive
- 28% c. Private
- 7% d. Public
- 21% e. Proxy
- 7% f. System

Because it is a proxy for sensitive variables, SAT score can lead to discrimination i.e. can be considered sensitive.

Article on SAT scores' correlation with race:
<https://www.brookings.edu/articles/race-gaps-in-sat-scores-highlight-inequality-and-hinder-upward-mobility/>

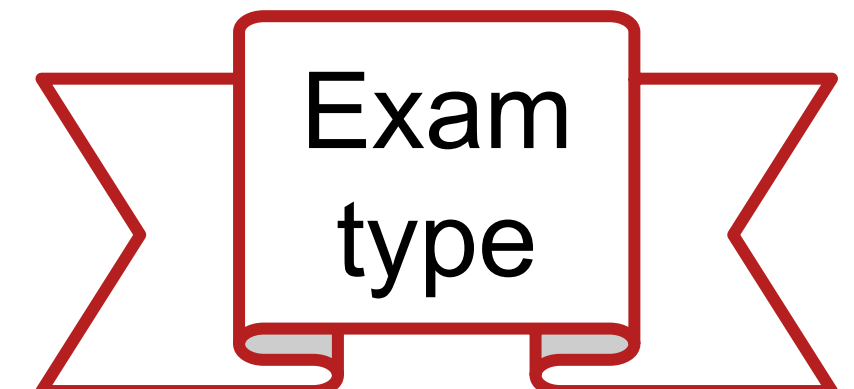
Bias - 1

URL: ttpoll.eu
Session ID: cs290

The city of Lozhann decides to deploy a smartphone app that allows residents to report potholes throughout the city to help with the identification of repair needs.

What bias will the data collected by the app probably exhibit?
(select one answer)

- 0% a. Confirmation bias
- 53% b. Representation bias
- 47% c. Measurement bias
- 0% d. Automation bias



Bias - 2

URL: ttpoll.eu
Session ID: cs290

In the new ArcFit fitness tracker, the calory burn feature uses the "metabolic equivalent of task" formula, which estimates the energy a body uses during a specific activity. The same calculation is used during walking and running.

What type of bias will the calory burn variable probably have?
(select one answer)

- 0% a. Confirmation bias
- 2% b. Representation bias
- 98% c. Measurement bias
- 0% d. Automation bias

Exam
type

Bias - 3

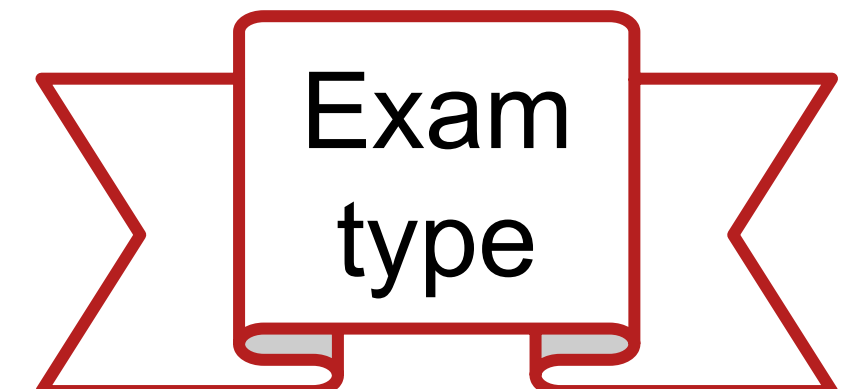
URL: ttpoll.eu
Session ID: cs290

A group of computer scientists with similar background, all experts in software development, are starting a new software project in the healthcare domain.

The question has been reframed compared to the in-class session to make clearer:
- they are not using AI to develop (was unclear)
- we do not consider here biases from the data but biases from humans

What type of bias will these scientists probably have?
(select one answer)

- 44% a. Confirmation bias
- 17% b. Automation bias
- 24% c. Pre-existing bias
- 15% d. Sunk cost fallacy



Case studies

Where to find the cases?

1. Go to **courseware**
2. Find **the case studies** for today: **Fairness 1**
3. Download:
 - The **instruction sheet**
 - 2 cheatsheets: Value Analysis, Inclusive Design
 - 1 template

Value Analysis

Part 1 – Your own values

1. **Fill the “own values questionnaire”** (template / appendix 1.1)
2. Discuss with your neighbor: what are your conclusions about the nature of values?

Part 2 – Values manifested in products

Individually, fill out the Artifact Values Questionnaire for Google Maps (template / appendix 1.2):

- Look at the advertisement website
<https://google.com/intl/en/maps/about>
- Explore the different features of Google Maps in a web browser (use a private window).

1. Indicate which values you identify
2. Indicate how they manifest

Values in Google Maps

URL: ttpoll.eu

Session ID: cs290

Select the values you have identified among these:

- 12% a. Power-Resources (POR)
- 11% b. Power-Dominance (POD)
- 11% c. Achievement (AC)
- 26% d. Stimulation (ST)
- 9% e. Universalism-Concern (UNC)
- 5% f. Universalism-Nature (UNN)
- 7% g. Security-Societal (SES)
- 12% h. Security-Personal (SEP)
- 7% i. Face (FAC)

Part 3a – Stakeholder values

1. Read the stakeholder description
2. For each **stakeholder**:
identify **2 values** that you can relate to Google Maps
3. Specify whether they are **benefited or harmed** by Google Maps and how in the table:

Stakeholder	Key values	Manifested	Benefits	Harms

Post your value harms/benefits

Which **value harms/benefits** did you identify?

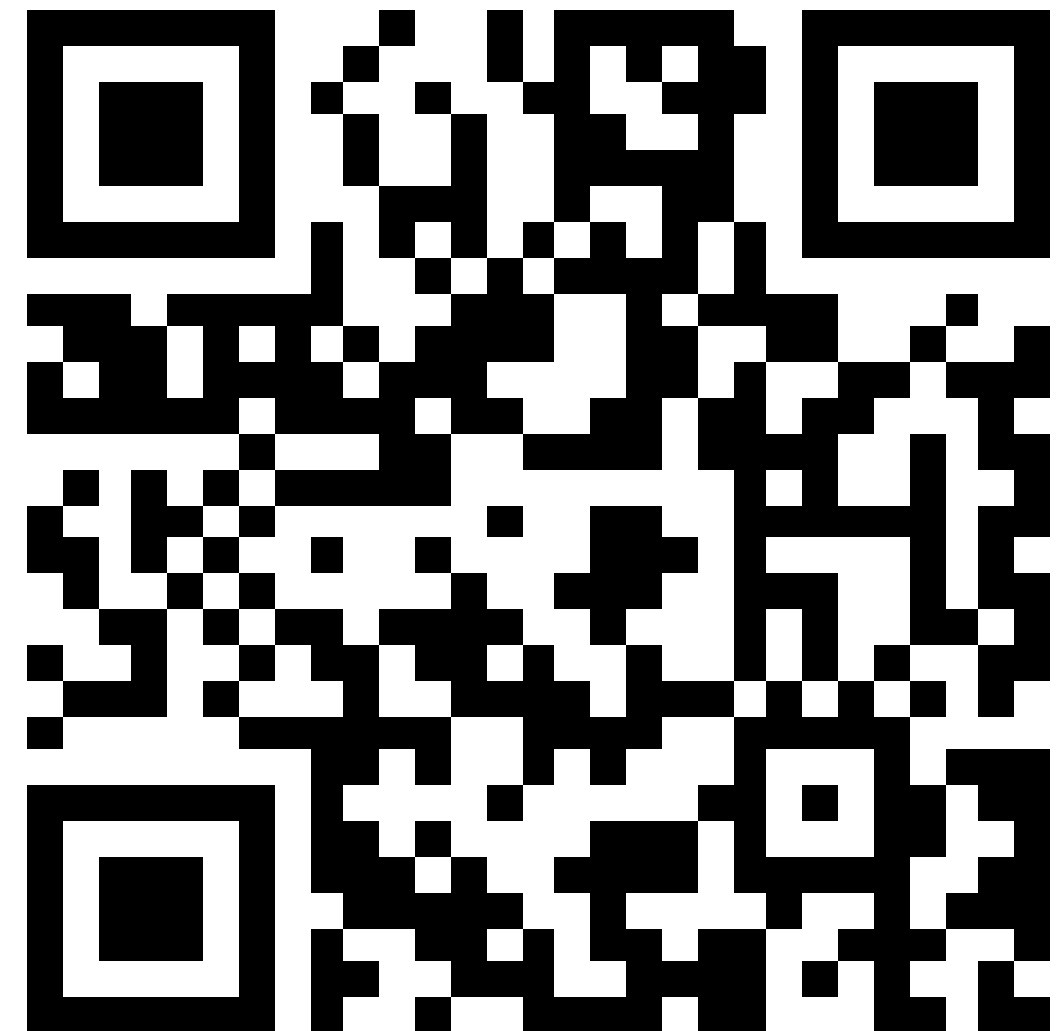
- 👉 1 post / value harm/benefit
 - Name of the **stakeholder**
 - How the value is **harmed/benefited**
(short description)

⚠️ If your value harm/benefit has already been posted, add a vote 👍

Post your ideas:

<https://speakup.epfl.ch>

Room key: **46824**



Part 3b – Value tensions

Draw the value tension map:

1. Place the **values**
2. Add the **stakeholders** concerned by each value and indicate if it's a harm (-) or benefit (+)
3. Add lines to indicate the **value tensions** i.e., harm vs. benefit

They can be:

- ◆ Between different stakeholders or for the same stakeholder
- ◆ Between values or for the same value

Overall debriefing of the strategy

- **Products are not neutral**: they embed and manifest values (of their designers, marketers...)
- **Stakeholders have values**, which may be harmed or supported by products
 - 👉 Value tensions help identify how the design of a product may negatively affect some stakeholders
- In a real project, there are many ways you can do value analysis, including interviews with stakeholders

Inclusive Design

We will go back to this case next week.

What's next?

Indicative evaluation

- The feedback system is open until **midnight Sunday 12 October**
 - 👉 log on ISA
 - Single question: “The running of the course enables my learning and an appropriate class climate”
 - Field for comments
- **Your feedback is important to improve the course!**
 - It is collected early so we have a chance to adjust what can be adjusted this year
 - Your comments are appreciated!

We start Fairness 2!

Tomorrow, Tuesday 7: notebook on **fairness in predicting recidivism** (COMPAS case)

By Monday 13:

- Watch **videos 4.1 to 4.5** + do the **quizzes**
- Finish the notebook
(and any other leftover from previous weeks)

On Monday 13:

- Interactive questions on the theory
- Work on the **case studies together in class**