# Class X: Liability & Personal Assistants | Lesson Plan

## 

## Class Breakdown

| **#** | **Step** | **Mode** | **Duration** |
| --- | --- | --- | --- |
| 1 | Input: Introduction to Legal Liability and AIs Challenges | Plenary | 10 |
| 2 | Presenting the case-study | Plenary | 10 |
| 3 | Activity 1: Preparing Client’s Arguments | Teams | 30 |
|  | Break | | 15 |
| 4 | Input: Introduction to Ethics and AI | Plenary | 15 |
| 5 | Additional Discoveries | Plenary | 5 |
| 6 | Activity 2: Updating Arguments | Teams | 40 |
|  | Break | | 15 |
| 7 | Team Presentations and Discussions | Plenary | 30 |
| 8 | Wrap-Up & Q&A | Plenary | 10 |

## 

## Core Legal Issues

The core legal issues on which we would like our students to focus on relate to the challenges, gaps, and vulnerabilities that AIs pose through their design, nature, use, and implementation.

The issues are sometimes cross-domained and could manifest in one or more sectors/fields of application. Many of the issues are common to all technology (e.g., privacy/data protection); many could be interrelated (e.g., transparency, fairness, accountability). However, the ability of AI to amplify and/or facilitate their adverse effects must not be underestimated at any time.

Some of the main legal aspects that students will be required to reason on are:

* Legal personhood issues
* Intellectual property issues
* Liability for damages
* Lack of accountability for harms

## Core Technical Issues

The core technical issue in question is whether or not AI assistants (and other AIs that give humans instructions or interact with them via speech, text, conversation, orders, etc…) can, after producing the expected response to the query at hand, determine whether or not that response could lead to a violation of certain pieces of legislation.

We can think of an example query where a question made to an AI like Siri, whose response could lead to a user killing themselves or harming others. Should Siri give that response? Or should it withhold the information since it will clearly lead to what could be a criminal outcome?

Furthermore, we want to inquire about how can we train a model to properly determine this fact? What are the limiting factors to its development and usage? Who takes responsibility if/when this layer of protection fails?

## Case Study Excerpt

Following are details about a road accident that occurred in the State of Israel. The case resulted in a civil lawsuit involving three parties: John, Jamie, and “Stark Inc”.

John owns a brand new "Stark" car. This top-of-the-line model had several high-tech features, including "Jessica", a smart virtual assistant made by "Stark Industries".

On the 19th of May 2019, John was driving his car in the suburbs when he accidentally hit Jamie, who was jogging across the road. When John saw Jamie lying on the ground, he panicked, and, remembering the fact that he had already been involved in a speeding accident 3 years ago, fled. He drove towards the forest, leaving Jamie unconscious, and bleeding, on the ground.

(. . .)

## Internal Documentation

Final Presentation - <https://bit.ly/3989IOd>

Lesson Plan - <https://bit.ly/3ynndUy>

Inputs Overview - <https://bit.ly/3FrHQAs>

Pre-Class Handout - <https://bit.ly/39CfdVq>

Case Study - <https://bit.ly/3KRaiwU>

Activity #1 Handout - <https://bit.ly/3M0RpsP>

Use of Personal Assistance AIs Evidence Summary - <https://bit.ly/3FukVoi>

Activity #2 Handout - <https://bit.ly/3L4XcMM>

Activity #2 Notebook - <https://bit.ly/3smdTN3>

AI Moral Compass Evidence Summary - <https://bit.ly/3yjsxZg>

DataSet.csv - <https://bit.ly/396C0s7>

## Outside References

Document Style Taken From - <https://bit.ly/3vZVI1G>

PowerPoint Style Taken From - <https://bit.ly/3kTcxFt>

Mock Trial Structure Inspired By - <https://bit.ly/3vZVI1G>

Heavy Influence for “Moral AI Filter” - <https://bit.ly/3PbMMOD>

Dataset Manually Collected From - <https://bit.ly/3PbMMOD>

Google Drive Data Load Method From - <https://bit.ly/3wk5BWW>

Personal Assistant AI Harm Incident #1 - <https://bbc.in/3L2QMNU>

Personal Assistant AI Harm Incident #2 - <https://bit.ly/3L0Q97D>

Personal Assistant AI Stats From - <https://bit.ly/38axmcN>

Final Project Considerations From:

* <https://bit.ly/3KXy1vu>
* <https://bit.ly/3Pa60nv>

Legal Aspects and AIs From:

* <https://bit.ly/3L1U9F1>
* <https://bit.ly/3KUeYCc>
* <https://bit.ly/3vXzPAk>
* <https://bit.ly/3wjKqEI>

## Special Thank Yous

The entirety of Team #9 would like to extend a special thank you to all the Course Staff, but specifically to Avi, Lillà, and Sivan, who, at several points during the development of this project, took the time to discuss the different topics with us, giving great suggestions and useful feedback, without which we could not have completed the assignment.