

# Ein Lehrplan zur Ethik der Künstlichen Intelligenz



## Teil 1: Einführung Künstliche Intelligenz (KI)

- An was denken wir beim Wort «Künstliche Intelligenz»? Und wo begegnen wir KI im Alltag?
- Die drei Teile einer KI: Datensatz, Lernalgorithmus, Vorhersage
- Datensatz: Bilder, Messungen, Text, Sprache, Musik, Video
- Die Teile eines Algorithmus: Eingabe, Schritte zum Ändern der Eingabe, Ausgabe
- KI-Experimente (Beispiele Internet): Was passiert? Was ist die Funktion von KI bei den Beispielen? Was ist die Vorhersage? Welcher Datensatz wird verwendet?

## Teil 2: Algorithmen & Meinungen

- Die Teile eines Algorithmus: Eingabe, Schritte zum Ändern der Eingabe, Ausgabe
- Ein Algorithmus um Algorithmus um das beste Sandwich zu machen. Was ist die Bedeutung von «beste»?
- Die Optimierung von Resultaten: Wer entscheidet wie ein Algorithmus optimiert werden soll?

## Teil 3: Einführung Ethik Matrix

- Welches sind die Ziele (Optimierung) des Sandwich-Algorithmus?
- Welche Stakeholder (Definition) sind am Resultat des Sandwich-Algorithmus interessiert?
- Die Ethik Matrix (Werte, Stakeholders und Ziele)
- Programmieren mit scratch: Vergleich Algorithmus / KI (Optional)

## Teil 4a: Einführung Supervised Machine Learning

- Wie lernen wir? Wie lernt eine KI?
- Klassifikation und Regression
- Neuronale Netzwerke
- Google's Teachable Machine I: KI zum Klassifizieren von Hunden und Katzen (KI-System, Klassen, Datensatz)
- Google's Teachable Machine III: KI zum Sortieren von Gemüse (Anwendung)

## Teil 4b: Algorithmische Voreingenommenheit

- Google's Teachable Machine II: die Rolle der Trainingsdaten in einem KI-System
- Algorithmische Voreingenommenheit (Definition)
- Beispiel Gender Shades: Die Folge von «schlechten» Trainingsdaten

## Teil 5: Spekulative Fiktion

- Die Folgen neuer Technologien für die Menschheit: Wie helfen sie uns? Wie können sie Schaden anrichten?
- Beispiel AlterEgo
- Spekulative Fiktion (Beispiele Internet): Technologie, Stakeholders, Möglicher Einsatz der Technologie (Hilfe, Schaden)

## Teil 6: Redesigning Youtube

- Youtube neu gestalten: Auf welche Art und Weise verwendet Youtube Künstliche Intelligenz? Stakeholder, Ethik Matrix, Ziel der Algorithmen
- Mock-up (Vorführmodell ) der Benutzeroberfläche (user interface) und der Benutzererfahrung (user's experience) für Vorgeschlagene Videos (recommendations), Kommentar-Teil (comments section), Resultate der Suche (search results)